Proceedings of the ***10th International Conference of the Hellenic Crystallographic Association (HeCrA),***
NCSR "Demokritos", Athens, Greece, 15-17 October 2021

Buried β–turns in Hydrophobic Cores: Structural and Functional Implications

Aikaterini – Alexandra Giannaki[1*], Nicholas M. Glykos[1]
[1]Department of Molecular Biology and Genetics, Democritus University of Thrace, Alexandroupolis, 68100, Greece
*Correspondence to: katherineynnk@gmail.com

***Abstract*— A turn is a protein secondary structure where the polypeptide chain reverses its direction. In particular, a β–turn consists of only four consecutive residues and a γ–turn, which is the second most commonly found turn after β–turns, involves three consecutive residues. Turns are usually located at the molecular surface. However, they can sporadically be found buried within the protein hydrophobic core and the precise role of such occurrences remains unknown. In this study, we identify such buried turns and we examine putative structural and functional implications of their presence. In order to search for these patterns, we developed a program which incorporates and systematically applies PROMOTIF's and Stride's algorithms and identifies buried turns in a large sample of structures obtained from the Protein Data Bank via the PISCES interface. Preliminary results obtained from this procedure are presented.**

I. INTRODUCTION

A β–turn is a secondary structure motif which involves four consecutive residues, R_i, R_{i+1}, R_{i+2}, R_{i+3}, where the distance between the atoms Ca_i and Ca_{i+3} is less than 7 Å and the two central residues are not in a helical conformation[1]. β–turns are the most commonly found turns and are classified into nine different types (Type I, Type I', Type II, Type II', Type IV, Type VIa1, Type VIa2, Type VIb, Type VIII) using the φ and ψ angles of residues R_i and R_{i+2}.[2] The second most commonly found turns are γ–turns. A γ–turn consists of three consecutive residues, R_i, R_{i+1}, R_{i+2}, with a hydrogen bond forming between the backbone CO_i and NH_{i+2}.[3] γ–turns are divided into two different types, classic and inverse. The main chain atoms of the two forms can be found in two possible enantiomers[4].

Turns are more frequently found on the solvent–exposed surface of the proteins, however, there have been reports where turns have been located buried within the hydrophobic core of the protein[5]. The precise structural and functional significance of such occurrences is currently unknown. Here we identify such structures in a large sample of structures obtained from the PDB, and we attempt to identify common structural and functional features shared between them.

II. METHODS

24951 protein entries were downloaded from the Protein Data Base Archive[6],[7],[8] with a list obtained from the protein culling server PISCES[9] with the following criteria: maximum 2.2 Å resolution, maximum R-factor 1.0 and 70% identity cut–off. The PROMOTIF[10] algorithm was then used in order to identify the β– and γ–turns found in the proteins. Moreover, the Stride[11] algorithm was incorporated for the purpose of calculating the solvent accessible area of the residues involved in the turns found by PROMOTIF[10]. The maximum solvent accessibility number of all residues involved in the found turns was calculated and Gnuplot was used for plotting the density distributions. Some of the turns found with a maximum solvent accessibility of zero were visualized using PyMOL[12]. For various computational scripts the Perl programming language was used.

III. RESULTS

A total of 427 β– and γ–turns with a maximum solvent accessibility number of the residues involved equal to 0.0 were found. These 427 turns belonged to a total of 402 proteins (out of the 24921 proteins in the initial sample). Four histograms with bin numbers 0.2, 0.4, 0.6 and 0.8 showing the distribution of maximum solvent accessibility in the found turns were plotted. All of the histograms plotted revealed a peak at 0.0, a drop at around 1.0, a plateau from 2.0 till 60.0 and then a greater peak at around 100.0.

TABLE I.

Turn Type per Turn			
β-turns		***γ-turns***	
I	39	CLASSIC	7
I'	4	INVERSE	163
II	21		
II'	8		
IV	144		
VIa1	1		
VIa2	0		
VIb	3		
VIII	37		

a. This table showcases the types of the buried turns we obtained from our sample

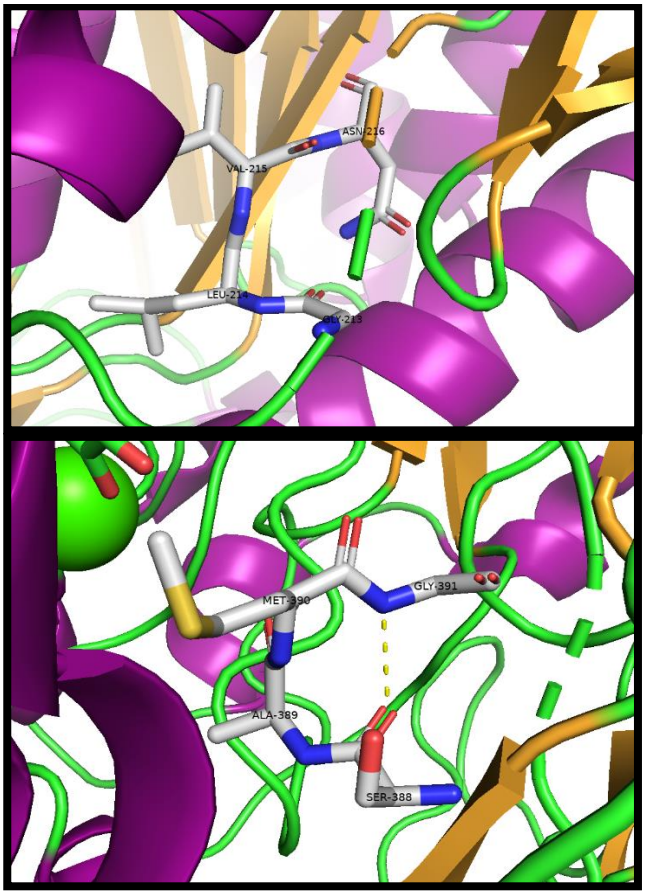

Figure 1. Buried β-turns, 1st: Betaine Aldehyde Dehydrogenase (1a4s), 2nd: Methanol Dehydrogenase (1w6s)

IV. CONCLUSIONS

We have identified a significant number of turns buried inside the hydrophobic cores of a non-redundant set of proteins. Although this is still a work in progress, we have established the presence of buried turns, we have examined -using molecular graphics- a large number of these turns, and we have identified the most prominent types of turns involved in such motifs.

V. FUTURE WORK

Our future intentions involve the construction of an RMSD- or TM-score-based[13] distance matrix with the aim of further reducing the structural redundancy that may be present in our data set. Furthermore, we will analyze our data in terms of the function of the proteins that we have identified with the aim of characterizing any putative functional significance of the buried turns. Additionally, we will analyze the structural context within which these turns are located with the aim of identifying any systematic trend in terms of their structural environment.

REFERENCES

[1] K. C. Chou and J. R. Blinn, "Classification and prediction of β-turn types," *J. Protein Chem.*, vol. 16, no. 6, pp. 575–595, 1997, doi: 10.1023/A:1026366706677.

[2] E. G. Hutchinson and J. M. Thornton, "A revised set of potentials for β-turn formation in proteins," *Protein Sci.*, vol. 3, no. 12, pp. 2207–2216, 1994, doi: 10.1002/pro.5560031206.

[3] C. Fang, Y. Shang, and D. Xu, "Improving Protein Gamma-Turn Prediction Using Inception Capsule Networks," *Sci. Rep.*, vol. 8, no. 1, Dec. 2018, doi: 10.1038/S41598-018-34114-2.

[4] E. J. Milner-White, "Situations of gamma-turns in proteins. Their relation to alpha-helices, beta-sheets and ligand binding sites," *J. Mol. Biol.*, vol. 216, no. 2, pp. 385–397, 1990, doi: 10.1016/S0022-2836(05)80329-8.

[5] G. D. Rose, W. B. Young, and L. M. Gierasch, "Interior turns in globular proteins," *Nature*, vol. 304, no. 5927, pp. 654–657, 1983, doi: 10.1038/304654a0.

[6] B. H, H. K, and N. H, "Announcing the worldwide Protein Data Bank," *Nat. Struct. Biol.*, vol. 10, no. 12, p. 980, Dec. 2003, doi: 10.1038/NSB1203-980.

[7] B. H, H. K, N. H, and M. JL, "The worldwide Protein Data Bank (wwPDB): ensuring a single, uniform archive of PDB data," *Nucleic Acids Res.*, vol. 35, no. Database issue, Jan. 2007, doi: 10.1093/NAR/GKL971.

[8] "Protein Data Bank: the single global archive for 3D macromolecular structure data," *Nucleic Acids Res.*, vol. 47, no. D1, pp. D520–D528, Jan. 2019, doi: 10.1093/NAR/GKY949.

[9] W. G and D. RL, "PISCES: a protein sequence culling server," *Bioinformatics*, vol. 19, no. 12, pp. 1589–1591, Aug. 2003, doi: 10.1093/BIOINFORMATICS/BTG224.

[10] E. G. Hutchinson and J. M. Thornton, "PROMOTIF--a program to identify and analyze structural motifs in proteins.," *Protein Sci.*, vol. 5, no. 2, p. 212, 1996, doi: 10.1002/PRO.5560050204.

[11] M. Heinig and D. Frishman, "STRIDE: A web server for secondary structure assignment from known atomic coordinates of proteins," *Nucleic Acids Res.*, vol. 32, no. WEB SERVER ISS., pp. 500–502, 2004, doi: 10.1093/nar/gkh429.

[12] "PyMOL | pymol.org." https://pymol.org/2/ (accessed Sep. 12, 2021).

[13] Z. Y and S. J, "Scoring function for automated assessment of protein structure template quality," *Proteins*, vol. 57, no. 4, pp. 702–710, Dec. 2004, doi: 10.1002/PROT.20264.