Laboratory of Protein Evolution

Three subfamilies of fungal GH13 α-amylases – their evolutionary relatedness and ancestry

Zuzana Janíčková[1,2] & Štefan Janeček[1,2]

[1] *Department of Biology, Faculty of Natural Sciences, University of SS. Cyril and Methodius, Trnava, Slovakia*

[2] *Institute of Molecular Biology, Slovak Academy of Sciences, Bratislava, Slovakia*

INTRODUCTION

Amylolytic enzymes, in general, along with other glycoside hydrolases (GHs) based on similarities in their amino acid sequences have been classified into the GH families within the CAZy database (Carbohydrate-Active EnZymes) [1]. α-Amylases (EC 3.2.1.1) are important industrial enzymes used in starch hydrolysis and many other biotechnologies.The α-amylase family GH13 was established in 1991 [2]. With more than 60,000 sequences and about 30 different enzyme specificities it belongs to the largest GH families [3]. The basic characteristics of the family GH13 include the retaining reaction mechanism, TIM-barrel structure of the catalytic domain, 4-7 conserved sequence regions (CSRs) and the same catalytic machinery [4] consisting of the catalytic nucleophile (aspartic acid), proton donor (glutamic acid) and transition state stabiliser (aspartic acid) positioned at strands β4, β5 and β7, respectively, of the catalytic TIM-barrel domain. In 2006, the α-amylase family GH13 was divided into 35 subfamilies by CAZy curators [5]; the current number of GH13 subfamilies being 42 [1]. The division, based on the existence of groups of enzymes in the family GH13, reveals a higher degree of similarity within each subfamily than within the entire family, either in specificity, taxonomy or in sequences [4]. The presented work delivers the *in silico* analysis of α-amylases from fungal origin, several of which being known as psychrophiles, thermophiles, halophiles and even xerophiles. Originally, fungal α-amylases were classified in the subfamily GH13_1 [4,5], but recently, such enzymes were identified as close homologues of the rather actinobacterial subfamily GH13_32 [6]. Members of the subfamily GH13_1 are typical extracellular α-amylases produced by molds and yeasts with the Taka-amylase A, i.e. the α-amylase from *Aspergillus oryzae*, as the main representative [4]. On the other hand, the subfamily GH13_32 was considered for a long time to be a typical subfamily of bacterial α-amylases, originating mainly from actinobacteria and exhibiting a close relatedness to animal α-amylases from subfamilies GH13_15 and GH13_24 [6,7]. Some intracellular α-amylases of fungal origin, involved in the synthesis of cell wall α-1,3-glucan, have been classified also in the subfamily GH13_5 that originally covered bacterial liquefying α-amylases [4,8].

AIM OF THE STUDY

The main goal of the present study was therefore to perform a detailed *in silico* analysis of fungal α-amylases that exist in two major sequentially different, yet related forms – subfamilies GH13_1 and GH13_32. In addition, α-amylases of fungal origin from the subfamily GH13_5 have also been included. The study was undertaken in an effort to: (i) reveal eventual intermediates between the three α-amylase subfamilies GH13_1, GH13_5 and GH13_32; (ii) contribute to the correct annotation of hypothetical GH13 proteins obtained from genome sequencing projects; and (iii) map potential GH13 α-amylases from extremophilic fungi.

Table 1. α-Amylases from the family GH13 used in the present study.

Subfamily	GH13_1			GH13_5			GH13_32		
Source	Molds	Mushrooms	Yeasts	Molds	Mushrooms	Yeasts	Molds	Mushrooms	Yeasts
Number	254	11	61	33	6	7	88	61	1
Total	**326**			**46**			**150**		

The set was created based on sequences classified in the CAZy family GH13 completed by α-amylases obtained with BLAST.

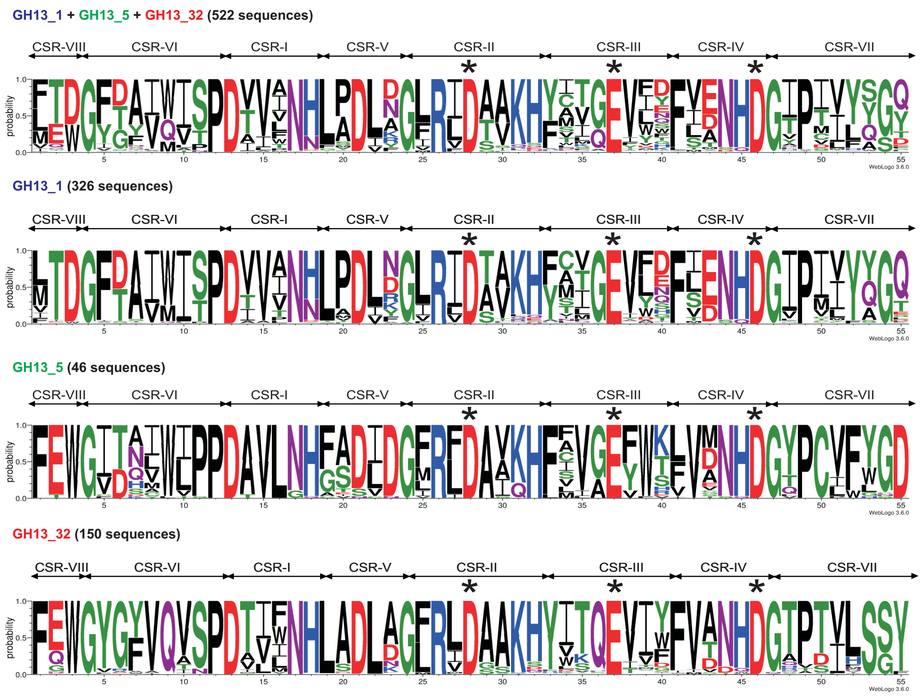

Figure 1. Sequence logos of α-amylases from subfamilies GH13_1, GH13_5, GH13_32. CSR-I, residues 13-18; CSR-II, residues 24-32; CSR-III, residues 33-40; CSR-IV, residues 41-46; CSR-V, residues 19-23; CSR-VI, residues 4-12; CSR-VII, residues 47-55; CSR-VIII, residues 1-3. The catalytic triad, i.e. the catalytic nucleophile (No. 28, aspartic acid), the proton donor (No. 37, glutamic acid) and the transition-state stabiliser (No. 46, aspartic acid) are indicated by asterisks.

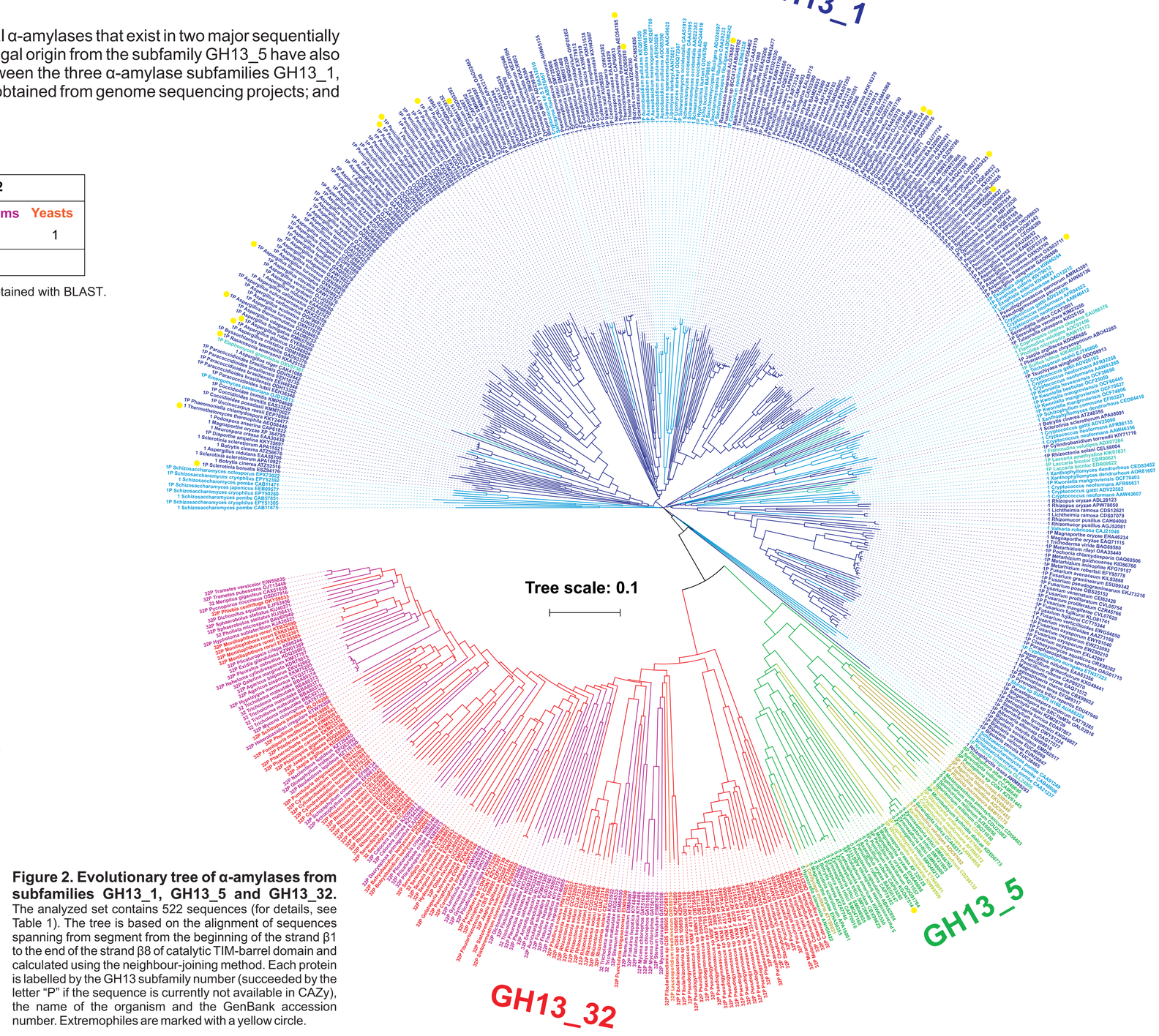

Figure 2. Evolutionary tree of α-amylases from subfamilies GH13_1, GH13_5 and GH13_32. The analyzed set contains 522 sequences (for details, see Table 1). The tree is based on the alignment of sequences spanning from segment from the beginning of the strand β1 to the end of the strand β8 of catalytic TIM-barrel domain and calculated using the neighbour-joining method. Each protein is labelled by the GH13 subfamily number (succeeded by the letter "P" if the sequence is currently not available in CAZy), the name of the organism and the GenBank accession number. Extremophiles are marked with a yellow circle.

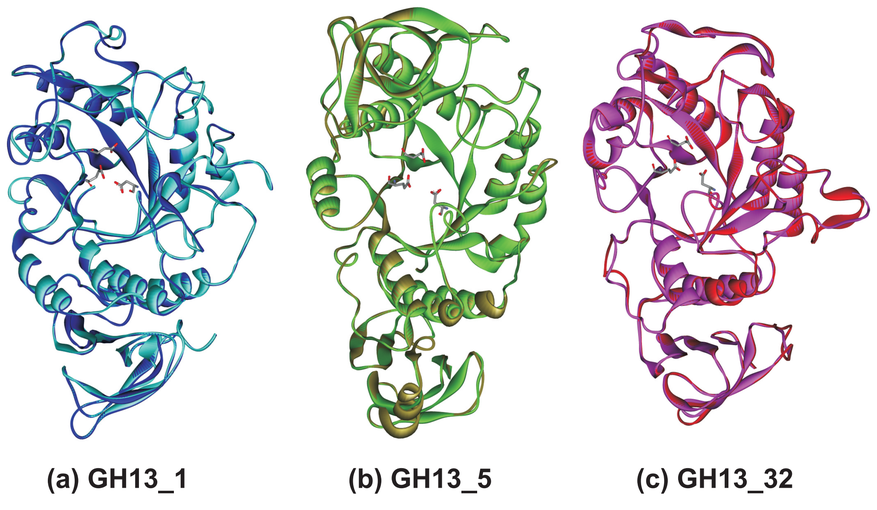

Figure 3. Tertiary structures of selected α-amylases from subfamilies GH13_1, GH13_5 and GH13_32. Superimposition of a real structure (i.e. template) of an α-amylase from a given GH13 subfamily with a modelled structure of its counterpart originating always from *Pholiota microspora*. (a) GH13_1 – *Aspergillus oryzae* (PDB: 2TAA; blue) and *P. microspora* (UniProt: A0A1L7MPN9; cyan) – 460 Cα-atoms and 0.48 Å RMSD; (b) GH13_5 – *Geobacillus stearothermophilus* (PDB: 1HVX; green) and *P. microspora* (UniProt: A0A1L7MPN7; olive) – 477 Cα-atoms and 0.40 Å RMSD; (c) GH13_32 – *Pseudoalteromonas haloplanctis* (PDB: 1JD7; red) and *P. microspora* (UniProt: A0A1E1ERR9; magenta) – 430 Cα-atoms and 0.48 Å RMSD. The side-chains of the catalytic triad are displayed in each α-amylase model .

CONCLUSIONS

(1) The present study delivers the *in silico* analysis of fungal α-amylases from subfamilies GH13_1 and GH13_32 including those from GH13_5. Overall, 522 GH13 proteins of fungal origin were collected – first from the CAZy database and then completed by a BLAST search – as follows: (i) subfamily GH13_1: 326 sequences; (ii) subfamily GH13_32: 150 sequences; and (iii) subfamily GH13_5: 46 sequences (Table 1). Each of 522 studied α-amylases was selected to contain all CSRs and complete catalytic triad characteristic for the α-amylase family GH13 [4].

(2) Four sequence logos were calculated based on CSRs identified in all α-amylases (Fig. 1). While the first sequence logo covers all the three GH13 subfamilies (522 α-amylases), the three additional logos created for the three individual subfamilies indicate the features dicriminating the subfamilies from each other [3,4]. It is of interest that although each subfamily may exhibit its own unique sequence features, there are also some positions within the CSRs that are shared by two subfamilies, e.g. FEW and FeW (positions 1-3) in CSR-V**III** for GH13_5 and GH13_32, respectively, as well as YyGq and FyGD (positions 52-55) in CSR-V**II** for GH13_1 and GH13_5, respectively.

(3) The evolutionary relationships among the fungal α-amylases from the three subfamilies are depicted in the evolutionary tree (Fig. 2). It is obvious that, despite the identical catalytic machinery and significant similarities in all CSRs, each subfamily retains its uniqueness and independency. Nevertheless, fungal α-amylases from the subfamily GH13_32 seem to be evolutionarily more closely related to those from the subfamily GH13_5 than to their counterparts from the subfamily GH13_1. The α-amylases from the subfamily GH13_1 originating from *Rhizophlyctis rosea* (soil unicellular fungus), *Neolecta irregularis* (a mushroom known as *irregular earth tongue)* and *Schizosaccharomyces pombe* (a yeast) occupy an intermediary position in the tree indicating a remarkable evolutionary relatedness to their counterparts from subfamilies GH13_5 and GH13_32.

(4) With regard to extremophiles, it seems that, according to our current knowledge, only microscopic fungi (molds) produce family GH13 α-amylases. They have been found mostly within the subfamily GH13_1 with one representative in the subfamily GH13_5 (Fig. 3). In order to draw a more conclusive view, a more detailed analysis is necessary with regard to both fungal taxonomy and properties of produced α-amylases.

(5) The preliminary analysis of tertiary structures has confirmed the very high similarity between the well-established α-amylases from the three GH13 subfamilies and their homologues, as documented by the comparison of real structures of α-amylases from *Aspergillus oryzae* (GH13_1), *Geobacillus stearothermophilus* (GH13_5) and *Pseudoalteromonas haloplanctis* (GH13_32) with their modelled *Pholiota microspora* counterparts (a mushroom called *nameko*) possessing the α-amylases from all the three respective GH13 subfamilies (Fig. 3).

Acknowledgements. This work was supported by grants No. FPV-02-2018 from the University of SS. Cyril and Methodius in Trnava and No. 2/0146/17 from the Slovak Grant Agency VEGA.

References

[1] Lombard V., Ramulu H.G., Drula E., Coutinho P.M. & Henrissat B. (2014) *Nucleic Acid Res.* **42:** D490-D495.
[2] Henrissat B. (1991) *Biochem. J.* **280:** 309-316.
[3] Janecek S. & Gabrisko M. (2016) *Cell. Mol. Life Sci.* **73:** 2707-2725.
[4] Janecek S., Svensson B. & MacGregor E.A. (2014) *Cell. Mol. Life Sci.* **71:** 1149-1170.
[5] Stam M.R., Danchin E.G., Rancurel C., Coutinho P.M. & Henrissat B. (2006). *Protein Eng. Design Select.* **19:** 555-562.
[6] Da Lage J.L., Binder M., Hua-Van A., Janecek S. & Casane D. (2013) *BMC Evol. Biol.* **13:** 40.
[7] Da Lage J.L., Feller G. & Janecek S. (2004) *Cell. Mol. Life Sci.* **61:** 97-109
[8] van der Kaaij R.M., Janecek S., van der Maarel M.J.E.C. & Dijkhuizen L. (2007) *Microbiology* **153:** 4003-4015.