

On the turning away

Alexandre de Brevern

▶ **To cite this version:**

Alexandre de Brevern. On the turning away. Joint 25th Annual International Conference on Intelligent Systems for Molecular Biology (ISMB) and 16th European Conference on Computational Biology (ECCB) 2017, Aug 2017, Prague, Czech Republic. , 2017. inserm-01575565

HAL Id: inserm-01575565

https://inserm.hal.science/inserm-01575565

Submitted on 21 Aug 2017

HAL is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L'archive ouverte pluridisciplinaire **HAL**, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.

On the turning away

Alexandre G. de Brevern[1]

[1] INSERM, U 1134, DSIMB, Univ Paris Diderot, Univ. Sorbonne Paris Cité, INTS, Laboratoire d'Excellence GR-Ex, F-75739 Paris, France.

Abstract

β-turn IV, *i.e.* the miscellaneous category, represents near 1/3rd of β-turn residues in protein structure, and is the second most frequent β -turn. An innovative clustering approach was able to underline the existence of different new turns not previously described. The four most occurring clusters defined the new β-turn types. They exhibit interesting sequence – structure relationships.

Introduction

The classical secondary structures are composed of α-helices and β-strands connected by coil. Two other repetitive structures also exist, namely the PolyProline II and the β-turns. These last have been characterized by a hydrogen bond between N-H and C=O of residues *i* and *i*+3 by Venkatachalam[1]. He also characterized the first β-turn types. Later novel turns were defined, some being discarded, leading to a final collection of type I, I', II, II', IV, VIa1, VIa2, VIb, and VIII β-turns. Turns that do not fit any of the above criteria are classified as type IV [2]. β-turn IV, *i.e.* the miscellaneous category, represents near 1/3rd of β-turn residues in protein structure, and is the second most frequent β -turn. An automatic clustering approach based on the rules of β-turn type assignment was designed to search for recurrent new turns inside this miscellaneous type. A comparison with related studies and amino acid sequence over- and underrepresentation was performed underlying interesting features[3].

Material & Methods

Different non-redundant dataset of protein structures were taken from low to higher redundancy to analyse potential bias of the dataset (none was found). From these datasets, were taken only the type IV β-turns encoded as their series of central dihedral angles.

A specific clustering approach was designed to cluster type IV β-turns by using the classical rule, allowing +/- 30° for all angles, with the exception of one at +/-45° for the defined values. The clustering derived from Self-Organizing Maps (SOM, without diffusion between the clusters)[4]. The training was carried out in 2 successive parts; the first one limited the potential bias of initialization,

and the second refined the clustering by using the specific rules for β-turn types. The type IV β-turns were selected from a dataset D. Thus, each dataset was associated with T type IV β-turns.

Results & Conclusions

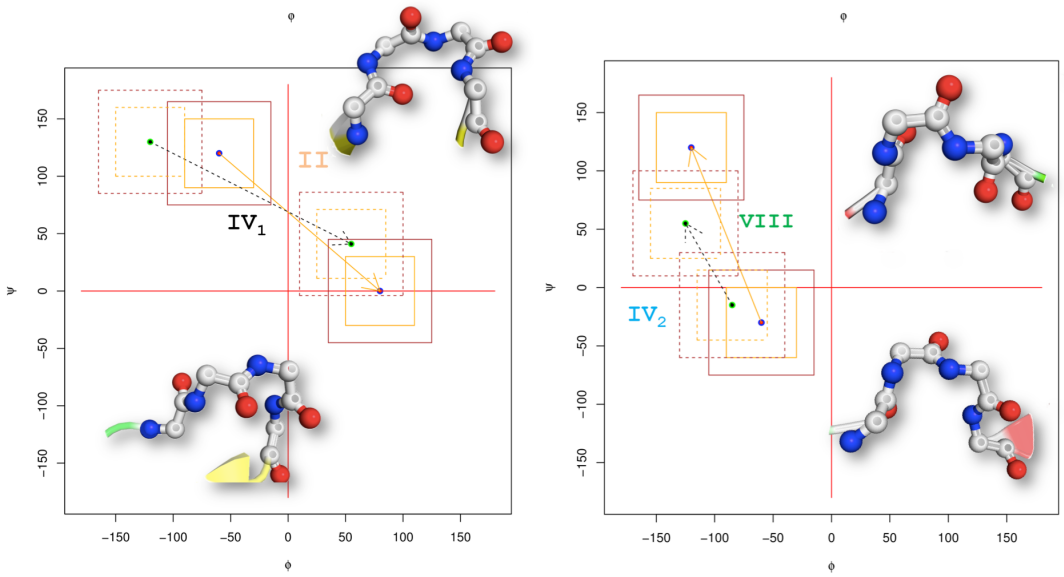


FIGURE 1.
Ramachandran plot of β-turns. (Left) a close-up of type II and IV_1 β -turns, and (right) on type VIII and IV_2 β-turns, the first square corresponds to the +/– 30° rule, and the second one to the +/– 45° rule.

Surprisingly, with 10 different datasets, the unsupervised training was highly robust, producing always the same four major clusters (with negligible variation). These types, named IV_1, IV_2, IV_3 and IV_4, represent half of the type IV β-turns, and are more frequent that many established ones. Figure 1 shows a direct comparison of the two most frequent new turns with their closest relatives underlying their relationship but also their differences. Type IV_1, is in the neighbourhood of type II but with very different amino acid composition, while IV_2 is close to type VIII with related amino acid content. Types IV_3 and IV_4 are in the same dihedral angle region than frequent β-turn type I, but with distinct dihedral angle values.

Comparisons with the previous alternative classification proposed by Efimov [5] and Thornton's group[6] emphasized the uniqueness of the approach. Notably, the most frequent new turn (type IV_1 β-turn) was not highlighted, although it is the 5th most occurring turn (including type IV_{misc} β-turns (including rest of type IV β-turns). Only the type IV_2 β-turns were previously included.

References

1 Venkatachalam CM. Biopolymers 6:1425-1436 (1968).
2 Hutchinson E.G. & Thornton J.M. Protein Sci 5:212-220 (1996).
3 de Brevern A.G. Sci Rep. 6:33191 (2016).
4. Kohonen, T. Biol. Cybern **43**, 59-69 (1982).
5. Efimov, A. V. Prog Biophys Mol Biol 60, 201-239 (1993).
6. Wilmot, C. M. & Thornton, J. M. Protein Eng 3, 479-493 (1990).