

Charged single alpha-helices: identification, role and evolution

<u>Zoltán Gáspári¹, Dániel Süveges², András Perczel³,</u> László Nyitray⁴, Gábor Tóth⁵

Coiled-coil



¹ Pázmány Péter Catholic University, Faculty of Information Technology, Budapest, Hungary ² Department of Cellular and Molecular Pharmacology, University of California, San Francisco, USA ³ Structural Chemistry and Biology Laboratory, Institute of Chemistry, Eötvös Loránd University, Budapest, Hungary ⁴ Department of Biochemistry, Eötvös Loránd University, Budapest, Hungary ⁵ Bioinformatics Group, Agricultural Biotechnology Center, Gödöllő, Hungary

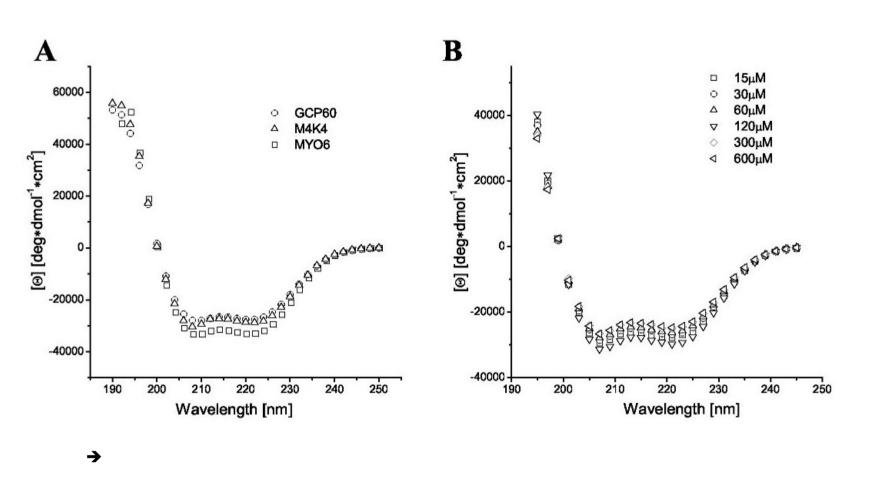
PfamA

http://users.itk.ppke.hu/~gaszo | gaspari.zoltan@itk.ppke.hu

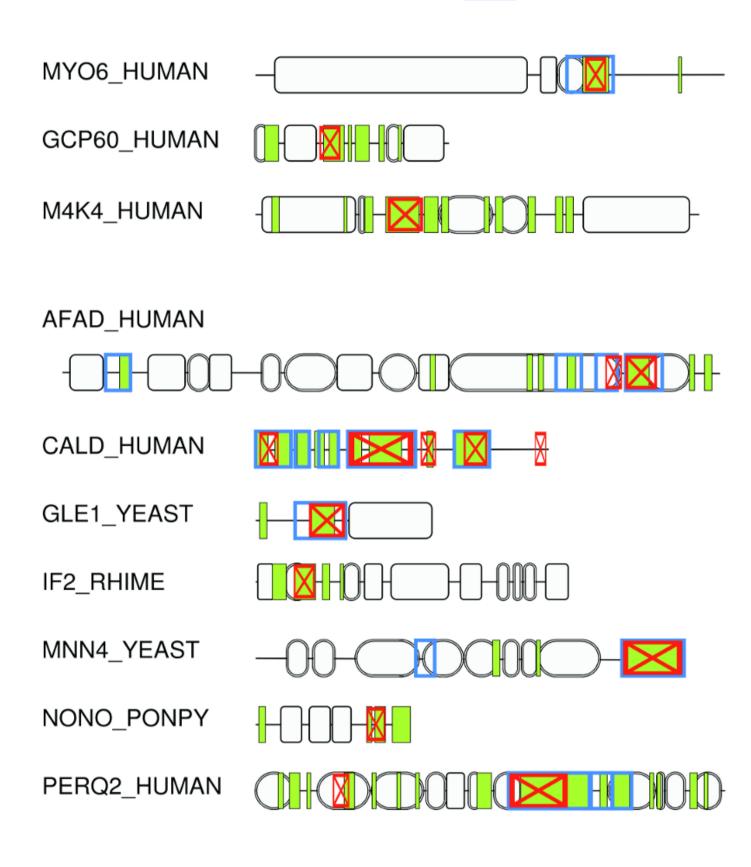
Charged single α -helices (CSAHs)

Consensus detection: csahserver.chem.elte.hu

- A recently identified protein structural motif [1]
- → High Arg, Lys and Glu content
- Regularly alternating charges along the sequence
- Such semgents are helical under a wide range of conditions, no cooperative unfolding



CSAHs often coincide with predicted coiled-coils and disordered segments [2]



PfamB Low complexity CSAH





Detect Charged Single α -Helices in protein sequences					
<u>Description</u>	<u>Example</u>	CSAH Server	<u>Download</u>		

CSAHs detected in the input sequence by the consensus of *scan4csah* and *ft_charge*

mddferrrelrrqkreemrleaeriayqrndddeeeaarerrrrarqerlrqkqeeeslg qvtdqvevnaqnsvpdeeakttttntqvegddeaaflerlarreerrqkrlqealerqke fdptitdaslslpsrrmqndtaenettekeeksesrqeryeieetetvtksyqkndwrda eenkkedkekeeeeeekpkrgsigenqvevmveekttesqeetvvmslkngqisseepkq eeereqgsdeishhekmeeedkeraeaerarleAEERERIKAEQDKKIADERARIEAEEK AAAQERERREAEERERMREEEKRAAEERQRIKEEEKRAAEERQRIKEEEKRAAEERQRIK EEEKRAAEERQRARAEEEEKAKVEEQKRNKQLEEKKHAMQETKIKGEkveqkiegkwvne kkaqedklqtavlkkqgeekgtkvqakreklqedkptfkkeeikdekikkdkepkeevks fmdrkkgftevksqngefmthklkhtentfsrpggrasvdtkeaegapqveagkrleelr rrrgeteseefeklkqkqqeaalelEELKKKREERRKVLEEEEQRRKQEEADRKLREEEE KRRLKEEIERRRAEAAEKRQKMPEDGLSDDKKpfkcftpkgsslkieeraeflnksvqks sgvksthqaaivskidsrleqytsaiegtksakptkpaasdlpvpaegvrniksmwekgn vfssptaagtpnketaglkvgvssrinewltktpdgnkspapkpsdlrpgdvsskrnlwe kqsvdkvtsptkv



Full proteomes in UniProt with CSAH segments

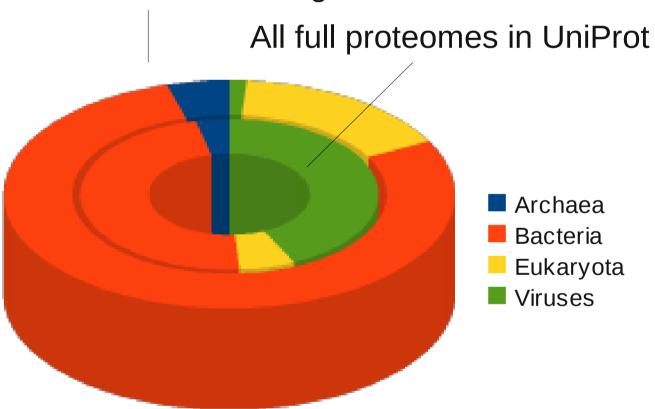
CSAH list

CSAHs in UniProt: a rare motif

→ Consensus approach:

→ SwissProt: 586 CSAHs in 533 proteins

- → TrEMBL: 8908 CSAHs in 8150 proteins
- → 0.2 % of human proteins contains CSAHs
- → Remarkably abundant in *Plasmodium falciparum*



Function of CSAH-containing proteins

	All SwissProt70 proteins			
GO term	expected	observed	obs/exp	P-value
Biological process				
symbiosis, encompassing mutualism through parasitism	0.06	7	115.212.	3.13E-13
nuclear-transcribed mRNA catabolic process, nonsense-mediated decay	0.12	4	33.24	7.16E-06
cytokinesis	0.20	4	19.79	5.48E-05
Molecular function				
actin filament binding	0.11	6	54.05	1.83E-09
translation initiation factor activity	1.24	37	29.85	2.20E-16
motor activity	0.51	3	5.85	1.00E-02
Cellular localization				
chromosome, centromeric region	0.17	3	17.71	6.82E-04
spindle	0.31	5	16.27	1.60E-05
cell cortex	0.22	3	13.8	1.40E-03

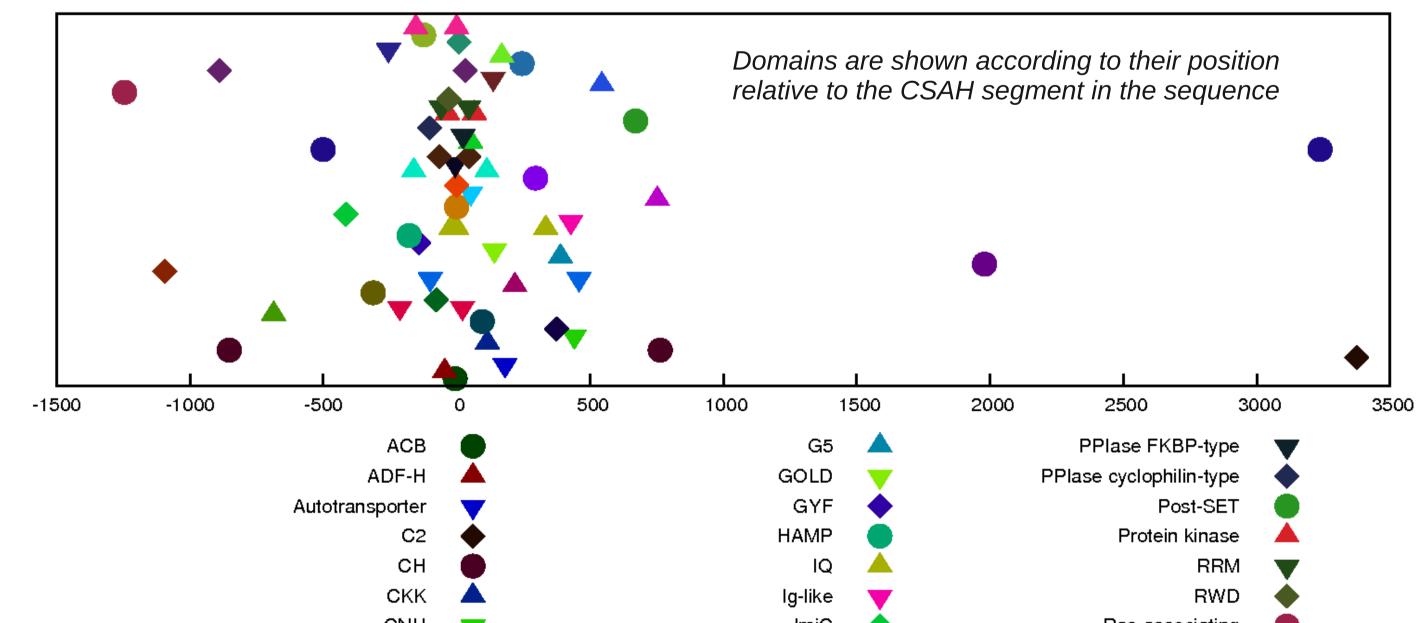
Data refer to the 70% similartiy-filtered sets from SwiProt and UniProt, release 2011_05

CSAH No.	Consensus	scan4csah	ft_charge
1	274 - 407	256-407	274-419
2	566 - 632	566-632	523-645

Combines two independent methods for CSAH detection: SCAN4CSAH & FT CHARGE [3]

Aimed at minimizing the number of false positives

Domains found in CSAH-containing proteins



CSAHs' charge pattern can be easily disrupted by mutations

MYO1O_HUMAN RRRFLHLKKAAIVFQKQLRGQIARRVYRQLLAEKREQEEKKKQEEEEKKKREEEEREREREREREAELRAQQEEETRKQQELEALQK-SQKEAELTRELEKQ

Summary

CSAHs can reliably be identified by our consensus approach aimed at minimizing the number of false positives

CSAH is a rare but ubiquitious motif involved in a number of RNA-related and motility processes

→ Found in a wide variety of proteins, shows little conservation

Acknowledgements

This work was supported by grants from the Hungarian Scientific Research Fund (OTKA F68079, K72973, K61784, NI68466, NK81950), and the International Centre for Genetic Engineering and Biotechnology (CRP/HUN09-03). The European Union and the European Social Fund have provided financial support to the project under the grant agreement no. TÁMOP 4.2.1./B-09/1/KMR-2010-0003. A János Bolyai Research Fellowship to Z.G. and Ferenc Deák Fellowship to D.S. is also acknowledged.

Ras-associating CNH DH LEM SE1 DZF LRRCT Dilute MIF4G SH3 STI1 EF-hand Methyl-accepting transducer EGF-like TSP type-1 MyTH4 EΗ Myb-like F-box Myosin head-like UBX FERM PCI WH₂ FHA PDZ Zinc-hook PH Fido

References

- [1] Süveges. D., Gáspári, Z., Tóth, G. and Nyitray, L. (2009) Charged single α-helix: a versatile protein structural motif. Proteins 74, 905-916.
- [2] Szappanos, B., Süveges, D., Nyitray, L., Perczel, A. and Gáspári, Z. (2010) Folded-unfolded cross-predictions and protein evolution: the case study of coiled-coils. FEBS Lett. 584, 1623-1627.
- [3] Gáspári, Z., Süveges, D., Perczel, A., Nyitray, L. and Tóth, G. (2012) Charged single alpha-helices in proteomes revealed by a consensus prediction approach. Biochem. Biophys. Acta Proteins and Proteomes 1824, 637-646.