IN SILICO STRUCTURAL AND FUNCTIONAL ANALYSIS OF *Bacillus megaterium* ARGINASE

A

Thesis report

Submitted in partial fulfillment of the requirement for the degree of

MASTER OF SCIENCE

IN

BIOTECHNOLOGY

BY

DISHA ROHAL (197812)

UNDER THE GUIDANCE OF

DR SAURABH BANSAL

MAY – 2021

DEPARTMENT OF BIOTECHNOLOGY AND BIOINFORMATICS

JAYPEE UNIVERSITY OF INFORMATION TECHNOLOGY,

WAKNAGHAT, SOLAN, HIMACHAL PRADESH -173234.

DECLARATION

I do hereby declare that this dissertation entitled **"In silico structural and functional analysis of *Bacillus megaterium* arginase"** submitted towards fulfillment for the award of the degree of **Masters of Science in Biotechnology from the Department of Biotechnology and Bioinformatics, Jaypee University of Information Technology Waknaghat** is wholly based on the study and results carried out under the guidance of **Dr Saurabh Bansal.** Also, till now, this work has not been submitted anywhere for any additional degree or diploma. Therefore, the declaration made is true and genuine.

Disha Rohal (197812)

Department of Biotechnology and Bioinformatics

Jaypee University of Information Technology Wakhnaghat

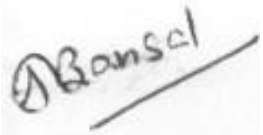
Solan-173234, H.P.

Date: 21th May, 2021

CERTIFICATE

This is to certify that the work which is being presented in the thesis report titled **"In silico structural and functional analysis of *Bacillus megaterium* arginase"** in partial fulfillment of the requirements for the award of the degree of **Master of Science in Biotechnology** submitted to the **Department of Biotechnology and Bioinformatics, Jaypee University of Information Technology, Waknaghat** is an authentic record of work carried out by **Ms Disha Rohal (197812)** during a period from July 2020 to May 2021 under my supervision at Department of Biotechnology and Bioinformatics, Jaypee University of Information Technology, Waknaghat. This work has not been submitted partially or wholly to any other University or Institute for the award of this or any other degree or diploma**.**

Date: 25-May-2021



Dr SAURABH BANSAL

Assistant Professor

Department of Biotechnology and Bioinformatics

Jaypee University of Information Technology

Wakhnaghat, Distt-Solan, H.P. – 173234

saurabh.bansal@juit.ac.in

ACKNOWLEDGEMENT

There was a part of individuals who were without a doubt exceptionally supportive, kind, agreeable and affable along with the advancement of this project as well. I would like to express my acknowledgement to those concerned.

This extends and the report would have been outlandish in the event without the kind help, extraordinary back and perpetual instruction of our scholarly project guide. I am exceedingly respect–bound to **Dr Saurabh Bansal**. His unfavorably offer supervision, consistent support, constant encouragement, liberal nature are among the exceptionally few favoring that they offered upon me from time to time from the initiation of this project to the day of its completion.

I am very thankful to **Dr Saurabh Bansal** for his continuous assistance in the viable work that came to fulfilment.

Date:

Disha Rohal - 197812

TABLE OF CONTENTS

LIST OF FIGURES

LIST OF TABLES

ABBREVIATIONS

pI	Isoelectric point
GRAVY	Grand Average Hydropathicity Index
B. megaterium	*Bacillus megaterium*
SmF	Submerged fermentation
SSF	Solid-State Fermentation

ABSTRACT

Arginase is a ubiquitously occurring metalloenzyme that shows an anticancer property. It can be isolated from various species of plants, animals, bacteria and fungi. Although several bacterial species have already been explored to get new arginases, they are not up to the mark for application in cancer treatment. As a result, there is a need for a new arginase with a longer half-life, improved stability, and efficacy. Therefore, the present study is aimed to conduct an in silico study of the arginase of *Bacillus megaterium* to analyze its physicochemical properties and their correlations with their structure. For performing the current study, full-length sequences of *Bacillus* arginases were procured from the UNIPROT database and then analyzed the physicochemical and secondary structures and active site composition using various online bioinformatics tools. PROTOPARAM is used to analyze physicochemical characteristics. SWISS-Model and I-Tasser server were used for making the 3D structure of the *B. megaterium* arginase (BmA). The current study suggests that *B. megaterium* arginase is a thermostable enzyme with a slightly acidic pH. All other properties of *Bacillus megaterium* arginase are in the line as for from different *Bacillus* sp. Hence, the current studies suggest that *Bacillus megaterium* arginase also has similar potential as other arginases of various *Bacillus* sp. Further experimental validation is needed for the characterization of BmA to establish its potential application in cancer treatment.

CHAPTER 1

INTRODUCTION

1. INTRODUCTION

1.1 What is arginase

Arginase takes stayed and rummage-sale to extravagance arginine reliant on cancers. Arginase (EC 3.5.3.1, arginine amidinase, canavanase, L- arginase, arginine transamidase) is a manganese encompassing enzyme that catalyzes the deamination of L- arginine to L- ornithine and urea. Arginase is an enzyme that catalyzes the last phase of the urea cycle, which will help to marshal noxious constituents by altering L – arginine to L – ornithine and urea [1]. Arginase is one of the important medical enzymes [2]

There are two forms of arginase[3]

a) Type 1 is liver type
b) Type 2 is mitochondrial type

Arginase type 1 is mainly present in the mammalian liver and the mammary glands.

Arginase type 2 is a mitochondrial enzyme, and it is articulated in the kidney.

1.2 Arginase applications

Arginase affect the arginine auxotrophic malignancies, such as melanoma, lung cancer, renal cell carcinomas and hepatocellular carcinomas, and acute neurological disorders, supervisor of penial and vaginal movement, therefore in concert an imperative protagonist the masculine besides feminine erotic provocation, neural regeneration pathways, maintenance of semen quality, treatment of Hepatitis -B, biosensors for monitoring arginine levels in juice samples, prostate cancer, human T- lymphoblastic leukaemia, and osteosarcoma, rheumatoid arthritis therapy, an unintended controller of penial and vaginal movement playing an imperative character in man and feminine sexual stimulation [4].

Presently recombinant human arginase (rhArg1) is one of the arginases which is under clinical trials. But it has several side effects such as abdominal pain, bloating, diarrhoea, worsening of asthma, and low blood pressure. It can also cause nausea, vomiting and blood abnormalities. Sometimes it may cause liver failure. Therefore, we scan the Uniprot database for available and unexplored arginase sequences to search for novel arginase. Out of many unexplored arginase sequences, the *Bacillus megaterium* arginase sequence was selected for further analysis as the *Bacillus* is known for synthesizing various industrial and medicinal

importance enzyme production. So in the same line, *B. megaterium* can be a good source of arginase and also it is known to be a GRAS organism. Also, there is no in silico and wet-lab studies available for *B. megaterium* arginase (BmA). Therefore, the current work is aimed for in silico analysis of BmA using various bioinformatics tools to understand its functional and structural properties, including sequence analysis, physicochemical characterization, secondary structure prediction, functional analysis, and structure classification. This work will provide an initial understanding of the enzyme's properties and its potential application in cancer treatment.

CHAPTER 2

REVIEW OF LITERATURE

2. REVIEW OF LITERATURE

2.1 Background

Malignancy is connected through a large number of impermanence tolls granting high amounts of improvements that have been made in early recognition and dealing. Major techniques involved in cancer treatment are radiotherapy, chemotherapy and surgery; these depend upon the type and requirement of the tumor. But these types of therapy have maximum chances, which can cause damage to the normal cells [5]. Scientists try to find out different kinds of treatments to cause minimal damage to the normal cells.

It is a well-known fact that cancer cells require a higher amount of nutrition than normal cells. Cancer cells get their nutrition from external sources; on the other hand. Normal cells can capability grow from their nutrition.

Arginine is a semi-essential amino acid and the major sources of arginine are dietary products and synthesized in the kidney and liver. Roughly tumor cells require amino acid arginine for growth. Several studies show that the exhaustion of arginine can exaggerate tumor cells. Normal cells can synthesize their arginine, but tumor cells require their arginine from the external environment.

2.2 History of arginase

There were two scientists named Krebs and Henseleit who showed a sequence of experimentations by exploitation of the portions of liver and manometric assays to illustration down the occurrence of arginase; ornithine produced urea [6]. There were different events related to arginase its first discovery, preparation, inhibitor investigation that is shown in (Table 1)

Table1 Some events related to the development of arginase with Scientist's name

YEAR	EVENT	SCIENTIST
1931	Rudimentary arrangements of arginase were testified	Salaskin and Solowjew
1935	First initiation by divalent metallic ions counting cobalt, nickel, manganese, and iron	Hellerman and Perkins

1940	Improved purification development exhibited that the commotion might be re-established to pH – incapacitated arginase by Mn^{2+} and Fe^{2+}	Richards and Hellerman
1956	The limited decontamination was supplementarily amended	Robbins and Shields
1975	Investigating inhibitors	Rosenfeld
1977	Deliberate the consequence of the merchandise/ inhibitor orthenine, and anticipated an allosteric archetypal for a directive of enzyme's commotion	Bedino
2006	Current investigates the protagonists of arginase in vascular illness, pulmonary illness, infectious illness, and cancer.	Zimmerman and Rothenberg

2.3 Different roles of arginase

2.3.1 Role in the skeletal – muscle system

High intensity dynamic human muscles performance is enhanced by arginine, and amino acid is now being promoted throughout the healthy living sector involving dietary supplements. Arginine can diminish the paraphernalia of sepsis and infection, concluded the NO alleyway.

2.3.2 Role in nervous systems

Arginine plays a role in regenerating dented axons of neurons, substitute as a negotiator for undignified proteins scratched through axon injury. It has been claimed that it also acts as a memory enhancer and hormonal control modulator. It is convenient to handle Alzheimer's disease through its ability to increase polyamide levels and help repair damaged axons.

2.3.3 In skin

Arginine plays a role in healing wounds by stimulating the proliferation of fibroblasts, the proclamation of anthropological evolution hormone and collagen manufacture. Treatment increases vascular endothelial growth factor release, which helps to keep the skin young and improve scleroderma. A high risk of infection after cardiac surgery can be reduced where the healing of wounds is dramatically accelerated.

2.3.4 In cardiovascular

Arginase is an abundant enzyme that legalizes polyamine- and nitric-oxide-necessitating vascular purposes. It is fine-conventional that, in creatures, arginase overactivation subsidizes endothelial dysfunction, a symbol of cardiac sicknesses [7].

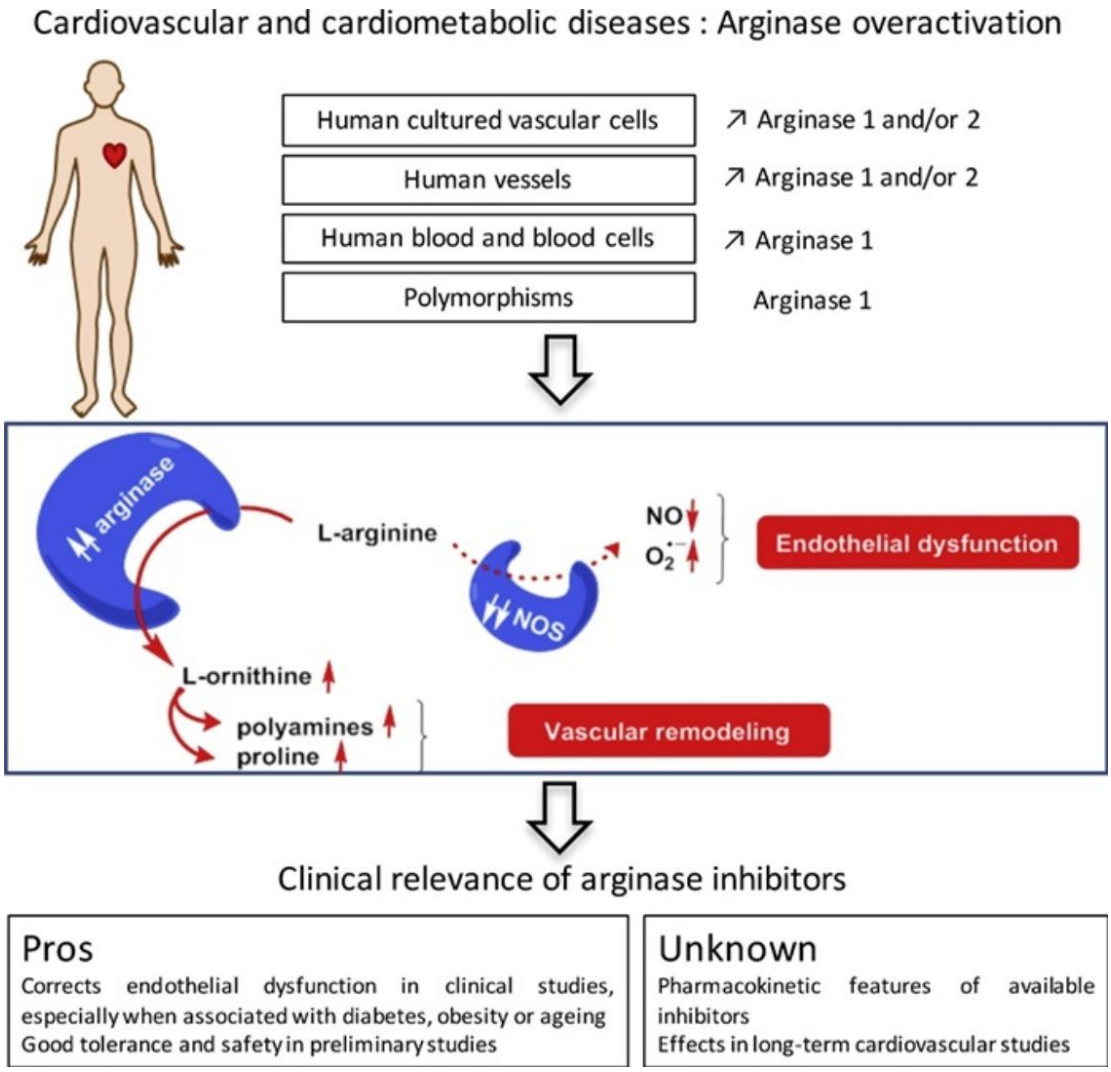
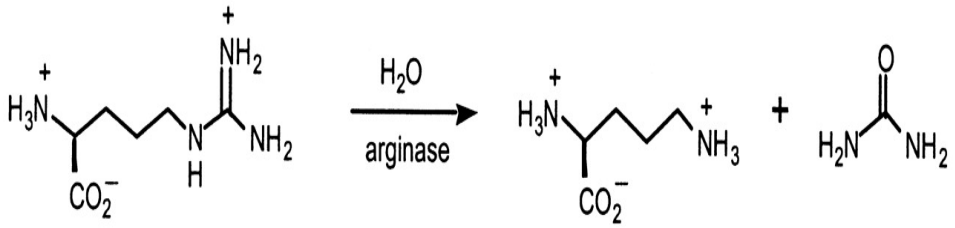


Figure 1 Role of Arginase in Cardiovascular and cardiometabolic diseases imported from [7], [8]

2.4 L- Arginase reaction

The arginase catalyzes the bivalent positive ion reliant on hydrolysis of L- arginine to arrange the non – protein amino acid L- ornithine and urea.



Arginine Ornithine Urea

2.5 Mechanism of arginase as anticancer drugs

When arginase is converted into ornithine and urea, the normal cells will not be affected because they can manufacture arginine from citrulline using enzymes argininosuccinate synthase (ASS) and argininosuccinate lyase (ASL), which are not expressed in the cancerous cells shown in (Figure 2).

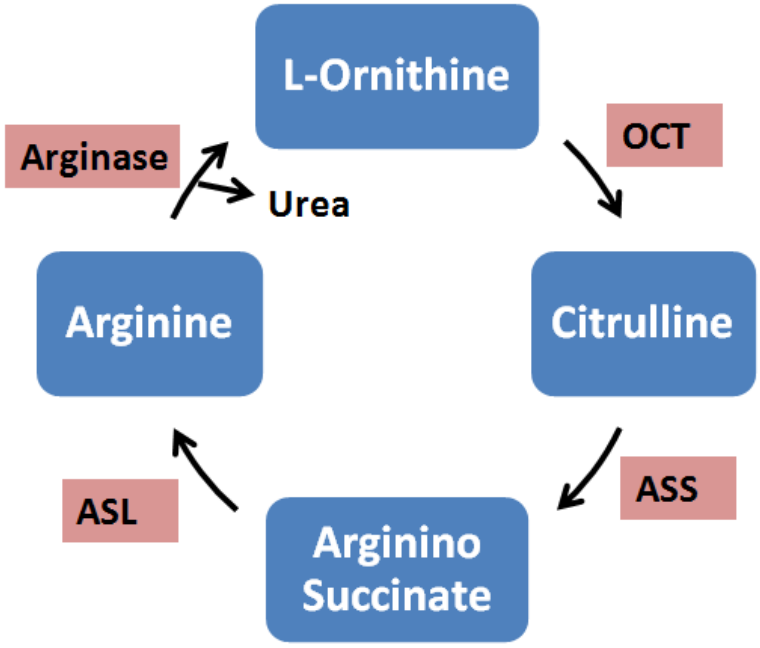


Figure 2 Arginine metabolism in humans. Here OCT: Ornithine Carbamoyl Transferase, ASS: Argininosuccinate synthase, ASL: Argininosuccinate lyase.

2.6 Different Sources of L- arginase

L- arginase has been identified and characterized from microbial, plants and animal sources. Microbial sources are subdivided into bacteria, fungi, yeasts and protozoa.

2.6.1 Bacterial sources

Helicobacter pylori roc F gene is the gene that encodes the arginase, which plays a vital protagonist in the pathogenesis of its infection. Arginase has similar molecular parameters

and kinetic constants to that of bacteria *Saccharomyces cerevisiae.* Other bacterial sources are listed below (Table 2).

Table 2 Bacterial sources of L – arginase

BACTERIAL SOURCES	REFERENCE
Bacillus licheniformis	[9]
Bacillus subtilis KY 3281	[10]
Agrobacterium Ti plasmid C58	[11]
Bacillus caldovelox	[12]
Streptomyces clavuligerus	[13]
Nocardia lactamdurans	[14]
Bacillus brevis	[15]
Helicobacter pylori	[8]
Rhodobacter	[16]
Cyanobacteria	[17]
Bacillus anthracis	[18]
Chlamydia pneumonia	[19]
Arthrobacter sp.	[20]

2.6.2 Fungal sources

Neurospora crassa, this fungus, does not possess a urea cycle and has a uricotelic arginase. Glycosylated arginase was secreted on or after *Xanthoria parietina* thallus persuades forfeiture cytoplasmatic quantifiable as of its photobionts. According to studies, Arginase is

also isolated from *Neurospora crassa* as isoforms and lots of fungal sources are available according to studies (Table3).

Table 3 Fungal sources of L – arginase

FUNGAL SOURCES	REFERENCES
Aspergillus nidulans	[21]
Xanthoria parietina	[22]
Trichoderma sp.	[23]
Agaricus bisporus	[24]
Neurospora crassa	[25]

2.6.3 Yeast sources

S. cerevisiae arginase forms multienzyme composite through ornithine transcarbamoylase on which arginase will act as an undesirable allosteric effecter of ornithine transcarbamoylase. Different types of yeast sources are listed below (Table 4).

Table 4 Sources of L – arginase from yeast

YEAST SOURCES	REFERENCE
Schizosaccharomyces pombe strain 972	[26]
Xanthoria prientina	[27]
Evernia prunastri	[28]
Peltigera canina	[29]
Leptogium corniculatum	[30]

2.6.4 Plant sources

Arginase is the enzyme that has been reported to isolate from Xylem sap of apple shoots. Germinating pumpkin cotyledons shows an increased arginase activity[31]. Arginase is also present in tulips in large an amount. The Arginase of cherry tomatoes protects them from chilling injury. Arginase can be isolated from many other plant sources (Table 5).

Table 5 Sources of L – arginase from plants

PLANT SOURCES	REFERENCES
Lathyrus sativus seeds	[32]
Vicia faba	[33]
Vitis vinifera	[34]
Canavalia ensiformis	[35]
Glycine max	[36]
Actinidia deliciosa var. deliciosa	[37]
Pisum sativum	[38]
Lycopersicon esculatum	[39]
Saccharum officinarum	[40]
Arabidopsis	[41]
Panax ginseng	[42]
Pinus taeda	[43]

2.6.5 Animal sources

Arginase can also be isolated from the gut of earthworms. Gills and bottom influence tissues of *Chiton latus,* a common mollusc, has been used as a source of arginase. The haemolymph of the giant African land snail is another source that has been reported to extract the arginase enzyme. There are several other sources from where arginase can be isolated (Table 6).

Table 6 Sources of L – arginase from animals

ANIMAL SOURCES	REFERENCES
Human heart	[44]
Human erythrocytes	[45]
Genypterus maculates	[46]
Clarias batrachus	[47]
Enthrocytes of pigs	[48]
Mus booduga	[49]
Human saliva	[50]
Buffalo liver	[51]
Notothenia neglecta	[52]

Fruit bat	[53]
Felis catus (tissues of cat)	[54]
Notothenia rossii	[55]
Rana temporaria	[56]

2.7 Isolation of arginase

For isolation of the arginase, the first step is the identification of a suitable source of arginase. The bacterial strains can be isolated from nature in different ways. Several samples need to be collected from various sites such as soil or some watersides for the same. The samples are then serially diluted and streaked or poured off over the nutrient agar plate. The isolates are then allowed to grow by incubating at 37°C for 48 hours. The evolved isolates are then analyzed for arginase production and its desired characteristics.

2.8 Production of arginase

Many past years, human beings used different types of microorganisms and their enzymes to produce bread, cheese, and wine but nowadays, enzymes also used as a therapeutic agent, so many studies claim that those enzymes were produced by controlled and contained fermentation in closed fermentation tanks, by using proper fermentation tanks, and exploitation of glowing demarcated production strain. There are various types of methods from which we produce arginase by fermentation, including solid-state fermentation (SSF) and submerged fermentation (SmF). The enzyme production by fermentation has different advantages, such as it ensures a consistent quality of the product with a very high production yield. SSF has a high production yield compared to SmF. Bacteria are the primary sources for L-arginase production (Table 7). A general schematic flow chart for the Production and Purification of L-arginase is given in Figure 3.

Table 7 Production method of bacterial L–arginase

S. No.	Organism	Production Method	Medium	Opt. Temp	Opt. pH	Nitrogen and carbon sources	Yield	References
1	***Bacillus firmus***	SSF	Nutrient Broth	37°C	5	Glucose	7 – 15 IU/ml	[57]
2	***B. subtilis***	SSF	Nutrient Broth	45°C	8.4	Yeast extract / Glucose		[58]

3	***B. cereus***	SSF	Nutrient Broth	33°C	5	Glucose	31µg/ml	[57]
4	***B. licheninformis***	SSF	Nutrient Broth	25°C	8	Yeast extract / maltose		[57]

Cultural isolation

↓

Screening of isolate (*B. megaterium*)

↓

Preparation of production media

↓

Production by submerged fermentation

↓

Determination of arginase activity

↓

Partial purification of the enzyme

↓

Kinetics of the partially purified enzymes

↓

Bioprocessing of arginase enzyme under solid-state fermentation

↓

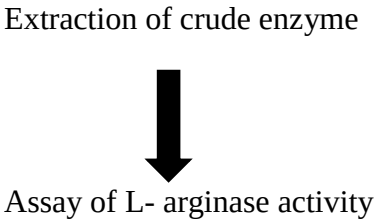


Figure 3 Flow chart of Production and Purification of arginase from *B. megaterium* [59]

2.9 Purification of arginase

Arginase can be purified from different microorganisms, plant and animal species by different purification methods and that is listed below in (Table 8)

Table 8 Purification methods and characterization of arginase from different sources

Sr. No	Name of organism	Purification methods	K_m	V_{max}	References
1	*Bacillus thuringiensis*	Gel filtration	15.6 mM	538.9 U/mg	[60]
2	Human Arginase I expressed in *Saccharomyces cerevisiae*	Metal affinity chromatography	21.8 mM	1600 U/mg	[61]
3	*Pseudomonas aeruginosa* IH2	Ion Exchange, Gel filtration chromatography	-	-	[62]
4	*Rummeliibacillus pycnus*	Column chromatography	0.212 mM	1111 U/mg	[63]

2.10 Arginase and cancer

Several surveys have revealed that affected role through enduring provocative illnesses or conditions such as tumor, autoimmunity, contagions, and corporeal ordeal, among others, have lessened T lymphocyte answers. Upsurges in L-arginine breakdown, specifically through polyp cells, myelic cells, and freshly trendy distinctive lymphoid cells, yield occurred as per a chief intermediary in the intonation of T-cell rejoinders through pathologies

related to enduring tenderness. Though this vigorous ground is still emerging, empathetic of the consequence of L-arginine absorption purpose of invulnerable cells in the tumor.

Tumor cells need nutrition from the outer environment to get extensive growth. Arginine is a nutritionally essential amino acid and plays very different types of functions in its cell activity. When a tumor cannot synthesize arginine vulnerably, it will be named as arginine auxotrophic. After that, arginine diminution will be attributed to Achille's heel in cancer and used to treat arginine auxotrophic cancers.

2.10.1 Advantages of arginase in cancer

As a multipurpose amino acid, arginine plays an imperative role in physique well-being. Arginine is tangled in fatal expansion, brand-new development, tissue injury and chronic metabolic diseases.

The request of arginine on a renal grievance is contingent on NO since disproportionate NO subsidized to the renal grievance. Lump cells have to inflate nourishing desires to encounter exhaustive evolution. Certain tumours mislay the capability to manufacture arginine vulnerably, explicitly as arginine auxotrophic. Therefore, arginine diminution was attributed as Achilles' heel in cancer treatment for arginine auxotrophic tumors.

2.10.2 Drawbacks of arginase in cancer

The premature arginase inhibitors had numerous adjacent paraphernalia due to their low-slung effectiveness and broad-spectrum arrangements. For example, (DFMO) is a scrawny inhibitor of arginase but is broad-based because the situation is correspondingly a compelling inhibitor intended for ODC. Thus, while DFMO conduct can intensify NO creation in mock-ups of preeminent arginase commotion, this consequence might remain facilitated through the situation accomplishment in endorsing orthenine accumulations known to constrain arginase.

2.10.3 Recent work in arginase enzyme

It is cumulative vibrant that arginine has pedestals of biotic purposes in cellular commotion together with tumor pathogenesis. Arginine scarcity treatment unfastens up a new-fangled realm of meticulousness and embattled cancer dealing for certain cancer patients. More and more biomarkers of compassion and resistance to arginine scarcity will be acknowledged to afford methodical pulverized for forthcoming medical expansion.

2.11 Therapeutic tactics

Initial arginase inhibitors have masses of lateral possessions due to their short effectiveness and non-explicit movements. For example, difluoro-methyl ornithine (DFMO) is considered as a feeble inhibitor of arginase nonetheless then broad-based in that it is correspondingly an active inhibitor of ODC. Therefore, though DFMO behaviour can intensify the NO manufacture in the presence of high arginase commotion, this result may remain influenced by the ability to promote ornithine accretion, which is acknowledged to constrain arginase. The embarrassment of ODC will similarly circumscribe polyamine construction, which may perhaps diminish oxidative pressure by averting the retrograde breakdown of spermine and spermidine by polyamine oxidases.

2.12 Novel arginase for treating cancer

Deuce isoforms of enzyme occur Arginase 1 is a cytosolic enzyme that mainly originates in hepatocytes, wherever it acts as a grave part in eliminating NH^3 over urea amalgamation. Arginase 2, a mitochondrial enzyme extremely articulated in organ meats, elaborate in the creation of ornithine, an ancestor for polyamines and prolines significant for cell propagation and collagen manufacture.

Lumps practice manifold invulnerable oppressive devices towards dodge resistant classification. Ace of these is the lessening of L-arginine finished augmented heights of mingling arginase, augmented countenance, and excretion of arginase by growth cells, and enrolment of arginase articulating and concealing myeloid resulting suppresser gene cells. Pharmacological reserve of arginase motion has been reported to counteract the low L-arginine-persuaded protected destruction in animal replicas. Thus, it is essential to discriminate arginase inhibitors to prevent invulnerable destruction and re-energize anticancer protection in patients.

The contemporary tender designates a sequence of original arginase inhibitors intended to handle growth and respirational provocative sicknesses. Respirational provocative diseases comprise asthma or chronic obstructive pulmonary disease (COPD). Additionally, the submission unveils amalgams, their grounding, usage, pharmacological configuration, or handling.

2.13 In Silico Study

In silico term is devised in 1989. In Silico is an expression used to mean "performed on computer or via computer simulation". For in silico studies, various bioinformatics tools can

be used for the structural and functional analysis of the protein based on their primary structure (Table 9).

Table 9 Tools used in the protein prediction

Sr. No.	Tools	Basic function
1	UniPtrotkb	Sequence retrieval
2	BLASTp	Sequence analysis
3	Clustal-w	Multiple sequence alignment
4	Expasy-ProtParam	Physicochemical characterization
5	PSIPRED or CFSSPS	Secondary structure prediction
6	Motif search tool	Functional analysis
7	CATH server	Structural classification
8	SWISS MODEL/ I- TASSER	Homology Modelling/ Threading
9	QMEAN and UCLA	Evaluation of predicted protein

2.13.1 UniProt

UniProt is the universal protein resource. An essential source (storage area) for protein data. UniProt is an inclusive (large scope), high-value, spontaneously reachable catalogue.

UniProt remains formed by the UniProt Association - a collaboration between

- EBI
- SIB
- PIR.

Use of UniProt = Useful in finding protein classification and efficient data.

- Useful for inferring evolutionary information.
- We also treasure the role of proteins by
    - Perusal possessions such as the biotic methods they are elaborate in.
    - Their column-translational alterations.
    - Their collaboration by extra particles.
    - Place in cells and creatures.

Components = UniProt archive (UniParc). UniParc contains only protein sequences.

2.13.2 BLASTp

BLAST: Basic Local Alignment Search Tool (BLAST) compares genetic factors and protein arrangements in contradiction of others in communal catalogues.

BLAST is a usual categorization assessment procedure second hand to exploration records for best local alignments to a query.

Protein-protein BLAST (BLASTp) gives the protein query and proceeds the maximum alike protein sequences since the operator stipulates the protein catalogue.

2.13.3 ClustalW

Multiple sequence alignment: MSA is the arrangement of 3 or additional biological orders of parallel extent. From the production of MSA requests, homology can be indirect and the evolutionary connection between the sequences premeditated.

A popular heuristic algorithm is ClustalW by Des Higgins and Paul Sharp in 1988.

ClustalW makes multiple global alignments using a progressive alignment approach.

The ClustalW processing: First, it computes all pairwise alignment and calculates sequence similarity between pairs

1. Then align the most similar pairs of sequences (this gives us an alignment of 2 sequences called 'profile').
2. Align the next closest pair of sequence (or pair of profiles or sequence and profile).
3. Align the nest closest pair of sequences/profile

2.13.4 ExPASy (PROTOPARAM)

ProtParam is an implement tool that permits the totalling of numerous corporal and chemical structures for a given protein stored in Swiss-Prot or TrEMBL or for a user-entered sequence.

The calculated structures comprise the molecular weight, theoretical pI, amino acid composition, atomic composition, extinction coefficient, estimated half-life, instability index, aliphatic index or GRAVY.

2.13.5 CFSSP

In the Chou Fasman method, the propensity value is important.

- R- groups attached to the protein chain are responsible for the propensity value.
- The main terms used in the Chou Fasman method was
    - Alpha helix or Beta sheet makers
    - Alpha helix or Beta sheet breakers
    - Propensity Value

Many online and offline server tools are available to envisaging the subordinate structure by the Chou Fasman method.

2.13.6 MOTIF search tool

Defined as a nucleotide or amino acid sequence pattern that is widespread and is associated with a biological function.

- A sequence motif = A structural Motif.
- A sequence motif residing in the coding region may encode a structural motif.
- Non-coding nucleotide motifs may have a regulatory role. May have recognition sites for DNA binding proteins.

2.13.7 CATH

The CATH catalogue runs a classified organization of protein.

Provinces are attained after protein edifices placed in Protein Data Bank.

Mutually area credentials and consequent classification use guide as well as automatic dealings**.**

Database construction

- The information in CATH is attained since PDB records placed in the Protein Data Bank.
- The structures can be determined only through a determination of 4Å or improved are encompassed.

Structural Classification

Class: Highest level-placed the selected protein into 1 of 4 categories of secondary structure.

Architecture: Description of the cross arrangement of secondary structure, independent of topology.

Topology: Indication of complete outline and connectivity of protein's secondary structures.

Homologous superfamily: Homologous Proteins of known structure (share a common ancestor) to a selected protein.

2.13.8 SWISS MODEL

Homology modelling: Homology modelling, similarly recognized as qualified modelling of protein, is the procedure that permits to hypothesis an anonymous atomic-resolution prototypical of the "target" protein from:

1. Its amino acid classifications

SWISS-MODEL: It is an entirely automatic protein homology modelling server.

Reachable through the ExPASy Molecular Biology web server (http://www.expasy.org/)

Also, accessible as of Deep view (Swiss Pdb-Viewer) program

Deep view provides manual control during homology modelling and tools for analyzing generated models,

Swiss-model provides two ways of doing homology modelling

First approach mode fully automatic procedure (http://www.expasy.org/swissmod/SWISS-MODEL.html)

Optimize Mode manually optimized sequence alignment via DeepView program (Swiss Pdb-Viewer)

2.14 My work

From the literature survey, we identified that arginase has therapeutic potential for cancer treatment, particularly arginine auxotrophic cancers. The already studies arginases have their limitations, such as low stability, less efficacy and short half-life. Therefore, we tried to identify an alternative source of arginase, which was not yet explored, neither in silico nor experimentally.

During that search, we find out that arginase has been isolated, produced and purified from five different *Bacillus* sp., those species are:

- *Bacillus licheniformis*
- *Bacillus subtilis*

- *Bacillus brevis*
- *Bacillus caldovelox*
- *Bacillus anthracis*

As we study further, we learnt that the arginase from *Bacillus megaterium* is not explored yet. To date, the enzyme isolated from various *Bacillus* sp. are very promising for producing different enzymes. *B. megaterium* is also a GRAS (Generally Referred As Safe) organism. Therefore, we selected *B. megaterium* arginase for our study.

Bacillus megaterium is a rod-like Gram-positive aerobic bacteria. These bacteria are mainly found in soil and considered as saprophytic organisms. It is a tremendously huge bacteria; it is approximately 100 epochs as great as *E. coli*. Owing to its massive magnitude, which is near about 60 micrometres diced. *B. megaterium* takes remained used to study its assembly, protein localization and membranes of the bacteria. While studying the *B. megaterium, a new study* known as the lysogeny classification of *B. megaterium* can be shown in (Table 10).

Table 10 Classification of *B. megaterium*.

Domain	Bacteria
Phylum	Firmicutes
Class	Bacilli
Order	Bacillales
Family	Bacillaceae
Genus	*Bacillus*
Species	*megaterium*

OBJECTIVE

OBJECTIVE

1. Identification of a suitable arginase producing bacterial source
2. To analyze the primary structure of *B. megaterium* arginase for physiochemical characteristics
3. To analyze the protein structure and its correlation with its function using bioinformatics tools

CHAPTER – 3

MATERIALS AND METHODS

3 MATERIALS AND METHODS

3.1 Literature survey

The literature review survey is carried out by reading some erudite foundations such as book chapters, journal articles, and books on bioinformatics, arginase enzyme, and *Bacillus megaterium*.

3.2 Sequence retrieval

Sequence retrieval is the procedure to retrieve or gain specific direction of a gene in DNA, RNA and protein[64]. The complete *B. megaterium* arginase sequence was recognized and recovered since the operational **UNIPROT/ Swiss Prot** catalogue server, the knowledge base of universal protein resource[65]. The protein sequences of arginase of B. megaterium and other *Bacillus* sp. (*Bacillus licheniformis*, *B. subtilis*, *B. brevis*, *B. caldovelox*, *B. anthracis*) were copied in the FASTA format for further study.

3.3 Sequence analysis

The protein (arginase) sequence of B. megaterium was used as a query sequence and other species sequences as a subject sequence for carrying out BLASTp (protein-protein BLAST). The BLASTp has given information about amino acid length, similarity/positivity, identity, query coverage and E- values of the sequence [66].

3.4 Multiple sequence alignment

MSA remained done using **ClustalW**. All the arginase sequences of various species in FASTA format were aligned and got the MSA results and the evolutionary relationship in the phylogenetic tree format[67].

3.5 Physico – Chemical characterization

A server is used to calculate the Physico – Chemical characterization of all the retrieved sequences that software is **Expasy and ProtParam** [68]. The server provides various information about the protein based on their primary structure. The report includes protein accession number, theoretical pi, extinction coefficient (Ɛ), instability index, aliphatic index, and grand average of hydropathicity (GRAVY).

3.6 Secondary structure prediction

Secondary structure prediction consists of the secondary elements: helix, turn and sheet, which is predicted using the P**SIPRED** and **CFSSP** [69]. It defines the functions of proteins.

3.7 Functional analysis

Function analysis of arginase was done using the **Motif search tool** [70]. The SOSUI server was used to identify whether BmA is membrane-bound or a soluble protein by exploiting to envisage transmembrane helices amino acid arrangements through high meticulousness and accuracy [71].

3.8 Structural classification

Structural classification of arginase was done using the CATH server. The CATH catalogue assemblages the protein edifices based on the resemblance or a communal evolutionary derivation using manual curation and particular processes. This cataloguing technique comprises the dispersal of proteins into 4 clusters that are (C) class, (A) architecture, (T) topology and (H) homologous family[72]. The class refers to the contents of the secondary structure of the proteins such as alpha, beta, mixed alpha/beta mixed etc. In contrast, architecture describes the overall arrangements of the secondary structures, regardless of their connectivity, for example, an alpha/beta-sandwich. Topology also called the 'fold level', considers the connectivity of the secondary structures within the chain and the last hierarchical classification, i.e., homologous superfamily proteins to a group of domains that share a common ancestor.

3.9 Homology modelling

The comparative protein model of *B. megaterium* Arginase prepared using **SWISS-Model** (2) using the selected template. If two proteins have a high sequence similarity, then they have a similar 3D structure. The prophesied protein model was appraised concluded by both QMEAN and UCLA—DOE LAB SAVES [73] (The Structure Analysis and Verification Server), which runs six programs for inspection and confirmative protein assemblies throughout and subsequently model modification.

3.10 I – Tasser

I-TASSER (Iterative Threading Assembly Refinement) is a classified tactic to protein building prophecy and structure-based function explanation.

CHAPTER 4

RESULTS AND DISCUSSIONS

4 RESULTS AND DISCUSSIONS

4.1 Sequence retrieval

For the current studies, five different *Bacillus* species (*B. licheniformis, B. subtilis, B. brevis, B. caldovelox and B. anthracis*) were identified and their arginase sequences were retrieved along with *Bacillus megaterium* arginase (BmA) in FASTA format (Figure 4) from UniProt server (1). The five species have been selected based on their comparative arginase length with respect to *BmA* obtained in UniProt through an equivalent number of amino acids (length) to illustrate homogeneity in the study (Table 11).

>tr|D5DKK5|D5DKK5_BACMD Arginase OS:*Bacillus megaterium* (strain DSM 319) OX:592022GN:argIPE:3SV:1MKKDISIIGVPMDYGQTRRGVDMGPSAIRYAGMNTRLE ALGYTVHDEGDIKVEIKERADVDKNTNLKNLAAVASGNEQLAARVEEVREQDRFPL ILGGDHSIAIGTLAGVAKGSENLGVIWYDAHGDLNTAETSPSGNIHGMPLAVSLGIGH PVLLNIGGYTPKIKPENLVIIGARSLDDGEKELIKEKGIKVYTMHEIRLGMTQVMKETI DYLSGTDGVHLSLDLDGLDPDDAPGVGTPVKGGISYRESHLEMLAEADIVTSAEFVE VNPILDQHNKTAEVAVALMSSLFGDKLL

Figure 4 FASTA sequence of *Bacillus megaterium* arginase from UniProtKb tool

Table 11 Details of selected sequences of *Bacillus sp.*

S. No.	Protein Accession number	Name of The Bacteria	Length
1	A0A109FWS0	*Bacillus megaterium*	299
2	Q65DS7	*Bacillus licheniformis*	297
3	P39138	*Bacillus subtilis*	296
4	Q65DS7	*Bacillus brevis*	299
5	P53608	*Bacillus caldovelox*	299
6	D8GXL7	*Bacillus anthracis*	299

4.2 Sequence analysis

While carrying out the NCBI BLASTp (protein-protein BLAST), the amino acid sequence of *BmA* was used as query and other species sequences as the subject. The BLASTp gave the data of amino acid length, similarity/positivity, identity, query coverage and E- value of the arginase sequences of *B. megaterium* and other *Bacillus species*. BmA shows higher homology (>68%) with the selected arginase sequences (Table 12).

Table 12 Pairwise sequence alignment of BmA using BLASTp

Bacteria	Amino Acid Length	Similarity / Positivity	Identity	Query Coverage	E - Value
Bacillus licheniformis	297	243/300(81%)	213/300 (71%)	100%	5e - 154
Bacillus subtilis	296	248/299(82%)	216/299 (72%)	100%	1e - 159
Bacillus brevis	299	236/300(78%)	204/300 (68%)	100%	5e - 146
Bacillus caldovelox	299	253/298(84%)	222/298 (74%)	99%	9e - 165
Bacillus anthracis	299	253/298(84%)	217/298 (73%)	99%	1e - 160

Multiple sequence alignment is done with the software named **ClustalW.** It also generates **a** phylogenetic tree of all the species to check the evolutionary relationship between the *B. megaterium* and other *Bacillus species* (Figure 4). The result we get is *Bacillus megaterium* is ancestrally very similar to another *bacillus species*.

Figure 4 Phylogenetic classification of *B. megaterium* with other *Bacillus species*

4.3 Physico – Chemical characterization

A server ExpASY ProtoParam was used to calculate the Physico-chemical characterization of all the retrieved sequences. The server provides information about protein accession number, theoretical pi, extinction coefficient (Ɛ), instability index, aliphatic index, GRAVY (Table 13).

a) Theoretical pI: This is the pH of protein at which protein posses a net charge of zero. The proteins with more basic amino acids will have a higher pI value; acidic protein will have a lower pI value.
b) Extinction coefficient: It measures the strongly a chemical species or substrate absorbs light at a particular wavelength.
c) Instability index: It quantitates the protein's stability. It determines whether the protein is altering in the test tube or not. If the index is less than 40, it is stable in the test tube; if it is greater than 40, then probably not stable.
d) Aliphatic index: This restriction signifies the overall measurements of a protein engaged by aliphatic chains. Studies have revealed that a high aliphatic index may deliberate a grade of thermostability to a prearranged protein. The aliphatic index of a given protein characterizes the capacity of a protein encompassed by aliphatic amino acids.
e) GRAVY: This stricture is illustrative of the general hydrophilic or aquaphobic attractiveness of a specified protein. An undesirable GRAVY value implies that

the protein is aquaholic though an optimistic charge designates that the protein is aquaphobic. All the classifications testified here were originate from having undesirable GRAVY values.

Table 13 Physico-chemical characteristics of arginases of selected *Bacillus* sp. using ExpASY PROTOPARAM tool.

Bacteria	Protein Accession number	Theoretical pI	Extinction coefficients ($M^{-1}cm^{-1}$)	Instability index	Aliphatic index	Grand average of hydropathicity (GRAVY)
Bacillus megaterium	A0A109FWS0	4.85	17420	23.02	99.46	0.152
Bacillus licheniformis	Q65DS7	5.07	20400	35.33	104.01	-0.161
Bacillus subtilis	P39138	5.10	20400	35.62	102.74	-0.168
Bacillus brevis	Q65DS7	5.32	7450	27.92	99.77	0.024
Bacillus caldovelox	P53608	5.58	17420	31.86	98.83	-0.116
Bacillus anthracis	D8GXL7	4.84	18910	33.66	97.14	-0.081

4.4 Secondary structure prediction

In secondary structure prediction, the secondary elements are helix, turn and sheet prophesied via the **PSIPRED and CFSSP**. It defines the functions of the proteins. The server generates secondary structural elements (helices, turns and sheets) (Table 14); it's been pragmatic that *B. megaterium* proteins remained confidential in the two main assemblies' helix and sheet. Numerous structural verdicts of *B. megaterium* have maintained this illustration from diverse microbes. Crystal assembly taxations have exposed a variability of organizational structures for several *B. megaterium* from CFSSP server (Figure 5).

Target Sequence:

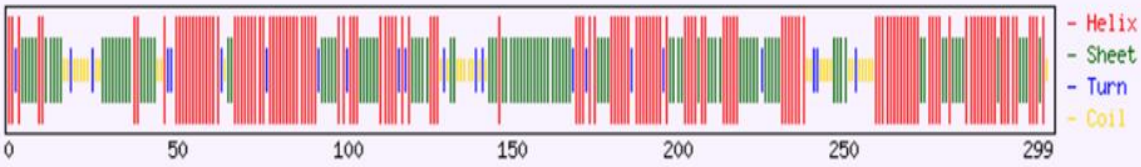
```
        10         20         30         40         50         60         70
MKKDISIIGV PMDYGQTRRG VDMGPSAIRY AGMNTRLEAL GYTVHDEGDI KVEIKERADV DKNTNLKNLA

        80         90        100        110        120        130        140
AVASGNEQLA ARVEEVREQD RFPLILGGDH SIAIGTLAGV AKGSENLGVI WYDAHGDLNT AETSPSGNIH

       150        160        170        180        190        200        210
GMPLAVSLGI GHPVLLNIGG YTPKIKPENL VIIGARSLDD GEKELIKEKG IKVYTMHEID RLGMTQVMKE

       220        230        240        250        260        270        280
TIDYLSGTDG VHLSLDLDGL DPDDAPGVGT PVKGGISYRE SHLAMEMLAE ADIVTSAEFV EVNPILDQHN

       290
KTAEVAVALM SSLFGDKLL
```



Secondary Structure:

Figure 5 Built secondary structure of *B. megaterium* from Chou and Fasman server showing the percentage of different secondary arrangements

Table 14 Secondary structures of selected sequences.

S. No.	Protein Acession Numer	Helix (%)	Sheet (%)	Turn (%)
1	A0A109FWS0	68.6	55.9	12.0
2	Q65DS7	72.1	64.0	9.1
3	P39138	75.0	60.1	10.8
4	Q65DS7	72.8	58.1	8.1
5	P53608	77.6	30.4	9.7
6	D8GXL7	77.1	60.6	9.1

4.5 Functional analysis

For function analysis, we use the **Motif search tool** casing and solvable proteins. It is illustrious via the SOSUI server that was able to envisage transmembrane helices as amino

acid sequences with tall exactness and correctness. There is one motif present in the complete BmA sequence shown in (Figure 6).

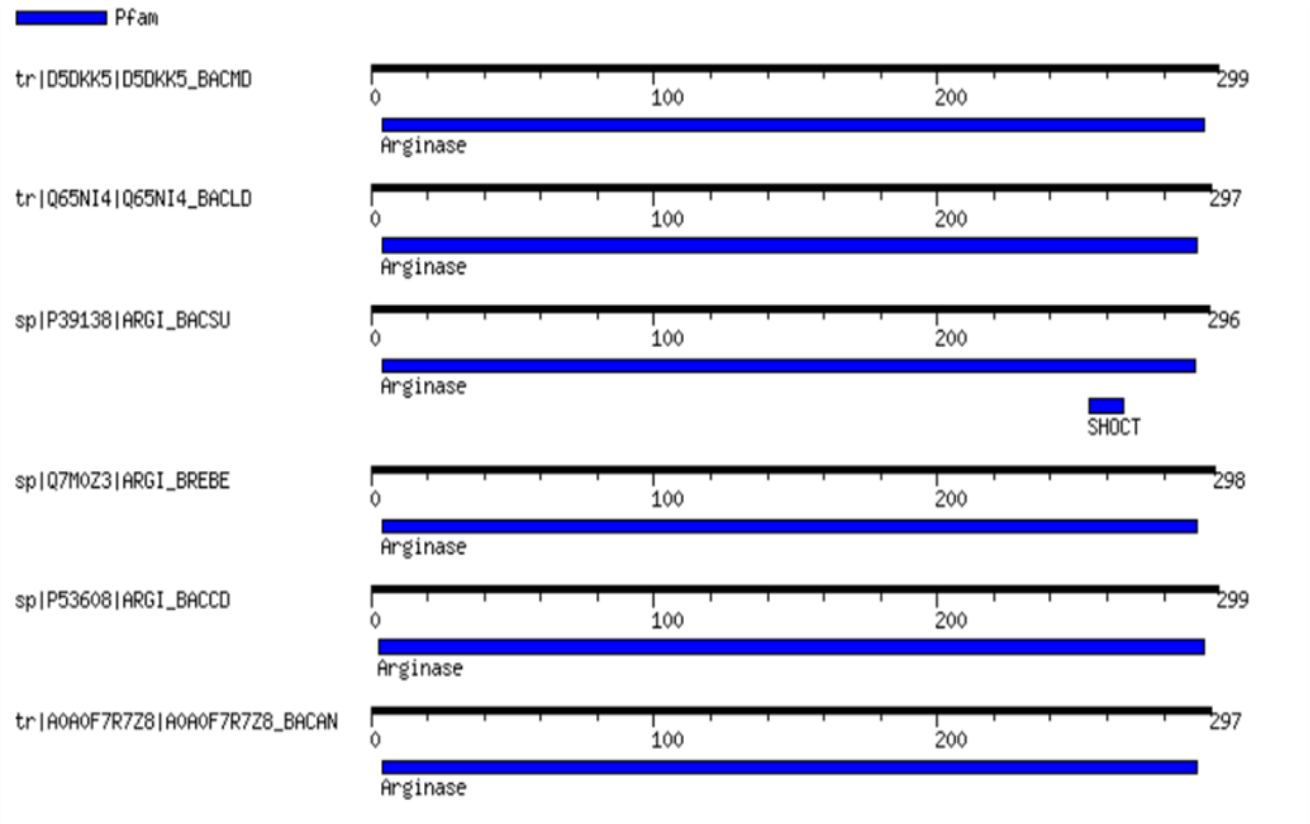


Figure 6 Identified motif graphical details.

SOSUI

Bacillus megaterium*:* The amino acid sequence is of a soluble protein.

Length: 356 AA

Average of hydrophobicity: -0.137302

4.6 Structural classification

For structural classification, the CATH server remained cast off. CATH catalogue assemblages the protein edifices on the source of resemblance or a communal evolutionary derivation through manual curation lengthways with convinced procedures. This cataloguing technique embraces the circulation of proteins hooked on four assemblages that are (C) class, (A) architecture, (T) topology and (H) homologous family. It gives the 3D structure of arginase (Figure 7)

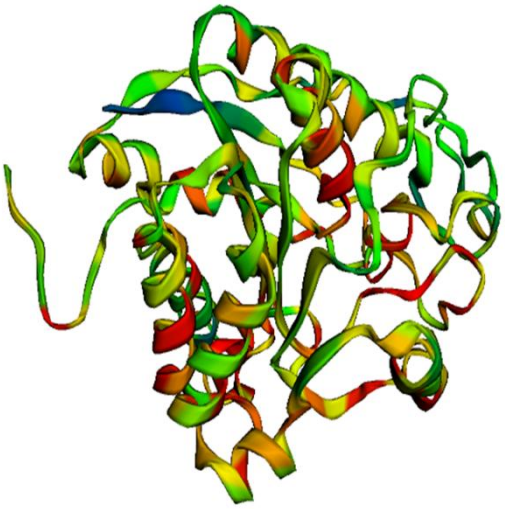


Figure 7 Structure of *B. megaterium* arginase

CATH CLASSIFICATION

C: Alpha Beta

A: 3 – Layer (aba) sandwich

T: Arginase; Chain A

H: Ureohydrolase domain

4.7 Homology modelling

The comparative protein model of *Bacillus megaterium* in Arginase was built through the SWISS-Model using the selected template. It gives different types of structures ribbon-shaped, cartoon shaped that is shown in (Figure 8). *Bacillus cereus* has 71.96% similarity with *B. megaterium* arginase sequence. It has homo – trimer state and they have 0.88 GMQE values and 0.91 QMEAN values, respectively. The data is given in (Table 15).

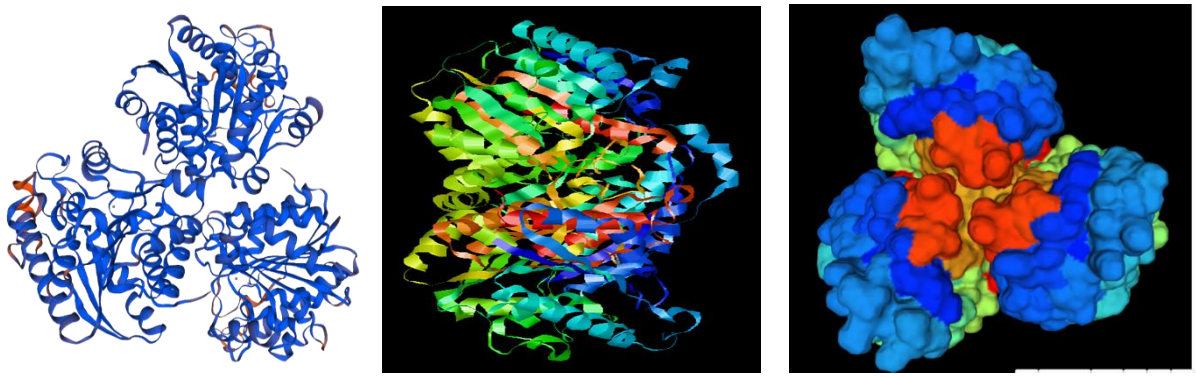


Figure 8 Predicted protein structures of arginase based on sequence homology modelling using SWISS MODEL

Table 15 Homology modelling data of primary sequence of *B. megaterium* arginase.

S. No	Template	Coverage	GMQE	QMEAN	Identity	Oligo state
1	*Bacillus cereus* (crystal structure)	0.99	0.88	0.91	71.96	Homo - trimer

Quality estimation of predicted protein

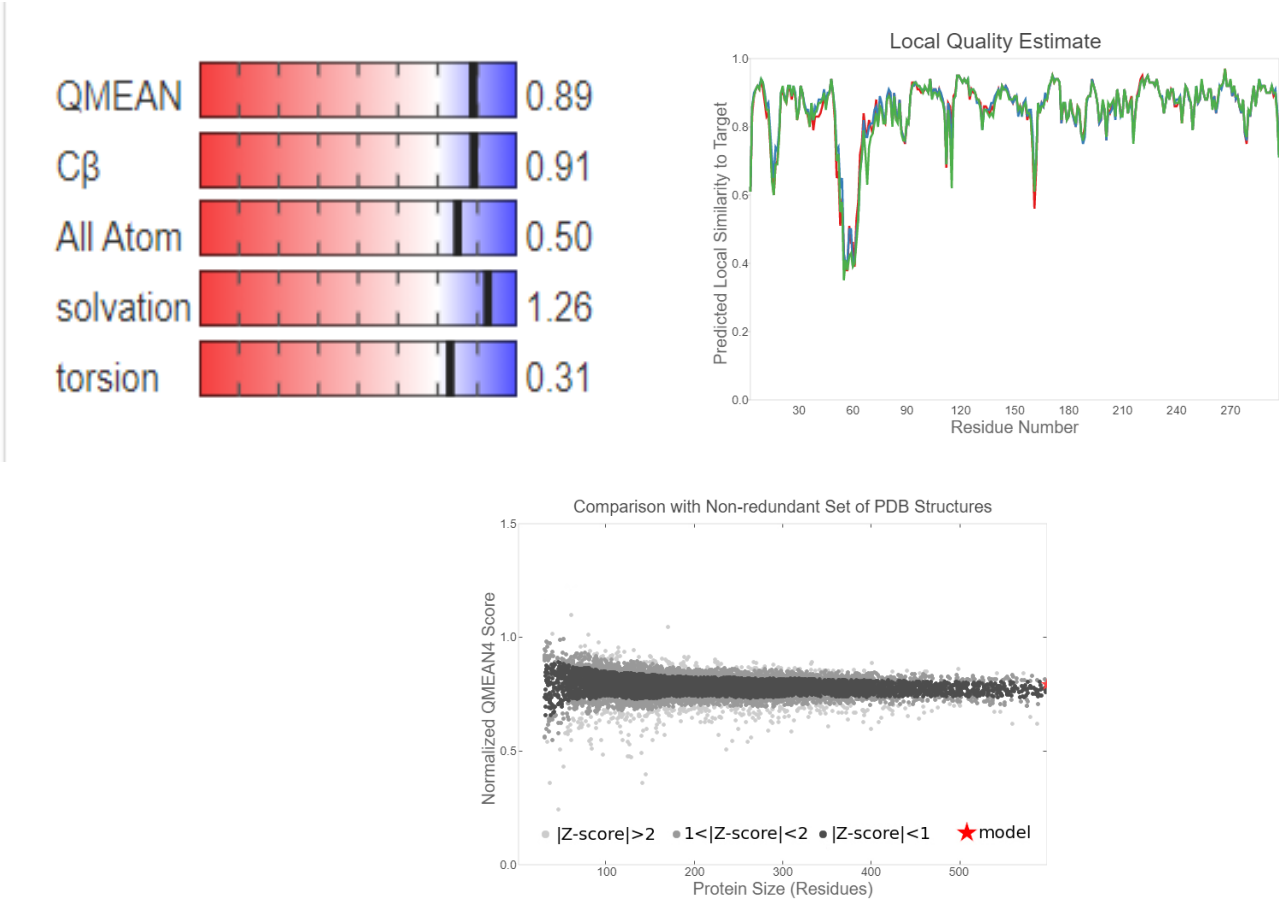


Figure 9 Quality analysis of predicted protein from different sources: (1) QMEAN score of predicted protein,(B) Local quality estimate of QMEAN server, (C) Comparison of built model with non – reductant set of structures from QMEAN server.

Ramachandran plot: Used to check chemical properties of protein structure. This plot is generated by RAMPAGE (Figure 10)

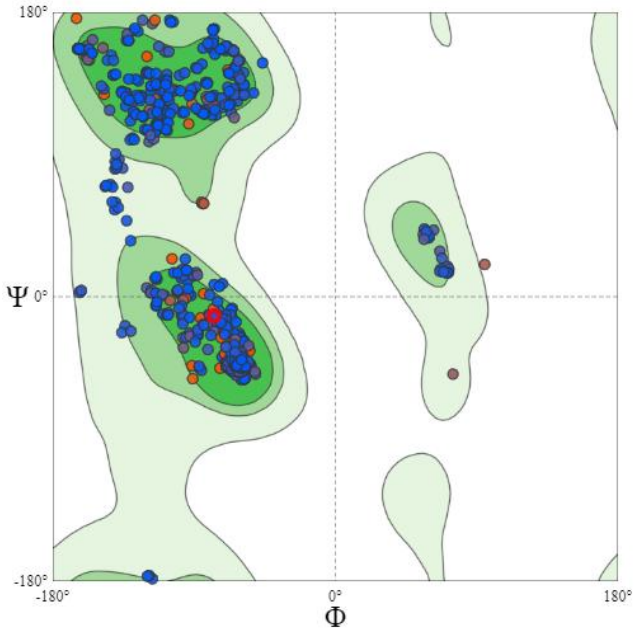


Figure 10 Ramachandran plot

4.8 I-TASSER

I-TASSER (Iterative Threading Assembly Refinement) is a classified tactic to protein construction prophecy and structure-based function explanation. C – score of the structure is 1.45, which is considered a positive score and the model is correct according to I-Tasser tool (Figure11).

Figure 11 Predicted 3D model of arginase using I-Tasser tool. C – score = 1.45

4.8.1 Predicted secondary structure

There are lots of helices in this structure; strands and coils are also present, which is shown in (Figure 12). When the value of the amino acid score is close to 9 the structure is reliable. When the score is close to 0 – 3 then the structure is unreliable; in Figure 12, only a few residues are unreliable mainly, all residues are reliable.

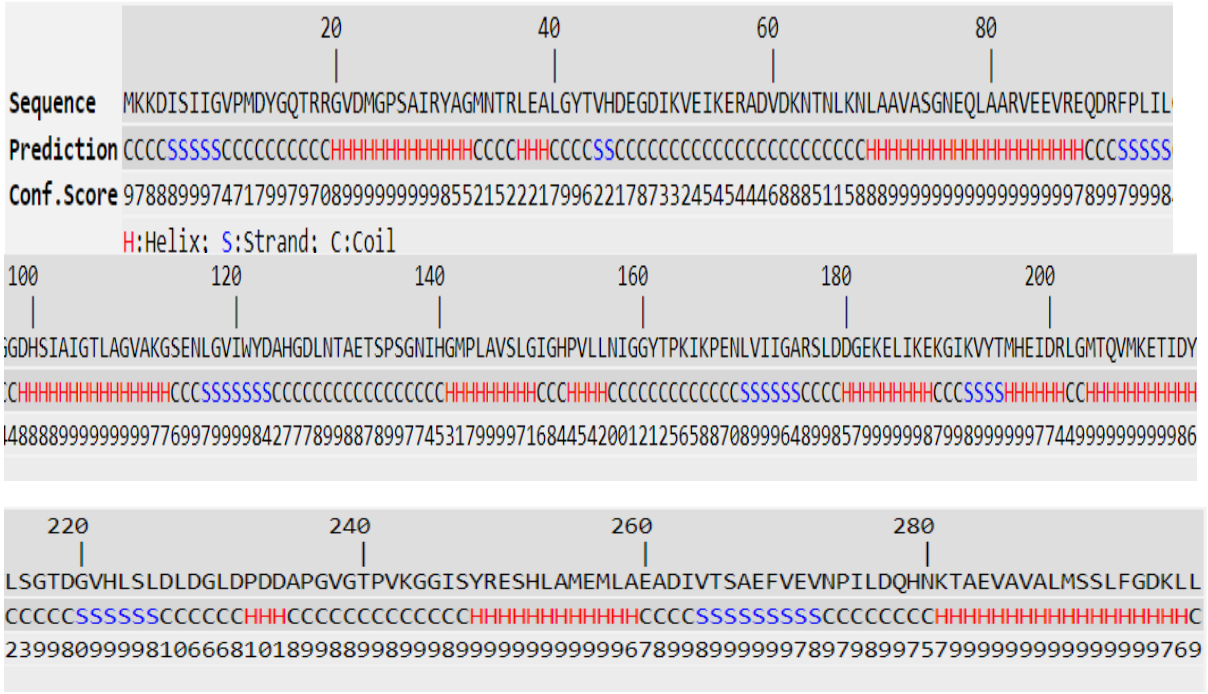


Figure 12 Sequence-based prediction of secondary structure

Predicted normalized B – factor

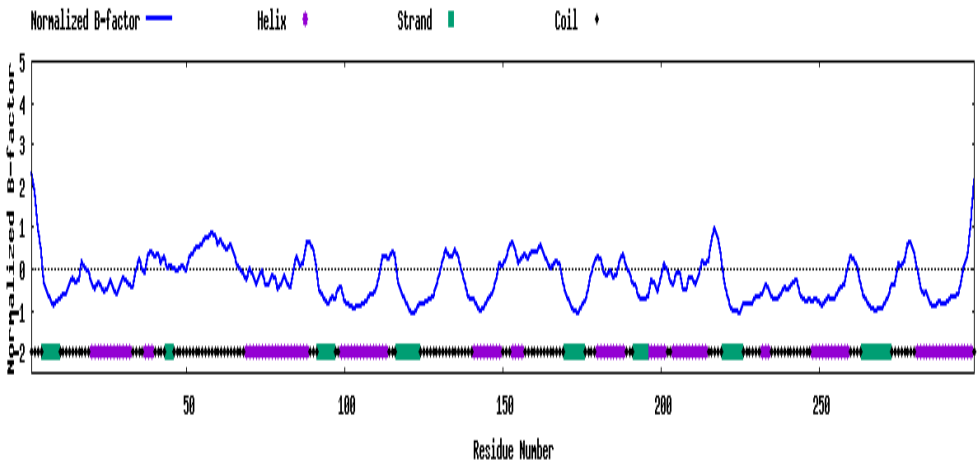


Figure 13 Predicted normalized B - factor

4.9 QMEAN and UCLA—DOE LAB SAVES (For evaluation of predicted protein)

4.9.1 ERRAT SERVER OUTPUT

Conferring to SAVES ERRAT, a decent high tenacity (>3 A˚) building commonly produce standards around 95% or advanced. As the general eminence feature determined from SAVES ERRAT was 97.905, the tenacity of the current erected 3D model is upright than 3 A˚ which is required that is shown in (Figure 14).

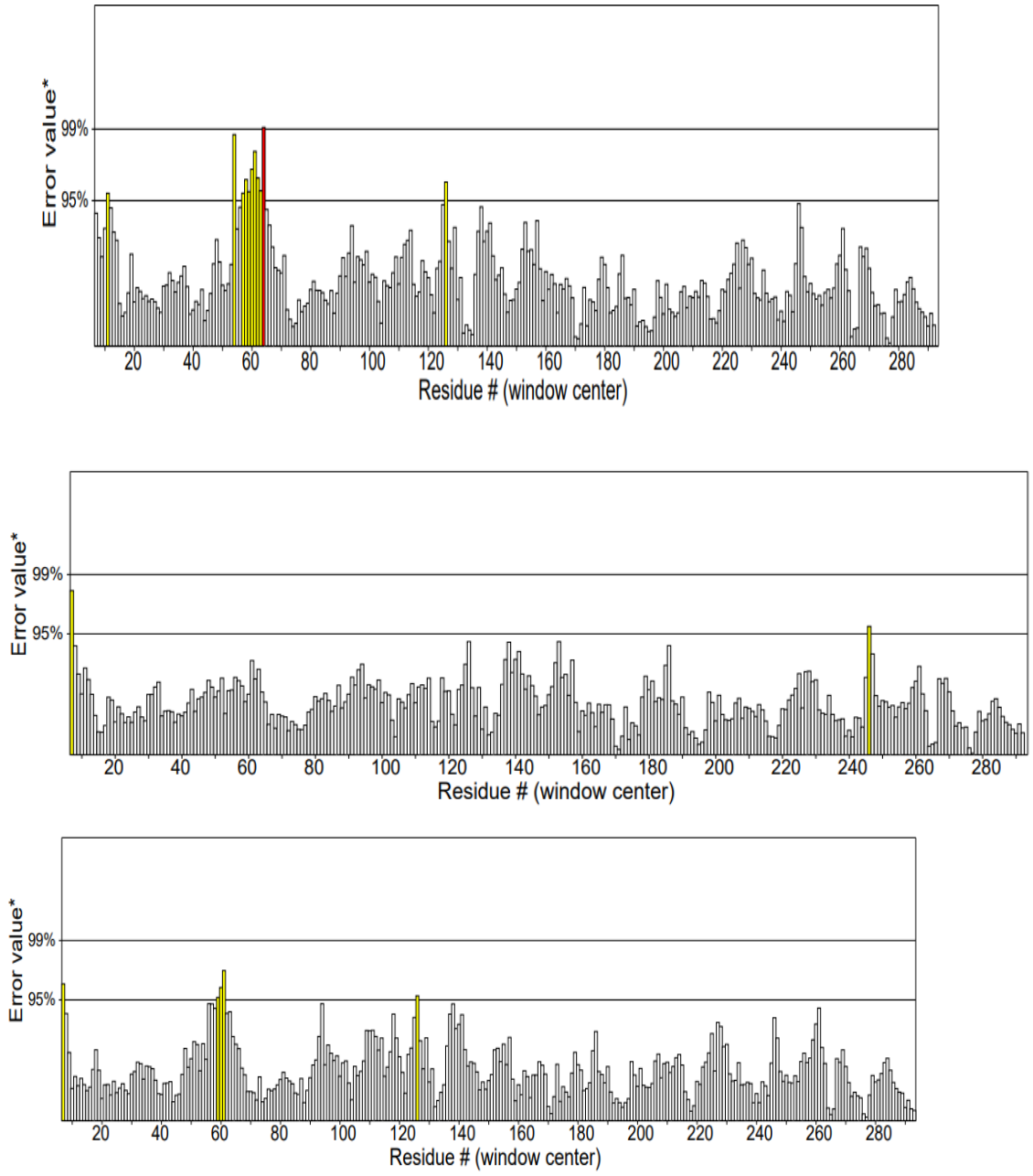


Figure 14 Overall quality check by ERRAT (SAVES server)

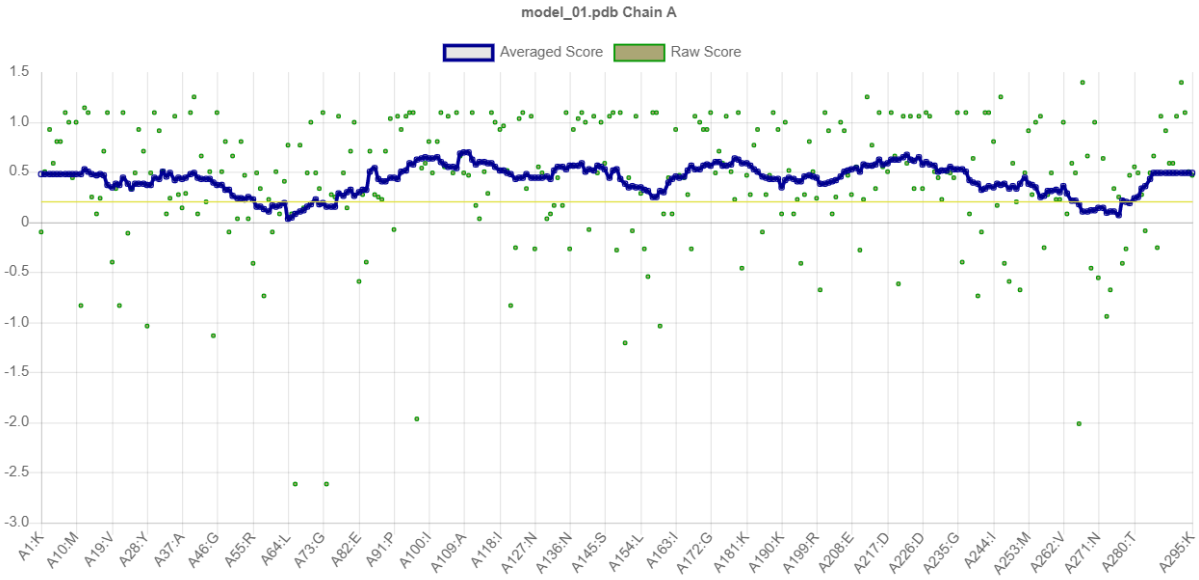
Overall quality factor = 97.905

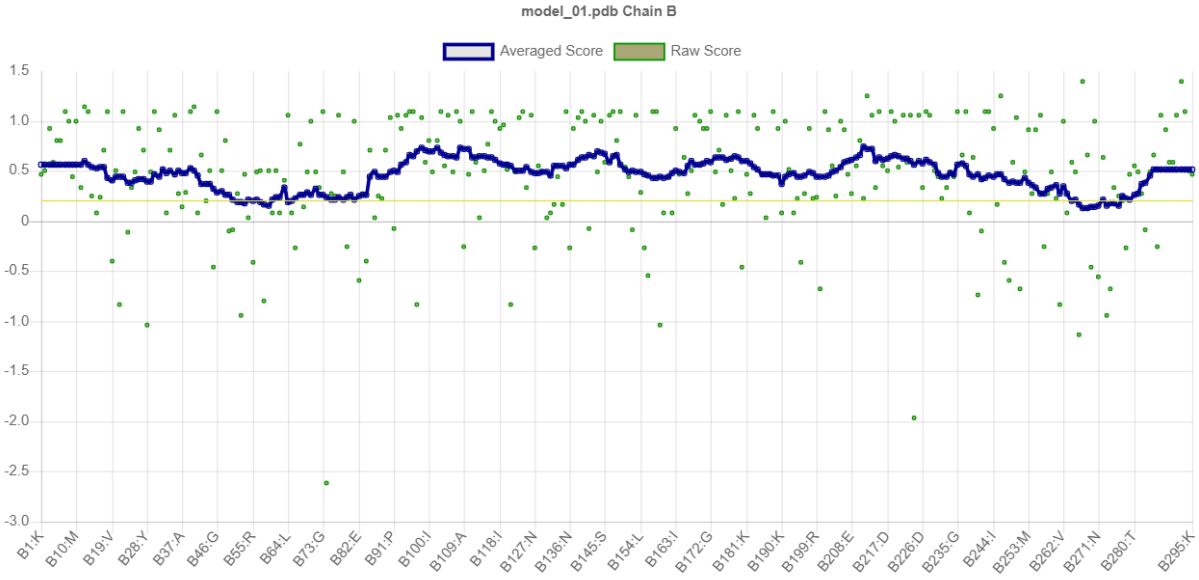
4.9.2 Verify 3D output

Giving to SAVES Verify3D, the tiniest 80% of the amino acids would score C 0.2 in the 3D/1D outline to get a pass for existence a normal good quality edifice, which is here 89.83%

At least 80% of the amino acids have scored >= 0.2 in the 3D/1D profile**.**

The model is pass because it is an 89.93% score, as shown in (Figure 15).





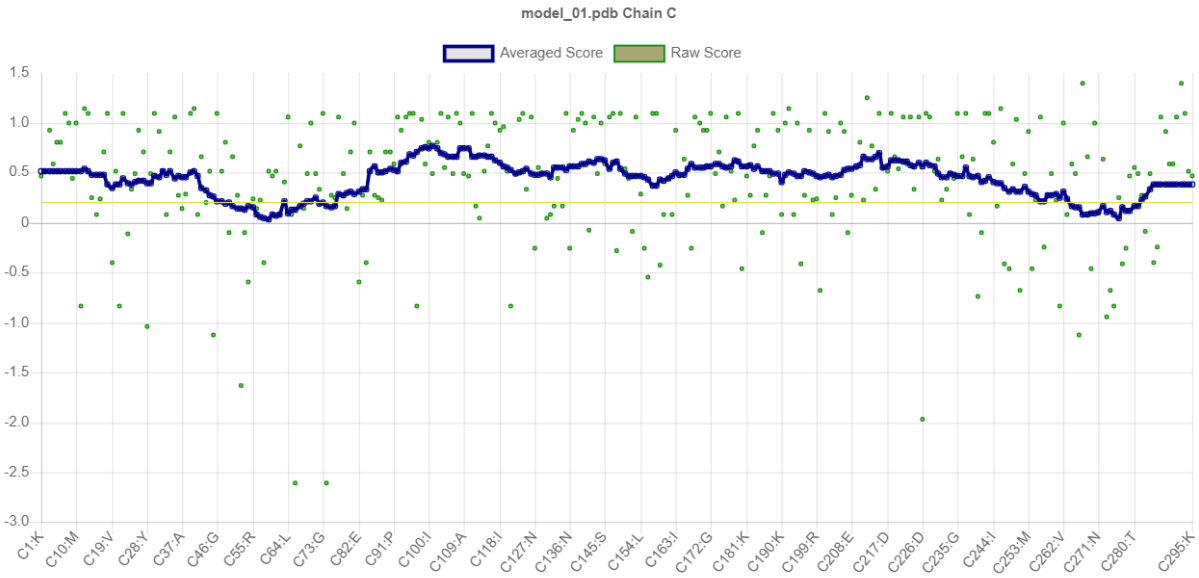


Figure 15 Average score and raw score of predicted *B. megaterium* arginase structure by VERIFY3D 89.93% of the residues have averaged 3D – 1D score > = 0.2

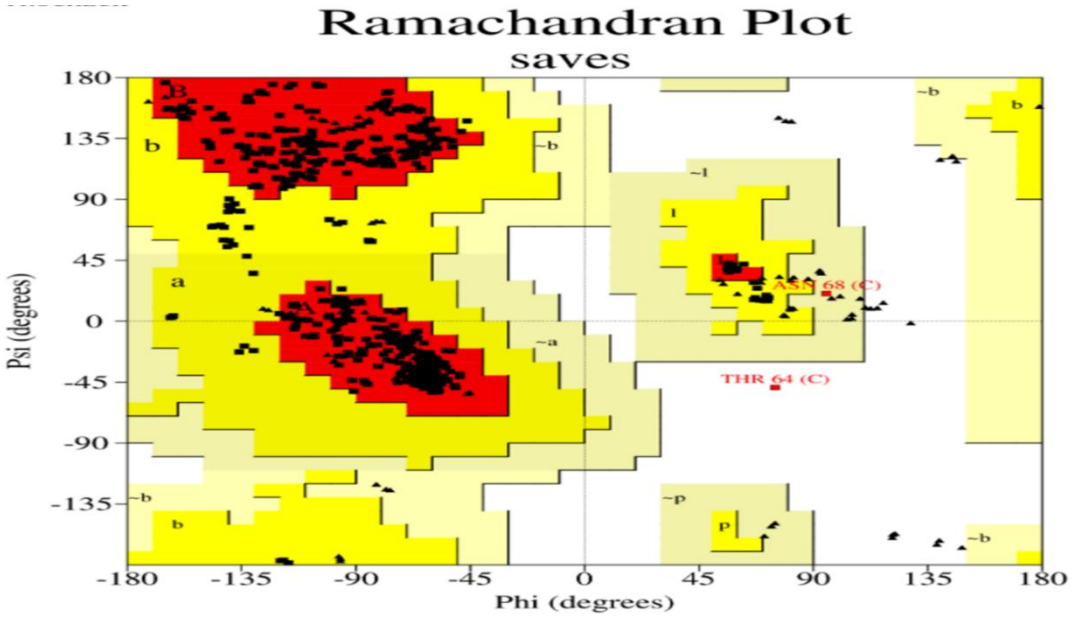


Figure 16 Ramachandran plot of matrix protein of BmA generated using Procheck software

4.9.3 PROCHECK

Ramachandran plot: 93.3% core, 6.5% allow, 0.1% general, 0.1% disallow,

Overall = - 0.06

This plot shows the psi – psi torsion angles for all deposits in the collaborative.

Discussion: From the literature survey, it was determined that the arginase enzyme has some antineoplastic possessions, which is convenient to treat arginase auxotrophic cancers such as prostate cancer, melanoma, lung cancer, T- lymphoblastic leukaemia, and osteosarcoma. From the UniProt kb database, a total of 6 arginase sequences of *B. species* were retrieved based on the number of amino acid residues length. The sequence was analyzed using BLASTp. The result suggested that *B. megaterium* arginase has great similarity with other *Bacillus* species arginases. Multiple sequence alignment of selected arginase sequences was done using ClustalW. This data is then used to get the phylogenetic tree which showed very similar ancestral characters of *Bacillus megaterium* arginase with that of other *Bacillus* species.

Primary structures of BmA and other selected arginases were used to characterize physicochemical properties of the protein using Expasy PROTPARAM Tools, which detailed the information about amino acid composition, identification, quality, purity and stability of all the *Bacillus sp.* The data suggested all the arginase proteins are stable, having pI near to acidic range.

The secondary structure of the BmA was predicted using the PSIPRED server that suggested the protein is an αβ type. Then another step is structural classification. It gives the protein structure of the *Bacillus megaterium*, which is required for the purpose of future studies. Since the BmA protein crystal structure is not available, the 3D structure was homology modelled using SWISS-MODEL. The overall quality factor determined from SAVES ERRAT was 97.905% which is desirable hence the model is passed. According to SAVES Verify3D, which is here 89.89%.and considered as good score hence our model is passed.

According to all the results, it is concluded that the *B. megaterium* is a good source for isolating the arginase enzyme.

REFERENCES

[1] L. C. Burrage *et al.*, "Human recombinant arginase enzyme reduces plasma arginine in mouse models of arginase deficiency," *Human Molecular Genetics*, vol. 24, no. 22, pp. 6417–6427, Nov. 2015, doi: 10.1093/hmg/ddv352.

[2] "L-Arginase: a Medically Important Enzyme-Indian Journals." https://www.indianjournals.com/ijor.aspx?target=ijor:rjpt&volume=6&issue=12&article=022 (accessed Jun. 08, 2021).

[3] B. C. Tennant and S. A. Center, "Hepatic Function," in *Clinical Biochemistry of Domestic Animals*, Elsevier Inc., 2008, pp. 379–412. doi: 10.1016/B978-0-12-370491-7.00013-1.

[4] R. W. Caldwell, P. C. Rodriguez, H. A. Toque, S. Priya Narayanan, and R. B. Caldwell, "Arginase: A multifaceted enzyme important in health and disease," *Physiological Reviews*, vol. 98, no. 2. American Physiological Society, pp. 641–665, Apr. 01, 2018. doi: 10.1152/physrev.00037.2016.

[5] R. W. Caldwell, P. C. Rodriguez, H. A. Toque, S. Priya Narayanan, and R. B. Caldwell, "Arginase: A multifaceted enzyme important in health and disease," *Physiological Reviews*, vol. 98, no. 2. American Physiological Society, pp. 641–665, Apr. 01, 2018. doi: 10.1152/physrev.00037.2016.

[6] "Arginase Deficiency: Background, Pathophysiology, Epidemiology." https://emedicine.medscape.com/article/941838-overview (accessed Jun. 08, 2021).

[7] J. Moretto, M. Pudlo, and C. Demougeot, "Human-based evidence for the therapeutic potential of arginase inhibitors in cardiovascular diseases," *Drug Discovery Today*, vol. 26, no. 1. Elsevier Ltd, pp. 138–147, Jan. 01, 2021. doi: 10.1016/j.drudis.2020.11.005.

[8] D. J. McGee, J. Zabaleta, R. J. Viator, T. L. Testerman, A. C. Ochoa, and G. L. Mendz, "Purification and characterization of Helicobacter pylori arginase, RocF: Unique features among the arginase superfamily," *European Journal of Biochemistry*, vol. 271, no. 10, pp. 1952–1962, May 2004, doi: 10.1111/j.1432-1033.2004.04105.x.

[9] E. J. Laishley and R. W. Bernlohr, "Regulation of arginine and proline catabolism in Bacillus licheniformis.," *Journal of bacteriology*, vol. 96, no. 2, pp. 322–329, 1968, doi: 10.1128/jb.96.2.322-329.1968.

[10] N. Nakamura, M. Fujita, and K. Kimura, "Purification and properties of l-arginase from bacillus subtilis," *Agricultural and Biological Chemistry*, vol. 37, no. 12, pp. 2827–2833, Dec. 1973, doi: 10.1080/00021369.1973.10861080.

[11] Y. Dessaux, A. Petit, J. Tempe, M. Demarez, C. Legrain, and J. M. Wiame, "Arginine catabolism in Agrobacterium strains: Role of the Ti plasmid," *Journal of Bacteriology*, vol. 166, no. 1, pp. 44–50, 1986, doi: 10.1128/jb.166.1.44-50.1986.

[12] M. C. Bewley, P. D. Jeffrey, M. L. Patchett, Z. F. Kanyo, and E. N. Baker, "Crystal structures of Bacillus caldovelox arginase in complex with substrate and inhibitors reveal new insights into activation, inhibition and catalysis in the arginase superfamily," *Structure*, vol. 7, no. 4, pp. 435–448, Apr. 1999, doi: 10.1016/S0969-2126(99)80056-2.

[13] J. J. Coque, F. J. Pérez-Llarena, F. J. Enguita, J. L. Fuente, J. F. Martin, and P. Liras, "Characterization of the cmcH genes of Nocardia lactamdurans and streptomyces clavuligerus encoding a functional 3'-hydroxymethylcephem O-carbamoyltransferase for cephamycin biosynthesis," *Gene*, vol. 162, no. 1, pp. 21–27, Aug. 1995, doi: 10.1016/0378-1119(95)00308-S.

[14] R. B. Hamed, J. R. Gomez-Castellanos, L. Henry, C. Ducho, M. A. McDonough, and C. J. Schofield, "The enzymes of β-lactam biosynthesis," *Natural Product Reports*, vol. 30, no. 1. Royal Society of Chemistry, pp. 21–107, Dec. 10, 2013. doi: 10.1039/c2np20065a.

[15] M. Kanda, K. Ohgishi, T. Hanawa, and Y. Saito, "Arginase of Bacillus brevis Nagano: Purification, properties, and implication in gramicidin S biosynthesis," *Archives of Biochemistry and Biophysics*, vol. 344, no. 1, pp. 37–42, Aug. 1997, doi: 10.1006/abbi.1997.0174.

[16] C. MORENO-VIVIÁN, G. SOLER, and F. CASTILLO, "Arginine catabolism in the phototrophic bacterium Rhodobacter capsulatus E1F1: Purification and properties of arginase," *European Journal of Biochemistry*, vol. 204, no. 2, pp. 531–537, 1992, doi: 10.1111/j.1432-1033.1992.tb16664.x.

[17] E. Flores and A. Herrero, "Nitrogen assimilation and nitrogen control in cyanobacteria," in *Biochemical Society Transactions*, Feb. 2005, vol. 33, no. 1, pp. 164–167. doi: 10.1042/BST0330164.

[18] R. J. Viator, R. F. Rest, E. Hildebrandt, and D. J. McGee, "Characterization of Bacillus anthracis arginase: Effects of pH, temperature, and cell viability on metal preference," *BMC Biochemistry*, vol. 9, no. 1, p. 15, 2008, doi: 10.1186/1471-2091-9-15.

[19] S. Hartenbach, M. Daoud-El Baba, W. Weber, and M. Fussenegger, "An engineered L-arginine sensor of Chlamydia pneumoniae enables arginine-adjustable transcription control in mammalian cells and mice," *Nucleic Acids Research*, vol. 35, no. 20, Nov. 2007, doi: 10.1093/nar/gkm652.

[20] N. Arakawa, M. Igarashi, T. Kazuoka, T. Oikawa, and K. Soda, "D-arginase of Arthrobacter sp. KUJ 8602: Characterization and its identity with Zn2+-guanidinobutyrase," *Journal of Biochemistry*, vol. 133, no. 1, pp. 33–42, Jan. 2003, doi: 10.1093/jb/mvg016.

[21] S. Keni and N. S. Punekar, "Contribution of arginase to manganese metabolism of Aspergillus Niger," *BioMetals*, vol. 29, no. 1, pp. 95–106, Feb. 2016, doi: 10.1007/s10534-015-9900-6.

[22] M. C. Molina, C. Vicente, M. M. Pedrosa, and M. E. Legaz, "Isoforms of Arginase in the Lichens Evernia prunastri and Xanthoria parietina: Physiological Roles and Their

Implication in the Controlled Parasitism of the Mycobiont," in *Eukaryotism and Symbiosis*, Springer Berlin Heidelberg, 1997, pp. 477–483. doi: 10.1007/978-3-642-60885-8_41.

[23] "Partial purification and some characteristics of arginase of trichoderma sp | Virtual Health Sciences Library." https://vlibrary.emro.who.int/imemr/partial-purification-and-some-characteristics-of-arginase-of-trichoderma-sp/ (accessed Jun. 08, 2021).

[24] M. J. M. Wagemaker, W. Welboren, C. van der Drift, M. S. M. Jetten, L. J. L. D. van Griensven, and H. J. M. Op Den Camp, "The ornithine cycle enzyme arginase from Agaricus bisporus and its role in urea accumulation in fruit bodies," *Biochimica et Biophysica Acta - Gene Structure and Expression*, vol. 1681, no. 2–3, pp. 107–115, Jan. 2005, doi: 10.1016/j.bbaexp.2004.10.007.

[25] G. Vaca and J. Mora, "Nitrogen regulation of arginase in Neurospora crassa," *Journal of Bacteriology*, vol. 131, no. 3, pp. 719–725, 1977, doi: 10.1128/jb.131.3.719-725.1977.

[26] "Regulation of arginine metabolism in Saccharomyces cerevisiae. Association of arginase and ornithine transcarbamoylase - PubMed." https://pubmed.ncbi.nlm.nih.gov/3528164/ (accessed Jun. 08, 2021).

[27] "(PDF) Purification and Characterization of Two Isolectins with Arginase Activity from the Lichen Xanthoria parietina." https://www.researchgate.net/publication/228549233_Purification_and_Characterization_of_Two_Isolectins_with_Arginase_Activity_from_the_Lichen_Xanthoria_parietina (accessed Jun. 08, 2021).

[28] M. Sacristán, A. M. Millanes, M. E. Legaz, and C. Vicente, "A lichen lectin specifically binds to the α-1,4-polygalactoside moiety of urease located in the cell wall of homologous algae," *Plant Signaling and Behavior*, vol. 1, no. 1, pp. 23–27, 2006, doi: 10.4161/psb.1.1.2276.

[29] E. M. Díaz, M. Sacristán, M. E. Legaz, and C. Vicente, "Isolation and characterization of a cyanobacterium-binding protein and its cell wall receptor in the lichen Peltigera canina," *Plant Signaling and Behavior*, vol. 4, no. 7, pp. 598–603, Jul. 2009, doi: 10.4161/psb.4.7.9164.

[30] M. E. Legaz, B. Fontaniella, A. M. Millanes, and C. Vicente, "Secreted arginases from phylogenetically far-related lichen species act as cross-recognition factors for two different algal cells," *European Journal of Cell Biology*, vol. 83, no. 8, pp. 435–446, 2004, doi: 10.1078/0171-9335-00384.

[31] H. J. Hwang, E. H. Kim, and Y. D. Cho, "Isolation and properties of arginase from a shade plant, ginseng (Panax ginseng C.A. Meyer) roots," *Phytochemistry*, vol. 58, no. 7, pp. 1015–1024, Dec. 2001, doi: 10.1016/S0031-9422(01)00392-2.

[32] S. L. N. Rao, P. R. Adiga, and P. S. Sarma, "The Isolation and Characterization of β-N-oxalyl-l-α,β-Diaminopropionic Acid: A Neurotoxin from the Seeds of Lathyrus sativus," *Biochemistry*, vol. 3, no. 3, pp. 432–436, Mar. 1964, doi: 10.1021/bi00891a022.

[33] C. Kollöffel and H. D. van Dijke, "Mitochondrial Arginase Activity from Cotyledons of Developing and Germinating Seeds of Vicia faba L ," *Plant Physiology*, vol. 55, no. 3, pp. 507–510, Mar. 1975, doi: 10.1104/pp.55.3.507.

[34] K. A. Roubelakis and W. M. Kliewer, " Enzymes of Krebs-Henseleit Cycle in Vitis vinifera L ," *Plant Physiology*, vol. 62, no. 3, pp. 344–347, Sep. 1978, doi: 10.1104/pp.62.3.344.

[35] V. M. Loyola-Vargas, M. E. Román, J. Quiroz, C. Oropeza, M. L. Robert, and K. N. Scorer, "Nitrogen Metabolism in Canavalia ensiformis L. DC. I. Arginase and Urease Ontogeny," *Journal of Plant Physiology*, vol. 132, no. 3, pp. 284–288, Apr. 1988, doi: 10.1016/S0176-1617(88)80106-8.

[36] J. H. Kang and Y. D. Cho, "Purification and properties of arginase from soybean, Glycine max, axes," *Plant Physiology*, vol. 93, no. 3, pp. 1230–1234, 1990, doi: 10.1104/pp.93.3.1230.

[37] C. A. Hale, C. J. Clark, H. H. Petach, and R. M. Daniel, "Arginase from kiwifruit: Properties and seasonal variation," *New Zealand Journal of Crop and Horticultural Science*, vol. 25, no. 3, pp. 295–301, 1997, doi: 10.1080/01140671.1997.9514019.

[38] H. de Ruiter and C. Kollöffel, "Arginine Catabolism in the Cotyledons of Developing and Germinating Pea Seeds," *Plant Physiology*, vol. 73, no. 3, pp. 525–528, Nov. 1983, doi: 10.1104/pp.73.3.525.

[39] S. Chen, Z. Vaghchhipawala, W. Li, H. Asard, and M. B. Dickman, "Tomato phospholipid hydroperoxide glutathione peroxidase inhibits cell death induced by bax and oxidative stresses in yeast and plants," *Plant Physiology*, vol. 135, no. 3, pp. 1630–1641, Jul. 2004, doi: 10.1104/pp.103.038091.

[40] A. M. Millanes, B. Fontaniella, M. E. Legaz, and C. Vicente, "Glycoproteins from sugarcane plants regulate cell polarity of Ustilago scitaminea teliospores," *Journal of Plant Physiology*, vol. 162, no. 3, pp. 253–265, Mar. 2005, doi: 10.1016/j.jplph.2004.05.017.

[41] L. Palmieri *et al.*, "Molecular identification of an Arabidopsis S-adenosylmethionine transporter. analysis of organ distribution, bacterial expression, reconstitution into liposomes, and functional characterization," *Plant Physiology*, vol. 142, no. 3, pp. 855–865, Nov. 2006, doi: 10.1104/pp.106.086975.

[42] H. J. Hwang, E. H. Kim, and Y. D. Cho, "Isolation and properties of arginase from a shade plant, ginseng (Panax ginseng C.A. Meyer) roots," *Phytochemistry*, vol. 58, no. 7, pp. 1015–1024, Dec. 2001, doi: 10.1016/S0031-9422(01)00392-2.

[43] "Nitrogen recycling in Pinus taeda during active phenylpropanoid... | Download Scientific Diagram." https://www.researchgate.net/figure/Nitrogen-recycling-in-Pinus-taeda-during-active-phenylpropanoid-metabolism-Enzymes-are-as_fig2_7617304 (accessed Jun. 08, 2021).

[44] N. Shiono *et al.*, "L-arginine protects human heart cells from low-volume anoxia and reoxygenation," *American Journal of Physiology - Heart and Circulatory Physiology*, vol. 282, no. 3 51-3, 2002, doi: 10.1152/ajpheart.00594.2001.

[45] M. M. Balach, C. H. Casale, and A. N. Campetelli, "Erythrocyte plasma membrane potential: past and current methods for its measurement," *Biophysical Reviews*, vol. 11, no. 6. Springer, pp. 995–1005, Dec. 01, 2019. doi: 10.1007/s12551-019-00603-5.

[46] N. Carvajal, C. Torres, E. Uribe, and M. Salas, "Interaction of arginase with metal ions: studies of the enzyme from human liver and comparison with other arginases," *Comparative Biochemistry and Physiology -- Part B: Biochemistry and*, vol. 112, no. 1, pp. 153–159, Sep. 1995, doi: 10.1016/0305-0491(95)00027-6.

[47] A. K. Pati, R. Maheshwari, and S. Gupta, "Opercular activity and temporal organization of surfacing behaviour in Indian catfishes, Clarias batrachus and Heteropneustes fossilis," *Biological Rhythm Research*, vol. 29, no. 1, pp. 75–85, 1998, doi: 10.1076/brhm.29.1.75.3046.

[48] G. Wu and D. A. Knabe, "Arginine synthesis in enterocytes of neonatal pigs," *American Journal of Physiology - Regulatory Integrative and Comparative Physiology*, vol. 269, no. 3 38-3, 1995, doi: 10.1152/ajpregu.1995.269.3.r621.

[49] S. Singh, N. Cheong, G. Narayan, and T. Sharma, "Burrow characteristics of the co-existing sibling species Mus booduga and Mus terricolor and the genetic basis of adaptation to hypoxic/hypercapnic stress," *BMC Ecology*, vol. 9, no. 1, pp. 1–7, Apr. 2009, doi: 10.1186/1472-6785-9-6.

[50] N. Özmeriç, S. Elgün, and A. Uraz, "Salivary arginase in patients with adult periodontitis," *Clinical oral investigations*, vol. 4, no. 1, pp. 21–24, 2000, doi: 10.1007/s007840050108.

[51] S. Dabir, P. Dabir, and B. Somvanshi, "The kinetics of inhibition of Vigna catjang cotyledon and buffalo liver arginase by L-proline and branched-chain amino acids," *Journal of Enzyme Inhibition and Medicinal Chemistry*, vol. 21, no. 6, pp. 727–731, Dec. 2006, doi: 10.1080/14756360600862317.

[52] E. Rodrigues, A. C. M. T. Ribeiro, and M. Bacila, "L-arginine metabolism in mitochondria isolated from the liver of antarctic fish Notothenia rossii and Notothenia neglecta," *Brazilian Archives of Biology and Technology*, vol. 49, no. 5, pp. 825–833, 2006, doi: 10.1590/S1516-89132006000600017.

[53] "(PDF) Human Erythrocyte Arginase Purification and Characterization." https://www.researchgate.net/publication/325847807_Human_Erythrocyte_Arginase_Purification_and_Characterization (accessed Jun. 09, 2021).

[54] M. Aminlari, H. R. Shahbazkia, and A. Esfandiari, "Distribution of arginase in tissues of cat (Felis catus)," *Journal of Feline Medicine and Surgery*, vol. 9, no. 2, pp. 133–139, Apr. 2007, doi: 10.1016/j.jfms.2006.10.001.

[55] E. Rodrigues, A. C. M. T. Ribeiro, and M. Bacila, "L-arginine metabolism in mitochondria isolated from the liver of antarctic fish Notothenia rossii and Notothenia neglecta," *Brazilian Archives of Biology and Technology*, vol. 49, no. 5, pp. 825–833, 2006, doi: 10.1590/S1516-89132006000600017.

[56] S. D. Nikolaeva, V. T. Bakhteeva, E. M. Fock, E. A. Lavrova, and R. G. Parnova, "Arginase activity in frog urinary bladder epithelial cells and its involvement in

regulation of nitric oxide production," *Journal of Evolutionary Biochemistry and Physiology*, vol. 44, no. 3, pp. 275–282, Jun. 2008, doi: 10.1134/S0022093008030022.

[57] K. W. Shimotohno, J. Iida, N. Takizawa, and T. Endō, "Purification and Characterization of Arginine Amidinohydrolase from Bacillusï¿¿brevis TT02-8," *Bioscience, Biotechnology, and Biochemistry*, vol. 58, no. 6, pp. 1045–1049, 1994, doi: 10.1271/bbb.58.1045.

[58] N. Nakamura, M. Fujita, and K. Kimura, "Agricultural and Biological Chemistry Purification and Properties of l-Arginase from Bacillus subtilis," 2014, doi: 10.1080/00021369.1973.10861080.

[59] P. Rudrapati and A. v Audipudi, "Production and Purification of Anticancer Enzyme L-Asparaginase from Bacillus firmus AVP 18 of Mangrove Sample through Submerged Fermentation." Accessed: Jun. 09, 2021. [Online]. Available: http://www.ijcmas.com

[60] T. Zhang, Y. Guo, H. Zhang, W. Mu, M. Miao, and B. Jiang, "Arginase from Bacillus thuringiensis SK 20.001: Purification, characteristics, and implications for l-ornithine biosynthesis," *Process Biochemistry*, vol. 48, no. 4, pp. 663–668, Apr. 2013, doi: 10.1016/j.procbio.2013.02.023.

[61] A. E. Zakalskiy *et al.*, "Overexpression of (His) 6-tagged human arginase i in Saccharomyces cerevisiae and enzyme purification using metal affinity chromatography," *Protein Expression and Purification*, vol. 81, no. 1, pp. 63–68, Jan. 2012, doi: 10.1016/j.pep.2011.09.001.

[62] I. Husain, K. Bala, A. Wani, U. Makhdoomi, F. Malik, and A. Sharma, "Arginase purified from endophytic Pseudomonas aeruginosa IH2: Induce apoptosis through both cell cycle arrest and MMP loss in human leukemic HL-60 cells," *Chemico-Biological Interactions*, vol. 274, pp. 35–49, Aug. 2017, doi: 10.1016/j.cbi.2017.07.001.

[63] K. Huang, T. Zhang, B. Jiang, W. Mu, and M. Miao, "Characterization of a thermostable arginase from Rummeliibacillus pycnus SK31.001," *Journal of Molecular Catalysis B: Enzymatic*, vol. 133, pp. S68–S75, Aug. 2016, doi: 10.1016/j.molcatb.2016.11.020.

[64] C. T. Gonçalves, R. Camacho, and E. Oliveira, "From sequences to papers: An information retrieval exercise," in *Proceedings - IEEE International Conference on Data Mining, ICDM*, 2011, pp. 1010–1017. doi: 10.1109/ICDMW.2011.184.

[65] R. Apweiler, "The Universal Protein resource (UniProt)," *Nucleic Acids Research*, vol. 36, no. SUPPL. 1, p. D190, Jan. 2008, doi: 10.1093/nar/gkm895.

[66] T. Madden, "The BLAST Sequence Analysis Tool."

[67] M. Chatzou *et al.*, "Multiple sequence alignment modeling: Methods and applications," *Briefings in Bioinformatics*, vol. 17, no. 6. Oxford University Press, pp. 1009–1023, Nov. 01, 2016. doi: 10.1093/BIB/BBV099.

[68] E. Gasteiger *et al.*, "Protein Analysis Tools on the ExPASy Server 571 571 From: The Proteomics Protocols Handbook Protein Identification and Analysis Tools on the

ExPASy Server." Accessed: Jun. 09, 2021. [Online]. Available: http://www.expasy.org/tools/.

[69] M. A. Zervou, E. Doutsi, P. Pavlidis, and P. Tsakalides, "Structural classification of proteins based on the computationally efficient recurrence quantification analysis and horizontal visibility graphs," *bioRxiv*, p. 2020.10.23.350736, Oct. 2020, doi: 10.1101/2020.10.23.350736.

[70] T. L. Bailey *et al.*, "MEME Suite: Tools for motif discovery and searching," *Nucleic Acids Research*, vol. 37, no. SUPPL. 2, 2009, doi: 10.1093/nar/gkp335.

[71] J. Luo, "Applied bioinformatics tools," in *Basics of Bioinformatics: Lecture Notes of the Graduate Summer School on Bioinformatics of China*, vol. 9783642389511, Springer-Verlag Berlin Heidelberg, 2013, pp. 271–301. doi: 10.1007/978-3-642-38951-1_9.

[72] M. Knudsen and C. Wiuf, "The CATH database," *Human Genomics*, vol. 4, no. 3, p. 207, Feb. 2010, doi: 10.1186/1479-7364-4-3-207.

[73] P. Benkert, M. Künzli, and T. Schwede, "QMEAN server for protein model quality estimation," *Nucleic Acids Research*, vol. 37, no. SUPPL. 2, 2009, doi: 10.1093/nar/gkp322.