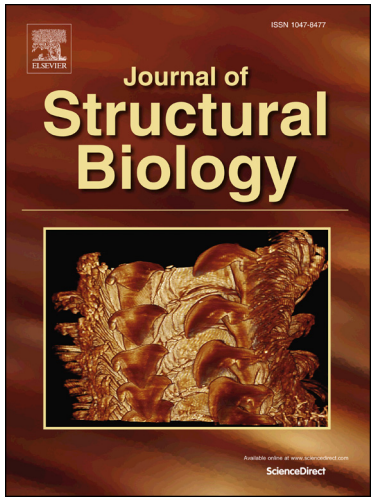

Detection of single alpha-helices in large protein sequence sets using hardware acceleration

Ákos Kovács, Dániel Dudola, László Nyitray, Gábor Tóth, Zoltán Nagy, Zoltán Gáspári

PII: S1047-8477(18)30143-6
DOI: https://doi.org/10.1016/j.jsb.2018.06.005
Reference: YJSBI 7206

To appear in: *Journal of Structural Biology*

Received Date: 30 December 2017
Revised Date: 11 June 2018
Accepted Date: 12 June 2018

Please cite this article as: Kovács, A., Dudola, D., Nyitray, L., Tóth, G., Nagy, Z., Gáspári, Z., Detection of single alpha-helices in large protein sequence sets using hardware acceleration, *Journal of Structural Biology* (2018), doi: https://doi.org/10.1016/j.jsb.2018.06.005

Detection of single alpha-helices in large protein sequence sets using hardware acceleration

Ákos Kovács[1,*], Dániel Dudola[1,*], László Nyitray[2], Gábor Tóth[3], Zoltán Nagy[1], Zoltán Gáspári[1]

[1]Faculty of Information Technology and Bionics, Pázmány Péter Catholic University, Budapest, Hungary
[2]Department of Biochemistry, Eötvös Loránd University, Budapest, Hungary
[3]Department for Research and Development, National Research, Development and Innovation Office, Budapest, Hungary

[*]These authors contributed equally to this work.

Correspondence to: nagy.zoltan@itk.ppke.hu, gaspari.zoltan@itk.ppke.hu

Abstract

Single alpha-helices (SAHs) are increasingly recognized as important structural and functional elements of proteins. Comprehensive identification of SAH segments in large protein datasets was largely hindered by the slow speed of the most restrictive prediction tool for their identification, FT_CHARGE on common hardware. We have previously implemented an FPGA-based version of this tool allowing fast analysis of a large number of sequences. Using this implementation, we have set up of a semi-automated pipeline capable of analyzing full UniProt releases in reasonable time and compiling monthly updates of a comprehensive database of SAH segments. Releases of this database, denoted CSAHDB, is available on the CSAHserver 2 website at *csahserver.itk.ppke.hu*. An overview of human SAH-containing sequences combined with a literature survey suggests specific roles of SAH segments in proteins involved in RNA-based regulation processes as well as cytoskeletal proteins, a number of which is also linked to the development and function of synapses.

Keywords: Single alpha-helix; FPGA; protein structure prediction; cytoskeleton; RNA processing