Insilico Characterization of Proteins of Annona Muricata and Assessment of Nanoparticle Synthesising Ability of Closely Related Plant Species

Nandhini[1], pakutharivu[2], rubalakshmi.G[3], nirubama K[4], prabhakaran.S[5], Dr.T.Pakutharivu[6]
[1, 2] Dept of Biochemistry
[1, 2]MGR College, Hosur -30, Tamilnadu, India
[3, 4, 5]GRD Bio clinical Research, Rasipuram, Namakkal, Tamilnadu, India
[6]MGR College, Dr. M.G.R. Nagar, Krishnagiri District, Hosur, Tamil Nadu- 635130,India .

Abstract- *Natural products have been the starting point for the discovery of many important searches for pharmacologically important substance from plant source. Medicinal plants maintain the health and vitality of individuals, and also cure various diseases, including cancer without causing toxicity. In this present study, 10 proteins of Annona muricata were analysed using bioinformatics tools. Structural prediction and functional characterization of proteins of Annona muricata were done using Expasy Protparm server, 3D structure was done using SWISS MODEL. Plants of different family showing identity 80% and above were selected and its sequences retrieved, aligned using Clustal Omega. phylogenetic tree was constructed for the aligned sequence. Structure prediction showed that α – helix, random coil, β – turn and extended strand predominates. Transmembrane region was found in cytochrome b, NADH ubiquinone oxidoreductase, NADH dehydrogenase subunit proteins. NAD(P)H quinone oxidoreductase possess the abilty to reduce metals and hence Annona muricata was predicted for its ability to synthesis nanoparticles. Phylogenetic analysis of NAD(P)H quinone oxidoreductase of Annona muricata reveals that the plants of Annonaceae family are closely related. Annona muricata(graviola) which has been called as one of the nature's finest medicines and known for its nutritive value and it has to be analyzed further for identifying its medicinal properties.*

Keywords- Annona muricata, Bioinformatics, NAD(P)H quinone oxidoreductase, Phylogeny, medicinal properties.

I. INTRODUCTION

India is the birth place of renewed system of indigenous medicine such as Siddha, Ayurvedha and Unani. Traditional systems of medicines are prepared from a single plant or combinations of number of plants. The efficacy depends on the use of proper plant part and its biological potency which in turn depends upon the presence of required quantity and nature of secondary metabolite in a raw drug. The medicinal plants are rich in secondary metabolites and are often termed as medicinal or officinal plants. These secondary metabolites or products exert a profound physiological effect on the mammalian system and thus are known as active principles of the plant [1].

Once such a plant is Guyabano*(Annona muricata*) which belongs to the family of Annonaceae is an evergreen tree species used as traditional medicines. Extracts and metabolites from this plant exhibit pharmacological properties such as anti-inflammatory, antiulcer, anthelmintic, antibacterial, and free radical scavenging activity [2]. *Annona muricata* L (Annonaceae), commonly known as soursop has a long, rich history in herbal medicine with a lengthy recorded indigenous use [3].

Nanotechnology is a branch of science which deals with the ability to control and manipulate matter ranging from a scale less than a nano metre up to 100 nm [4]. Nanoparticles are increasingly used in a variety of fields, including medical, food, health care, consumer, and industrial purposes, due to their unique physical and chemical properties. Thus nanoparticles are used widely in modern medicine. They are used for gene and drug delivery to treat serious illness. Thus this study is to analyse the ability of *Annona muricata* and its related species to synthesis nanoparticles [5].

Bioinformatics is the field of science in which biology, computer science and information technology merges to form a single discipline. It focuses more on hypothesis testing and discovery in the biological domain. The task used in bioinformatics ranges from the creation and maintenance of database of biological information to the analysis of sequence information. The wide range of application of bioinformatics include molecular medicine, gene therapy, drug development, waste cleanup, forensic analysis of microbes, evolutionary studies, comparative studies etc. [6].

However there remains still a huge scope for use of modern scientific methods - genomics, proteomics and bioinformatics in this plant. Bioinformatics shall facilitate analysis and integration of information from these related fields to enable the identification of genes and gene products and elucidate the functional relationships between genotype and observed phenotype. This research report provides a state-of-the-art overview of bioinformatics study of *Annona muricta* with emphasis on the current progress and future directions, which shall provide tools and resources necessary to understand and promote advances in this important field.

II. MATERIALS AND METHODS

Sequence Retrieval

The FASTA sequence of the proteins [TABLE: 1] were retrieved from Genbank database hosted by the NCBI (http://www.ncbi.nlm.nih.gov) [7].

Primary Structure Prediction:- For Physio-chemical characterization, theoretical Isoelectric Point (pI), molecular weight, total number of positive and negative residues, extinction coefficient, instability index, aliphatic index and grand average of hydropathy (GRAVY) were computed using the Expasy Protparm server. [8] **(**http://us.expasy.org/tools/protparam.html).

Secondary structure prediction

SOPMA (Self Optimized Prediction Method with Alignment) was used for the secondary structure prediction.

Functional characterization

SOSUI and TMHMM v.2.0 tools were used to characterize whether the protein is soluble or trans membrane in nature. Inter Pro is an integrated resource for protein families, domains and functional sites. Inter Pro incorporates the major protein signature databases into a single resource. These include: PROSITE, which uses regular expressions and profiles, PRINTS, which uses Position Specific Scoring Matrix-based (PSSM-based) fingerprints, ProDom, which uses automatic sequence clustering, and Pfam, SMART, TIGRFAMs, PIRSF, SUPERFAMILY, Gene3D and PANTHER, all of which use hidden Markov models (HMMs). Superfamily and molecular function were predicted by Inter pro protein sequencing and classification. [9]. (http://www.ebi.ac.uk/interpro/).

Sequence Alignment

Sequence alignment of NADPH was performed using pair wise sequence alignment tool (NCBI- BLAST) (http://blast.ncbi.nlm.nih.gov/Blast.cgi) and multiple sequence alignment was done using the EBI-CLUSTAL OMEGA (http://www.ebi.ac.uk/Tools/msa/clustalo/) tool. Clustal Omega also has powerful features for adding sequences to and exploiting information in existing alignments, making use of the vast amount of precomputed information in public databases like Pfam [10].The emphasis of this work was to find the regions of sequence similarity, which in other words allows us to yield functional and evolutionary relationships among the proteins considered in this study.

Phylogenetic Analysis

The phylogentic analysis of NADPH ubiquinone oxido reductase was performed to determine the number of proteins that share common structural and functional features. As an input to Clustal Omega all sequences in Fasta formats were supplied with default options. The output was analyzed for sequences that are aligned for the complete length, scores, alignment, conserved residues, substitutes and semi conserved substituted residue patterns. The phylogenetic tree was constructed based on the bootstrap Neighbour Joining (NJ) method [11].The stability of the internal nodes was assessed by bootstrap analysis with 1000 replicates.

III. RESULTS AND DISCUSSION:

Table 1: **PRIMARY STRUCTURE OF PROTEINS OF *ANNONA MURICATA***

S.NO	ACCESSION NUMBER	PROTEIN	LENGTH
1	AKA59062.1	Ribulose bisphosphate carboxylase	159
2	AAF17027.1.	ATP synthase Alpha	414
3	CAB90103.1.	ATP synthase Beta	403
4	ABK91377.1	Cytochrome b	347
5	AAQ11840.1.	Maturase K	507
6	ABD78205.1	NADH Ubiquinone oxidoreductase	387
7	ABS45073.1.	NADH dehydrogenase subunit F	677
8	ADL63753.1.	Ribosomal Protein S3	429
9	AML79553.1	Putative LOV domain	787
10	ABD93745.1	RNA polymerase beta chain	827

TABLE 2: EXPASY PROTOPARAM PROTEINS OF ANNONA MURICATA

1	Ribulose bisphosphate carboxylase	AKA59062.1	159	17558.105	8.44	16	18	24535	33.03	85.28	-0.192
2	ATP synthase Alpha	AAF17027.1	414	44570.96	6.55	46	45	20985	37.52	98.74	-0.094
3	ATP synthase Beta	CAB90103.1	403	43388.87	5.61	45	40	16390	33.39	96.55	-0.017
4	Cytochrome b	ABK91377.1	347	38479.97	7.09	16	15	81485	35.97	110.66	0.518
5	Maturase K	AAQ1	507	60514.05	9.71	40	66	68105	57.30	91.32	-0.213
6	NADH Ubiquinone oxidoreductase	ABD78205.1	387	41919	8.87	18	23	43235	39.78	95.84	0.483
7	NADH dehydrogenase subunit F	AB845073.1	677	75424.61	9.16	35	48	108680	22.89	109.47	0.513
8	Ribosomal Protein S3	ADL63753.1	429	49349.49	10.	34	87	37610	55.67	73.43	-0.546

The primary structure prediction was done with the help of protparam tool (Table). The parameters were computed using Expasy's protparam tool which revealed that the molecular weights for ten different proteins as 17558.105 (Ribulose bisphosphate carboxylase), 44570.96 (ATP synthase Alpha), 433.88.87 (ATP synthase Beta), 38479.97 (Cytochrome b), 60514.05 (Maturase K), 41919 (NADH ubiquinone), 75424.61 (NADH Deydrogenase subunit F), 49349.49 (Ribosomal protein S3), 89149.15 (Putative LOV domain), 94351.78 (RNA polymerase beta chain). The pI of two protein was less than 7 which indicated that they are acidic and eight protein was greater than 7 which showed that it is basic in character. The proteins are found to be compact and stable at their pI. Among the two proteins one is showed instability index lesser than 40, indicating that the protein are stable [12].

Aliphatic index of the proteins ranged between85.28-96.25. The computed extinction coefficients help in the quantitative study of protein–protein and protein–ligand interactions in solution. The range of GRAVY (Grand Average of Hydropathicity) of Annona muricata proteins was found to be -0.192 to -0.344. The lowest value of GRAVY indicates the possibility of better interaction with water [13].

TABLE 3: SECONDARY STRUCTURE OF PROTEINS OF *ANNONA MURICATA*

S.no	Secondary sutures	Proteins									
		AKA59062.1	AAF17027.1	CAB90103.1	ABK91377.1	AAQ1 1840.1	ABD78205.1	AB845073.1	ADL63753.1	AML79553.1	ABD9345.1
1	Alpha helix	38.36	34.30	41.94	35.73	42.01	27.65	37.67	32.63	35.63	26.12
2	Extended strand	20.13	20.77	17.37	21.90	22.88	29.46	26.00	20.75	19.31	26.72
3	Bend region	11.95	11.35	9.68	7.49	7.89	6.98	8.86	10.72	8.77	10.64
4	Random coil	29.56	33.57	31.02	34.87	27.22	35.92	27.47	35.90	36.47	36.52

TABLE 4: TRANSMEMBRANE REGION PREDICTED BY SOSUI SERVER

1.Cytochrome-b

This amino acid sequence is of a MEMBRANE PROTEIN which have 6 transmembrane helices.

No.	N terminal	transmembrane region	C terminal	type	length
1	24	FGPLAGISLVIQIVTGVFLAMHH	46	PRIMARY	23
2	64	VEGGWLLRYMHANGASMFLIVVH	86	SECONDARY	23
3	103	EFVRCLGVVIFLLMIVTAFTGYV	125	PRIMARY	23
4	135	GATVITSLASAIPVVGDTIVTWL	157	PRIMARY	23
5	223	VGRVASAIFSSIWIFYAPNVLGH	245	SECONDARY	23
6	318	FWLLLADRLLLGWIGQQPVEAPF	340	SECONDARY	23

2.NADH ubiquinone oxidoreductase

This amino acid sequence is of a MEMBRANE PROTEIN which have 11 transmembrane helices.

No.	N terminal	transmembrane region	C terminal	type	Length
1	1	IVVTSISSLVHLYSISYMSED	21	PRIMARY	21
2	27	FMCYLSIPTFFMPMLVTGDNSLQ	49	SECONDARY	23
3	51	FLGWEGVGLASYLLIHFWFTRLQ	73	SECONDARY	23
4	126	NMRLNAITLICILLLIGAAGKSA	148	PRIMARY	23
5	167	SASIHAATTVTAGVFMIARCSPL	189	SECONDARY	23
6	194	PTALIVITSAGATTSFLAATTGI	216	SECONDARY	23
7	231	SQLGYMIFACGISNYSVSVFHLM	253	SECONDARY	23
8	255	HAFFKALLFLSAGSVIHAMSDEQ	277	SECONDARY	23
9	290	PLTYAMMLMGSLSLIGSPFPTG	311	SECONDARY	22
10	332	FAFWLGSVSVLFTSYYSFRSLFL	354	SECONDARY	23
11	366	DILRCHDAPIPMAIPSILLALG	387	SECONDARY	22

3.NADH dehydrogenase subunit-F

This amino acid sequence is of a MEMBRANE PROTEIN which have 14 transmembrane helices.

No.	N terminal	transmembrane region	C terminal	Type	length
1	1	TISIGLELLLIPTAAKNA	18	SECONDARY	18
2	21	MWAFPSISLLSIVMVFSADLSIH	43	PRIMARY	23
3	72	DPLTSIMSILITTIGIVVLIYS	93	PRIMARY	22
4	102	GYLRFFASMGFSTISMLGLVTST	124	SECONDARY	23
5	128	QIYVFWELVGMCSYLLIGFWFTR	150	PRIMARY	23
6	164	NRVGDFGLLLGILGFYWITGSFE	186	SECONDARY	23
7	202	GVNSLFVTLAFLLFVGAVAKSAQ	224	PRIMARY	23
8	235	MEGPTPISALIHAATMVAAGVFL	257	SECONDARY	23
9	269	PRIMGLISFIGIITVLLGATLSL	291	PRIMARY	23
10	301	AYSTMSQLGYIMLAPGIGSYRAA	323	SECONDARY	23
11	377	TTFYVGTLSLCGIPPLACFWSKD	399	SECONDARY	23
12	414	IIACSAAGLTAFYMFRIYLLIFE	436	PRIMARY	23
13	524	MLLSLLVLFLFTLFIGSIGIPFG	546	PRIMARY	23
14	580	VTNSIFSASIACFGIFIASLFYG	602	PRIMARY	23

PREDICTION OF TRANSMEMBRANE HELICES IN PROTEINS

Figure- 1: Cytochrome--b

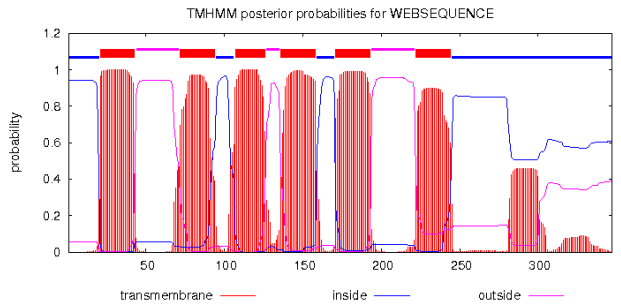


Figure-2:NADH UBIQUINONE OXIDO REDUCTASE

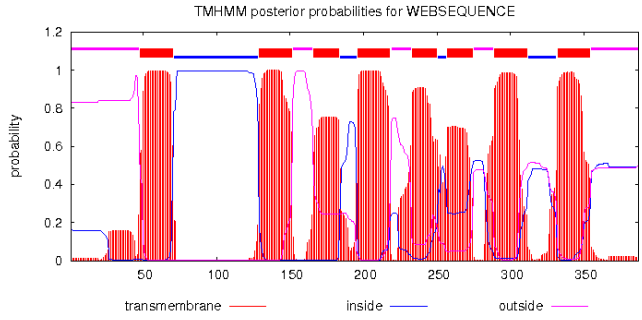
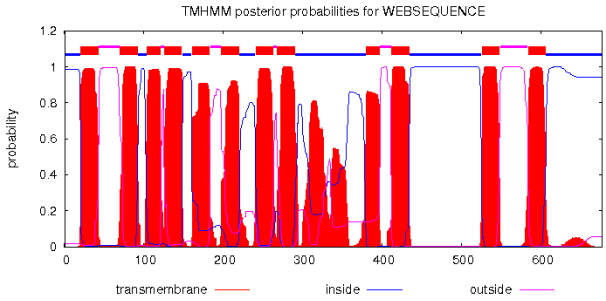


Figure-3:NADH DEHYDROGENASE SUB UNIT-F



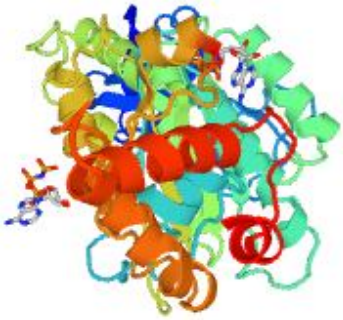
The secondary structure prediction of *Annona muricata* proteins (Table-3) was analyzed by SOPMA which revealed that alpha helix, extended strand, beta turn and random coil, were more predominant. In all the ten proteins alpha helix dominates which is followed by random coil, extended strand and beta turn. The secondary structure were predicted by using default parameters (Window width: 17, similarity threshold: 8 and number of states: 4). TMHMM v.2.0 and SOSUI predicted that 6 proteins were soluble protein remaining four proteins are Transmembranes proteins.

TABLE5 : **INTERPRO RESULTS OF PROTEINS OF ANNONA MURICATA**

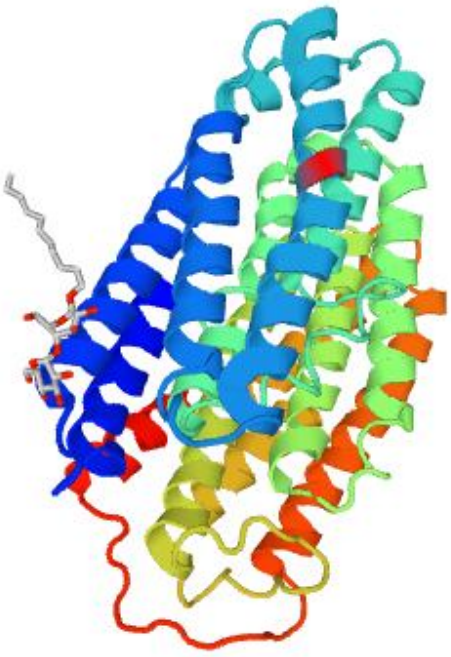
S.NO	ACCESSION NUMBER	SUPER FAMILY	MOLECULAR FUNCTION
1	AKA39062.1	Rubisco-C terminal domain	Mg ion binding, Ribulose bis po4 carboxylase activity
2	AAF17027.1.	C-terminal domain of α & β subunits of F1 ATP synthase	ATP binding ATP synthase transporting ATPase activity rotational mechanism
3	CAB90103.1.	C-terminal domain of α & β subunits of F1 ATP synthase	ATP binding ATP synthase transporting ATPase activity rotational mechanism
4	ABK91377.1	Transmembrane di-heme cytochromes	Electron carrier activity Metal ion binding
5	AAQ11840.1	-	RNA binding
6	ABD78205.1	-	NADH dehydrogenase (ubiquinone) activity
7	AB845075.1.	-	Quinone binding NADH dehydrogenase (ubiquinone) activity
8	ADL63753.1.	Ribosomal protein C- terminal domain S3	Structural constituent of ribosome ATP binding, Phosphorelay sensor kinase activity, Protein serine/ threonine kinase activity
9	AML79553.1	PYP – like sensor domain (PAS) domain	DNA binding , DNA – directed RNA polymerase activity
10	ABD95745.1	-	

Fig 4: TERTIARY STRUTURE OF *ANNONA MURICATA*

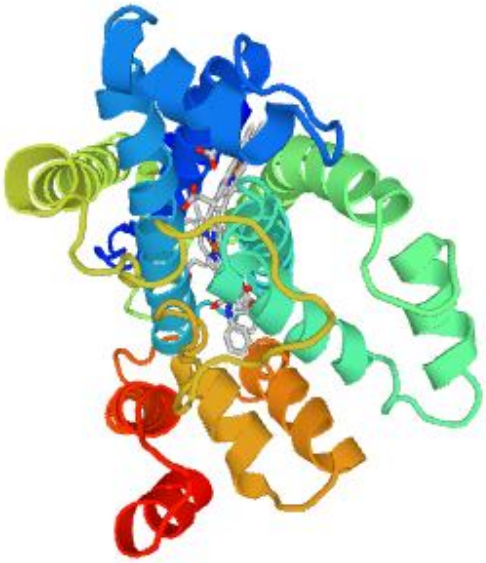
ATP synthase alpha



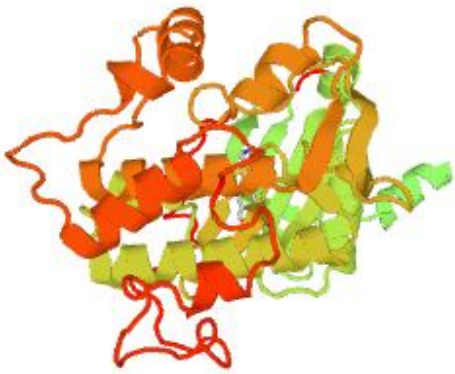
structure of cytochrome b:



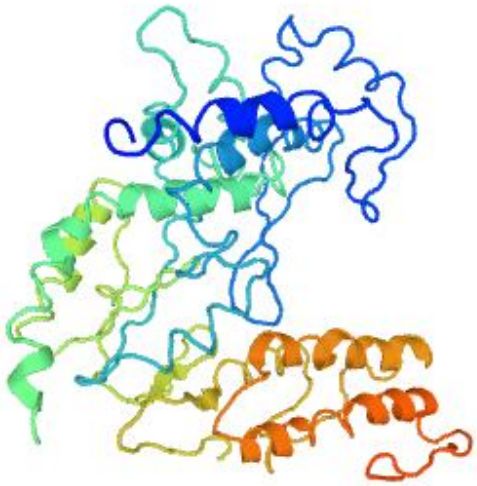
Putative LOV domain



Maturase K



RNA beta chain polymerase



NADH oxido reductase

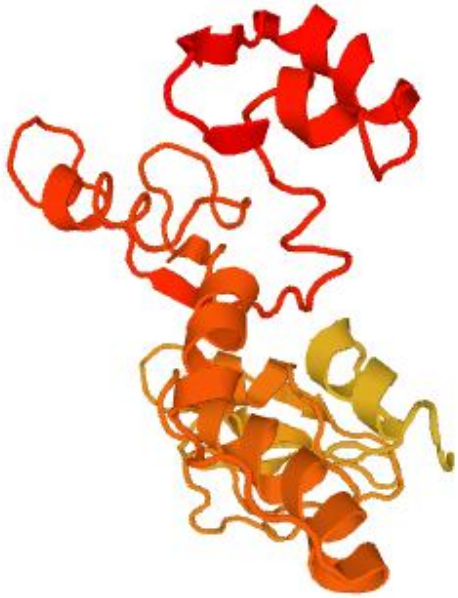
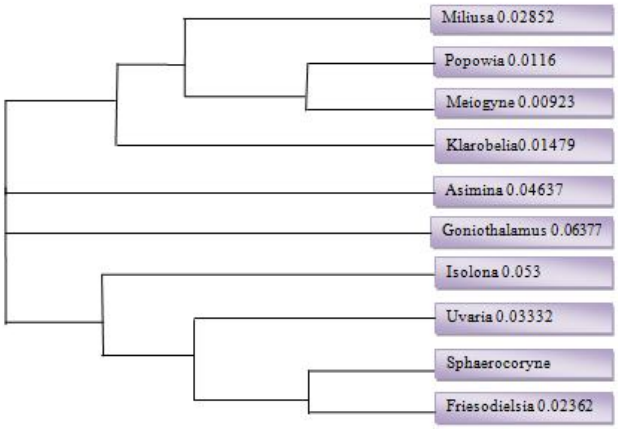


TABLE 6: LISTS OF PLANT SPECIES SHOWING SIMILARITY OF 85% AND ABOVE WITH THE (NADPH UBIQUINONE OXIDO REDUCTASE)

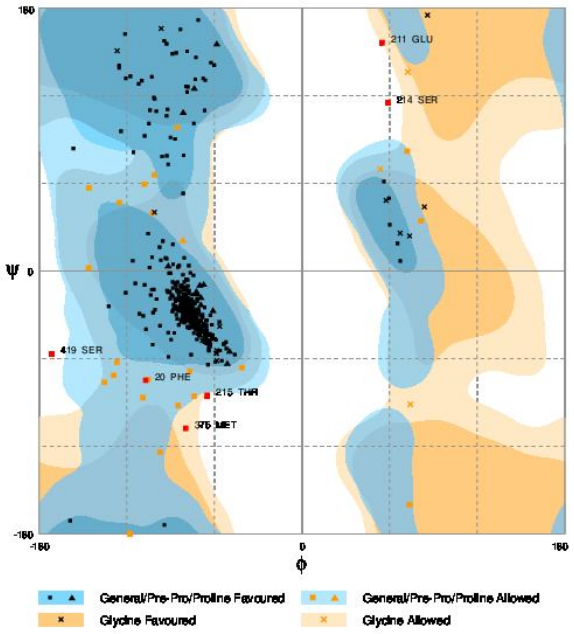
S.NO	Plant Species containing NADPH protein	Family	Accession Number	Identity
1	Popowia hirta	Annonaceae	AFW05355.1	86%
2	Asimina triloba	Annonaceae	AAO85891.1	92%
3	Goniothalamus macrophyllus	Annonaceae	AJF44256.1	88%
4	Sphaerocoryne affinis	Annonaceae	AKA97839.1	87%
5	Uvaria zeylanica	Annonaceae	AKA97846.1	87%
6	Friesodielsia obovata	Annonaceae	AFM94242.1	87%
7	Isolona campanulata	Annonaceae	ABY56051.1	87%
8	Meiogyne stenopetala	Annonaceae	AGA12904.1	86%
9	Klarobelia inundata	Annonaceae	AAW29488.1	86%
10	Miliusa velutina	Annonaceae	AFW98676.1	87%

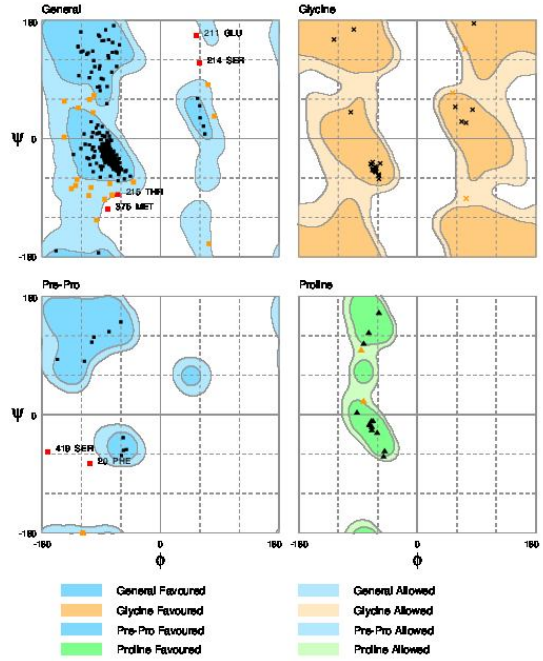
Fig 5: PHYLOGENETIC TREE OF PROTEIN NADPH UBIQUINONE OXIDO REDUCTASE



Insilico methods to compare the closeness in sequence of NAD(P)H quinone oxidoreductase, a protein responsible for nanoparticle synthesis. The phylogenetic tree obtained, revealed the distant relationship of *Annona muricata* with *Asimina triloba* (above 90%) whereas remaining all they are closely related to *Uvaria zeylanica*, *Meiogyne stenopetala* etc. A large number of medicinal plants are being exploited and evaluated from nature for the profitable production of drugs. Green nanotechnology refers to the process of synthesising nanoparticles from plants and microbes. Protein assays revealed that NADH dependant oxidoreductase is the main reason for nanoparticles synthesis [14]. The NADH is oxidised to NAD + by reductases and so it gains ability to reduce metal ions. The protein NAD(P)H quinone oxidoreductase of *Annona muricata* are concentrated in this study because it is responsible for synthesis of nanoparticles [15]. On performing sequence similarity and Phylogenetic analysis of a number of plant species were found to be similar to *Annona muricata*. Therefore, these plants including *Annona muricata* can be used to synthesis nanoparticles and therapeutic agents as they are abundant and available free of cost.

Fig 6: RAMACHANDRAN PLOT OF NADPH UBIQUINONE OXIDO REDUCTASE





The Ramachandran plot analyses reveals that the CPH model has 98.0% of the residues are in favoured region, 2.0% of the residues are in allowed region and 1.4% of the residues are in outlier location (Figure -6) [16].

CONCLUSION

The use of *in silico* methods for drug discovery in natural products has increased during the previous decade. The appearance of new chemo- and bioinformatics methods, along with a growing range of OMICS data and data on phytochemical structures opens vast perspectives in the study of pharmacological activity of plant preparations. Nevertheless, scientists should consider the quality of the data and computational models used. Therefore, despite the increasing number of known phytoconstituents, not all pharmacological effects of medicinal plants may be modeled by their action on drug targets. However, even the application of currently available chemo- and bioinformatics resources and approaches provides valuable information for discovery of novel applications of medicinal plants beyond their traditional use. *Anona muricata* were selected. Expasy's ProtParam tool predicted the physio-chemical characters of the proteins. Further analyses are required for drug target identification.

The present study is a cost effective method to find related plant species possessing nanoparticles synthesizing ability when provided with suitable precursors which can be further carried out for laboratory experimental methods to obtain significant nanoparticles by green synthesis from plants.

REFERENCES

[1] Y. Vaghasiya, H. Patel, and S. Chanda, *Afric. J. Biotech*, **2011,** 10(70): 15788-15794.

[2] Mishra S., Ahmad S., Kumar N., Sharma B.K. *Annona muricata* (the cancer killer): A review. Glob. J. Pharm. Res. 2013;2:1613–1618.

[3] Nawwar M., Ayoub N., Hussein S., Hashim A., El-Sharawy R., Wende K., Harms M., Lindequist U. Flavonol triglycoside and investigation of the antioxidant and cell stimulating activities of *Annona muricata* linn. Arch. Pharm. Res. 2012;35:761–767. doi: 10.1007/s12272-012-0501-4.

[4] Mansoori, G. A., "Principles of Nanotechnology: Molecular Based Study of Condensed Matter in Small Systems," World Scientific Pub., New York, 2005.

[5] Yih, T.C.; Al-Fandi, M. Engineered nanoparticles as precise drug delivery systems. *J. Cell. Biochem.* **2006**, *97*, 1184–1190.

[6] T. T. H. Dao, H. J. M. Linthorst, and R. Verpoorte (2011). Chalcone synthase and its functions in plant resistance. Phytochemistry Review. 10 (3): 397–412.

[7] Gasteiger, C.Hoogland, A.Gattiker, S.Duvaud, M.R.Wilkins, R.D. Appel, A.Bairoch

[8] (2005). Protein Identification and Analysis Tools on the ExPASy Server, (In) John

[9] M.Walker (ed): The Proteomics Protocols Handbook, Humana Press. 571-607.

[10] Fabina Sievers et al (2011). Fast, Scalable Generation of high-quality protein multiple
[11] sequence alignments using Clustal Omega. Molecular Systems Biology. 7:539.
[12] 9. G.Perriere and M. Gouy. (1996). WWW-Query: An on-line retrieval system for biological sequence banks. Biochimie.78: 364-369.
[13] Upgade Akhilesh, Bhaskar, Anusha, Issar Sakshi and Senthamarai Selvi V (2012). In Silico Characterization of Keratitis Causing Herpes Simplex Virus (HSV 1) Membrane Proteins using Computational Tools and Servers. Research Journal of Recent Sciences. 1(11): 27-31.
[14] T. Shohaib, M.Shafique, N. Dhanya, Madhu.C.Divakar. Importance of flavonoides in therapeutics. Hygeia. Journal for Drug and Medicne. 2001; 3(1): 1-18.
[15] G.Perriere and M. Gouy. (1996). WWW-Query: An on-line retrieval system for biological sequence banks. Biochimie.78: 364-369.
[16] Kebila Venkatasamy. In silico Analysis and Homology Modeling of Putative Hypothetical Protein Q4QH83 of *Leishmania* major. Advanced BioTech.2013; 13 (3): 1-4.
[17] Kumari A, Yadav SK, Pakade YB, Singh B, Yadav SC. Development of biodegradable nanoparticles for delivery of quercetin. Colloids Surf B Biointerfaces 2010; 80: 184-92.
[18] Li, R., Bianchet, M. A., Talalay, P. & Amzel, L. M. (1995). The three-dimensional structure of NAD(P)H:quinone reductase, a ¯avoprotein involved in cancer chemoprotection and chemotherapy: mechanism of the two-electron reduction. *Proc NatlAcad Sci USA* **92**, 8846±8859.
[19] Ramachandran, G.N.; Ramakrishnan, C.; Sasisekharan, V. (1963). "Stereochemistry of polypeptide chain configurations". Journal of Molecular Biology. **7**: 95–9.