ULS Timisoara
Multidisciplinary Conference on Sustainable Development
25-26 May 2023

In silico characterization of an endoglucanase from *Actinoalloteichus hoggarensis*

Mahfoud BAKLI [1,2], Noureddine BOURAS [3,4], Raul PAŞCALĂU [5], and Laura ŞMULEAC [5]

[1] *Département de Biologie, Faculté des Sciences et Technologie, Université d'Aïn Temouchent, B.P 284, 46000, Aïn Temouchent, Algeria.*
[2] *Physiologie, Physiopathologie et Biochimie de la Nutrition, Université de Tlemcen, Tlemcen, Algeria.*
[3] *Laboratoire de Valorisation et Conservation des Écosystèmes Arides (LVCEA), Faculté des Sciences de la Nature et de la Vie et Sciences de la Terre, Université de Ghardaïa,*
[4] *Laboratoire de Biologie des Systèmes Microbiens (LBSM), Ecole Normale Supérieure de Kouba, Alger, Algeria.*
[5] *Banat's University of Life Sciences "King Michael I of Romania", 119 Calea Aradului, 300645, Timisoara, Romania.*

Abstract:

Actinoalloteichus hoggarensis is a rare bacterial species that was isolated from the Algerian Saharan desert and is known for producing biologically active compounds. Despite its potential, little is understood about the enzymes it produces, including endoglucanases. These cellulase enzymes break down cellulose, the primary structural component of plant cell walls that provides strength and rigidity. The breakdown of cellulose by endoglucanases has numerous biotechnological applications, such as the production of biofuels, bioplastics, and paper. This study involves an *in silico* characterization of an endoglucanase from *A. hoggarensis* to gain insight into its structural and functional properties, with the goal of informing the development of novel biotechnological applications. Our study represents a major milestone in understanding the potential of this rare bacterial species and its enzymes, opening up exciting new avenues for further research and development.

Keywords : Endoglucanase, *Actinoalloteichus hoggarensis*, structural and functional properties, *in silico* characterization.

• Introduction

Cellulose is the primary structural component of plant cell walls that provides strength and rigidity and is the most abundant biopolymer on earth. Cellulase enzymes encompass endoglucanases (EC 3.2.1.4), β-glucosidase (EC 3.2.1.21), and exoglucanases (EC 3.2.1.91). Endoglucanase (EG) is the essential component for hydrolyzing β-1,4-glycosidic bonds of the carbon skeleton in cellulose to produce celloligosaccharides. Endoglucanases are enzymes that break down cellulose by hydrolyzing the internal bonds between glucose units, and have numerous biotechnological applications, such as the production of biofuels, bioplastics, and paper. Despite their potential, the industrial-scale production of endoglucanases using traditional methods, such as bacterial or fungal fermentation, is costly and inefficient.

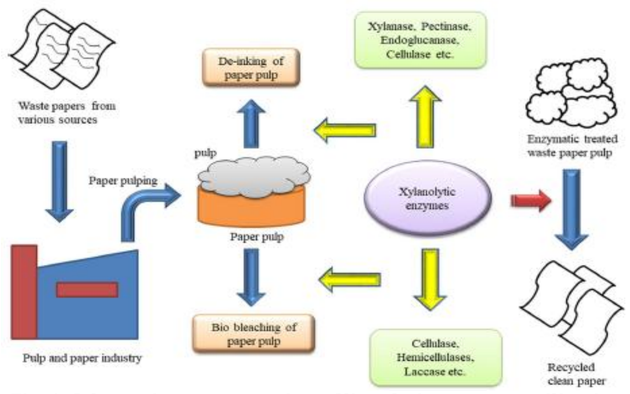


Fig. 1. Schematic representation of lignolytic enzymes utilization in pulp and paper industry (Gupta *et al.*, 2022).

Actinoalloteichus hoggarensis is a rare actinobacterial species that was isolated from the Algerian Saharan desert and is known for producing bioactive secondary metabolites. Little is understood about the enzymes it produces, including endoglucanases. However, its potential for biotechnological applications is promising due to its unique properties, such as thermotolerance, halotolerance, and the ability to grow in extreme conditions. The study of endoglucanases from rare bacterial species like *A. hoggarensis* is still in its early stages, leaving a significant knowledge gap in understanding their properties. Further research is needed to fill this gap and explore the potential applications of these enzymes. However, The experimental expression and purification of proteins can pose challenges and require significant time and effort.
Computational approaches, such as homology modeling, molecular dynamics simulation, mutational analysis, etc., provide a promising alternative to design endoglucanase enzymes with desirable characteristics for industrial applications. In this study, we use *in silico* methods to perform a structural and functional characterization of an endoglucanase from *A. hoggarensis*, with the goal of gaining insights into its properties and potential for biotechnological applications.

• Materials and methods

In silico methods	*In silico* tools
1. Amino acid sequence retrieval	Uniprot (ID: A0A221VZV6)
2. Physicochemical characterization and protein solubility	ExPASy ProtParam and SOSUI
3. Secondary structure prediction	PSIPRED 4.0 and SOPMA
4. Tertiary structure prediction, structure validation, and quality prediction	SWISS-MODEL, Raptor X, $(PS)^2$-V2, $Phyre^2$, MODELLER, LOMETS, I-TASSER, ModRefiner, PyMOL, and PROCHECK
5. Ligand binding site prediction and protein localization	COACH, PSORT-B v3.0.3, SignalP-6.0, and TMHMM
6. Funtional analysis	MotifFinder

Table1. The physiochemical properties of the predicted endoglucanase.

Property	Value
Number of amino acids residues (AA)	875
Molecular weight (Da)	92883.92
Theoretical pI	4.08
Total number of negatively charged residues (Asp + Glu)	124
Total number of positively charged residues (Arg + Lys)	33
Extinction coefficient (EC)	172855
half-life	30 hours
Instability index (II)	33.03
Aliphatic index (AI)	78.22
Grand average of hydropathicity (GRAVY)	-0. 176

Table2. Ramachandran plot calculation of the endoglucanase models computed with PROCHECK tool.

Model	Type of model	Favored region (%)	Additional allowed region (%)	Generously allowed region (%)	Disallowed region (%)
SWISS-MODEL	Initial	88.0	0.2	0.2	0.4
	Refined	89.6	0.4	0.4	0.6
Raptor X	Initial	86.8	1.9	1.9	1.0
	Refined	86.8	1.9	1.9	1.0
$(PS)^2$-V2	Initial	81.0	1.4	1.4	0.0
	Refined	91.8	0.0	0.0	0.0
$Phyre^2$	Initial	84.5	1.3	1.3	0.6
	Refined	89.2	0.8	0.8	0.8
MODELLER	Initial	90.0	1.5	1.5	0.4
	Refined	90.5	0.8	0.8	0.4
LOMETS	Initial	73.0	5.5	5.5	2.2
	Refined	78.9	4.1	4.1	2.5
I-TASSER	Initial	59.6	7.4	7.4	3.3
	Refined	78.6	3.9	3.9	2.1

• Results and discussions

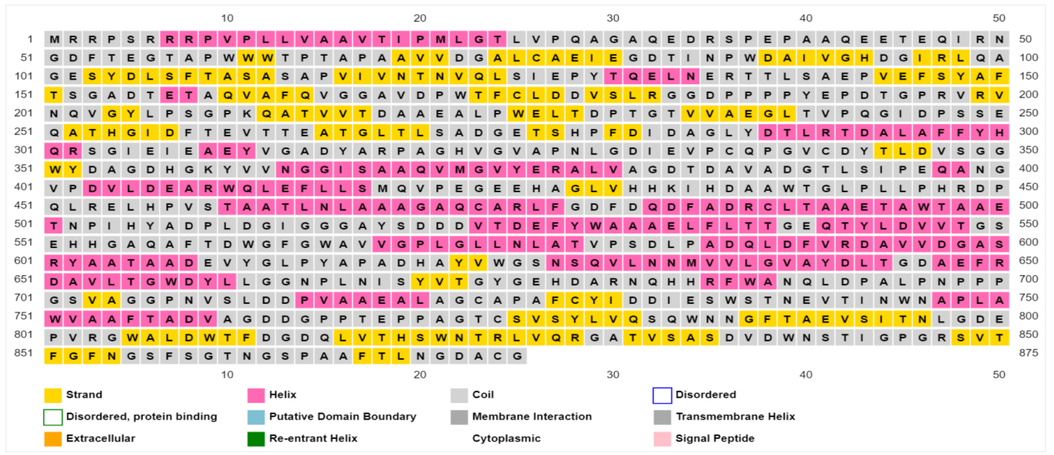


Fig. 2. Sequence annotation plot analysing the secondary structure prediction of endoglucanase from *A. hoggarensis* by PSIPRED. (yellow for β-strands, pink for α-helix and grey for coil structures).

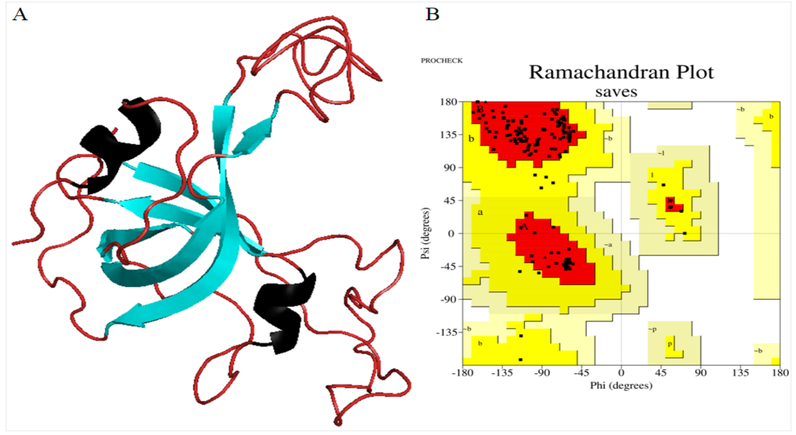


Fig. 3. Tertiary structure prediction and validation for the endoglucanase from *A. hoggarensis*.
(A) Tertiary structure prediction of the protein was done by homology modeling using $(PS)^2$-v2. This structural model was refined by ModRefiner and visualised by PyMOL. α-helix (black), β-sheet (cyan), and loop (red).
(B) The predicted and refined structure's Ramachandran plot was validated by the PROCHECK program, with residues classified based on their conformational preferences as per Ramachandran plot analysis. Regions with favoured (A, B, and L), additional allowed (a, b, l, and p), generously allowed (~a, ~b, ~l, and ~p), and disallowed conformations are highlighted in red, yellow, beige, and white, respectively. Non-glycine and non-proline residues are represented by filled black squares, while glycines (excluding those at the ends of the polypeptide chains) are depicted as filled black triangles.

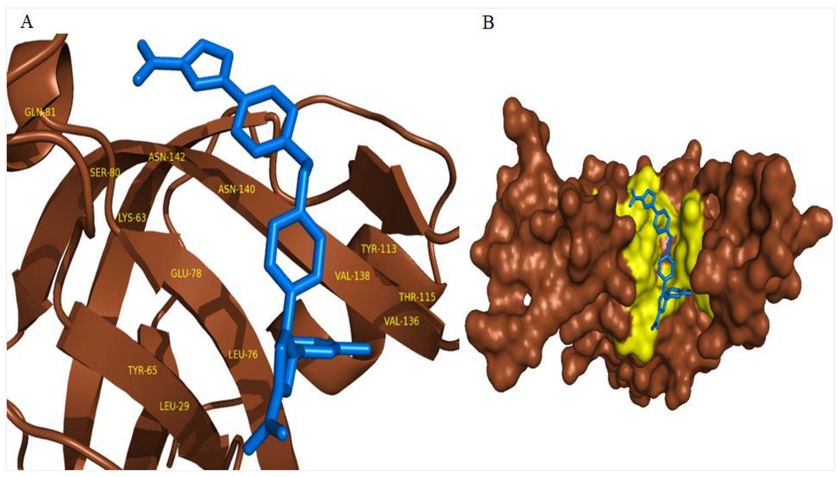
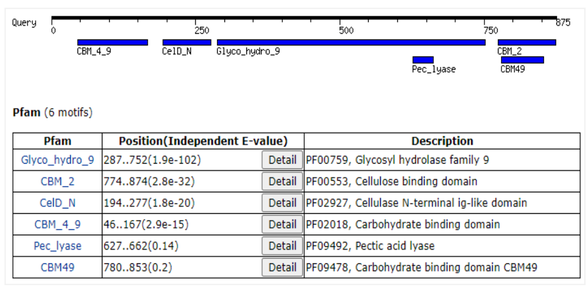


Fig. 4. Ligand binding sites of endoglucanase from *A. hoggarensis* predicted by COACH. (A) Ligand binding sites (Leu 29, Lys 63, Tyr 65, Leu 76, Glu 78, Ser 80, Gln 81, Tyr 113, Thr 115, Val 136, Val 138, Asn 140, and Asn 142). **(B)** Surface view of the protein with ligand in pocket exposed active site. Sticks represent the ligand (blue). Ligand binding sites (yellow) within the protein structure (brown).



Pfam (6 motifs)

Pfam	Position(Independent E-value)		Description
Glyco_hydro_9	287..752(1.9e-102)	Detail	PF00759, Glycosyl hydrolase family 9
CBM_2	774..874(2.8e-32)	Detail	PF00553, Cellulose binding domain
CelD_N	194..277(1.8e-20)	Detail	PF02927, Cellulase N-terminal ig-like domain
CBM_4_9	46..167(2.9e-15)	Detail	PF02018, Carbohydrate binding domain
Pec_lyase	627..662(0.14)	Detail	PF09492, Pectic acid lyase
CBM49	780..853(0.2)	Detail	PF09478, Carbohydrate binding domain CBM49

Fig. 5. The motif analysis of *A. hoggarensis* endoglucanase determined by MotifFinder.

• Conclusions

In silico analysis of the endoglucanase from *A. hoggarensis* revealed several key properties, including its acidic nature, thermostability, hydrophilicity, and extracellular location. The protein's secondary structure was found to be predominantly random coil, with some alpha-helix and beta-sheet elements. Additionally, we conducted both structural homology modeling and functional analyses. Nevertheless, further experimental studies, such as enzyme kinetics, substrate specificity, and protein expression, are needed to fully understand the potential of this endoglucanase and its applications in biotechnology.

• References

- Gupta, G. K., Dixit, M., Kapoor, R. K., & Shukla, P. (2022). Xylanolytic enzymes in pulp and paper industry: new technologies and perspectives. *Molecular biotechnology*, 1-14.
- Boudjelal, F., Zitouni, A., Bouras, N., Schumann, P., Spröer, C., Sabaou, N., & Klenk, H.-P. (2015). *Actinoalloteichus hoggarensis* sp. nov., an actinomycete isolated from Saharan soil. *International journal of systematic and evolutionary microbiology, 65*(6), 2006-2010.
- Mandeep, Liu, H., & Shukla, P. (2021). Synthetic biology and biocomputational approaches for improving microbial endoglucanases toward their innovative applications. *ACS omega, 6*(9), 6055-6063.