

Original Research Article

In silico Prediction and Analysis of Domains, DNA Binding and Antioxidant Proteins of *Aloe vera* (*A. barbadensis*)

Asad Amir^{1*}, Virta Tomar¹, Neelesh Kapoor², Hirdesh Kumar¹ and Mohd. Asif Siddiqui¹

¹Department of Biotechnology and Microbiology, MIET, Meerut, U.P., India

²Department of Fingerprinting, College of Biotechnology, SVPUAT, Meerut, U.P., India

*Corresponding author

ABSTRACT

DNA binding proteins significantly works for controlling protein production, regulating cell division and growth, and storing DNA inside the nucleus. Antioxidant proteins have been found closely associated to disease control for its ability to eliminate excess free radicals. The objective of present study is to perform *In silico* analysis and find out possible DNA binding and antioxidant proteins of *Aloe vera* (*A. barbadensis*). The *Aloe vera* plant has been identified for its health, beauty, medicinal and skin care properties. *A. barbadensis* is well known for its various medicinal uses. Despite enormous pharmacological applications, very less genomic and molecular information is available for this plant. Still huge information of *Aloe vera* need to study and identify its hidden and valuable information. So, one need to reveal its genetic constitution. Therefore, *In silico* analysis and functional characterization of *A. barbadensis* was carried out. Total 133 proteins were found from *A. barbadensis* using UniProtKB protein database. Functional characterization of these proteins using ProtParam tool revealed 13 proteins, subjected to DNA binding and Anti-oxidant properties. Further analysis includes domain identifications and analysis using Prosite, secondary structure prediction using GOR.

Keywords

A. barbadensis,
UniProtKB,
Domain, Prosite,
ProtParam, GOR

Introduction

The *Aloe vera* plant has been known and used for centuries for its curative and therapeutic, health, beauty and skin care properties, over 75 active ingredients of its gel have been identified (Habeeb *et al.*, 2007). Most of the medicinal effects of aloe leaf extracts have been recognized to the polysaccharides found in the inner leaf parenchymatous tissue (Ni *et al.*, 2004).

Aloe vera is the most commercialized aloe species and processing of its leaf pulp has become a large global industry. It has been

used in the food industry, as a source of functional foods and as an ingredient in other food products, for the production of gel-containing health drinks. In the toiletry and cosmetic industry, for the production of soaps, shampoos, creams, facial cleansers, lotions, and other products. In the drug industry, it has been used for the making of topical products such as ointments and gel preparations, as well as in the production of tablets and capsules (Eshun and He, 2004) (He *et al.*, 2004). Significant pharmaceutical properties that have recently been discovered for both the *Aloe vera* gel and its leaf extract (Vinson *et al.*, 2005). Due to its absorption

enhancing effects, *A. vera* gel may be used to efficiently deliver poorly absorbable drugs through the oral route of drug administration. The parched powder obtained from *Aloe vera* gel was effectively used in the production of directly compressible matrix type tablets. These tablets gradually released a model compound over an extended period of time and thereby showing potential to be used as an excipient in modified release dosage forms (Jani *et al.*, 2007).

The genus *Aloe* (family *Asphodelaceae*) includes around 360 species. *Aloe barbadensis* Miller, *A. arborescence*, *A. ferox* and *A. perryi* are known to reveal medicinal properties. *A. barbadensis*, predominantly, is well known tropical luscious plant that has been known and grown for their enormous pharmaceutical importance since thousands of years (Pugh *et al.*, 2001; Chen *et al.*, 2012). *A. barbadensis* have been known for their anti-hyperglycemic, anti-ulcer, antioxidant, anti-asthmatic, skin burn healing, hepatoprotective, anti-tumor, anti-inflammatory and antihypercholesterolemic properties (Borrelli and Izzo 2000; Langmead *et al.*, 2004; Lin *et al.*, 2006; Chandan *et al.*, 2007; Mikadi and Damak, 2008; Qadir, 2009; Tarameshloo *et al.*, 2012; Huseini *et al.*, 2012; Kumar *et al.*, 2013).

Huge amount of potential applications and very little genomic, proteomic and gene content related information is available for *A. barbadensis*. Such restrictions stop bio-prospecting at molecular or gene level of this plant. Important findings such as DNA binding, antioxidant proteins and prediction of proteins functional region can aid in revealing hidden unique information, which can utilize in developing the improved variety for the present food and drug industries. Present study is the analysis, identification and computational characterization of *A. barbadensis* proteins. *In*

silico functional characterization of respective proteins of transcript sequences and their Domain analysis were carried out.

Materials and Methods

Retrieval of all *Aloe vera* (*A. barbadensis*) Proteins

Protein sequences of *A. barbadensis* were retrieved from UniPort (www.uniprot.org) protein sequence database, we had retrieved all proteins of *A. barbadensis* by entered the name of aloe vera (*A. barbadensis*) in search box. The retrieved sequences were subjected to functional characterization (Figure 1).

Characterization of all proteins

Characterization of all protein was done to find their Molecular Function, Biological Process and Cellular Components (Figure 2a and 2b).

Selection of DNA Binding and Antioxidant Proteins

Molecular functions of each proteins provide the information whether a protein is DNA binding or Antioxidant (Figure 2a).

Retrieval of all protein sequences in fasta format

To perform analysis, Proteins sequences of all proteins of aloe vera (*A. barbadensis*) proteins, were retrieved in Fasta Format (Figure 3).

Prediction of Physico Chemical Parameter by using Prot Param (www.expasy.org/protparam)

Physico Chemical parameter involves: No. Of Amino Acids, Molecular Weight, Theoretical pI, Instability Index(Guruprasad

et al., 1990), Aliphatic Index (Ikai, 1980) and Gravy (Kyte and Doolittle, 1982) for each amino acids were predicted using online tool ProtParam (Gasteiger *et al.*, 2005) (Figure 4).

Protein Domains Prediction and Analysis

Prediction of protein Domains is important tool to predict the functional regions of proteins. Using Prosite (Online Domain prediction tool) (<https://prosite.expasy.org/cgi-bin/prosite>) all possible domains, with their functions were predicted (Christian *et al.*, 2009)(Figure 5).

Retrieval Secondary Structure elements

To inspect the secondary structure conformation of each sequence, we had predicted secondary structure of each sequence. Open EXPASY (<https://www.expasy.org>) select GOR to predict Secondary structure elements (Figure 6).

Results and Discussion

In silico analysis of Aloe vera (*A. barbadensis*) proteins has been done by retrieving all the proteins from UniProtKB (www.uniprot.org). Initially we had obtained 133 proteins of Aloe vera (*A. barbadensis*) (Table 1). Genomics information of each protein were retrieved by characterization of all 133 proteins was done using UniProtKB (<https://www.uniprot.org/help/uniprotkb>) by identifying their Molecular function, Biological Process, and Cellular components (Table 2). After analysis, proteins with accession number H2F9W4, H2F5S8, H2F9G7, A0A240EY87, H2F9B8, H2F9R5, H2F920, H2FB04, H6WR00, Q9MRT8, were found with unknown Molecular function, protein with accession number H2FD27, H2FDC0, H2FB04, A0A240SZ85, H2F5N1, P83233, B6V3J3, Q9MRT7, Q9MRT8,

Q84XV5, W6E8Y2 Biological Process, and proteins with accession number A0A240TNC9, A0A1U9ZPS6, A0A1U9ZPT0, A0A0U2SUV0, Q84XV5, W6E8Y2 were found to be unknown Cellular components. Proteins with accession number A0A2C9NK57, A0A2C9NKA5, A0A2C9NL35, A0A2C9NK43, A0A2C9NK44, A5HLY2, A0A126WZT7 and A0A126WYJ9 were found to be unknown Molecular function, Biological Process, and Cellular components (Table 2). DNA-binding proteins are the transcription factors binding proteins which control the transcription process, various polymerases, nucleases which cleave DNA molecules, and histones which are involved in chromosome packaging and transcription in the cell nucleus (*Pabo and Sauer, 1984*), An antioxidant protein is a molecule that can prevent free radicals from causing damages in organisms, So we filtered the DNA binding and antioxidant activity proteins, and found total nine DNA binding proteins and one antioxidant activity protein, proteins with accession number H2F6V7, H2F6Q9, A0A240EY83, A0A240EY84, A0A240EY84, H2F834, H2F747, Q002B8, B6V3J3 and D3W327, showed their molecular function as DNA binding proteins, proteins with accession number P83233, showed their molecular function as antioxidant activity (Table 3). To perform further analysis as protein sequence level, protein sequences of the selected 10 proteins were retrieved in Fasta Format (Table 4). Physicochemical characterization is crucial to characterize specific proteins. Further analysis was done by identifying their physicochemical parameters using Expasy's ProtParam tool, the results are shown in Table 5. Number of amino acids in a protein provides a clue about the complication of the 3D structure it attains. The maximum number of amino acids was found 1378, in the protein with accession no. H2F834 and least number of amino acid was 8, found in the protein

with accession no. P83233 (Table 5), similar results have been obtained by Rzeznicka *et al.*, 2010. Similarly proteins with accession no. H2F834 was found with maximum molecular weight i.e. 156638.90, and the protein with accession P83233 was found with least molecular weight i.e. 917.97 Dalton. The computed pI value of proteins H2F6V7, H2F6Q9, A0A240EY83, A0A240EY84, H2F834, H2F747, Q002B8, D3W327 is >7 which indicates that these proteins are basic, and the pI of all other proteins is <7 which reveals that these are acidic in character.

The retrieved instability index value suggest that protein whose instability index is less than 40 is consider as stable, and a value above 40 consider as unstable protein. This study result showed that the instability index of proteins H2F6V7, H2F6Q9, A0A240EY83, A0A240EY84, H2F747, Q002B8, and B6V3J3 is greater than 40, indicating that these proteins are unstable while the rest of proteins having the value less than 40, indicating that these proteins are stable.

The AI is a parameter for calculating thermal stability of a protein directly linking with the mole fraction of aliphatic side chains (Alanine, isoleucine, leucine, and valine) in the protein. In this work, high AI values of proteins (69.91–133.75) indicating high thermostability of these proteins.

The GRAVY value (Kyte and Doolittle, 1982) for a protein indicates the solubility of proteins, increasing positive score indicates a greater hydrophobicity. A low GRAVY value interprets that there is better interaction between protein and water. The GRAVY index of all proteins is ranging from -0.656 to 0.259.

Domain is the important position or small stretch of the sequence, which indicate the

functional regions of proteins. Present study involves the prediction and analysis of conserved domains present in each protein sequence (Table 6).

The results reveals that the proteins H2F6V7, A0A240EY83, A0A240EY84, H2F834, H2F747, P83233 did not contains domains, while the protein H2F6Q9 contains RNA polymerases domain, which have important role in transcription of genomic DNA into RNA, proteins Q002B8 and B6V3J3 contains AP2/ERF domain, which have specific role in plant growth and development, defence mechanism.

Protein secondary structure can be viewed as an information link that associates the primary sequence and tertiary structure.

Precise 8-state secondary structure prediction can significantly give more accurate and high resolution on structure-based properties analysis.

In the present study Secondary structure of all selected 10 proteins was predicted and analyzed using GOR (Secondary structure Prediction Tools). It is observed that in most of the proteins the percentage of the random coil was much higher than the percentage of other secondary structure elements such as helix, sheet, and turn.

This high coiled structural content might be due to the presence of flexible glycine and hydrophobic proline amino acids. Proline has a special property of creating links in polypeptide chains and disrupting ordered secondary structure, thus results in coiling (Vidhya *et al.*, 2018) (Table 7, Figure 6).

In silico technology is fast inexpensive techniques, which involves the use of online genomic and proteomics tools to perform biological data analysis.

Table.1 List of all proteins of Aloe vera (*A.barbadensis*).

S.NO	ENTRY	ENTRY NAME	PROTEIN NAME
1	H2FAV5	H2FAV5_ ALOVR	Photosystem II Protein D1(PSII D1Protein) (EC 1.10.3.9) (Photosystem II Q(B) Protein)
2	H2FBJ8	H2FBJ8_ ALOVR	Photosystem I P700 Chlorophyll a apoprotein A1 (EC 1.97.1.12) (PSI-A) (PsaA)
3	H2FBE9	H2FBE9_ ALOVR	Photosystem I P700 Chlorophyll a apoprotein A2 (EC 1.97.1.12) (PSI-B) (PsaB)
4	A0A240SZ31	A0A240EY80-ALOVR	Ribulose biphosphate carboxylase large chain (RuBisCO large subunit) (EC 4.1.1.39)
5	A0A240EY80	H2FAG0_ ALOVR	ATP synthase subunit c, chloroplastic (ATP synthase F(0) Sector subunit c)(ATPase subunit III) (F-type ATPasesubunit c) (Lipid-binding protein)
6	H2FAG0	H2F584_ ALOVR	Photosystem II D2 Protein (PSII D2 Protein) (EC 1.10.3.9) (Photosystem Q(A) Protein)
7	H2F584	H2F535_ ALOVR	Ribulose biphosphate carboxylase large chain (RuBisCO large subunit) (EC 4.1.1.39)
8	H2F535	H2FE10_ ALOVR	ATP synthase subunit c, chloroplastic (ATP synthase F(0) Sector subunit c) (ATPase subunit III) (F-type ATPasesubunit c) (Lipid-binding protein)
9	H2FE10	H2FAK8_ ALOVR	Acetyl-coenzyme A carboxylase carboxyl transferase subunit beta, chloroplastic (ACCcase subunit beta) (Acetyl-coA carboxylase carboxyltransferase subunit beta) (EC 6.4.1.2)
10	H2FAK8	A0A240SZ37_ ALOVR	Photosystem II CP43 reaction center protein (PSII 43 kDa protein) (protein CP-43)
11	A0A240SZ37	H2FBA1_ ALOVR	NAD(P)H-quinone oxidoreductase subunit K, chloroplastic (EC 1.6.5.-) (NAD(P)H dehydrogenase subunit K) (NADH-Plastoquinone oxidoreductase subunit K)
12	H2FBA1	H2FCD9_ ALOVR	Photosystem I iron-sulfur center (EC 1.97.1.12) (9 Kda polypeptide) (PSI-C) (Photosystem I subunitVII) (PsaC)
13	H2FCD9	H2FCT4_ ALOVR	NAD(P)H-quinone oxidoreductase subunit 2, chloroplastic (EC 1.6.5.-) (NAD(P)H dehydrogenase subunit 2) (NADH-Plastoquinone oxidoreductase subunit 2)
14	H2FCT4	H2F4E5_ ALOVR	NAD(P)H-quinone oxidoreductase subunit 4L, chloroplastic (EC 1.6.5.-) (NAD(P)H dehydrogenase subunit 4L) (NADH-Plastoquinone oxidoreductase subunit 4L)
15	H2F4E5	A0A240SZ77_ ALOVR	ATP synthesis subunit alpha, chloroplastic (EC 3.6.3.14) (ATP synthesis F1 sector subunit alpha) (F-ATPase subunit alpha)
16	A0A240SZ77	H2FC90_ ALOVR	NAD(P)H-quinone oxidoreductase subunit I, chloroplastic (EC 1.6.5.-) (NAD(P)H dehydrogenase subunit I) (NDH subunit I) (NADH-Plastoquinone oxidoreductase subunit I)
17	H2FC90	A0A240SZ58_ ALOVR	NAD(P)H-quinone oxidoreductase subunit I, chloroplastic (EC 1.6.5.-) (NAD(P)H dehydrogenase subunit I) (NDH subunit I) (NADH-Plastoquinone oxidoreductase subunit I)
18	A0A240SZ58	H2FC17_ ALOVR	NAD(P)H-quinone oxidoreductase subunit 1, chloroplastic (EC 1.6.5.-) (NAD(P)H dehydrogenase subunit I) (NDH subunit 1) (NADH-Plastoquinone oxidoreductase subunit 1)
19	H2FC17	H2FDG8_ ALOVR	Cytochrome b6

20	H2FDG8	H2FDL7_ALOVR	NAD(P)H-quinone oxidoreductase subunit 3, chloroplastic (EC 1.6.5.-) (NAD(P)H dehydrogenase subunit I) (NADH-Plastoquinone oxidoreductase subunit 3)
21	H2FDL7	AOA219Y1V4_ALOVR	NAD(P)H-quinone oxidoreductase subunit I, chloroplastic
22	AOA219Y1V4	H2FAB1_ALOVR	ATP synthase subunit b, chloroplastic
23	H2FAB1	H2FA62_ALOVR	Cytochrome b559 subunit alpha
24	H2FA62	H2F4J3_ALOVR	Cytochrome b559 subunit beta
25	H2F4J3	H2FD72_ALOVR	ATP synthase subunit beta, chloroplastic
26	H2FD72	A0A240SZ43_ALOVR	NAD(P)H-quinone oxidoreductase subunit 5, chloroplastic
27	A0A240SZ43	P84752_ALOVR	Cytochrome f
28	P84752	H2F4U0_ALOVR	Peroxidase A
29	H2F4U0	H2F6V7_ALOVR	ATP synthase subunit b, chloroplastic
30	H2F6V7	H2FAQ7_ALOVR	DNA-directed RNA polymerase subunit alpha
31	H2FAQ7	H2F6Q9_ALOVR	Photosystem II CP47 reaction center protein
32	H2F6Q9	H2FDR6_ALOVR	DNA-directed RNA polymerase subunit beta
33	H2FDR6	H2F4Y8_ALOVR	NAD(P)H-quinone oxidoreductase subunit J, chloroplastic
34	H2F4Y8	A0A240SZ48_ALOVR	ATP synthase subunit a, chloroplastic
35	A0A240SZ48	H2FBP6_ALOVR	Cytochrome b6-f complex subunit 5
36	H2FBP6	H2FBP6_ALOVR	Cytochrome b6-f complex subunit 8
37	A0A240EY83	A0A240EY83_ALOVR	DNA-directed RNA polymerase subunit beta"
38	A0A109P017	A0A109P017_ALOVR	Ribulose biphosphate carboxylase large chain
39	Q37166	Q37166_ALOVR	Ribulose biphosphate carboxylase large chain
40	Q7YJP1	Q7YJP1_ALOVR	Ribulose biphosphate carboxylase large chain
41	Q6VW13	Q6VW13_ALOVR	Ribulose biphosphate carboxylase large chain
42	H2F4P1	H2F4P1_ALOVR	ATP synthase epsilon chain, chloroplastic
43	P84753	P84753_ALOVR	Peroxidase B
44	A0A240SZ74	A0A240SZ74_ALOVR	NAD(P)H-quinone oxidoreductase chain 4, chloroplastic
45	A0A240EY84	A0A240EY84_ALOVR	DNA-directed RNA polymerase subunit beta'
46	H2F834	H2F834_ALOVR	DNA-directed RNA polymerase subunit beta"
47	A0A219Y210	A0A219Y210_ALOVR	Protein TIC 214
48	A0A240SZ52	A0A240SZ52_ALOVR	Cytochrome b6-f complex subunit 4
49	H2F747	H2F747_ALOVR	DNA-directed RNA polymerase subunit beta'
50	H2FA13	H2FA13_ALOVR	Photosystem II reaction center protein H
51	H2F9L6	H2F9L6_ALOVR	Photosystem II reaction center protein K
52	H2FD27	H2FD27_ALOVR	ATP-dependent Clp protease proteolytic subunit
53	A0A240SZ64	A0A240SZ64_ALOVR	50S ribosomal protein L22, chloroplastic
54	H2F6CO	H2F6CO_ALOVR	30S ribosomal protein S4, chloroplastic
55	H2F9W4	H2F9W4_ALOVR	Photosystem II reaction center protein I
56	H2F5S8	H2F5S8_ALOVR	Protein TIC 214
57	H2F9G7	H2F9G7_ALOVR	Photosystem II reaction center protein L

58	A0A240EY87	A0A240EY87_ALOVR	Photosystem II reaction center protein M
59	H2F9B8	H2F9B8_ALOVR	Protein PsbN
60	H2F9R5	H2F4E5_ALOVR	Photosystem II reaction center protein J
61	Q9TMV2	Q9TMV2_ALOVR	ATP synthase subunit beta
62	H2FBU5	H2FBU5_ALOVR	Cytochrome c biogenesis protein CcsA
63	H2FDC0	H2FDC0_ALOVR	NAD(P)H-quinone oxidoreductase subunit 6, chloroplastic
64	A0A240SZ51	A0A240SZ51_ALOVR	30S ribosomal protein S12, chloroplastic
65	H2F671	H2F671_ALOVR	30S ribosomal protein S7, chloroplastic
66	H2F8M5	H2F8M5_ALOVR	50S ribosomal protein L14, chloroplastic
67	H2F920	H2F920_ALOVR	Photosystem II reaction center protein Z
68	H2FB04	H2FB04_ALOVR	Photosystem I reaction center subunit IX
69	H2F5X8	H2F5X8_ALOVR	30S ribosomal protein S8, chloroplastic
70	H2F7T9	H2F7T9_ALOVR	30S ribosomal protein S14, chloroplastic
71	H2F8H9	H2F8H9_ALOVR	50S ribosomal protein L23, chloroplastic
72	A0A240TNC9	A0A240TNC9_ALOVR	50S ribosomal protein L2, chloroplastic
73	H2F8S4	H2F8S4_ALOVR	50S ribosomal protein L20, chloroplastic
74	H2FC42	H2FC42_ALOVR	Maturase K
75	H2FBZ3	H2FBZ3_ALOVR	envelope membrane protein, chloroplastic
76	H2F512	H2F512_ALOVR	Photosystem I assembly protein Ycf3
77	H2FB52	H2FB52_ALOVR	Photosystem I reaction center subunit VIII
78	H2F969	H2F969_ALOVR	Photosystem II reaction center protein T
79	H2F8X1	H2F8X1_ALOVR	50S ribosomal protein L2, chloroplastic
80	H2F5D3	H2F5D3_ALOVR	Photosystem I assembly protein Ycf4
81	Q7YJJ7	Q7YJJ7_ALOVR	Maturase K
82	Q6VVX6	Q6VVX6_ALOVR	Maturase K
83	H2F7E3	H2F7E3_ALOVR	30S ribosomal protein S19, chloroplastic
84	H2F7J2	H2F7J2_ALOVR	30S ribosomal protein S18, chloroplastic
85	H2F6G9	H2F6G9_ALOVR	30S ribosomal protein S3, chloroplastic
86	H2F7Y8	H2F7Y8_ALOVR	30S ribosomal protein S11, chloroplastic
87	H2F796	H2F796_ALOVR	50S ribosomal protein L16, chloroplastic
88	H2F7P0	H2F7P0_ALOVR	30S ribosomal protein S15, chloroplastic
89	R4NGH0	R4NGH0_ALOVR	Maturase K
90	A0A109P005	A0A109P005_ALOVR	Maturase K
91	A0A0X9CVS1	A0A0X9CVS1_ALOVR	Maturase K
92	A0A240SZ85	A0A240SZ85_ALOVR	Protein Ycf2
93	D3W166	D3W166_ALOVR	Maturase K
94	D3W167	D3W167_ALOVR	Maturase K
95	A0A1B1XY94	A0A1B1XY94_ALOVR	Maturase K
96	D3W163	D3W163_ALOVR	Maturase K

97	G3KG51	G3KG51_ALOVR	Maturase K
98	A0A1X9ZIA8	A0A1X9ZIA8_ALOVR	Maturase K
99	D3W164	D3W164_ALOVR	Maturase K
100	H2F5N1	H2F5N1_ALOVR	Protein Ycf2
101	H2F6L1	H2F6L1_ALOVR	30S ribosomal protein S2, chloroplastic
102	A0A240EY76	A0A240EY76_ALOVR	30S ribosomal protein S16, chloroplastic
103	H2F881	H2F881_ALOVR	50S ribosomal protein L36, chloroplastic
104	H2F8D0	H2F8D0_ALOVR	50S ribosomal protein L33, chloroplastic
105	P83233	P83233_ALOVR	Verectin
106	Q5MFB5	Q5MFB5_ALOVR	ATP synthase F0 subunit 9
107	Q002B8	Q002B8_ALOVR	Putative DRE-binding protein DREB1
108	A0A0C512T7	A0A0C512T7_ALOVR	HVA22-like protein
109	B6V3J3	B6V3J3_ALOVR	Dehydration-responsive element binding protein
110	H2FCN5	H2FCN5_ALOVR	NADH-plastoquinone oxidoreductase subunit 4
111	D3W1Y9	D3W1Y9_ALOVR	Ycf5
112	Q8W318	Q8W318_ALOVR	Phosphoenolpyruvate carboxylase, isoform 1
113	Q8W319	Q8W319_ALOVR	Phosphoenolpyruvate carboxylase, isoform 1
114	H2FCY2	H2FCY2_ALOVR	Ribosomal protein S16
115	D3W327	D3W327_ALOVR	RNA polymerase C
116	A0A1U9ZPS8	A0A1U9ZPS8_ALOVR	Ribulose-1,5-bisphosphate carboxylase/oxygenase large subunit
117	A0A1U9ZPS6	A0A1U9ZPS6_ALOVR	Ribulose-1,5-bisphosphate carboxylase/oxygenase large subunit
118	A0A1U9ZPT0	A0A1U9ZPT0_ALOVR	Ribulose-1,5-bisphosphate carboxylase/oxygenase large subunit
119	Q9MRT7	Q9MRT7_ALOVR	NADH dehydrogenase subunit 4
120	H6WR00	H6WR00_ALOVR	ATP synthase subunit b
121	A0A0U2SUV0	A0A0U2SUV0_ALOVR	Glycerol kinase-like protein
122	Q9MRT8	Q9MRT8_ALOVR	PSI 9 kDa protein
123	Q84XV5	Q84XV5_ALOVR	NADP-malic enzyme
124	W6E8Y2	W6E8Y2_ALOVR	Gag protease polyprotein
125	A0A2C9NK57	A0A2C9NK57_ALOVR	Ycf1
126	A0A2C9NKA5	A0A2C9NKA5_ALOVR	Ribosomal protein L16
127	A0A2C9NL35	A0A2C9NL35_ALOVR	Ycf3
128	A0A2C9NK43	A0A2C9NK43_ALOVR	PetL
129	A0A2C9NK44	A0A2C9NK44_ALOVR	ClpP
130	A5HLY2	A5HLY2_ALOVR	Ubiquitin
131	A0A126WYZ0	A0A126WYZ0_ALOVR	Putative LOV domain-containing protein
132	A0A126WZT7	A0A126WZT7_ALOVR	Putative LOV domain-containing protein
133	A0A126WYJ9	A0A126WYJ9_ALOVR	Putative LOV domain-containing protein

Table.2 Characterization of all 133 proteins of *A. barbadensis*

S.No.	Entry No.	Molecular Function	Biological Process	Cellular Component
1	H2FAV5	chlorophyll binding electron transporter, transferring electrons within the cyclic electron transport pathway of photosynthesis activity iron ion binding oxidoreductase activity, acting on diphenols and related substances as donors, oxygen as acceptor	photosynthetic electron transport in photosystem II protein-chromophore linkage response to herbicide	chloroplast
2	H2FBJ8	4 iron, 4 sulfur cluster binding chlorophyll binding electron transfer activity magnesium ion binding	photosynthesis protein-chromophore linkage	chloroplast
3	H2FBE9	4 iron, 4 sulfur cluster binding chlorophyll binding electron transfer activity magnesium ion binding	photosynthesis protein-chromophore linkage	chloroplast
4	A0A240SZ3	magnesium ion binding monooxygenase activity ribulose-bisphosphate carboxylase activity	photorespiration reductive pentose-phosphate cycle	chloroplast
5	A0A240EY80	lipid binding proton-transporting ATP synthase activity, rotational mechanism ATP hydrolysis coupled proton transport	ATP hydrolysis coupled proton transport ATP synthesis coupled proton transport	chloroplast
6	H2FAG0	chlorophyll binding electron transporter, transferring electrons within the cyclic electron transport pathway of photosynthesis activity iron ion binding	photosynthetic electron transport in photosystem II protein-chromophore linkage	chloroplast
7	H2F584	magnesium ion binding monooxygenase activity ribulose-bisphosphate carboxylase activity	photorespiration reductive pentose-phosphate cycle	chloroplast
8	H2F535	lipid binding proton-transporting ATP synthase activity, rotational mechanism	ATP hydrolysis coupled proton transport ATP synthesis coupled proton transport	chloroplast
9	H2FE10	acetyl-CoA carboxylase activity ATP binding zinc ion binding	fatty acid biosynthetic process malonyl-CoA biosynthetic process	chloroplast
10	H2FAK8	chlorophyll binding electron transporter, transferring electrons within the cyclic electron transport pathway of photosynthesis activity metal ion binding	photosynthetic electron transport in photosystem II protein-chromophore linkage	chloroplast
11	A0A240SZ37	4 iron, 4 sulfur cluster binding iron ion binding NADH dehydrogenase (ubiquinone) activity quinone binding	photosynthesis, light reaction transport	chloroplast
12	H2FBA1	4 iron, 4 sulfur cluster binding electron transfer activity metal ion binding	photosynthetic electron transport in photosystem I	chloroplast

13	H2FCD9	NADH dehydrogenase (ubiquinone) activity quinone binding	ATP synthesis coupled electron transport photosynthesis, light reaction transport	chloroplast
14	H2FCT4	oxidoreductase activity, acting on NAD(P)H, quinone or similar compound as acceptor quinone binding	ATP synthesis coupled electron transport photosynthesis, light reaction transport	chloroplast
15	H2F4E5	ATP binding proton-transporting ATP synthase activity, rotational mechanism	ATP synthesis coupled proton transport	chloroplast
16	A0A240SZ77	4 iron, 4 sulfur cluster binding iron ion binding NADH dehydrogenase (ubiquinone) activity quinone binding	photosynthesis, light reaction	chloroplast
17	H2FC90	oxidoreductase activity, acting on NAD(P)H, quinone or similar compound as acceptor quinone binding	photosynthesis, light reaction	chloroplast
18	A0A240SZ58	electron transporter, transferring electrons within cytochrome b6/f complex of photosystem II activity heme binding iron ion binding	photosynthesis respiratory electron transport chain	chloroplast
19	H2FCI7	NADH dehydrogenase (ubiquinone) activity S quinone binding	photosynthesis, light reaction transport	chloroplast
20	H2FDG8	NAD binding oxidoreductase activity, acting on NAD(P)H, quinone or similar compound as acceptor quinone binding	photosynthesis, light reaction transport	chloroplast
21	H2FDL7	4 iron, 4 sulfur cluster binding iron ion binding NADH dehydrogenase (ubiquinone) activity quinone binding	photosynthesis, light reaction	chloroplast
22	A0A219Y1V4	proton-transporting ATP synthase activity, rotational mechanism	ATP synthesis coupled proton transport	chloroplast
23	H2FAB1	electron transfer activity heme binding iron ion binding	photosynthetic electron transport chain	chloroplast
24	H2FA62	electron transfer activity heme binding iron ion binding	photosynthetic electron transport chain	chloroplast
25	H2F4J3	ATP binding proton-transporting ATP synthase activity, rotational mechanism	ATP synthesis coupled proton transport	chloroplast
26	H2FD72	NADH dehydrogenase (ubiquinone) activity quinone binding	ATP synthesis coupled electron transport transport	chloroplast
27	A0A240SZ43	electron transfer activity heme binding iron ion binding	photosynthesis	chloroplast

28	P84752	heme binding peroxidase activity	hydrogen peroxide catabolic process response to oxidative stress	chloroplast
29	H2F4U0	proton-transporting ATP synthase activity , rotational mechanism	ATP synthesis coupled proton transport	chloroplast
30	H2F6V7	DNA binding DNA-directed 5'-3' RNA polymerase activity protein dimerization activity	transcription, DNA- templated	chloroplast
31	H2FAQ7	chlorophyll binding electron transporter, transferring electrons within the cyclic electron transport pathway of photosynthesis activity	photosynthetic electron transport in photosystem II protein-chromophore linkage	chloroplast
32	H2F6Q9	DNA binding DNA-directed 5'-3' RNA polymerase activity ribonucleoside binding	transcription, DNA- templated	chloroplast
33	H2FDR6	NADH dehydrogenase (ubiquinone) activity quinone binding	photosynthesis, light reaction transport	chloroplast
34	H2F4Y8	proton-transporting ATP synthase activity, rotational mechanism	ATP synthesis coupled proton transport	chloroplast
35	A0A240SZ48	electron transporter, transferring electrons within cytochrome b6/f complex of photosystem II activity	cytochrome complex assembly photosynthesis	chloroplast
36	H2FBP6	electron transporter, transferring electrons within cytochrome b6/f complex of photosystem II activity	cytochrome complex assembly photosynthesis	chloroplast
37	A0A240EY83	DNA binding DNA-directed 5'-3' RNA polymerase activity	transcription, DNA- templated	chloroplast
38	A0A109P017	magnesium ion binding monooxygenase activity ribulose-bisphosphate carboxylase activity	photorespiration reductive pentose- phosphate cycle	chloroplast
39	Q37166	magnesium ion binding monooxygenase activity ribulose-bisphosphate carboxylase activity	photorespiration reductive pentose- phosphate cycle	chloroplast
40	Q7YJP1	magnesium ion binding monooxygenase activity ribulose-bisphosphate carboxylase activity	photorespiration reductive pentose- phosphate cycle	chloroplast
41	Q6VW13	magnesium ion binding monooxygenase activity ribulose-bisphosphate carboxylase activity	photorespiration reductive pentose- phosphate cycle	chloroplast
42	Q6VW13	ATP binding proton-transporting ATP synthase activity, rotational mechanism	ATP synthesis coupled proton transport	chloroplast
43	H2F4P1	ATP binding proton-transporting ATP synthase activity, rotational mechanism	ATP synthesis coupled proton transport	chloroplast
44	P84753	peroxidase activity	hydrogen peroxide catabolic process	chloroplast
45	A0A240SZ74	NADH dehydrogenase (ubiquinone) activity quinone binding	ATP synthesis coupled electron transport	chloroplast

46	A0A240EY84	DNA binding DNA-directed 5'-3' RNA polymerase activity	transcription, DNA-templated	chloroplast
47	H2F834	DNA binding DNA-directed 5'-3' RNA polymerase activity	transcription, DNA-templated	chloroplast
48	A0A219Y210	Unknown	protein transport	chloroplast
49	A0A240SZ52	electron transporter, transferring electrons within cytochrome b6/f complex of photosystem II activity electron transporter, transferring electrons within the cyclic electron transport pathway of photosynthesis activity	photosynthetic electron transport chain	chloroplast
50	H2F747	DNA binding DNA-directed 5'-3' RNA polymerase activity	transcription, DNA-templated	chloroplast
51	H2FA13	phosphate ion binding	photosynthesis protein stabilization	chloroplast
52	H2F9L6	Unknown	photosynthesis	chloroplast
53	H2FD27	serine-type endopeptidase activity	Unknown	chloroplast
54	A0A240SZ64	rRNA binding structural constituent of ribosome	translation	chloroplast
55	H2F6C0	rRNA binding structural constituent of ribosome	translation	chloroplast
56	H2F9W4	Unknown	photosynthesis	chloroplast
57	H2F5S8	Unknown	protein transport	chloroplast
58	H2F9G7	Unknown	photosynthesis	chloroplast
59	A0A240EY87	Unknown	photosynthesis, light reaction	chloroplast
60	H2F9B8	Unknown	photosynthesis	chloroplast
61	H2F9R5	Unknown	photosynthesis	chloroplast
62	Q9TMV2	ATP binding proton-transporting ATP synthase activity, rotational mechanism	ATP synthesis coupled proton transport	chloroplast
63	H2FBU5	heme binding	cytochrome complex assembly	chloroplast
64	H2FDC0	NADH dehydrogenase (ubiquinone) activity quinone binding	Unknown	chloroplast
65	A0A240SZ51	rRNA binding structural constituent of ribosome	translation	chloroplast
66	H2F671	rRNA binding structural constituent of ribosome	translation	chloroplast
67	H2F8M5	rRNA binding structural constituent of ribosome	translation	chloroplast
68	H2F920	Unknown	photosynthesis photosystem II stabilization	chloroplast
69	H2FB04	Unknown	Unknown	chloroplast
70	H2F5X8	rRNA binding structural constituent of ribosome	translation	chloroplast
71	H2F7T9	rRNA binding structural constituent of ribosome	translation	chloroplast

72	H2F8H9	rRNA binding structural constituent of ribosome	translation	chloroplast
73	A0A240TNC9	rRNA binding structural constituent of ribosome transferase activity	translation	Unknown
74	H2F8S4	rRNA binding structural constituent of ribosome	ribosomal large subunit assembly translation	chloroplast
75	H2FC42	RNA binding	mRNA processing RNA splicing tRNA processing	chloroplast
76	H2FBZ3	hydrogen ion transmembrane transporter activity	Unknown	chloroplast
77	H2F5I2	Unknown	photosynthesis	Unknown
78	H2FB52	Unknown	photosynthesis	chloroplast
79	H2F969	Unknown	photosynthesis	chloroplast
80	H2F8X1	rRNA binding structural constituent of ribosome transferase activity	translation	chloroplast
81	H2F5D3	Unknown	photosynthesis	chloroplast
82	Q7YJJ7	RNA binding	mRNA processing RNA splicing tRNA processing	chloroplast
83	Q6VVX6	RNA binding	mRNA processing RNA splicing tRNA processing	chloroplast
84	H2F7E3	rRNA binding structural constituent of ribosome	translation	chloroplast
85	H2F7J2	rRNA binding structural constituent of ribosome	translation	chloroplast
86	H2F6G9	rRNA binding structural constituent of ribosome	translation	chloroplast
87	H2F7Y8	rRNA binding structural constituent of ribosome	translation	chloroplast
88	H2F796	rRNA binding structural constituent of ribosome	translation	chloroplast
89	H2F7P0	structural constituent of ribosome	translation	chloroplast
90	R4NGH0	RNA binding	mRNA processing tRNA processing	chloroplast
91	A0A109P005	RNA binding	mRNA processing tRNA processing	chloroplast
92	A0A0X9CVS1	RNA binding	mRNA processing tRNA processing	chloroplast
93	A0A240SZ85	ATP binding	Unknown	chloroplast
94	D3W166	RNA binding	mRNA processing tRNA processing	chloroplast
95	D3W167	RNA binding	mRNA processing tRNA processing	chloroplast

96	A0A1B1XY94	RNA binding	mRNA processing tRNA processing	chloroplast
97	D3W163	RNA binding	mRNA processing tRNA processing	chloroplast
98	G3KG51	RNA binding	mRNA processing tRNA processing	chloroplast
99	A0A1X9ZIA8	RNA binding	mRNA processing tRNA processing	chloroplast
100	D3W164	RNA binding	mRNA processing tRNA processing	chloroplast
101	H2F5N1	ATP binding	Unknown	chloroplast
102	H2F6L1	structural constituent of ribosome	translation	chloroplast
103	A0A240EY76	structural constituent of ribosome	translation	chloroplast
104	H2F881	structural constituent of ribosome	translation	chloroplast
105	H2F8D0	structural constituent of ribosome	translation	chloroplast
106	P83233	antioxidant activity	Unknown	chloroplast
107	Q5MFB5	hydrogen ion transmembrane transporter activity	ATP hydrolysis coupled proton transport ATP synthesis coupled proton transport	chloroplast
108	Q002B8	DNA binding transcription factor activity, sequence-specific DNA binding	transcription, DNA- templated	chloroplast
109	A0A0C5I2T7	Unknown	response to abscisic acid response to stress	chloroplast
110	B6V3J3	DNA binding transcription factor activity, sequence-specific DNA binding	Unknown	chloroplast
111	H2FCN5	NADH dehydrogenase (ubiquinone) activity	ATP synthesis coupled electron transport	chloroplast
112	D3W1Y9	heme binding	cytochrome complex assembly	chloroplast
113	Q8W3I8	phosphoenolpyruvate carboxylase activity	carbon fixation tricarboxylic acid cycle	chloroplast
114	Q8W3I9	phosphoenolpyruvate carboxylase activity	carbon fixation tricarboxylic acid cycle	chloroplast
115	H2FCY2	structural constituent of ribosome	translation	chloroplast
116	D3W327	DNA binding DNA-directed 5'-3' RNA polymerase activity	transcription, DNA- templated	chloroplast
117	A0A1U9ZPS8	magnesium ion binding ribulose-bisphosphate carboxylase activity	carbon fixation	chloroplast
118	A0A1U9ZPS6	magnesium ion binding ribulose-bisphosphate carboxylase activity	carbon fixation	unknown
119	A0A1U9ZPT0	magnesium ion binding ribulose-bisphosphate carboxylase activity	carbon fixation	unknown
120	Q9MRT7	NADH dehydrogenase (ubiquinone) activity	unknown	chloroplast

121	H6WR00	unknown	ATP synthesis coupled proton transport	Mitochondrion
122	A0A0U2SUV0	kinase activity phosphotransferase activity, alcohol group as acceptor	carbohydrate metabolic process	Unknown
123	Q9MRT8	Unknown	Unknown	Mitochondrion
124	Q84XV5	malate dehydrogenase (decarboxylating) (NAD+) activity	Unknown	Unknown
125	W6E8Y2	peptidase activity	Unknown	Unknown
126	A0A2C9NK57	Unknown	Unknown	Unknown
127	A0A2C9NKA5	Unknown	Unknown	Unknown
128	A0A2C9NL35	Unknown	Unknown	Unknown
129	A0A2C9NK43	Unknown	Unknown	Unknown
130	A0A2C9NK44	Unknown	Unknown	Unknown
131	A5HLY2	Unknown	Unknown	Unknown
132	A0A126WZT7	Unknown	Unknown	Unknown
133	A0A126WYJ9	Unknown	Unknown	Unknown

Table.3 DNA Binding proteins and antioxidant activity protein

S.No	Entry	Molecular Function	Biological Process	Cell Component
1	H2F6V7	DNA binding DNA-directed 5'-3' RNA polymerase activity protein dimerization activity	transcription, DNA-templated	chloroplast
2	H2F6Q9	DNA binding DNA-directed 5'-3' RNA polymerase activity ribonucleoside binding	transcription, DNA-templated	chloroplast
3	A0A240EY83	DNA binding DNA-directed 5'-3' RNA polymerase activity	transcription, DNA-templated	chloroplast
4	A0A240EY84	DNA binding DNA-directed 5'-3' RNA polymerase activity	transcription, DNA-templated	chloroplast
5	H2F834	DNA binding DNA-directed 5'-3' RNA polymerase activity	transcription, DNA-templated	chloroplast
6	H2F747	DNA binding DNA-directed 5'-3' RNA polymerase activity	transcription, DNA-templated	chloroplast
7	P83233	antioxidant activity	unknown	chloroplast
8	Q002B8	DNA binding transcription factor activity, sequence-specific DNA binding	transcription, DNA-templated	chloroplast
9	B6V3J3	DNA binding transcription factor activity, sequence-specific DNA binding	Unknown	chloroplast
10	D3W327	DNA binding DNA-directed 5'-3' RNA polymerase activity	transcription, DNA-templated	chloroplast

Table.4 Proteins sequences of selected 10 Proteins in fasta format were retrieved.

S. No.	Entry	SEQUENCE
1	H2F6V7	<p>>tr H2F6V7 H2F6V7_ALOVR DNA-directed RNA polymerase subunit alpha OS=Aloe vera GN=rpoA PE=3 SV=1 MVREEIAGSIRTLQWKCVESRIDSKRLYYGRFILSPLMKGQADTIGIAMRRALLGEMEGTCITRAKSEKVPHEYSTIVGIEESVHEILINLKEIVL RSNLYGVRDASICVRGPRYVTAQDIISPPSVEIVDTTQHIANLMEPIDLRIELQIKRDRGYRMNPTKNYQDGSYPIDAVSMPVRNANHGIHSYG NGNEKQEILFLEIWTNGSLTPKEALYEASRNLDLFIPLHAEEDIHFEENKNRFTLPFFTFQNRFTNLKKNKKGIPKLCIFIDQSELPSRTYNCL KRSNIHTLLDLLNKSQEDLMRIEYFRIEDVKQILDTLRKHFAINLPKNKFSI</p>
2	H2F6Q9	<p>>tr H2F6Q9 H2F6Q9_ALOVR DNA-directed RNA polymerase subunit beta OS=Aloe vera GN=rpoB PE=3 SV=1 MLRDGNEGMSAIPGFSQIQFEGFCRFINQGLSEEFHKFPKMEDRDQEMEFQLFVERYQLVEPLIKERDAVYESLTYSSSELYVPAGLIWKTGRD MQQQTVFIGNIPLMNSLGTIVNGIYRIVINQILQSPGIYRSELDHNGNSVYTSTIISDWGGRLELEIDRRTRIWARVSRKQKISILVLSSAMGL NLREILDNVCYPEIFLSFSNDKEKKKIGSKENAILEFYQQFACVGGDTVFSESLCKELQKKFFQRCCELGRVGRNMNRRLNLDIPQNTFLLP RDLLAAADHLIGMKFGMGTLDLDMNHLKKNRIRSVADLLQDQFGLALVRLNNAIRGTICAAIRHKLTPTPQNLVTSTSLTTTYESFFGLHPLSQ VLDRTNPLTQTVHGRKLSYLGPGGLTGRTASFRIRDIHPSHYGRICPIDTSEGINVGLIGSLAIHVRIHGWGSIESPFEISERSKEAQIYLSPNRD EYYMVAAGNSLALNRGIQEEQVVPARYRQEFRTIAWERIHFRSIFPFQYFSIGASLIPFIEHNDANRALMSSNMQRQAVPLSRSEKIVGTGLER QTALDSGVS VIAEREGKIIYTDTHKIISSNGDTISIPLVYQRSNKNTCMHQKPVQRQGCCKKQILADGAATVGGELALGKNVLVA YMPW EGYNSEDA VLISERLVYEDIYTSFHIRKYEIQTHVTSQGPERRITKEIPHLEAHLRLNLDNRNGIVMLGSWVEAGDILVGKLTPTANESSYAPEDR LLRAILGIQVSTAKETSLKLPIGGRGRVIDVRWIQKRWGSSYNPERIRVYISQKREIKVGDKVAGRHGNGKIISKILSRQDMPYLQDGTVPDMV FNPLGVPSRMNVGQIFECSLGLAGDLLKRHYRIVPFDERYEQEASRKLVFSELYEASKQTKNPWFPEPEYPGKSRIFDGRTPDFEQPVLIGKS YILKLIHQVDDKIHGRSSGHYALVTQQPLRGRAKQGGQRVGEVWALEGFGVAHILQEMLTYSKDHIRARQELLGAMIIGGTIPKPEDAPE SFRLLVRELRLSLALELNHFLVSEKNFQINRKEV</p>

3	A0A240EY83	<p>>tr A0A240EY83 A0A240EY83_ALOVR DNA-directed RNA polymerase subunit beta" OS=Aloe vera GN=rpoC2 PE=3 SV=1</p> <p>MAERADLVFHNKAIDGTAMKRLISRLIDHFGMAYTSHILDQVKTLGFGQATATSISLGIDDLLTIPSKGWLVDQAEQQSFILEKHHHCNVA VEKLRQSIEIWAYATSEYLRQEMNPNFRMTDPSNPVHLSFSGARGNASQVHQLVGMRLMSDPQGMIDLPIQSNLREGLSLTEYIISCYGAR KGVVDTAVRTSDAGYLTRRLVEVVQHIIVRKTDGCTIRGISVSPRNGMTEKIFVQSLIGRVLADDIYMGIRCIATRNDIGIGLANRFITFRAQPI YIRTPFCRNTSWICQLCYGRSPHGDVVELGEAVGIIAGQSIGEPGTQLTLRTRFHTGGVFTGGTAEHVVRAPSNGKIKFNEYWVHPTRTRRHGHP AFLCSIDLVTIESRDILHSVNIPPKSLILVQNDQYVESEQVIAEIRAGTSTFHFKERVRKHIIYSESEGEMHWSTDVYHAPEYTYSNVHLVPKTS HLWILAGGPCRSSIVSFLHKDQDQMNVHSFVNERYSITNNGRVRHKLLGPFKGDKEILDYSRLDRIISNGHWNFSPILQGNDFLAKR RRNRFLIPLQYDQEGEKELIPSGISIEIPINGILRRNSIFAYLDDPRYRKKSSGITKYGTVEVDSIVKKEDLIEYRGAKFNPKYQKYQMKVDRF FFIPEEVHILPGSSSIMVQNNIIGVNTQLALNTRSRVGGVLRVERKKKSIELKIFSGDIHFPGETDKISRHSILPPGTEKNSKESKNWKNWIY VQRITLTKKKYFVFRPVVITYE IADGINLAKVFPQDLLQEKNVQLRVVNYILYGNKSIREIYHTSIQLVRTCLVLNWDQEKNGSIEEMHSSFVEVRVNDLIRDFIRIELVKSTIS YTGKRYDTAGSGLIPDNGLDRTNINPFYSKVIPLFIQCQGTIGTFLNRNKEWQSLRILSSFNCYRVGPFNGSKYNNVTKESNPITPLTDLGLPL GTIVPKIVNFYSSYHLITHNQILLNKYWLLDNLKRTFQVLEIFKYCLIDENRRLYNPGPCNNILNPFHLNWCFLHNFSEETSTIISIGQFFCENI CLFKHGHPIKKSQILIVHVNSLVIRSAKPYLATLGATVHGHYGETLSEGDTLVTFIYAKSRSGDITQGLPKVEQILEVRSIGSISMNFERRFEG WNERIPGILGIPWESLIGAEITIAQSRISLVNKIQNVYRSQGVQIHNRIEIIIVRQVTSKVLVSEDGMSNVFSPPELIGLLRAERAGRALDEAICY RAILLGITRASLNTQSFISEASFQETARVLAKAALRGRIDWLKGLKENVVLGGVIPVGTGFKKLVRHRSQDKNIHLEIKKKNLFELEMRDILLH HREFFCSCVPNNFHDTLEKSFM</p>
4	A0A240EY84	<p>>tr A0A240EY84 A0A240EY84_ALOVR DNA-directed RNA polymerase subunit beta' OS=Aloe vera GN=rpoC1 PE=3 SV=1</p> <p>MIDRYKHQQLRIGLVSPQQIRAWANKILPNGEIVGEVTKPYTFHYKTNKPEKDGLFCERISGPIKSGICSCGNRYRVIGAEKEDPKFCEECGVEF VDSRIRRYQMGYIKLACPVTHVWYLRKRLPSYIANLLDKPLKELEGLVYCDVYLNFSFARPIAKKPTFLRLRGSFEYEIQSWKYSIPLFFTTPGFE KFRNREISTGAGAIQELANLDRRIIDNSLVEWKELGDEESAGNEWEDRKIRRRKDFLVRMELAKHFIRTNVEPERMVLCLLPVLPPELRPII QIDGGKLMSSDINELYRRVIYRNNTLTDLLATSRSSTPGELVMCQEKLVQEAVDTLDDNGIRGQPMRDGHNKVYKSFSDVIEGKEGRFRETLL GKRVDYSGRSVIVVGPSSLHQCGLPREIAIELFQTFVIRGLIRQHIASNIGIAKSKIREKEPIVWEILQEVMQHPVLLNRAPTLHRLGIQAFQPI LVEGRAICLHPLVCKGFNADFDGDQMAVHVPLSLEAQAEARLLMFSHMNLLSPAIGDPISVPTQDMLIGLYVLTMGNRRGICANRYNPSNCG NYQNKTVDNNNYQYTKKKEPYFCSSYDALGAYRQKRINLNSPFWLRWRLDQRGIGSTREVPVIEVQYESLGTYHEIYGHYLIVGSVKKEIRCIY IRTTVGHVSFYREIEEAIQGFCAYSYAI</p>

5	H2F834	>tr H2F834 H2F834_ALOVR DNA-directed RNA polymerase subunit beta" OS=Aloe vera GN=rpoC2 PE=3 SV=1 MEVLM AERADLVFHNK AIDGTAMKRLISRLIDHFGMAYTSHILDQVKT LGFQQATATSISLGIDDLTIPSKGWL VQDAEQQSFILEKHHHC NVH AVEKLRQSIWIYATSEYLRQEMNPNFRMTDPSNPVHLSFSGARGNASQVHQLVGM RGLMSDPQGMIDLPIQSNLREGLSLTEYIIS CYGARKGVVDTAVRTSDAGYLTRRLVEVVQHIIVRKTD CGTIRGISVSPRNGMTEKIFVQSLIGRVLADDIYMGIRCIATRNDIGIGLANR FIT FRAQPIYIRTPFTCRNTSWICQLCYGRSPTHGDLVELGEAVGIIAGQSIGEPGTQLTRTFHTGGVFTGGTAEHVRAPSNGKIKFNEYWVHPTR TRHGHPAFLCSIDLVTIESRDILHSVNIPPKSLILVQNDQYVESEQVIAEIRAGTSTFHFKERVRKHIYSESEGEMHWSTDVYHAPEYTYSNVH LVPKTSHLWILAGGPCRSSIVSFLSHKDQDQMNVHSFVNERYS DLSITNGRVRHKLLGPF GKDKKEILDYSRLDRIISNGHWNFPSILQGNS DFLAKRRRNRF LIPLQYDQEGEKELIPSGISIEIPINGILRRNSIFAYLDDPRYRKKSSGITKYGTVEVDSIVKKEDLIEYRGAKEFNPKYQKYQ MKVDRFFFIPEEVHILPGSSSIMVQNNIIIGVNTQLALNTRSRVGG LVRVERKKKSIELKIFSGDIHFGETDKISRHS GILIPPGETEKKNSKESKN WKNWIYVQRITLTKKKYFVVRPVVTYE IADGINLAKVFPQDLLQEKDNVQLRVVNYILYGN GKSIREIYHTSIQLVRTCLVLNWDQEKNGSI EEMHSSFVEVRVNDLIRDFIRIELVKSTISYTGKRYDTAGSGLIPDNGLDR TNINPFYSKVKIPLFIQCQGTIGTFLNRNKEWQSLRILSSFCYR VGPFGSKYNNVTKESNPITPLTDLGLPLGTIVPKIVNFYSSYHLITHNQILLNKYWLLDNLKRTFQVLEIFKYCLIDENRRLYNPGPCNNIILNP FHLNWCFLHNFSEETSTIISIGQFFCENICLFKHGPHIKKSGQILIVHNSLVIRSAKPYLATLGATVHGHYGETLSEGD TLVTFIYAKRSRGDIT QGLPKVEQILEVRSIGSISMNFERRFEGWNERIPGILGIPWESLIGAELTIAQSRISLVNKIQNVYRSQGVQIHN RHIEIIVRQVTSKVLVSEDGMS NVFSPEELIGLLRAERAGRALDEAICYRAILLGITRASLNTQSFISEASFQETARVLAKAALRGRIDWLKGLKENVVLGGVIPVGTGFKKL VHR SGQDKNIHLEIKKKNLFELEMRDILLH HREFFCSCVPNNFHD TLEKSFM
6	H2F747	>tr H2F747 H2F747_ALOVR DNA-directed RNA polymerase subunit beta' OS=Aloe vera GN=rpoC1 PE=3 SV=1 MIDRYKHQQLRIGLVSPQQIRAWANKILPNGEIVGEVTKPYTFHYKTNKPEKDGLFCERISGPIKSGICSCGNYRVIGAEKEDPKFCEECGVEF VDSRIRRYQMGYIKLACPVTHVWY LKRLPSYIANLLDKPLKELEGLVYCDYLNFSFARPIAKKPTFLRLRGSFEYEIQSWKYSIPLFFTPGF EK FRNREISTGAGAIREQLANLDLRIIIDNSLVEWKELGDEESAGNEWEDRKIRRRKDFLVRMELAKHFIRTNVEPERMVLCLLPVLPPELRPIIQI DGGKLMSSDINELYRRVIYRNNTLTDLLATSRS TP GELVMCQEKLVQEA VDTLLDNGIRGQPMRDGHNKVYKSFSDVIEGKEGRFRETLLGK RVDYSGRSVIVVGPSSLHQCGLPREIAIELFQTFVIRGLIRQH IASNIGIAKSKIREKEPIVWEILQEV MQGHPVLLNRAPTLHRLGIQAFQPI LV EGRAICLHPLVCKGFNADFDGDQMAVHVPLSLEAQAEARLLMFSHMNLLSPAIGDPISVPTQDMLIGLYVLTMGNNRRGICANRYNPSNCGNY QNKTVDNNNYQYTKKKEPYFCSSYDALGAYRQKRINLNSPFWLRWRLDQRGIGSTREVP IEVQYESLGTYHEIYGHYLIVG SVKKEIRCIYIR TTVGHVSFYREIEEAIQGF CRAYSIAI
7	P83233	>sp P83233 VTIN_ALOVR Verectin (Fragment) OS=Aloe vera PE=1 SV=1 DEDNVLLT
8	Q002B8	>tr Q002B8 Q002B8_ALOVR Putative DRE-binding protein DREB1 OS=Aloe vera GN=DREB1 PE=2 SV=1 MISRAMQARNYNLGVRSQPMKLA AAAAPPQPKPTKLYRGVQRHWGKWVAEIRLPKNRTRLWLGTFTAE EAALAYDKAAYKLRGDYA RLNFPHLKHTGAHLAPGGPLHSSVDAKLQAICQSLEQNKSSNSNSSKKEKRGDAVEEKSDKVVVVVAEGEESCSSSMNTGSASSPSSEIESL DFTEVPWDESEDFVLRKYPSWEIDWDAILS
9	B6V3J3	>tr B6V3J3 B6V3J3_ALOVR Dehydration-responsive element binding protein OS=Aloe vera GN=TINY PE=2 SV=1 MASEIEITKKNKGARPKRPRDTGNQTPCYRGV RMRAWGKWVSEIREPRKKSRIWLGTFPTPEMAARAH DV AALSIGPAAILNFPDLAASL PRPVLSPRDVQAAAASAAAMEPEGGGNAAAVEELGEIVELPKLDGELFDNLESSEFVCYDSSFECSFA YYHPPWMEGADDLGLLWDYKF
10	D3W327	>tr D3W327 D3W327_ALOVR RNA polymerase C (Fragment) OS=Aloe vera GN=rpoC1 PE=4 SV=1 ETLLGKRV DYSGRSVIVVGPSSLHQCGLPREIAIELFQTFVIRGLIRQH IASNIGIAKSKIREKEPIVWEILQEV MQGHPVLLNRAPTLHRLGIQA FQPILVEGRAICLHPLVCKGFNADFDGDQMAVHVPLSLEAQAEARLLMFSHMNLLSPAIGDPISV

Table.5 Identification physiochemical parameter using Protparam tool.

S.NO	ENTRY	PROTEIN	PARAMETER					
			NO. Of Amino Acids	Molecular Weight	Theoretical pI	Instability Index	Aliphatic Index	Gravy
1	H2F6V7	DNA-directed RNA polymerase subunit alpha	339	39231.22	8.66	50.66	95.49	-0.378
2	H2F6Q9	DNA-directed RNA polymerase subunit beta	1070	121375.14	8.97	43.89	92.20	-0.325
3	A0A240EY83	DNA-directed RNA polymerase subunit beta"	1374	156166.30	9.22	40.02	94.95	-0.248
4	A0A240EY84	DNA-directed RNA polymerase subunit beta'	687	78992.31	9.00	40.06	93.35	-0.301
5	H2F834	DNA-directed RNA polymerase subunit beta"	1378	156638.90	9.20	39.93	95.17	-0.243
6	H2F747	DNA-directed RNA polymerase subunit beta'	686	78893.18	9.00	40.21	93.06	-0.308
7	P83233	Verectin	8	917.97	3.49	32.83	133.75	-0.362
8	Q002B8	Putative DRE-binding protein DREB1	211	23447.33	8.45	55.12	69.91	-0.656
9	B6V3J3	Dehydration-responsive element binding protein	183	20059.73	5.33	46.52	70.55	-0.392
10	D3W327	RNA polymerase C	162	17882.03	7.29	39.26	119.75	0.259

Table.6 Prediction of Domains and their Analysis using Prosite

S.No.	Accession No.	Predicted Domain Accession No.	Domain Name	Domain Sequence	Function
1	H2F6V7	Not found	Not found	Not found	Not found
2	H2F6Q9	PS01166	RNA polymerases beta chain	GdKVAGrHGNGKI	necessary for the transcription of genomic DNA into RNA
3	A0A240EY83	Not found	Not found	Not found	Not found
4	A0A240EY84	Not found	Not found	Not found	Not found
5	H2F834	Not found	Not found	Not found	Not found
6	H2F747	Not found	Not found	Not found	Not found
7	P83233	Not found	Not found	Not found	Not found
8	Q002B8	PS51032	AP2/ERF domain	LYRGVQRHWGKWWAEIRLP- KNRTRLWLGTFDTAEEAALAYDKAAYKLRGDYARLNFP	plant growth and development, defence
9	B6V3J3	PS51032	AP2/ERF domain	CYRGVRMRAWGKWWSEIREP- RKKSRIWLGTFPTPEMAARAHDAALSIGPAAILNFP	plant growth and development, defence
10	D3W327	Not found	Not found	Not found	Not found

Table.7 Prediction Secondary Structure elements using GOR.

S.NO.	ACCESSION NO.	PROTEIN NAME	ALPHA HELIX (Hh)%	BETA BRIDGE (Bb)%	RANDOM COIL (Cc)%
1	H2F6V7	DNA-directed RNA polymerase subunit alpha	40.71%	0.00%	43.66%
2	H2F6Q9	DNA-directed RNA polymerase subunit beta	36.54%	0.00%	46.17%
3	A0A240EY83	DNA-directed RNA polymerase subunit beta"	25.55%	0.00%	50.00%
4	A0A240EY84	DNA-directed RNA polymerase subunit beta'	30.86%	0.00%	49.93%
5	H2F834	DNA-directed RNA polymerase subunit beta"	25.98%	0.00%	49.64%
6	H2F747	DNA-directed RNA polymerase subunit beta'	30.90%	0.00%	50.29%
7	P83233	Verectin	0.00%	0.00%	62.50%
8	Q002B8	Putative DRE-binding protein DREB1	22.75%	0.00%	56.40%
9	B6V3J3	Dehydration-responsive element binding protein	26.23%	0.00%	53.01%
10	D3W327	RNA polymerase C	46.30%	0.00%	43.83%

Fig.1 Retrieval of all alooe vera (*A. barbadensis*) Proteins using UniProt.

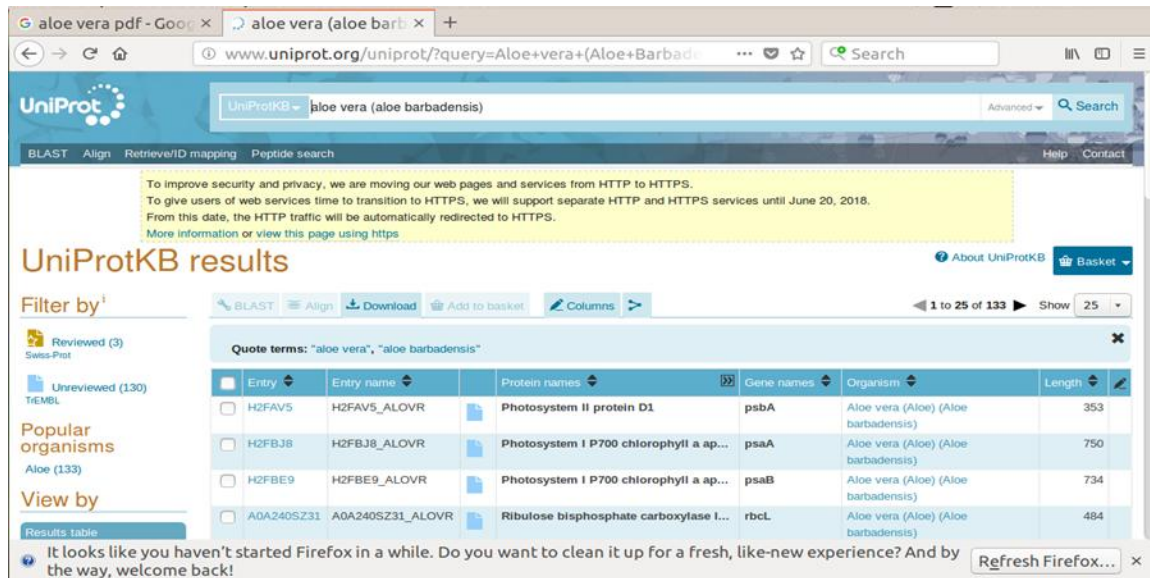


Fig.2a Showing Molecular functions and Biological process of alooe vera (*A. barbadensis*) Proteins.

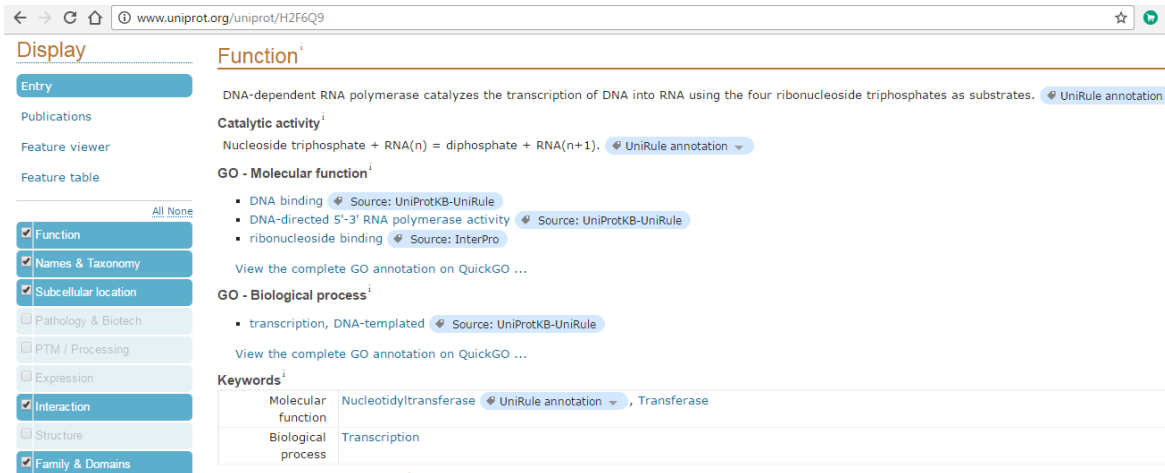


Fig.2b Showing Biological Process and Cellular Components of Aloe vera Proteins.

matK - Maturase K - Aloe vera (Aloe) - matK gene & protein - Mozilla Firefox

www.uniprot.org/uniprot/D3W166

Names & Taxonomy

Protein names¹ Recommended name: **Maturase K** SAAS annotation

Gene names¹ Name: **matK** Imported

Encoded on¹ Plastid: Chloroplast Imported

Organism¹ Aloe vera (Aloe) (Aloe barbadensis) Imported

Taxonomic identifier¹ 34199 [NCBI]

Taxonomic lineage¹ Eukaryota › Viridiplantae › Streptophyta › Embryophyta › Tracheophyta › Spermatophyta › Magnoliophyta › Liliopsida › Asparagales › Asphodelaceae › Asphodeloideae › Aloe

Subcellular location¹

UniProt annotation GO (Gene Ontology) annotation

Chloroplast

chloroplast SAAS annotation

Keywords - Cellular component¹

Fig.3 Protein Sequence in Fasta Format

www.uniprot.org/uniprot/H2FAV5.fasta

```
>tr|H2FAV5|H2FAV5_ALOVR|Photosystem II protein D1_05-Aloe vera|GN=psbA|PE=3|SV=1
MTATLERRESTSLWRFQWITSTENRLLYIGNFVLMIPITLLTATSVFIIAFAAPPDI
DGIREPVSGLLYGNLISGAIIP TSAIIGLHFYPTNEAASVDEMLYNGGPELVLVHFL
LGVACVQREWELSERLQMRPWIAVAYSAPVAATAVFLIYPIGOGSEFSDGMPGLISGTF
NFMILVQAEHLILMHPHMLGVAGVFGGSLFSAMHGLVTSLSLRETTENSAKNGYRFG
QEEETVNLVAHGYPGRLLIFQYASFNRSRSLHFFLAAMPVWGIWFTALGISTMAFNLNGF
NFWNSVDSOGRVINTWADLINRANLQMEVHERNAHNFPLDLAAVEVPSTNG
```

Fig.4 Prediction of Physico Chemical Parameter by using ProtParam

ExpASY - ProtParam tool - Mozilla Firefox

www.uniprot.org/uniprot/ H2FAV5.fasta

ProtParam

ProtParam tool

ProtParam (References / Documentation) is a tool which allows the computation of various physical and chemical parameters for a given protein stored in **Swiss-Prot** or **TrEMBL** or for a user entered protein sequence. The computed parameters include the molecular weight, theoretical pI, amino acid composition, atomic composition, extinction coefficient, estimated half-life, instability index, aliphatic index and grand average of hydropathicity (GRAVY) (Disclaimer).

Please note that you may only fill out one of the following fields at a time.

Enter a Swiss-Prot/TrEMBL accession number (AC) (for example **P05130**) or a sequence identifier (ID) (for example **KPC1_DROME**):

Or you can paste your own amino acid sequence (in one-letter code) in the box below:

```
MTATLERRESTSLWRFQWITSTENRLLYIGNFVLMIPITLLTATSVFIIAFAAPPDI
DGIREPVSGLLYGNLISGAIIP TSAIIGLHFYPTNEAASVDEMLYNGGPELVLVHFL
LGVACVQREWELSERLQMRPWIAVAYSAPVAATAVFLIYPIGOGSEFSDGMPGLISGTF
NFMILVQAEHLILMHPHMLGVAGVFGGSLFSAMHGLVTSLSLRETTENSAKNGYRFG
QEEETVNLVAHGYPGRLLIFQYASFNRSRSLHFFLAAMPVWGIWFTALGISTMAFNLNGF
NFWNSVDSOGRVINTWADLINRANLQMEVHERNAHNFPLDLAAVEVPSTNG
```

RESET Compute parameters

Fig.5 Domain Prediction using Prosite.

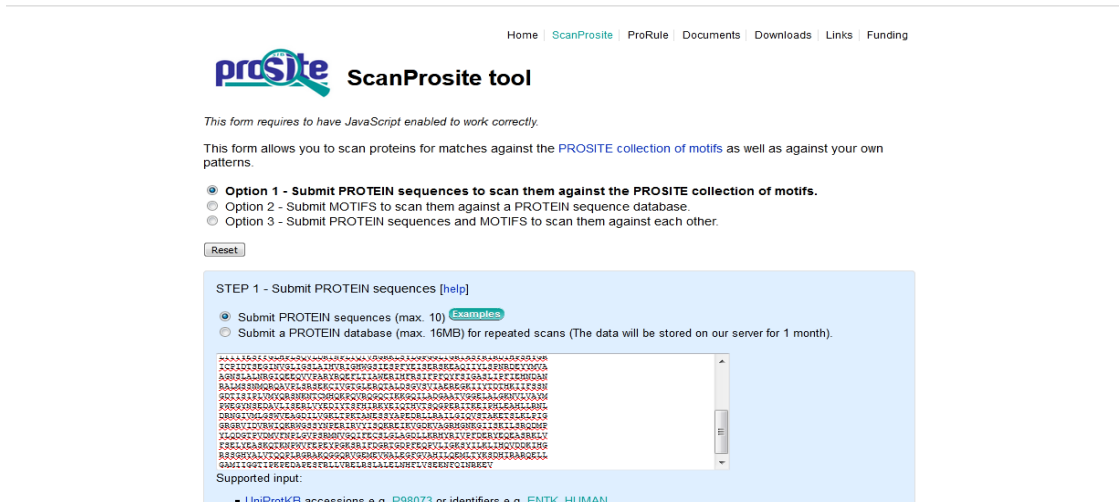


Fig.6 Retrieval of Proteins Secondary Structure.

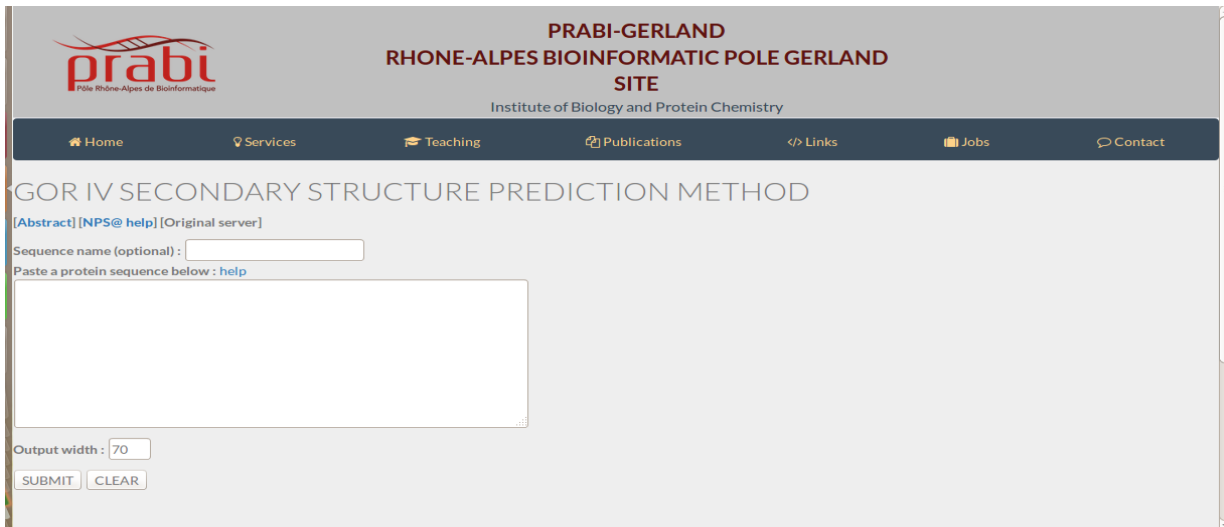
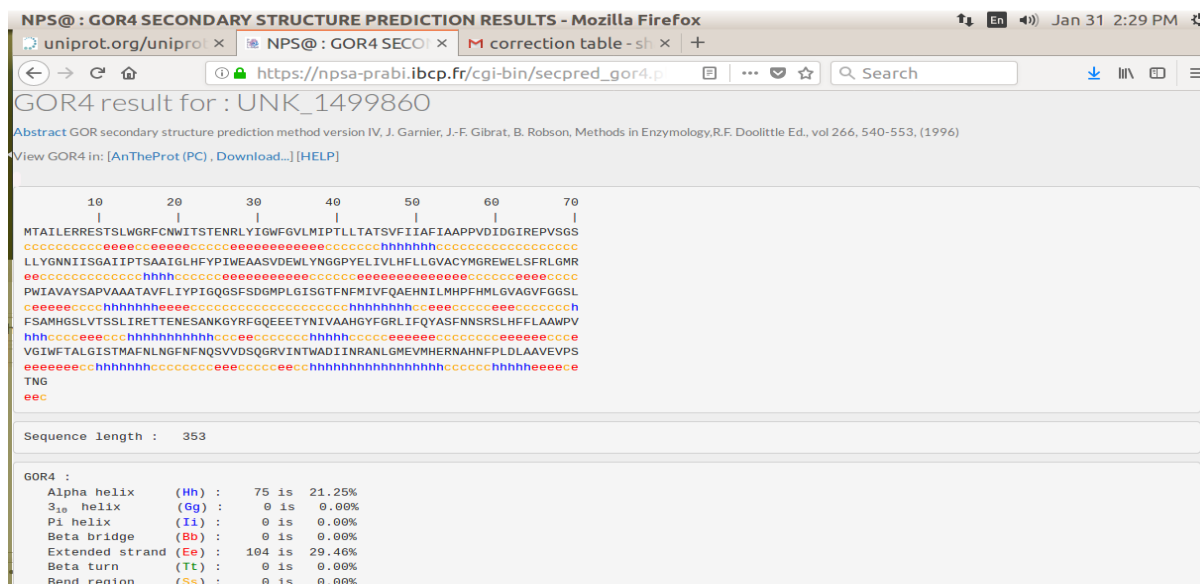


Fig.7 Secondary Structure elements GOR result



In the present study we have performed the analysis of *A. barbadensis* and find out important its important proteins such as DNA binding and Antioxidant proteins which have crucial role in growth, development and defence mechanism.

We have also predicted functional regions of proteins called protein domains, again they have vital role in the transcription and protection against various diseases.

Findings such proteins and prediction of proteins functional region can aid in revealing hidden unique information, which ultimately are the important candidate of present food and drug industries.

Acknowledgement

Authors are highly grateful to Chairman Shri. Vishu Sharan Agarwal and Vice Chairman Shri Puneet Agrawal MIET groups of institutions Meerut, for providing encouragements, platform and support to complete this work

References

Borrelli F, Izzo AA2000. The plant kingdom as a source of anti-ulcer remedies. *Phytotherapy Research* 14:581–591.

Chandan B K, Saxena A K, Shukla S, Sharma N, Gupta D K, Suri K A, Suri J, Bhadauria M. Singh B2007. Hepatoprotective potential of *Aloe barbadensis* mill against carbon tetrachloride induced hepatotoxicity. *Journal of Ethnopharmacology* 111:560–566.

Chen W, Van Wyk B E, Vermaak I, Viljoen A M 2012. Cape aloes—a review of the phytochemistry, pharmacology and commercialisation of *Aloe ferox*. *Phytochemistry Letters* 5:1–2.

Christian J A, Sigrist, Lorenzo C, Edouard de Castro, Petra S. L, Virginie B, Amos B, and Nicolas H 2009. PROSITE, a protein domain database for functional characterization and annotation. *Nucleic Acids Research* 166-166.

Eshun K, He Q 2004. *Aloe vera*: A valuable ingredient for the food, pharmaceutical and cosmetic industries – A review. *Critical Review in Food Science and Nutrition* 44: 91-96.

Gasteiger E, Hoogland C, Gattiker A, Duvaud S, Wilkins M R 2005. Protein identification

- and analysis tools on the ExPASy server. In: Walker JM (ed), *The proteomics protocols handbook*. Humana Press, Totowa 571-607.
- Guruprasad K, Reddy B V P, Pandit M W 1990. Correlation between stability of a protein and its dipeptide composition: a novel approach for predicting in vivo stability of a protein from its primary sequence. *Protein Engineering* 4: 155-164.
- Habeeb, F, Shakir E, Bradbury F, Cameron P, Taravati M R, Drummond A J, Gray A I, Ferro V A, Screening methods used to determine the anti-microbial properties of Aloe vera inner gel 2007. *Methods* 42: 315-320.
- He Q, Changhong L, Kojo, E Tian, Z. 2005. Quality and safety assurance in the processing of Aloe vera gel juice. *Food Control* 16: 95-104.
- Huseini H F, Kianbakht S, Hajiaghaee R, Dabaghian FH2012. Anti-hyperglycemic and anti-hypercholesterolemic effects of Aloe vera leaf gel in hyperlipidemic type 2 diabetic patients: a randomized double-blind placebo-controlled clinical trial *Planta Medica* 78: 311–316.
- Ikai A J 1980. Thermostability and aliphatic index of globular proteins. *Journal of Biochemistry* 88: 1895-1898.
- Jani, G K, Shah D P, Jain V C, Patel M J, Vithalan D A 2007. Evaluating mucilage from Aloe Barbadensis Miller as a pharmaceutical excipient for sustained-release matrix tablets. *Pharmaceutical Technology* 31: 90-98.
- Kumar M, Rakesh S, Nagpal R, Hemalatha R, Ramakrishna A, Sudarshan V, Ramagoni R, Shujaiddin M, Verma V, Kumar A, Tiwari A 2013. Probiotic lactobacillus rhamnosus GG and Aloe vera gel improve lipid profiles in hypercholesterolemic rats *Nutrition* 29:574–9.
- Kyte J, Doolittle R F 1982. A simple method for displaying the hydrophobic character of a protein. *Journal of Molecular Biology* 157: 105-132.
- Langmead L, Makins, R J, Rampton D S 2004. Anti-inflammatory effects of aloe vera gel in human colorectal mucosa in vitro *Aliment. Pharmacology & Therapeutics* 19:521–527.
- Miladi S, Damak M 2008. In vitro antioxidant activities of Aloe vera leaf skin extracts. *Journal of Social Chimique Tunisie* 10:101–109.
- Ni Y, Turner D, Yates K M, Tizard I. Isolation and characterisation of structural components of *Aloe vera* L. leaf pulp 2004. *International Journal of Immunopharmacology* 4: 1745-1755.
- Pabo C O, Sauer R T 1984. Protein-DNA recognition. *Annual Review of Biochemistry* 53 (1): 293–321.
- Pugh N, Ross S A, ElSohly M A, Pasco D S 2001. Characterisation of aloeride, a new highmolecular- weight polysaccharide from Aloe vera with potent immunostimulatory activity. *Journal of Agricultural and Food Chemistry*. 49:1030-1034.
- Qadir M I 2009. Medicinal and cosmetological importance of Aloe vera. *International Journal of Natural Therapy*2: 21–26.
- Rzeznicka K, Schatzle S, Bottcher D, Klein J, Bornscheuer U T 2010. C expression of a nitrilehydratase (NHase) from coccusequi TG328-2 in *Escherichia coli* purification and biochemical characterization *Applied Microbiology and Biotechnology*. 85: 1417-1425.
- Tarameshloo M, Norouzian M, Zarein-Dolab S, Dadpay M, Mohsenifar J. Gazor R. 2012. Aloe vera gel and thyroid hormone cream may improve wound healing in Wistar rats. *Anatomy & Cell Biology* 45:170–177.
- Vidhya V G, Swarnalatha Y, Anusha B 2018. An *In Silico* Analysis of Physicochemical Characterization and Protein-Protein Interaction Network Analysis of Human Anti-apoptotic Proteins. *Asian Journal of Pharmaceutics*12 (4): 1397-1407.
- Vinson J A, Al K, Andreoli L 2005. Effect of Aloe vera preparations on the human bioavailability of vitamins C and E. *Phytomedicine* 12:760-765.