Journal of Engineering Research and Applied Science
Available at www.journaleras.com
Volume 10 (1), June 2021, pp 1757-1763
ISSN 2147-3471 © 2021

In silico analysis of alcohol dehydrogenase from *Bacillus thuringiensis serovar israelensis* using bioinformatics tools

Fatih Aktas

Düzce University, Faculty of Engineering, Environmental Engineering, Düzce, Turkey.

Accepted 20 April 2021

Abstract
Alcohol dehydrogenases (ADH) (EC 1.1.1.1) that belong to dehydrogenase enzymes found in many organisms carry out the interconversion of alcohols and aldehydes or ketones using nicotinamide cofactor (NAD+ or NADH). Alcohol dehydrogenase from *Bacillus thuringiensis serovar israelensis* (BtADH) was cloned, expressed, and characterized in our previous study. This study provides the detailed comprehensive characterization of BtADH protein by using in silico tools to estimate physicochemical properties, structural characteristics, and motif patterns consisting of secondary structure prediction and functional analysis. The prediction of the secondary structure of BtADH protein was carried out by the PSIPRED protein analysis workbench and defining functional sites of BtADH was determined by InterProScan data resource. Motif scan analysis of BtADH was conducted in different databases by using the MyHits tool.

Keywords: Alcohol dehydrogenase, zinc-containing ADH, *Bacillus thuringiensis serovar israelensis*, in silico analysis, bioinformatics

1. Introduction

Alcohol dehydrogenases (ADHs) are a versatile biocatalyst belonging to the group of oxidoreductases found in many living beings from simple organisms such as bacteria [1], yeast [2], to higher organisms such as plants [3], animals [4], and humans [5]. ADHs have been reported to have three main classes according to their size, which is: (I) metal-free short-chain enzymes (250 residues per chain) [6] (II) Medium-chain enzymes containing mostly zinc (350 -375 residues per chain) [7] (III) long-chain enzymes (more than 700 residues) [8]. *Bacillus thuringiensis* is an effective biological agent and important microbial insecticide, which is widely used today against insect pests [9]. *Bacillus thuringiensis serovar israelensis* is a strain one of the most effective among this species [10].

Alcohol dehydrogenases are the most studied enzyme group in the dehydrogenase family, due to their important roles in both the metabolism of organisms and industrial biotechnological applications. Because it catalyzes the interconversion of alcohol aldehydes and ketones and has the potential to produce pharmaceutical fine chemicals, it has been cloned and characterized from many organisms for investigation and use in bioprocesses. Zinc-containing, NADH-dependent alcohol dehydrogenase was cloned from *Bacillus thuringiensis serovar israelensis* (BtADH) and characterized biochemically and enzymatically in our previous study [11].

In this previous study, it was shown that BtADH contains zinc and catalyzes from short-chain primary alcohols to long-chain ones using the nicotinamide cofactor. In sequence similarity comparisons, it was found that although BtADH is a mesophilic enzyme, it also shows high similarity with thermophilic ADHs. Due to this different and interesting character, it is seen that it is important to study and reveal the structure-function relationships in silico in detail and to increase the cofactor and substrate selectivity and thermal stability of the enzyme.

The BtADH enzyme is a $NAD(P)^+$ dependent dehydrogenase in the cofactor utilization classification and is capable of interconverting alcohols to the corresponding aldehydes and ketones. BtADH is also in the MDR2 superfamily [12], which is characterized by its short QSDL and catalytic zinc.

Corresponding author: fatihaktas@duzce.edu.tr

It is seen in the study of crystal structure similarities that BtADH has a quaternary conformation with two dimers [13].

Advancement and extensive use of DNA sequencing technology in the last few decades provide many organism genomes and protein sequences to researchers in different databases. Increasing the genomic information in databases, effective bioinformatics tools have begun to be presented for researchers and users to process huge information. Currently, many databases and bioinformatics tools supply detailed analysis of gene and protein sequences and their 3D forms for amateur or professional users. Bioinformatics tools used for in silico analysis of proteins provide important data for the development of drugs and treatments. Proteins increase the catalytic activity of biological processes and reactions in cellular activities. Therefore, analysis and knowledge of proteins are vital in determining the structure and functions of proteins in the metabolism of organisms [14].

In this study, molecular, physical properties, sequence-structure, and structure-function relationships of BtADH enzyme were determined by in silico methods and tools.

2. Material and Methods

2.1. Physico-chemical properties of BtADH

BtADH Protein sequence and fasta data file were obtained from UniProtKB database (RBTH_04909 - Alcohol dehydrogenase - Bacillus thuringiensis serovar israelensis ATCC 35646 - RBTH_04909 gene & protein (uniprot.org) (UniProtKB - Q3ETI8). The Physico-chemical properties of BtADH protein were calculated by using bioinformatics tools. The isoelectric point, molecular weight, charge of the protein, and average residue weight were defined by Emboss Pepstats [15]. Protparam tool was used for calculating amino acid composition, atomic composition, and instability index [16].

2.2. Secondary Structure and Function Analysis of BtADH

The PSIPRED protein analysis workbench (http://bioinf.cs.ucl.ac.uk/psipred/) was used for the prediction of the secondary structure of BtADH protein [17]. InterProScan was used for defining functional sites of BtADH [18].

2.3. Motif scan analysis of BtADH

Motif scan analysis of BtADH was performed in HAMAP profiles [hamap], PROSITE patterns [pat], PROSITE patterns (frequent match producers) [freq_pat], Pfam HMMs (global models) [pfam_ls], Pfam HMMs (local models) [pfam_fs], and PROSITE profiles [prf] databases by using MyHits tool developed by the Swiss Institute of Bioinformatics (SIB) (Motif Scan (sib.swiss)) [19].

3. Results and discussions

3.1. Physico-chemical analysis of BtADH

BtADH protein isoelectric point determined as 5.2, the molecular weight of 36758.3 g, atom numbers of 5172, and the molecular formula of ($C_{16}0_7H_{26}0_4N_{442}O_{498}S_{21}$) by used Emboss Pepstats tool. BtADH has 345 amino acids with 45 (Glu + Asp) negative charge residues and 36 (Lys + Arg) positive charge residues, furthermore, this protein has higher contents of Valine, Glycine, Alanine, and the instability index (II) was computed to be 22.90 (that means the protein as stable) were demonstrated by ProtParam tool. The Atomic and amino acid composition of BtADH were displayed in Table 1.

3.2. Structural and functional analysis of BtADH

Secondary structure analysis of BtADH was conducted by using the PSIPRED server. BtADH is mainly composed of alfa helix, beta-strands, and coils as shown in Figure 1. Besides, each subunit has a total of 11 α-helices and 15 β-strands. The functional analysis, predicting domains, and important sites of BtADH protein were performed by the InterPro database and shown in Figure 2. BtADH displayed good homology with NAD(P)-binding domain superfamily (IPR036291) (in between 174-297 as C terminal domain) [4], and GroES-like superfamily (IPR011032) (in between 27-133 as N terminal domain) domain known as the catalytic domain of alcohol dehydrogenase that contains an inserted zinc-binding domain and highly conserved hydrophobic core [20]. It has also an Alcohol dehydrogenase catalytic activity, zinc type, conserved site (IPR002328), zinc ion binding (GO:0008270) site, and oxidoreductase activity (GO:0016491) site. This analysis indicated that BtADH has also Medium-chain alcohol dehydrogenases, a catalytic domain (G3DSA:3.90.180.10), and NAD(P)-binding Rossmann-like Domain (G3DSA:3.40.50.720). Besides, BtADH resembled the Cinnamyl alcohol dehydrogenases (CAD-cd08297) family according to Interpro analysis.

Table 1. Amino acid and atomic composition of BtADH protein calculated by ProtParam tool.

Amino acid composition	
Ala (A) 33	9.6%
Arg (R) 12	3.5%
Asn (N) 15	4.3%
Asp (D) 23	6.7%
Cys (C) 13	3.8%
Gln (Q) 11	3.2%
Glu (E) 22	6.4%
Gly (G) 34	9.9%
His (H) 5	1.4%
Ile (I) 22	6.4%
Leu (L) 19	5.5%
Lys (K) 24	7.0%
Met (M) 8	2.3%
Phe (F) 9	2.6%
Pro (P) 10	2.9%
Ser (S) 14	4.1%
Thr (T) 14	4.1%
Trp (W) 1	0.3%
Tyr (Y) 8	2.3%
Val (V) 48	13.9%
Pyl (O) 0	0.0%
Sec (U) 0	0.0%
Atomic composition	
Carbon C	1607
Hydrogen H	2604
Nitrogen N	442
Oxygen O	498
Sulfur S	21

this family belongs to members of the medium-chain dehydrogenase/reductase and zinc-dependent alcohol dehydrogenases (ADHs). These ADHs generally have 2 tightly bound zinc atoms per subunit. One of these zinc atoms plays a structural role and is coordinated by four cysteines and the other one is located at an active site that is coordinated by a water molecule, a histidine, two cysteines. This analysis information is similar to the data in the cloning and characterization study of BtADH previously [11].

3.3. Motif scan analysis of BtADH

Motif scan analysis of BtADH protein in different databases was carried out by using MyHits. All potential motifs are listed in Fıgure 3 in our queried sequence. According to the results of motif scanning analysis, 28 matches were determined in the databases. Three of these matches are CK2_PHOSPHO_SITE, seven are MYRISTYL, four are PKC_PHOSPHO_SITE, two each are ADH_zinc_N and ADH_N [!], BRCA2, and one each is ADH_ZINC, NAD_BINDING, SAM_BIND, PPR, 2-Hacid_dh_N, motifs 2-Hacid_dh_C, TDH_MF_00627, shown in Figure 3. Strong matches are demonstrated with a [!] meanwhile weak matches are with [?]. The motifs ADH_ZINC, ADH_N, and ADH_zinc_N are strongly match motifs. ADH_ZINC pattern (GHEGVGIVTKIADDV) is known as zinc-containing alcohol dehydrogenases and was matched on 59-73 residues, ADH_N pattern is known as *Alcohol dehydrogenase GroES-like domain* and was matched on 26-133 residues and ADH_zinc_N pattern is known as *zinc-binding dehydrogenase* was matched on 163-302 residues in BtADH protein is shown in Figure 4.

The obtained data from the motif scanning analysis showed that the BtADH protein is a zinc-binding, MDR superfamily, a tetrameric alcohol dehydrogenase with a GroES-like domain. All the results obtained are in agreement with the data in our previous cloning and characterization study of BtADH. There have been many studies and researches on Medium-chain dehydrogenase /reductase (MDR) superfamily proteins, which have a large group of enzymes [21]. However, adequate studies have not yet been conducted on the zinc-containing MDR superfamily proteins [22]. There are many areas of bioinformatics and biochemical studies to be investigated, especially regarding zinc-containing ADH proteins.

4. Conclusion

In this study, bioinformatics tools and databases were utilized to determine some biophysical and structural properties of BtADH protein. Physico-chemical analysis of BtADH showed that this protein has more negative charged residues and higher contents of Valine, Alanine, and Glycine amino acids.

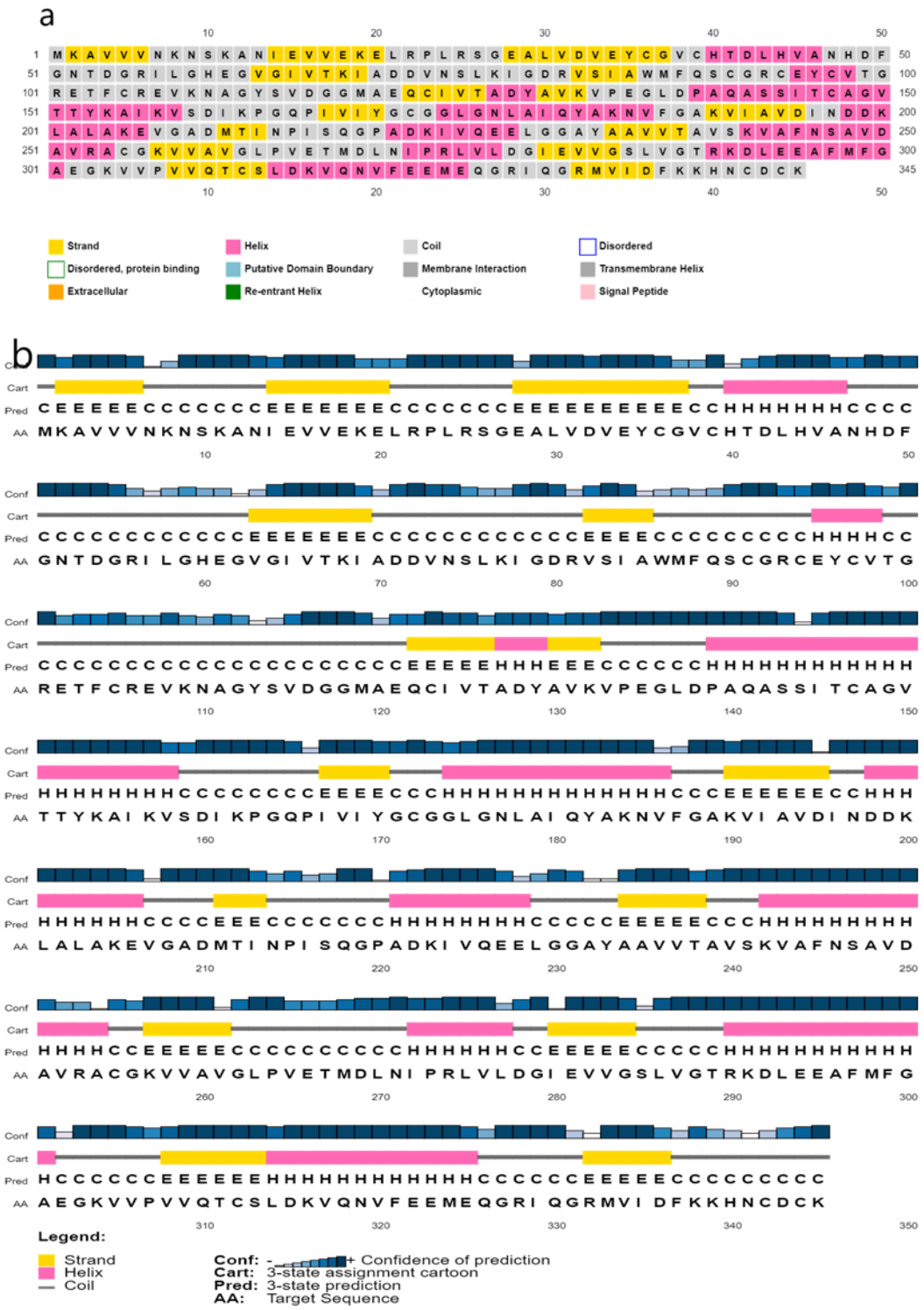

Figure 1. The secondary structure (α-helix, β-strand, and coil structures) prediction of BtADH protein was performed by using PSIPred. A: Amino acid residues in each secondary structure (β-strand (yellow), α- helix (pink), coil (grey)) B: 3-state assignment cartoon exhibition (β-strand (yellow), α- helix (pink), coil (grey)).

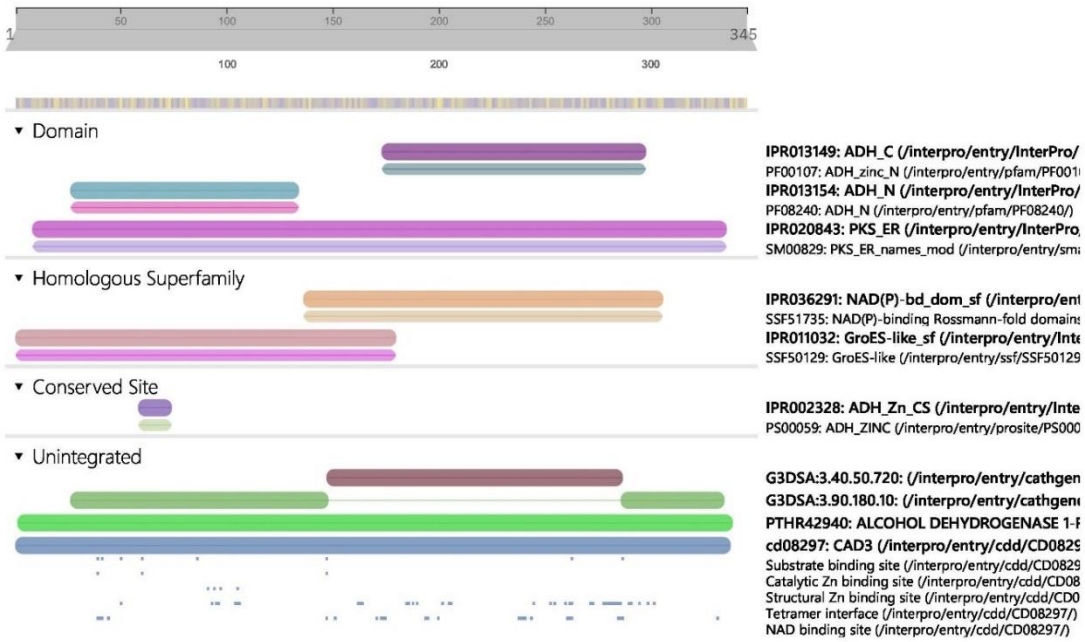

Figure 2. The functional analysis, predicting domains, and important sites of BtADH protein were performed by InterPro database.

```
FT   MYHIT        99    102       freq_pat:CK2_PHOSPHO_SITE [?]
FT   MYHIT       247    250       freq_pat:CK2_PHOSPHO_SITE [?]
FT   MYHIT       289    292       freq_pat:CK2_PHOSPHO_SITE [?]
FT   MYHIT        37     42       freq_pat:MYRISTYL [?]
FT   MYHIT        51     56       freq_pat:MYRISTYL [?]
FT   MYHIT       136    141       freq_pat:MYRISTYL [?]
FT   MYHIT       173    178       freq_pat:MYRISTYL [?]
FT   MYHIT       208    213       freq_pat:MYRISTYL [?]
FT   MYHIT       230    235       freq_pat:MYRISTYL [?]
FT   MYHIT       284    289       freq_pat:MYRISTYL [?]
FT   MYHIT        75     77       freq_pat:PKC_PHOSPHO_SITE [?]
FT   MYHIT        99    101       freq_pat:PKC_PHOSPHO_SITE [?]
FT   MYHIT       152    154       freq_pat:PKC_PHOSPHO_SITE [?]
FT   MYHIT       289    291       freq_pat:PKC_PHOSPHO_SITE [?]
FT   MYHIT         1    337       hamap:TDH_MF_00627 [!?]
FT   MYHIT        59     73       pat:ADH_ZINC [!]
FT   MYHIT       167    196       pre:NAD_BINDING [?]
FT   MYHIT       174    206       pre:SAM_BIND [?]
FT   MYHIT        80    115       prf:PPR [?]
FT   MYHIT       164    209       pfam_fs:2-Hacid_dh_C [?]
FT   MYHIT        26    133       pfam_fs:ADH_N [!]
FT   MYHIT       163    302       pfam_fs:ADH_zinc_N [!]
FT   MYHIT       301    326       pfam_fs:BRCA2 [?]
FT   MYHIT       171    210       pfam_fs:Methyltransf_11 [?]
FT   MYHIT       178    215       pfam_fs:TrkA_N [?]
FT   MYHIT        26    133       pfam_ls:ADH_N [!]
FT   MYHIT       163    302       pfam_ls:ADH_zinc_N [!]
FT   MYHIT       293    326       pfam_ls:BRCA2 [?]
```

Figure 3. Motif scan results of BtADH protein in HAMAP profiles [hamap], PROSITE patterns [pat], PROSITE patterns (frequent match producers) [freq_pat], Pfam HMMs (global models) [pfam_ls], Pfam HMMs (local models) [pfam_fs], and PROSITE profiles [prf] databases in MyHits tool developed by the Swiss Institute of Bioinformatics (SIB) Motif Scan Results (sib.swiss) . Legends**:** 1, freq_pat:CK2_PHOSPHO_SITE [?]; 2, freq_pat:MYRISTYL [?]; 3, freq_pat:PKC_PHOSPHO _SITE [?]; 4, pat:ADH_ZINC [!]; 5, pre:NAD_BINDING [?]; 6, pre:SAM_BIND [?]; 7, prf:PPR [?]; 8, pfam_fs:2-Hacid_dh_C [?]; 9, pfam_fs:BRCA2 [?]; 10, pfam_fs:Methyltransf_11 [?]; 11, pfam_fs:TrkA_N [?]; 12, pfam_ls:BRCA2 [?].

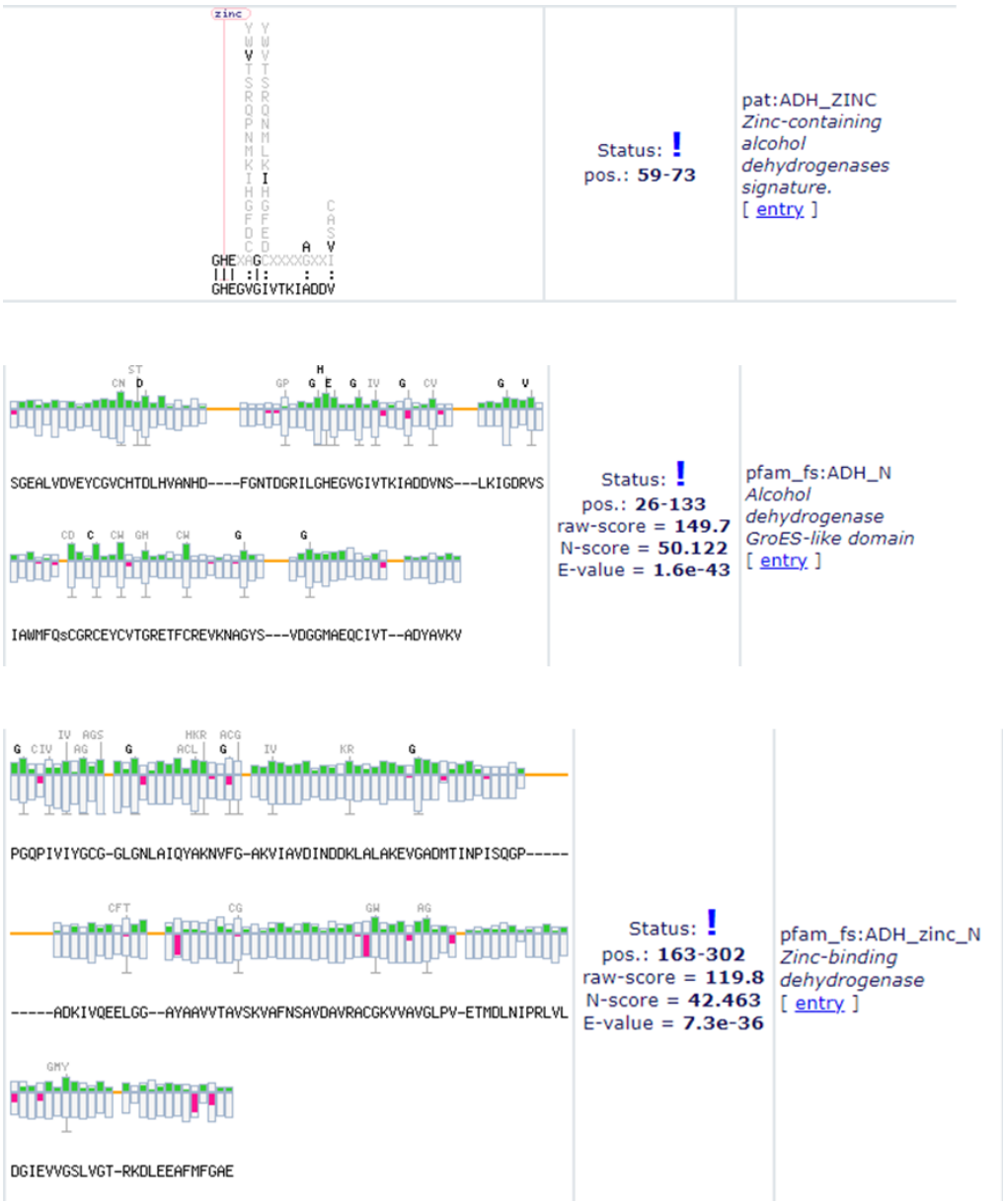

Figure 4. The motifs ADH_ZINC, ADH_N, and ADH_zinc_N are strongly matched with the BtADH protein sequence. Position, raw-score, N-score, and E-score values are given for all matches.

Secondary structure analysis of BtADH indicated that it has 11 α-helices and 15 β-strands. In functional analysis, it has been shown good homology with NAD(P)-binding domain superfamily (IPR036291), GroES-like superfamily (IPR011032), zinc ion binding (GO:0008270) site, and oxidoreductase activity (GO:0016491) site. Furthermore, In motif scan analysis of BtADH protein, it was found 3 strong match motifs as known as ADH_ZINC, ADH_N, and ADH_ZINC patterns.

Further protein structure-function study and immobilization to the appropriate matrix are planned with the site-directed mutations of the BtADH protein in the future.

References

1. Reid, M.F. and C.A. Fewson, *Molecular Characterization of Microbial Alcohol Dehydrogenases.* Critical Reviews in Microbiology, 1994. **20**(1): p. 13-56.
2. Shain, D.H., C. Salvadore, and C.L. Denis, *Evolution of the alcohol dehydrogenase (ADH) genes in yeast: characterization of a fourth ADH*

in Kluyveromyces lactis. Molecular and General Genetics MGG, 1992. **232**(3): p. 479-488.
3. Schwartz, D. and T. Endo, *ALCOHOL DEHYDROGENASE POLYMORPHISM IN MAIZE—SIMPLE AND COMPOUND LOCI.* Genetics, 1966. **53**(4): p. 709-715.
4. Sun, H.W. and B.V. Plapp, *Progressive sequence alignment and molecular evolution of the Zn-containing alcohol dehydrogenase family.* Journal of Molecular Evolution, 1992. **34**(6): p. 522-535.
5. Satre, M.A., M. Zgombić-Knight, and G. Duester, *The complete structure of human class IV alcohol dehydrogenase (retinol dehydrogenase) determined from the ADH7 gene.* J Biol Chem, 1994. **269**(22): p. 15606-12.
6. Schwartz, M., J. O'Donnell, and W. Sofer, *Origin of the multiple forms of alcohol dehydrogenase from Drosophila melanogaster.* Archives of Biochemistry and Biophysics, 1979. **194**(2): p. 365-378.
7. Jörnvall, H., J. Hempel, and B. Vallee, *Structures of Human Alcohol and Aldehyde Dehydrogenases.* Enzyme, 1987. **37**: p. 5-18.
8. Inoue, T., et al., *Cloning and sequencing of the gene encoding the 72-kilodalton dehydrogenase subunit of alcohol dehydrogenase from Acetobacter aceti.* 1989. **171**(6): p. 3115-3122.
9. Doggett, N.A., et al., *Complete Genome Sequence of Bacillus thuringiensis Serovar Israelensis Strain HD-789.* Genome announcements, 2013. **1**(6): p. e01023-13.
10. Méndez-López, I., R. Basurto-Ríos, and J.E. Ibarra, *Bacillus thuringiensis serovar israelensis is highly toxic to the coffee berry borer, Hypothenemus hampei Ferr. (Coleoptera: Scolytidae).* FEMS Microbiology Letters, 2003. **226**(1): p. 73-77.
11. Çelik, A. and F. Aktaş, *A new NADH-dependent, zinc containing alcohol dehydrogenase from Bacillus thuringiensis serovar israelensis involved in oxidations of short to medium chain primary alcohols.* Journal of Molecular Catalysis B: Enzymatic, 2013. **89**: p. 114-121.
12. Jörnvall, H., M. Landreh, and L.J. Östberg, *Alcohol dehydrogenase, SDR and MDR structural stages, present update and altered era.* Chemico-Biological Interactions, 2015. **234**: p. 75-79.
13. Karlsson, A., et al., *Tetrameric NAD-dependent alcohol dehydrogenase.* Chemico-Biological Interactions, 2003. **143-144**: p. 239-245.
14. Pruess, M. and R. Apweiler, *Bioinformatics Resources for In Silico Proteome Analysis.* J Biomed Biotechnol, 2003. **2003**(4): p. 231-236.
15. Madeira, F., et al., *The EMBL-EBI search and sequence analysis tools APIs in 2019.* Nucleic acids research, 2019. **47**(W1): p. W636-W641.
16. Wilkins, M.R., et al., *Protein identification and analysis tools in the ExPASy server.* Methods in molecular biology (Clifton, N.J.), 1999.
17. Buchan, D.W.A. and D.T. Jones, *The PSIPRED Protein Analysis Workbench: 20 years on.* Nucleic Acids Research, 2019. **47**(W1): p. W402-W407.
18. Zdobnov, E.M. and R. Apweiler, *InterProScan – an integration platform for the signature-recognition methods in InterPro.* Bioinformatics, 2001. **17**(9): p. 847-848.
19. Pagni, M., et al., *MyHits: improvements to an interactive resource for analyzing protein sequences.* Nucleic Acids Res, 2007. **35**(Web Server issue): p. W433-7.
20. Taneja, B. and S.C. Mande, *Conserved structural features and sequence patterns in the GroES fold family.* Protein engineering, 1999. **12**(10): p. 815-818.
21. Jörnvall, H., et al., *Origin and evolution of medium chain alcohol dehydrogenases.* Chemico-Biological Interactions, 2013. **202**(1): p. 91-96.
22. Tiwari, M.K., et al., *Role of Conserved Glycine in Zinc-dependent Medium Chain Dehydrogenase/Reductase Superfamily*.* Journal of Biological Chemistry, 2012. **287**(23): p. 19429-19439.