CHARACTERIZATION OF *ATP* GENE IN *Calotropis procera* MITOCHONDRIAL GENOME

S. E. HASSANEIN[1,2]

1. *Bioinformatics and Computer Networks Department, Agricultural Genetic Engineering Research Institute (AGERI), Agriculture Research Center (ARC), Giza, Egypt*
2. *College of Biotechnology, Misr University for Science and Technology (MUST), 6th October City, Egypt*

Although most DNA is packaged in the chromosomes within the nucleus, mitochondria also have a small amount of their own DNA. This genetic material is known as mitochondrial DNA or mtDNA. Mitochondria are structures within cells that convert the energy from food into a form that cells can use. Each cell contains hundreds to thousands of mitochondria, which are located in the fluid that surrounds the nucleus (the cytoplasm). The central role of mitochondria is providing ATPs to cover energy expenses of an organism, which are required for survival, growth and reproduction (Abumourad *et al*., 2013). Mitochondria produce energy through a process called oxidative phosphorylation. This process uses oxygen and simple sugars to generate adenosine triphosphate (ATP), the cell's main energy source. A set of enzyme complexes, designated as complexes I-V, carry out oxidative phosphorylation within mitochondria. Mitochondrial DNA contains about 37 genes, all of which are essential for normal mitochondrial function. Thirteen of these genes provide instructions for making enzymes involved in oxidative phosphorylation. The remaining genes provide instructions for making transfer RNAs (tRNAs) and ribosomal RNAs (rRNAs).

ATP is a substance present in all living cells that provides energy for many metabolic processes and is involved in making RNA. Most ATPases break down ATP to provide energy for molecule transport. ATP gene family provides instructions for making transporter proteins called ATPases, which carry many types of molecules, such as fats, sugars, charged atoms or molecules (ions), and drugs, across cell membranes. ATPases use energy from ATP to move substances across the cell membranes. Within ATPase, there are four subfamilies that are distinguished by their location within the cell and how they transport molecules. The F-types are located on the membranes of mitochondria, and instead of breaking down ATP to transport molecules, these types make ATPs, and so, they are called ATP synthase.

ATP synthase (or complex V) is the enzyme of aerobic ATP production. It is located in the inner mitochondrial membrane of eukaryotic cells together

Egypt. J. Genet. Cytol., 43:257-269, July, 2014
Web Site (*www.esg.net.eg*)

with four respiratory chain enzymes that generate the proton motive force, which in turn drives ATP synthesis. ATP synthase comprises a rotary catalytic portion, F_1-ATPase, whose structure has been characterized (Abrahams *et al.*, 1994) a transmembrane portion F_0, and two stalks that link F_1 and F_0 (Nijtmans *et al.*, 2001). The F_0 portion in *Saccharomyces cerevisiae* contains at least six non-identical subunits: subunits 6, 8 and 9 are encoded by mitochondrial DNA, whereas three other polypeptides with molecular masses of 25, 21 and 19 kDa are always associated with ATP synthase, nuclear DNA encodes these proteins (Velours *et al*., 1984). Little is known about either their structure or their roles in the holoenzyme. Velours *et al.* (1987) isolated the subunit 4 (molecular mass 25 kDa), which is the fourth polypeptide of the complex when classifying subunits in order of decreasing molecular mass.

The ATP synthase F_0 portion subunit 4 is located at the F_0-F_1 interface, since its cross-linkings shape enhances rotary mechanism to proton translocation, was performed between this subunit and F_1 α and β subunits and F_0 subunit 9 and the oligomycin-sensitivity-conferring protein (OSCP). (Genetics Home Reference http://ghr.nlm.nih.gov/).

Calotropis procera (Taxonomy ID: 141467) is a drought-tolerant wild plant. It belongs to *Asclepiadaceae* family and is characterized as a sustainable evergreen medicinal and toxic shrub. Seed spreads mainly by wind and can be transmitted by animals as well. *C. procera* is native to West and East Africa, and South Asia, while naturalized in Australia, Central and South America, and the Caribbean islands. Its bark and leaves are used for the treatment of leprosy and asthma (Duke *et al*., 2002).

In the present study, we aimed to uncover and characterize mitochondrial *ATP4* gene in this medicinal plant from the *de novo* assembled transcriptome contigs of a high-throughput sequencing dataset. We intend to compare the sequence as well as the three-dimensional (3D) structure of the obtained ATP4 protein with those of other plant species.

MATERIALS AND METHODS

Sample collection and isolation of total RNA

Five leaf discs of *C. procera* were sampled. They were frozen in liquid nitrogen (50 mg tissue each) and total RNA extraction was performed using RNeasy Plant Mini Kit (Qiagen, cat. no. 74903). To remove DNA contaminants 3 μl of 10 mg/ml DNase A, RNase and protease-free (Thermo Scientific cat no. EN0531) were added to the RNA samples and tubes were incubated at 30°C for 15 min. Estimation of the RNA concentration in different samples was done by measuring optical density at 260 nm according to the equation: RNA concentration (μg/ml) = OD260 X 40 X dilution factor. RNA samples were sent to Beijing Genomics Institute (BGI), Shenzhen, China, for deep se-

quencing and dataset were provided for analysis.

NGS sequence

Whole-RNAseq, paired-end short - sequence reads of *C. procera* were generated using the Illumina Genome AnalyserIIx (GAIIx) according to manufacturer's instructions (Illumina, San Diego, CA, USA).

Sequence filtering and bioinformatics analysis

The raw sequencing data were obtained using the Illumina python pipeline v. 1.3. For the obtained libraries, only high quality reads (quality >20) were retained. Then, *de novo* assembly of the obtained short (paired-end) read dataset was performed using assembler trinityrnaseq followed by creation of putative unique transcript (PUTs) with a combination of different *k*-mer lengths and expected coverage (Haas *et al*., 2013).

Twenty ATP4 sequences (Table 1 and Fig. 3) belonging to other plant species (*Silene vulgaris*, *Arabidopsis thaliana*, *Spirodela polyrhiza*, *Silene conica*, *Silene noctiflora*, *Bos taurus*, *Homo sapiens*, *Sorghum bicolor*, *Rattus norvegicus* and *Rhodomonas salina*) were obtained from GenBank and used as a reference for blasting (http://www.ncbi.nlm.nih.gov/ BLAST) our obtained library (the yielded EST assemblies from Velvet program) to identify coatings with ATP4 sequence.

Determination of phylogenetic relationships

The maximum likelihood method was used to build a dendrogram and CLC Genomics Workbench was used to allow doing bootstrap analysis. A bootstrap value is attached to each branch to indicate the confidence in this branch.

The 3D homology modeling

The functional domain was identified (acc. no. cl21478: Mt_ATP-synt_B Superfamily, Pfam acc. no. PF05405) from the NCBI Conserved Domain Database (CDD) (http://www. ncbi.nlm.nih. gov/Structure/cdd/cdd.shtml), which uses 3D-structure information to explicitly define domain boundaries and provide insights into sequence/ structure/function relationships, as well as domain models imported from a number of external source databases (Pfam, SMART, COG, PRK, TIGRFAM).

Homology modeling was carried out using Swiss-Model, protein-modeling server, accessible *via* the EXPASY (http://www.expasy.org/). Super-imposition of *ATP4* amino acid sequence model on other ATP4 proteins was constructed using RasMol (http:// www.umass.edu/ microbio/rasmol/), and Deep-View programs (http://spdbv.vital-it.ch/).

Structure alignment

The protein model was applied to pairwise comparison of protein structures using FATCAT server (Flexible structure

AlignmenT by Chaining Aligned fragment pairs allowing Twists; http://fatcat. burnham.org/fatcat/) and their alignments were used to identify conserved and diverse structure domains (Ye and Godzik, 2003). Root mean square deviation (RMSD), which measures the average distance between the backbones of superimposed proteins, (when given two sets of *n* points *v* and *ω*) was measured according to the following formula:

$$\mathrm{RMSD}\,(v, \omega) = \sqrt{\frac{1}{n}\sum_{i=1}^{n}(v_{ix} - w_{ix})^2 + (v_{iy} - w_{iy})^2 + (v_{iz} - w_{iz})^2}$$

Where *n* is the position of each equivalent atom, *v* and *ω* are two different sets of points

RESULTS AND DISCUSSION

The transport ATPase works on the F-type ATPases (F_0F_1) as a nano-motors. These ATPases driven in reverse by a proton gradient have the capacity to interconvert electro- chemical energy into mechanical energy and finally into chemical energy conserved in the terminal bond of ATP. In mammalian mitochondria these events occur on a larger complex or "nano-machine" called the "ATP synthasome" that consists of the ATP synthase in complex formation with carriers for Pi and ADP/ATP. (Pedersen, 2008).

Within our deduced amino acid sequence, to allocate protein domains, the protein sequence obtained from ORF analysis (Fig. 1) with a length of 198 aa was analyzed against the CDD database (conserved domain database, http:// www.ncbi.nlm.nih.gov/cdd) to detect protein domains. Domain analysis indicated the presence of the domain Mt_ATP-synt_B. acc. no. cl21478 (Fig. 2).

BLAST analysis

To identify sequence similarity with homologous proteins from other organisms PHI-BLAST and DELTA-BLAST tools (search performed using specific database, refseq_protein, and using Entrez query "mitochondria" to limit search) were performed to the obtained *C. procera* ATP4 protein (http://blast.ncbi. nlm.nih.gov/Blastp.cgi) The explanation of the score and sequence similarity from specialized BLAST searching eventually led to the identification of homologous protein sequences. Results for the most closely related protein to *C. procera* ATP4 protein indicated that the ATPase subunit 4 of *Silene vulgaris* has the lowest e-value (0.0) and high identity percent (88.89%). These results indicated that *C. procera* ATP4 has the same function.

Mt_ATP-synt_B

The F_0 sector of the ATP synthase is a membrane bound complex which mediates proton transport. It is composed of nine different polypeptide subunits (a, b, c, d, e, f, g F6, A6L. CDD ID: cl21478).

Mitochondrial membrane ATP synthase ($F_1 F_0$ ATP synthase or Complex V) produces ATP from ADP in the presence of a proton gradient across the membrane

that is generated by electron transport complexes of the respiratory chain. F-type ATPases consist of two structural domains, F_1 which contain the extramembraneous catalytic core, and F_0 which contains the membrane proton channel. ATP synthesis in the catalytic domain of F_1 is coupled *via* a rotary mechanism to proton translocation (Fig. 2).

Multi-sequence alignment (MSA) and phylogenetic analysis

Filters (Refseq_protein database, within "mitochondria" only) used to specify and limit the number of hits resulted in the specialized BLAST (Delta-Blast) search hits were used to perform multi-sequence alignment (MSA). This resulted in *Calotropis* ATP4 10 protein sequences from 10 different species (Table 2 and Figs 3 & 4). A multiple sequence alignment of the 11 sequences (10 sequences resulted for Delta-BLAST plus our deduced one) was obtained by gap open penalty of 10 and gap extension penalty of one. The results also showed that the closest sequence to the obtained *C. procera* ATP4 protein was *ATPase* subunit 4 (mitochondrion) with accession number YP_004935350.1 obtained from *Silene vulgaris*. These results support the obtained BLAST results. MSA results were used to perform phylogenetic tree for the 11 proteins and results (Fig. 5) were similar to those of previous analyses.

Structure analysis

Primary structure properties

Kyte-Doolittle hydropathy plots provide information about the possible structure of a protein, and can identify features such as transmembrane or surface regions. For surface region discovery in a globular protein, a window size of 9 is considered optimal, marked negative dips in the graph indicate possible surface regions. A region size of 19 is considered optimal for the discovery of transmembrane regions, which can be identified by peaks above a value of 1.8. (Kyte and Doolittle, 1982).

Figure (6) shows the hydrophobicity similarity between predicted ATP4 protein and its similar Bovine mitochondrial ATP synthase chain A (acc. no.: 2CLY_A, GI:110591026). Hydrophobic segment at the beginning of each proteins emphasize its location; A proton-transporting ATP synthase complex is found in the mitochondrial membrane. Its activities; first, ATPase activity, Catalysis of the reaction: ATP + H_2O = ADP + phosphate + $2H^+$. Then second, hydrogen ion transmembrane transporter activity, Catalysis of the transfer of hydrogen ions from one side of a membrane to the other (http:// www.uniprot.org/uniprot/P13619).

3D structure modeling

ATP4 signaling efficiency and specificity can be achieved through Subcomplex of the Stator of Bovine Mitochondrial Atp Synthase [Hydrolase, EC:

3.6.3.14] PDB ID: 2CLY_A, (http://www.ebi.ac.uk/pdbe-srv/view/ entry / 2CLY_A/ summary).

Based on structural alignment, a theoretical 3D model for *C. procera* ATP4 protein was created, corresponding to residues 1-198 of the primary structure (Fig. 7). The predicted model was created using the Swiss-Model, protein-modeling server.

Structure alignment

FATCAT rigid structure alignment online tool was applied on five *C. procera* ATP4 proteins 3D structure, which was created, based on structural alignment using Swiss-Model. 2CLY_A, Subcomplex of the Stator of Bovine Mitochondrial Atp Synthase, is the closest homologous protein sequence with available 3D structure to the obtained *C. procera* ATP4 protein.

To proof the accuracy of the theoretical 3D model, FATCAT server was used to compute optimal and suboptimal structural alignments between 2CLY_A 3D structure and the theoretical 3D model of *C. procera* ATP4 protein. The resulting superimposed figure is shown in Fig. (8) with Z-score of 3.54 and number of equivalent residues of 68 and RMSD of 2.84.

Aligned Fragment Pair (AFP) regarding to two protein structures, denoted a match of two fragments, one from each protein as an Aligned Fragment Pair (AFP). Each AFP can define a transformation of two structures. For the two aligned proteins, the two structures are significantly similar, with a P-value of $5.06e^{-05}$ and Afp-num 4421 Identity of 2.82%. Similarity was 26.76%, while Block 0 afp 6 score was 122.61 RMSD 2.84 gap 21 (0.30%) (Fig. 9).

Twist; indicate the secondary structure that has to be rearranged (a twist introduced at the hinge) so that the secondary structures can be better aligned. For the two aligned proteins, Twists 0 initial length was 48 aa and initial RMSD was 2.84.

Interpolating between 3D structure of ATP4 model and 2CLY_A model

Intermediate structures are calculated by linearly interpolating distance matrices of the aligned parts of superimposed conformers. Subsequently, the intermediate structures are optimized by energy gradient minimization employing a reduced representation force field. Thus, in contrast to a simple morphing, the structure changes can be interpolated or extrapolated while preserving protein-like geometry of intermediate structures and internal structure of the rigid elements of both proteins (Fig. 10).

CONCLUSION

In most biosystems, the ATP synthase sits in the membrane, and catalyzes the synthesis of ATP from ADP and phosphate driven by a flux of protons across the membrane down the proton gradient

generated by electron transfer. The flux goes from the protochemically positive (P^+) side (high proton electrochemical potential) to the protochemically negative (N^-) side. The reaction catalyzed by ATP synthase is fully reversible, so ATP hydrolysis generates a proton gradient by a reversal of this flux. The ATP synthase of the mitochondrial inner membrane is an enzymatic multi-subunit complex formed by two domains. The hydrophilic portion termed F_1 contains the catalytic site for ATPase activity. F_1 consists of five non-identical subunits (α, β, γ, δ, ε) these subunits are imported from the cytoplasm. The membranous portion, termed F_0, forms a proton channel. F_0 is composed of a variable number of polypeptides according to the organism, ranged from three to ten. (Velours *et al*., 1988).

These results support our finding that the obtained *C. procera* transcript sequences belong to ATP4 and possess the same functions regarding its functional domain and that the motif belong to ATP4 protein sequence. Also, the results prove the accuracy of our theoretical 3D modeling for the obtained *C. procera* ATP4 deduced protein from the sequence of the amino acids.

In conclusion, the present study provides an important gene involved in oxidative phosphorylation process in the arid land plant *Calotropis procera*, and suggests three important features. First, it exists in biological (mitochondrial) membranes. Second, being part of complex hydrolyzes ATPs (ATPase) which have a reverse activity makes it synthase ATPs (ATP syntheses), regarding to the existence of "Mt_ATP-synt_B" functional domain. And finally, it transports at least one substance across the biological membrane, at the expense of ATP hydrolytic activity.

SUMMARY

The drought-tolerant wild plant *C. procera* is important in medicine, industry and ornamental fields. Generally, its bark and leaves are used for many Folk medicine treatments. ATP4 (mitochondrion ATPase subunit 4), one of ATP gene family provides instructions for making transporter proteins called ATPases, which use energy from ATP molecule to move substances, such as fats, sugars, charged atoms or molecules (ions), and drugs, across the cell membranes. In this study, we uncovered and characterized ATP4 (ATP4, NCBI accession no. KP171515) gene in this medicinal plant from the *de novo* assembled transcriptome contigs of the high-throughput sequencing dataset. A number of GenBank accessions for ATP4 sequences were blasted with the recovered de novo assembled contigs. Homology modeling of the deduced amino acids was further carried out using Swiss-Model, accessible *via* the EXPASY. Superimposition of *C. procera* ATP4 full sequence model on Chain A, Subcomplex of the Stator of Bovine Mitochondrial ATP Synthase (PDB accession no. 2CLY_A) was constructed using RasMol and Deep-View programs. The functional domains of the novel ATP4 amino acids sequence were identified

from the NCBI conserved domain database (CDD, accession no. cl21478) that provide insights into sequence structure/function relationships, as well as domain models imported from a number of external source databases (Pfam, SMART, COG, PRK, TIGRFAM).

REFERENCES

Abrahams, J. P., A. G. Leslie, R. Lutter and J. E. Walker (1994). Structure at 2.8 Â resolution of F_1-ATPase from bovine heart mitochondria. Nature, 370: 621-628.

Abumourad, Iman M. K., S. M. Girgis and M. I. Hanna (2013). Identification and Expression of ATP Synthase F_0 Subunit 6 in Tilapia Fish during Temperature Acclimation. Research Journal of Pharmaceutical, Biological and Chemical Sciences (RJPBCS), 4: 288-298.

Duke, J. A., Mary Jo Bogenschutz-Godwin, Judi duCellier and P. K. Duke (2002). Handbook of medicinal herbs. CRC PRESS, 2nd ed.

Haas, B. J., A. Papanicolaou, M. Yassour, M. Grabherr, P. D. Blood, J. Bowden, M. B. Couger, D. Eccles, B. Li, M. Lieber, M. D. Macmanes, M. Ott, J. Orvis, N. Pochet, F. Strozzi, N. Weeks, R. Westerman, T. William, C. N. Dewey, R. Henschel, R. D. Leduc, N. Friedman and A. Regev (2013). *De novo* transcript sequence reconstruction from RNA-seq using the Trinity platform for reference generation and analysis. Nat. Protoc., 8: 1494-512.

Krzywinski, M. I., J. E. Schein, I. Birol, J. Connors, R. Gascoyne, D. Horsman, S. J Jones and M. A M. Circos (2009). An information aesthetic for comparative genomics. Genome Res., 19: 1639-1645.

Kyte, J. and R. Doolittle (1982). A simple method for displaying the hydropathic character of a protein. J. Mol. Biol., 157: 105-132.

Nijtmans, L. G. J., N. S. Henderson, G. Attardi and I. J. Holt (2001). Impaired ATP Synthase Assembly Associated with a Mutation in the Human ATP Synthase Subunit 6 Gene. J. Biol. Chem., 276: 6755-6762.

Nijtmans, L. G., P. Klement, J. Houstek and C. van den Bogert (1995). F_1-ATPase a-subunit made up from two fragments is stabilized by ATP and complexes containing it obey altered kinetics. Biochim. Biophys. Acta, 1272: 190-198.

Pedersen, P. L. (2008). Transport ATPases into the year 2008: A brief overview related to types, structures, functions and roles in health and disease. J Bioenerg Biomembr, 39: 349-355.

Velours, J., G. A. de Chateaubodeau, M. Galante and B. Gukrin (1987). Subunit 4 of ATP synthase (F_0F_1) from yeast mitochondria; Purification, amino-acid composition and partial N-terminal sequence. Eur. J. Biochem., 164: 579-584.

Velours, J., P. Durrens, M. Aigle and B. Guerin (1988). ATP4, the structural gene for yeast FoFl ATPase subunit 4. Eur. J. Biochem., 170: 637-642.

Velours, J., M. Esparza, J. Hoppe, W. Sebald and B. Guerin (1984). Amino acid sequence of a new mitochondrially synthesized proteolipid of the ATP synthase of *Saccharomyces cerevisiae*. EMBO J., 3: 207-212.

Ye Y. and A. Godzik (2003). Flexible structure alignment by chaining aligned fragment pairs allowing twists. Bioinformatics, 19: II246-II255.

Table (1): Accession numbers, description, score (Bites) and E. Value of the gene whose *ATP*4 gene was isolated.

Accession no.	Description	Score	E. value
XM_004152925.1	PREDICTED: *Cucumis sativus* ATP synthase p...	865	0
XM_004153363.1	PREDICTED: *Cucumis sativus* ATP synthase p...	865	0
XM_003550048.2	PREDICTED: *Glycine max* ATP synthase prote...	793	0
XM_004500030.1	PREDICTED: *Cicer arietinum* ATP syn, thase p...	782	0
XM_002535017.1	*Ricinus communis* conserved hypothetical p...	782	0
XM_003619113.1	Medicago truncatula ATPase subunit (MTR_6...	778	0
XM_004516932.1	PREDICTED: *Cicer arietinum* ATP synthase p...	773	0
XM_002874441.1	*Arabidopsis lyrata* subsp. lyrata hypothet...	773	0
XM_002863112.1	*Arabidopsis lyrata* subsp. lyrata hypothet...	754	0
XM_006409453.1	*Eutrema salsugineum* hypothetical protein ...	741	0
XM_006826107.1	*Amborella trichopoda* hypothetical protein...	699	0
XR_721311.1	REDICTED: *Eucalyptus grandis* uncharacterize...	628	5.0E-177
XM_004500048.1	PREDICTED: *Cicer arietinum* ATP synthase p...	586	3.0E-164
XM_002488961.1	Sorghum bicolor hypothetical protein (SOR...	507	2.0E-140
XM_003619105.1	*Medicago truncatula* ATP synthase subunit ...	313	6.0E-82
XM_003592480.1	*Medicago truncatula* ATPase subunit (MTR_1...	311	2.0E-81
XM_008246610.1	PREDICTED: *Prunus mume* uncharacterized mi...	302	1.0E-78
XM_008246609.1	PREDICTED: *Prunus mume* uncharacterized LO...	302	1.0E-78
XM_006396456.1	*Eutrema salsugineum* hypothetical protein ...	124	3.0E-25
XM_006396455.1	*Eutrema salsugineum* hypothetical protein ...	124	3.0E-25

Table (2): Accession number for each protein, description, organism name and the calculated e-value of homologous proteins to *C. procera* ATP4 amino acids sequence identified using specialized BLAST search programs.

Accession	Description	T.S.	Q.C.	E.value	Max ident.
YP_004935350.1	ATPase subunit 4 (mitochondrion) [*Silene vulgaris*]	202	88.89	1.00E-64	88.89
XP_002489006.1	hypothetical protein SORBIDRAFT_0506s002010…	110	85.59	7.00E-30	85.59
YP_006280941.1	ATPase subunit 4 (mitochondrion) [*Spirodela pol*…	166	83.05	2.00E-50	83.05
NP_085524.1	orf25 [*Arabidopsis thaliana*]	172	80.95	1.00E-52	80.95
YP_004935325.1	ATPase subunit 4 (mitochondrion) [*Silene nocti*…..	146	78.16	1.00E-42	78.16
YP_004935291.1	ATPase subunit 4 (mitochondrion) [*Silene conica*]	146	68.95	1.00E-42	68.95
NP_066455.1	Hypothetical protein RhsaoMp02 [*Rhodomonas sa*…	100	16.35	2.00E-25	15.09
NP_001033590.1	ATP synthase F(0) complex subunit B1, mitochon…..	132	17.37	2.00E-36	13.17
NP_001679.2	ATP synthase F(0) complex subunit B1, mitochon…..	122	14.97	9.00E-33	10.78
NP_599192.1	ATP synthase F(0) complex subunit B1, mitochon…..	113	13.77	1.00E-29	9.58

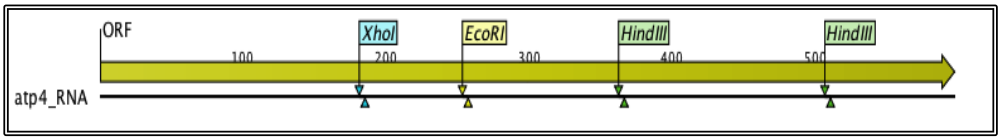

Fig. (1): ORF analysis for the obtained *ATP4* sequence.

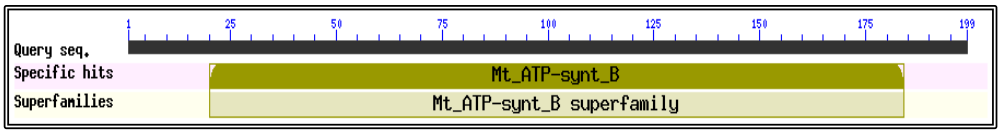

Fig. (2): Illustrate of protein domains of the deduced amino acid sequence of the obtained ATP4 protein.

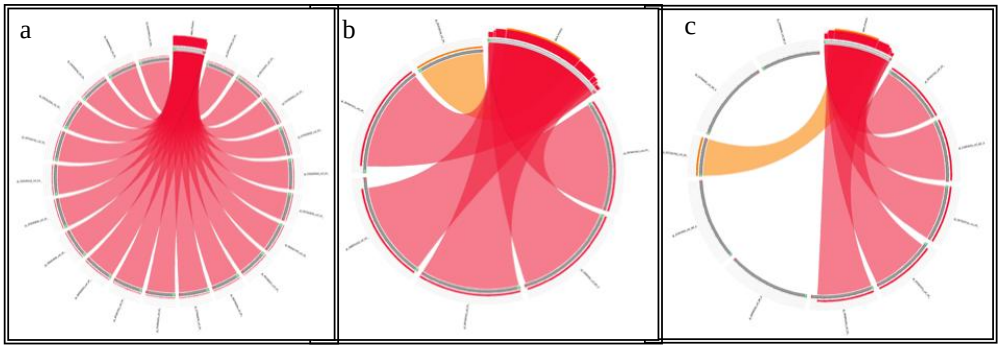

Fig. (3): Twenty ATP4 hit sequences belonging to other plant species (*Silene vulgaris*, *Arabidopsis thaliana*, *Spirodela polyrhiza*, *Silene conica*, *Silene noctiflora*, *Bos taurus*, *Homo sapiens*, *Sorghum bicolor*, *Rattus norvegicus* and *Rhodomonas salina*) were obtained by BLASTn (a), whereas specialized BLASTp programs, PSI-Blast (b) and Delta-BLAST (c) of homologous proteins to *C. procera* ATP4 amino acids sequence visualized by Circoletto, visualizing sequence similarity with Circos®, (Krzywinski *et al*., 2009). Hit sequences arranged clockwisely form query as shown in Table (2).

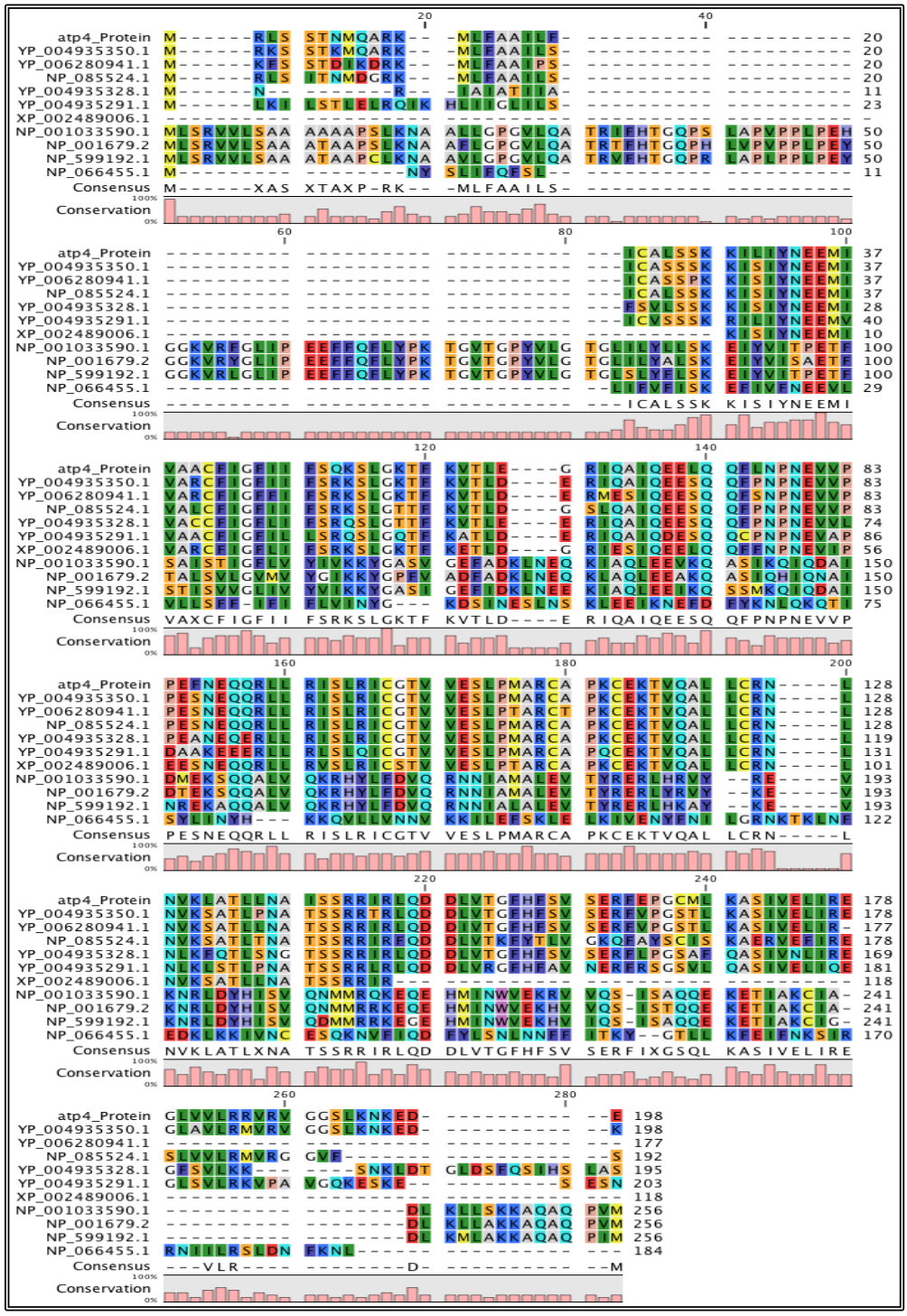

Fig. (4): Multiple sequence alignment of the 10 different ATP4 protein sequences with the obtained *C. procera* ATP4 protein sequence.

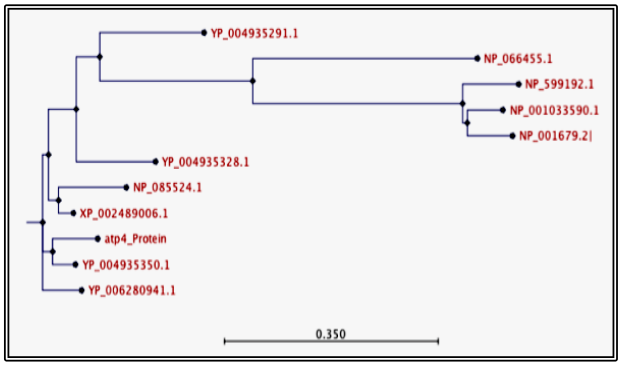

Fig. (5): Phylogenetic analysis of 10 different ATP proteins and *C. procera* ATP4 deduced amino acids sequence.

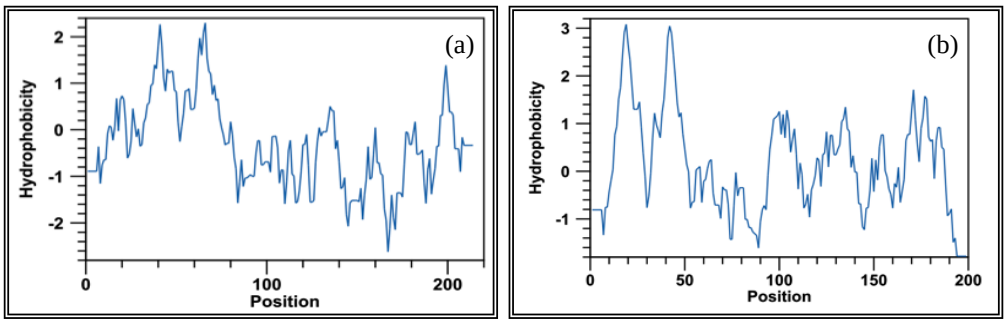

Fig. (6): Hydrophobicity plot Bovine mitochondrial ATP chain A synthase (a) *C. procera* ATP4 and (b) deduced amino acids sequence.

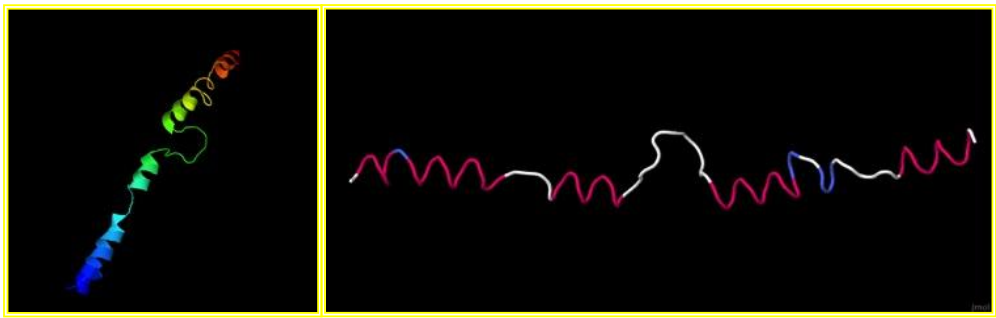

Fig. (7): Predicted 3D model of deduced amino acids sequence of *C. procera* ATP4 protein.

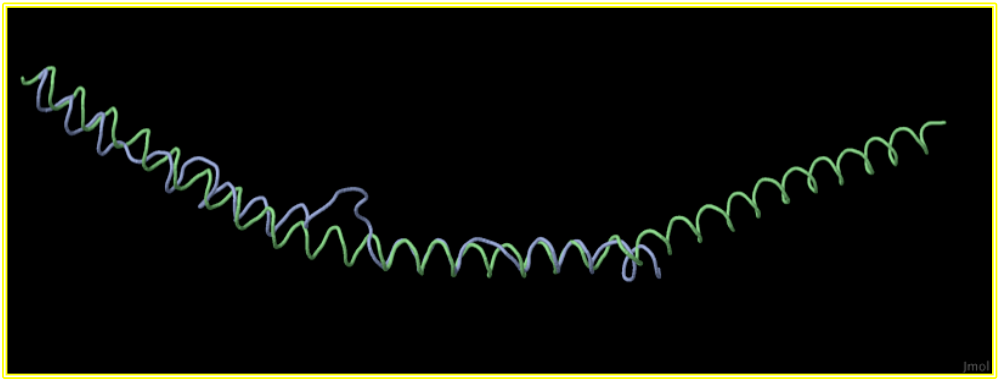

Fig. (8): Align of *C. procera* ATP4 3D model (68 aa) P5CS deduced amino acids (gray) that have almost the same coordinates of 2CLY_A (105 aa) 3D model (green) at functional regions.

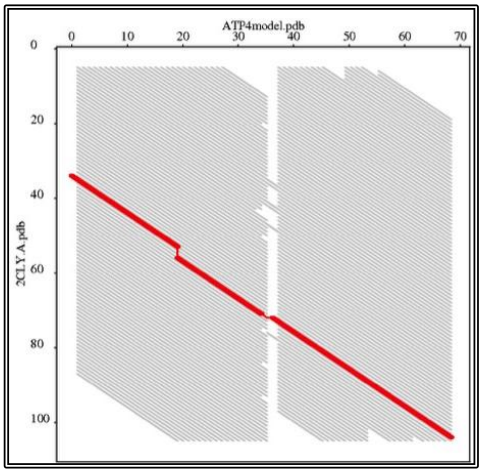

Fig. (9): Chaining diagram indicating the chaining process, each red line in the background represents an Aligned Fragment Pair (AFP).

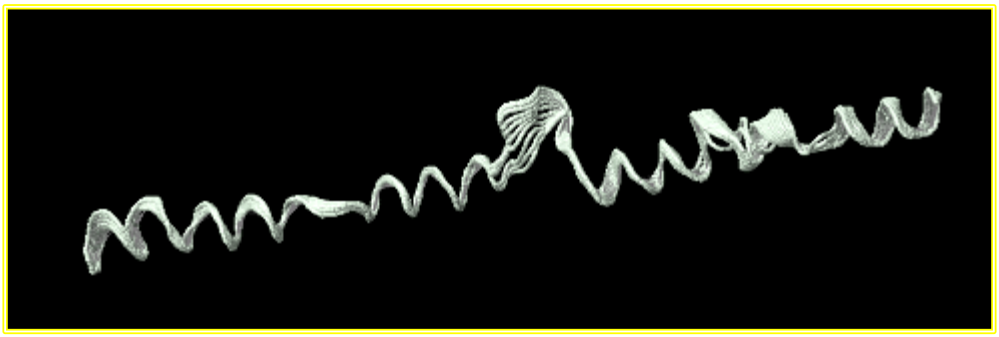

Fig. (10): Ten generated intermediate model structures.