Determination Of Helix Stability And Hydrophobicity In Folding Mechanism Of Chitinase Protein In The Plant Brassica Juncea

ARUNA AASIRVATHAM

ABSTRACT– The Brassica juncea plant have the chitinase protein which mainly helps the plant in their growth and development. Chitinase was described for the first time in 1911 by Bernard as anti-fungal factor. The chitinase protein helps plant to defend themselves against fungal attack. This protein can act directly on fungal and insect affected areas. Enzymes present in the protein can be used as free or in the immobilized form to kill fungi and insects affected areas in plant. Most of the chitinase enzyme classes of protein belongs to the family glycoside hydrolase 18 and glycoside hydrolase19. The chitinase protein was selected from the plant Brassica juncea. To study about the Helix stability, hydrophobicity, Best motif region, Force field, protein folding type it's classification, Number of folds, Half – life of protein and the functional site. These studies were undertaken to find the role of chitinase protein in the plant during the fungal attack and disease condition. The chitinase protein is expressed when the plant is affected by fungal disease. These protein kills the fungal cell wall in the plant and insects which cause disease. These protein helps the plant to defend themselves against the fungal attack.

KEYWORDS – Helix stability, Hydrophobic, Hydrophilic, Motif, Black leg disease, Protein folding mechanism, Chitinase protein

—————————— ◆ ——————————

INTRODUCTION

The Andaman and Nicobar group of islands are a vast repository of plant Brassica juncea situated 1200 km away from the mainland India in the emerald blue sea of Bay of Bengal. Indian sub continent account for over a thousand species of *Brassica japonica, Brasica arvensis, Brassica rapa* and *Brassica tournefortii* out of 10,000 to 12,000 Brassica species found world over. In Brassica plant about 150species are found in phyto-geographically interesting north eastern hill region. As per the present records, Andaman and Nicobar Islands consists about 132 species of mustards belonging to 59 genus and much more remain unreported. During the survey trips taken for the four genus and viz,sipadan, mabul, kapalai and borneo are found in andhaman and nicobar island,which have so far not been found in mainland India (Srivastava *et al*., 2000).

Brassica juncea, also known as mustard greens, Indian mustard, Chinese mustard, and leaf mustard, is a species of mustard plant. Sub-varieties include Southern Giant Curled Mustard, which resembles a headless cabbage such as Kale, but with a distinct horseradish-mustard flavor. It is also known as green mustard cabbage. It is a tetraploid having four set of chromosomes in each cell of the plant. The Seeds are globular, dark-brown and rarely yellow. This plant flowers in June-July and bears fruits in August-September,after ripening seeds don't germinate.Seeds begin to germinate in the spring. In deeper soil layers seeds remain viable for up to 5 years. As day length increases, mustard bolts up with a 3ft and (0.9 m) stalk supporting bright yellow flowers that soon develop into sickle-shaped green seed pods. It can withstand temperatures as low as 20ºF (-4ºC) without damage. It cannot stand temperatures above 85ºF (29 ºC). Tolerates a pH in the range 4.3 to 8.3. The parts of the plants like leaves, seed ,root and mustard oil are used in the treatment of disease (Balakrishnan *et al*., 2000).

- Seed - The plant seed have antibiotic effect which is used in the treatment of foot ache and tumor.
- Root - The roots are used in the treatment of cold, fever, and stomach disorders.
- Leaves - Leaves are applied in the fore head to relieve head ache.
- Mustard oil - The mustard oils are used to treat skin eruptions and ulcers.

In The Present Studies Following Objectives Were Undertaken

1. Identification of *Brassica juncea* chitinase protein.
2. Prediction of protein structure and finding the motif region in the chitinase protein.
3. Analysis of different types of motif in the chitinase protein.
4. Finding the hydrophobic and hydrophilic residues region in the protein.
5. Identifying the helix stability and the aminoacid residues involved to make the chitinase protein structure stable.

IJSER

ISSN 2229-5518

6. Determination of folding type, classification and it's residues of hydrophobic charged side chain(polar side chain,cysteine side chain and Interpretation of force field responsible residues atoms

REVIEW OF LITERATURE

The review of literature pertaining to the present study "Determination of Helix stability and Hydrophobicity in Folding mechanism of chitinase protein in the plant " **Brassica juncea** " is as follows.

2.1 Brassica juncea

The Brassica juncea Indian mustard is easier to grow than the brown mustard. Young tender leaves of mustard greens are used in salads or mixed with other salad greens. Older leaves with stems may be eaten fresh, canned or frozen, for potherbs, and to a limited extent in salads. Although widely and extensively grown as a vegetable, it is being grown more for its seeds which yield an essential oil and condiment since 1945. Mustard Oil is one of the major edible oils in India,the fixed oil content of rai varying between 28.6% and 45.7%. Oil is also used for hair oil, lubricants and, in Russia, as a substitute for olive oil. Adding 1.1–2.2% mustard oil to fresh apple cider retards fermentation. Seed residue is used as cattle feed and in fertilizers (Reed, 1976). Indian Mustard is a folk remedy for arthritis, footache, lumbago, and rheumatism (Duke and Wain 1981). Seed used for tumors in China and Root used as a galactagogue in Africa.Sun-dried leaf and flower are smoked in Tanganyika to "get in touch with the spirits". Ingestion may impart a body odor repellent to mosquitoes. Believed to be aperient and tonic, the volatile oil is used as a counterirritant and stimulant. The juncea plant is used as an antisyphilitic and emmenagogue. Leaves applied to the forehead are said to relieve headache (Burkill *et al*., 1966). In Korea, the seeds are used for abscesses, colds, lumbago, rheumatism, and stomach disorders. Chinese eat the leaves in soups for bladder, inflammation or hemorrhage.Mustard oil is used for skin eruptions and ulcers (Perry *et al*., 1980). Mustard greens are high in Vitamin A and C, and iron (Duke *et al* 1981).

Table 1: Requirement of Vitamin A and Vitamin C in Brassica Juncea plant

LEAF		ROOT	
REQUIREMENT	PERCENTAGE	REQUIREMENT	PERCENTAGE
WATER	91.8 MG	WATER	85.2 MG
PROTEIN	2.4 G	PROTEIN	0.3 G
FAT	0.4 G	FAT	8.8G
CARBOHYDRATE	4.3 G	CARBOHYDRATE	7.2 G
FIBER	1.0 G	FIBER	2.0 G
ASH	1.1 G	ASH	1.9 G
CALCIUM	160 MG	CALCIUM	120 MG
POTASSIUM	48 MG	POTASSIUM	34 MG
ASCORBICACID	2.7 MG	ASCORBICACID	2.0 MG
CAROTENE	73 MG	CAROTENE	45 MG

The **Table 1** mainly represents requirement of vitamin a and vitamin c with total acids and percentage present in root and leaf of the plant brassica juncea. The volatile mustard oil has strong antimicrobial (bacteria and fungi) properties.

IJSER

2.2 Scientific classification

Botanical name: Brassica juncea	Common name: Mustard

Kingdom	Plantae
Sub kingdom	Tracheobionta
Superdivision	Spermatophyta
Division	Magnoliophyta
Class	Magnoliopsida
Sub class	Dileniidae
Order	Capparales
Family	Brassicaceae
Genus	Brassica
Species	Brassica juncea

2.3 Nitrogen Fixing bacteria on Plant growth and yield of Brassica juncea

Reviewed the Role of Nitrogen Fixing Bacteria

Brassica juncea is the principle rabi oilseed crop in india which covers 22% of area and contributes 25% of production of total oil seed crops. The average productivity of the crop in india is nearly 1 tone. Soil nutrient supply system provides basic input for growth and development of plant. Biofertilizers have been mooted as ecofriendly poterntial fertilizer source for maintenance of soil health and sustainable crop production system(Geholt and Bohra,2001;Subha rao,1993). Free living Nitrogen fixing bacteria Azobacter has been considered as low cost biofertillizer in the agricultural production. Worldwide inoculation experiments carried out with strains of Azotobacter Chroococcum nitrogen fixing bacteria used for the plant growth and development and enhance the yield of crop in different soils and in

different climaological conditions. Azobacter is attributed to production of plant growth hormones, improved nutrient uptake and antagonistic effect on plant pathogens (Parmar and Dadarwal,1997). Azospirillum assimilates atmospheric nitrogen and fix it in soil and helps plant to save nitrogen. Azospirillum also secretes phytohormones in the plant root region, which in turn enhances the root growth. The results of the various experiments conducted throughout india have clearly shown that the Azospirillum can be used as a potential biofertillizer in both extensive and intensive agriculture. They found the the nitrogen fixing bacteria Azobacter and Azospirillum is used as biofertilizer in the plant Brassica juncea for the yield and plant growth. The aim of this study was to evaluate the efficiency of Azospirillum and Azobacter on plant growth and yield parameters of Brassica juncea. Azobacter inoculation of half dose of both the bacteria proved best in improving the plant growth and yield of the plant Brassica juncea (Irfan *et al*., 2000).

2.4 Copper stress in the plant Brassicajuncea

Reviewed the role of copper in the plant

Copper is an essential element for proper functioning of all living organisms including plants, but it can cause toxicity at elevated concentrations. In the present study, two varieties of Brassica juncea L. i.e. Pusa JK and TM 4 grown axenically were compared for Cu tolerance and accumulation ability. For further detailed biochemical studies, var.TM 4 was used because of its fast growth and better Cu accumulation in shoots. Toxic effects of Cu were manifested by a reduction in photosynthetic pigments and an increase in the levels of thiobarbituric acid reactive substances. The activities of antioxidant enzymes such as superoxide dismutase, ascorbate peroxidase, guaiacol peroxidase and catalase showed an increase in a concentration and exposure time dependent manner in roots of B. juncea exposed to copper, indicating that they play an important role in combating copper stress in this species (Singh *et al* ., 1999).

2.5 Brassica juncea Produces a Phytochelatin-Cadmium-Sulfide Complex

Reviewed the role of Phytochelatin – Cadmium – Sulphide complex in Brassica juncea.

Brassicajuncea,a natural allotetraploid derived from B.rapa and B.nigra,is a major oil seed crop of the Indian Sub – continent.B.napus is grown large amount in Australia and Canada because of it's desirable features Such as high-temperature tolerance, disease resistance and non-shattering (Woods *et.al.*,1991). Artificial resynthesis of B.juncea has been achieved through Sexual hybridization but the synthesis through somatic hybridization has not been well documented. So for the analysis of phytochelatin-cadmium sulphide complex in B.juncea, They developed the B.juncea lines through the protoplast fusion. Genetic diversity among B.juncea accessions is limited whereas there is a vast untapped diversity among the progenitor species. Therefore resynthesis of B.juncea is highly relevant for further improvement of this crop. The B.napus and B.carinata have

also been synthesized through protoplast fusion (Schenck and Robbelen et al.1982; Narasimhulu et al.1992). But the B.juncea plant from somatic hybridization were quite indistinguishable from nature. So for further analysis of presence of phytochelatin complex in the three species they selected B.juncea because of the presence of chitinase protein. This protein prevents the plant from the fungal attack. Phytochelatins (PCs) are enzymically synthesized peptides produced in higher plants and some fungi upon exposure to heavy metals. They examined that Phytochelatin Cadmium complex production in the Se-tolerant wild mustard Brassica juncea and found that it produces two types of Phytochelatin Cadmium Sulphide complexes. From this analysis they found that the Phytochelatins present in B.juncea plant can react with heavy metals and produce the more amount of Phytochelatin cadmium sulphide complex , then the other mustard species such as Brassica napus and Brassica carinata (David Speiser *et al.,* 1999).

2.6 *Brassica juncea* chitinase BjCHI1 inhibits growth of fungal phytopathogens and agglutinates Gram-negative bacteria

Reviewed the role of Fungal Phytopathogens

Brassica juncea BjCHI1 is a plant chitinase with two chitin-binding domains. Its expression, induced in response to wounding, methyl jasmonate treatment, *Aspergillus niger* infection, and caterpillar *Pieris rapae* feeding, suggests that it plays a role in defence.In this study, to investigate the potential of using BjCHI1 in agriculture, *Pichia*-expressed BjCHI1 and its deletion derivatives that lack one or both chitin-binding domains were tested against phytopathogenic fungi and bacteria. Transplastomic tobacco plant expressing BjCHI1 was also generated and its extracts assessed. In radial growth-inhibition assays, BjCHI1 and its derivative with one chitin-binding domain showed anti-fungal activities against phytopathogens, *Colletotrichum truncatum*, *C. acutatum*, *Botrytis cinerea*, and *Ascochyta rabiei*.BjCHI1 also inhibited spore germination of *C. truncatum*.Furthermore, BjCHI1, but not its derivatives lacking one or both domains, inhibited by the growth of Gram-negative bacteria (*Escherichia coli*, *Ralstonia solanacearum*, *Pseudomonas aeruginosa*) more effectively than Gram-positive bacteria (*Micrococcus luteus* and *Bacillus megaterium*), indicating that the duplicated chitin-binding domain, uncommon in chitinases, is essential for bacterial agglutination.Galactose, glucose, and lactose relieved agglutination, suggesting that BjCHI1 interacts with the carbohydrate components of the Gram-negative bacteria cell wall. Retention of chitinase and bacterial agglutination activities in transplastomic tobacco extracts by the implicates that BjCHI1 is potentially useful against both fungal and bacterial phytopathogens the field agriculture (Yuanfang Guan *et al*., 2001).

2.7 Transgenic Indian mustard (*Brassica juncea*) expressing tomato glucanase leads to arrested growth of *Alternaria brassicae*

Reviewed the role of fungal disease

Brassica juncea is an important oilseed crop of the Indian sub-continent. Yield loss due to fungal disease alternaria leaf spot caused by *Alternaria brassicae* is a serious problem in cultivation of this crop. Nonavailability of resistance genes within crossable germplasms of Brassica necessitates use of genetic engineering strategies to develop genetic resistance against this pathogen. The pathogenesis related (PR) proteins are group of plant proteins that are toxic to invading fungal pathogens, but are present in plant in trace amount. Thus, overexpression of PR proteins leads to increased resistance to pathogenic fungi in several crops. The PR protein glucanase hydrolyzes a major cell-wall component, glucan, of pathogenic fungi and acts as a plant defense barrier. We report the expression of a class I basic glucanase gene, under the control of CaMV 35S promoter, in Indian mustard and its genetic resistance against alternaria leaf spot.Southern and Northern hybridization confirmed stable integration and expression of the glucanase gene in mustard transgenics. Several independent transgenics were screened in vitro and under poly house conditions for their resistance against *Alternaria brassicae.* In an in vitro antifungal assay, transgenics arrested hyphal growth of *Alternaria brassicae* by 15–54%. Under pathogen-challenged conditions in poly house, the transgenics showed restricted number, size and spread of lesions caused by *Alternaria brassicae*. Also, the onset of disease was delayed in transgenics compared to untransformed parent plants. The results demonstrate potentiality of a PR protein from a heterologous source in developing alternaria leaf spot resistance in Indian mustard (kalyan *et al.*, 1989)

2.8 Metabolic engineering of fatty acid biosynthesis in Indian mustard (*Brassica juncea*) improves nutritional quality of seed oil .

Reviewed the role of nutritional quality of seed

The seed oil of *Brassica juncea*, a widely cultivated oil-seed crop, does not have properly balanced fatty acids required for human nutrition and energy. Moreover, all the cultivars in India produce oil with very high content of erucic acid (C22:1) which is nutritionally

IJSER

undesirable. To divert the carbon flux from erucic acid towards other potentially health beneficial fatty acids, two approaches were implemented. First, a novel FatB thioesterase from *Diploknema butyracea* was engineered into the *B. juncea* crop, driven by the napin promoter that is seed-specific. Second, the *B. juncea* fatty acid elongase was restricted at the genetic level by incorporation of hair-pin RNA known to cause post-transcriptional gene by the silencing. The fatty acid profile of the mature seed resulted in a 64–82% decrease in erucic acid production. Moreover, the altered seed fatty acid compositions in transgenic lines showed significant increases in the level of C18:1 and C16:0 or C18:0, along with enhancement in the ratio of C18:2/C18:3 and C18:1/C22:1. The reduction of C22 quantitatively accounted for the increase in the pool of C16 and C18 fatty acids in the seed oil. Interestingly, a significant finding was a 4–13% increase in oil content of the different transgenic lines developed by metabolic engineering involving both the plastidial and cytoplasmic enzyme approaches (SaheliSinha *et al*.,1996).

2.9 Chitinase-Mediated Inhibitory Activity of Brassica Transgenic on Growth of *Alternaria brassicae*.

Reviewed the role of inhibitory activity of chitinase

Chitinase, capable of degrading the cell walls of invading phytopathogenic fungi, plays an important role in plant defense response, particularly when this enzyme is over expressed through genetic engineering. In the present study, Brassica plant (*Brassica juncea* L.) was transformed with chitinase gene tagged with an overexpressing promoter 35 S CaMV. The putative transgenics were assayed for their inhibitory activity against *Alternaria brassicae*, the inducer of Alternaria leaf spot of Brassica both in vitro and under polyhouse conditions. In in vitro fungal growth inhibition assays, chitinase inhibited the fungal colony size by 12-56% over the non-trangenic control. The bioassay under artificial epiphytotic conditions revealed the delay in the onset of disease as well as reduced lesion number and size in 35S-chitinase Brassica as compared to the untransformed control plants (Kalyan et al., 2001).

2.10 BjCHI1 from Brassica juncea displays both chitinase and agglutination activity

Reviewed the role of agglutination activity in B.juncea

The proteins encoded by the Brassica juncea chitinase gene BjCHI1 and its derived genes BjCHI2 and BjCHI3 were expressed by Multi-copy Pichia expression system. The chitinase activity of FPLC purified BjCHI1, BjCHI2 and BjCHI3 were tested and the results showed that all the three proteins degraded both CM-chitin-RBV and colloidal chitin. The Km values of BjCHI1, BjCHI2 and BjCHI3 for CM-chitin-RBV were estimated as 0.799 mg/mL, 0.544 mg/mL and 0.793 mg/mL, respectively. When the colloidal chitin was used as substrate, the Km values were 0.281 mg/mL, 0.388 mg/mL and 1.643 mg/mL respeectively indicating chitin-binding domain can increase affinity of chitinase to insoluble substrate. In this the agglutination activity assay, only BjCHI1 shows activity when the protein concentration was more than the 33 micrograms/mL, while BjCHI2 and BjCHI3 without agglutination activity even when the concentration was increased as high as 800 micrograms/mL. This means that the two chitin-binding domains in BjCHI1 are essential for agglutination and BjCHI1 is the first protein which shows both chitinase and agglutination activity identified so far in plants (Ouyang *et al.,* 2001).

2.11 Inhibition of Fungal Growth by Combinations of Chitinase and β-1,3-Glucanase

Reviewed the role by combination of Chitinase and β -1,3-Glucanase

Chitinase and β-1,3-glucanase purified from pea pods acted synergistically in the degradation of fungal cell walls. The antifungal potential of the two enzymes was studied directly by adding protein preparations to paper discs placed on agar plates containing germinated fungal spores. Protein extracts from pea pods infected with Fusarium solani, Fusarium solani phaseoli, which contained high activities of chitinase and β-1,3-glucanase, inhibited growth of 15 out of 18 fungi tested. Protein extracts from uninfected pea pods, which contained low activities of chitinase and β-1,3-glucanase, did not inhibit fungal growth. Purified chitinase and β-1,3-glucanase,tested individually,which did not inhibit the growth of most of the test fungi. Only Trichoderma viride was inhibited by chitinase alone, and only Fusarium solani f.sp. pisi was inhibited by β-1,3-glucanase alone. However, combinations of purified chitinase and β-1,3-glucanase inhibited all fungi tested as effectively as crude protein extracts containing the same enzyme activities. The pea pathogen, Fusarium solani f.sp. pisi, and the nonpathogen of peas, Fusarium solani f.sp. phaseoli, were similarly strongly inhibited by chitinase and β-1,3-glucanase,indicating that the differential type by the pathogenicity of the two fungi is not due to the differential sensitivity to the pea enzymes. By the inhibition of fungal growth was caused by the lysis of the hyphal tips (Felix Mauch et al., 1999).

2.12 Expression of bacterial chitinase protein in tobacco leaves using two photosynthetic gene promoters

Reviewed the role of chitinase protein in Tobacco leaves.

A bacterial chitinase gene from *Serratia marcescens (chiA)* was fused to (i) a promoter of the ribulose bisphosphate carboxylase small subunit (*rbcS*) gene and (ii) two different chlorophyll a/b binding protein (*cab*) gene promoters from petunia. The resulting constructions were introduced into *Agrobacterium* Ti plasmid-based plant cell transformation vectors and used to generate multiple independent transgenic tobacco plants. *ChiA* mRNA and protein levels were measured in these plants. On average, the *rbcS*/*chiA* fusion gave rise to threefold more *chiA* mRNA than either *cab*/*chiA* fusion. We investigated the influence of sequences around the translational initiation ATG codon on the level of ChiA protein. The *rbcS*/*chiA* and *cab*/*chiA* fusions in which the sequence in the vicinity of the translational initiation codon is ACC ATGGC gave rise to transformants with higher levels of ChiA protein than those carrying a *cab*/*chiA* fusion with the sequence CAT ATGCG in the same region. This difference in translational efficiency is consistent with previous findings on preferred sequences in this region of the mRNA. In those transformants showing the highest level of ChiA expression, ChiA protein accumulated to about 0.25% of total soluble leaf protein. These plants contained significantly higher chitinase enzymatic activity than control plants (Jonathan Jones *et al.,* 2003).

2.13 Chitinase-**Like** Protein **CTL1 Plays a Role** in **Alter**ing **Root System Architecture** in **Response to Multiple Environmental Conditions**

Role of Chitinase protein in altering the root system

Plant root architecture is highly responsive to changes in nutrient availability. However, the molecular mechanisms governing the adaptability of root systems to changing environmental conditions is poorly understood. A screen for abnormal root architecture responses to high nitrate in the growth medium was carried out for a population of ethyl methanesulfonate-mutagenized Arabidopsis (*Arabidopsis thaliana*). The growth and root architecture of the *arm* (for *anion altered root morphology*) mutant described here was similar to wild-type plants when grown on low to moderate nitrate concentrations, but on high nitrate, *arm* exhibited reduced primary root elongation, radial swelling, increased numbers of lateral roots, and increased root hair density when compared to the wild-type control. High concentrations of chloride and sucrose induced the same phenotype. In contrast, hypocotyl elongation in the dark was decreased independently of nitrate availability. Positional cloning identified a point mutation in the *AtCTL1* gene that encodes a chitinase-related protein, although molecular and biochemical analysis showed that this protein does not possess chitinase enzymatic activity. CTL1 appears to play two roles in plant growth and development based on the constitutive effect of the *arm* mutation on primary root growth and its conditional impact on root architecture. We hypothesize that CTL1 plays a role in determining cell wall rigidity and that the activity is differentially regulated by pathways that are triggered by environmental condition (Christian Hermans and Daniel Bush 1992).

2.14 Crystal structures of a family 19 chitinase from *Brassica juncea* show flexibility of binding cleft loops

Role of crystal structure of family 19 Chitinase.

Brassica juncea chitinase is an endo-acting, pathogenesis-related protein that is classified into glycoside hydrolase family 19, with highest homology (50–60%) in its catalytic domain to class I plant chitinases. Here we report X-ray structures of the chitinase catalytic domain from wild-type (*apo*, as well as with chloride ions bound) and a Glu234Ala mutant enzyme, solved by molecular replacement and refined at 1.53, 1.8 and 1.7 Å resolution, respectively. Confirming our earlier mutagenesis studies, the active-site residues are identified as Glu212 and Glu234. Glu212 is believed to be the catalytic acid in the reaction, whereas Glu234 is thought to have a dual role, both activating a water molecule in its attack on the anomeric carbon, and stabilizing the charged intermediate. The molecules in the various structures differ significantly in the conformation of a number of loops that border the active-site cleft. The differences suggest an opening and closing of the enzyme during the catalytic cycle. Chitin is expected to dock first near Glu212, which will protonate it. Conformational changes then bring Glu234 closer, allowing it to assist in the following steps. These observations provide important insights into catalysis in family 19 chitinases (Wimal Ubhayasekera *et al.*, 2002).

3. MATERIALS AND METHODS

The following tools, softwares and database are discussed and they are used to find the proteomics analysis for chitinase protein in the plant Brassica juncea.

3.1 Retrieval of protein sequence

ISSN 2229-5518

3.1.1 Finding Deviation in fold

3.1.1.1 Protein Data Bank

3.2 Protein Structure Visualization

3.2.1 Antheprot 3D

3.3 Secondary Structure Classification.

3.3.1PDB SUM

3.3.2 PSI pred

3.3.3 Alpha pred

3.3.4 Chou and Fasmann

3.4 Visualizing the fold in the structure

3.4.1 YASARA

3.5 Hydrophobicity Analysis and the aminoacid residues involved with their position.

3.5.1 GENIOUS

3.5.2 HYDROMALC

3.5.3 CLC

3.5.4 mpex

3.5.5 molsoft

3.5.6 heliquest

3.5.7 HCA

3.5.8 BIOEDIT

Figure1: The fruit, flower and disease which affects the plant Brassica juncea

3.6 Finding polar and non – polar region

3.6.1 CHIMERA

IJSER

3.6.2 BMRD

3.7 Identifying motif region

3.7.1 motifscan

3.7.2 Motifsearch

3.7.3 Motif 3D

3.7.4 Improbizer

3.8 Finding Cysteine Side Chain in the structure with their position.

3.8.1 DIANNA 1.1

3.9 Protein Functional site prediction

3.9.1 efseek

3.10 Finding Pest region in the protein

3.10.1 PESTFIND

3.11 Fold Recognition

3.11.1 Desc – Fold

3.11.2 Foldindex

3.11.3 Foldnpro

3.11.4 MEMSAT

3.11.5 genthreader

3.12 Finding the type of fold in protein

3.12.1 PSCS

3.12.2 CSSP

3.12.3 PSIPRED

3.13 Analysing helix stability and mutation

3.13.1 mupro

3.13.2 Identification of Stabilizing residues

3.13.3 I- MUTANT

3.14 Calculating force field with responsible aminoacid residues involved

3.14.1 SPBDV

3.14.2 ANNOLEA

3.14.3 GROMOS

3.15 Physico chemical properties

3.15.1 protparam

The Materials and methods pertaining to the present study "Determination of Helix stability and Hydrophobicity in Folding mechanism of chitinase protein in the plant Brassica juncea" is as follows. The work involves online, offline tool, Databases and software's is also used.

3.1 Retrieval of protein Sequence:

3.1.1 Finding deviation in fold:

3.1.1.1 PROTEIN DATA BANK (PDB)

The **protein data bank (PDB)** (www.rcsb.org/pdb/home) is a repository for the 3-D structural data of large biological molecules, such as proteins and nucleic acids. The data typically obtained by X – ray crystallography or NMR spectroscopy and submitted by biologists and biochemists from around the world. The PDB is a key resource in areas of structural biology, Such as Structural genomics. The PDB database isupdated weekly and it contains 55,072 structure of protein and nucleic acids.

3.2 Protein Structure Visualization:

3.2.1 ANTHEPROT 3D

Antheprot 3D software is a molecule viewer that allows to look at PDB files. The program is developed by Pr Gilbert Deléage. Antheprot 3D is a molecular graphics program intended for the visualisation of proteins, nucleic acids from RCSB archive. Using this software the protein structure of chitinase protein was visualized. This program is also used to calculate the secondary structure content of proteins from their circular dichroism spectrum. All the information concerning analysis and results are given on a single screen.

3.3 Secondary structure classification:

3.3.1 PDBSUM

PDBsum provides an at-a-glance overview of every macromolecular structure deposited in the Protein Data Bank (PDB) and giving schematic diagrams of the molecules in each structure and the interaction between them. This database also provides a information about the secondary structure, class of protein and disulphide bonds. Using this the secondary structure of chitinase protein with their class and total number of disulphide bonds present was also analyzed.

3.3.2 PSIPRED

IJSER

PSIPRED protein structure prediction server allows users to submit a protein sequence, perform a prediction of their choice and receive the results of the prediction both textually via e-mail and graphically via the web. The user may select one of three prediction methods to apply to their sequence. PSIPRED, a highly accurate secondary structure prediction method; MEMSAT 2, a new version of a widely used transmembrane topology prediction method; or GenTHREADER, a sequence profile based fold recognition method. Using PSIPRED the fold was recognized with it's folding type for the chitinase protein Freely available to users at (http://globin.bio.warwick.ac.uk/psipred)

3.3.3 ALPHAPRED

AlphaPred prediction server predicts the alpha turn residues present in the protein sequence and it is based on the neural network training. Training. Irregular protein secondary structures are believed to be important structural domains involved in molecular recognition processes and protein folding. In this the tight turns are studied in detail. Depending on the number of residues forming the turn, the tight turns are classified as delta – turns, gamma turns, beta turns, alpha turns and pi turns. Using this the alpha turn residues are being analyzed.

3.3.4 Chou & Fasmann

Chou & Fasmann Secondary Structure Prediction Server. This server predicts secondary structure of protein from the amino acid sequence. In this server, Chou & Fasman algorithm has been implemented. The Chou-Fasman method are an empirical technique for the prediction of secondary structures in proteins, originally developed in the 1970s. The method is based on analyses of the relative frequencies of each amino acid in alpha helices, beta sheets, and turns based on known protein structures In which they are also being solved with X-ray crystallography. This server can be used by accessing freely by the users from the web page (http://www.biogem.org/tool/chou-fasman).

3.4 Visualizing the fold in the structure:

3.4.1 YASARA

YASARA is a software in which the the structure given by the SWISS – MODEL was visualized and analyzed. It is mainly used to analyze the secondary and tertiary structure of protein. YASARA is a molecular graphics, molecular modeling and molecular simulation program with an intuitive user interface, photorealistic graphics and support for affordable shutter glasses, autosteroscopic displays and input devices. Using Yasara the secondary structure of the chitinase protein was visualized to find the helix stability in the protein structure.

3.5 Hydrophobicity analysis and the amino acid residues involved with their position:

3.5.1 GENEIOUS

Geneious Pro is an integrated, cross platform bioinformatics software suite for manipulating, finding sharing, and exploring biological data such as DNA sequences or proteins, Phylogenies,3D structure information, publications, etc. Geneious tries to bundle various bioinformatics tool under one hood.

Geneious tool have the following festures

- Seuence alignment and sequence viewing
- Motif search and open reading frames(ORFs)
- Phylogenetic tree building UPGMA,Neighbor joining with Bootstrapping and consensus trees.
- Restriction analysis – finds and view restriction cut sites and performs restriction enzyme digests.
- Protein structure viewer
- Database integration – integrated searching with Genbank,Pubmed,BLAST and Uniprot.
- Teach bioinformatics – Create tutorials with direct links to material in geneious.

3.5.2 HYDROMCALC

HydroMCalc is a Java applet that calculates the mean hydrophobicity, mean hydrophobic moment and relative hydrophobic moment for the given protein sequence. This tool also gives the information about the Kyte-Doolittle and Eisenberg algorithm. The Amino acid residue, Hydrophobicity scale, and the position of the protein sequence was also being founded using HydroMCalc. The hydrophobicity scale and score for the chitinase protein was also predicted.

3.5.3 CLC

CLC - Combined Workbench was developed by CLC bio A/S. It has several workbenches such as genomic workbench, DNA workbench, RNA work bench, Sequence Viewer.

CLC workbench has the following features.

- View sequence.
- Genbank search and annotation.
- Align the protein sequence
- Create and modify a phylogenetic tree

IJSER

- Find restriction sites
- Proteolytic cleavage detection
- Sequence and genome assembly

3.5.4 MPex

Membrane Protein Explorer(**MPEx**), a tool for exploring the topology and other features of membrane proteins by means of hydropathy plots based on thermodynamic and biological principles. It is designed for examination of membrane protein sequences using hydropathy-plot methods popularized by Kyte and Doolittle (1982). The hydropathy profile of the entered sequence will be plotted in the graph window. The **black curve** is the actual profile; the superimposed green curve is a smoothed version of the profile.

3.5.5 MOLSOFT

MolSoft ICM Suite of Software provides an easy to use general environment for a biologist or chemist who is curious about protein structure. In just a few seconds we can browse hundreds of structures of interest, analyze and visualize sequences, alignments and binding sites. we can also perform molecular modeling, fully-flexible ligand and receptor docking, virtual ligand screening, chemical similarity Searching, chemical clustering and much more. The save ICM project capability is a very helpful tool and will save you time. Saving a project will allow you to quit from ICM and then return to the exact set-up at which you left off at a later date. A complete history of your ICM actions will be saved so that we can pick up exactly where we finished our previous ICM session. Using this the motif region involved in the aminoacid residues and their position were also founded.

- Load the PDB file 2z37 using the PDB Search tab.
- Move the structure around using the rotate and zoom options on the right hand side of the GUI
- Save the structure in an ICM Project .
- File/Save Project

3.5.6 HELIQUEST

HeliQuest calculates from an α-helix sequence its physicochemical properties and amino acid composition and uses the results to screen any databank in order to identify protein segments possessing similar features. The server is divided into 2 interconnected modules: the sequence analysis module and the screening module. The sequence analysis module is dedicated to characterizing helices known by the user. it determines properties as hydrophobicity, hydrophobic moment, Z and amino acid composition. The screening module allows user to screen large databases in order to find sequences that have the general physico-chemical features of a target sequence.It is availablefreely for the users at (http://heliquest.ipmc.cnrs.fr).

3.5.7 HCA

HCA is a new technique for protein sequence analysis. It is used to align and compare the protein sequence. The HCA is also used to find the clustered region in the protein secondary structure. And it is shown using different colours. Clusters of amino acid sequence are regarded as good markers for clustering. HCA is also used to find the fold region and secondary structure of protein(Helix). The protein structure was submitted in the database which gives the detailed information about the hydrophobicity clustered region present in the chitinase protein sequence.

3.5.8 BIOEDIT

BioEdit is a biological sequence editor that runs in windows 95/98/NT/2000 and is intended to provide basic functions for protein and nucleic sequence editing, alignment, manipulation and analysis. The main goal of BioEdit is to provide a useful tool for biologists who do not want to have to know much about a program to utilize it. BioEdit is intuitive, menu – driven, and highly graphical and offers a graphical interface for users to run external analysis programs. The main functions are intended to be visible by simply playing with menu options.

BioEdit software contains many features is following

- BioEdit translate DNA or RNA to protein, Find cDNA or the reverse complement, do a nice analysis with bar graph of the percentages of the four bases, and otherwise characterize your sequence.
- Automatically and manually annotate sequences with features such as introns, exons.
- Promoters, CDS, and all standard GenBank feature types Automatically annotate others.
- RNA comparative analysis tools, including co variation, potential pairings, and mutual
- User – defined motif searching using standard prosite nomenclature and utilizing IUPAC characters to allow searching in nucleic acid or amino acid sequences, as well as exact text searches including or ignoring gaps.

IJSER

- Six – Frame translation of nucleic acid sequences in to FASTA format ORF lists.

3.6 FINDING POLAR AND NON – POLAR REGION:

3.6.1 CHIMERA

Chimera software has been used to visualize the protein structure and to compare the 3D structure of macromolecules by superimposing one structure with another. The structure of chitinase protein was opened. Then select, secondary structure were chosen. Then the secondary structure of chitinase protein with their hydrophobic and hydrophilic region were seen in the chimera window.

3.6.2 BMRD

The **BMRB** database contains NMR data derived from biological molecules. These molecules include, peptides, proteins, and nucleic acids. NMR data about hemes, cofactors, ligands, polysaccharides, or any other biologically relevant molecules are accepted. The data should be assigned on an atom-specific basis and can be deposited using ADIT-NMR. The hydrophobic and hydrophilic region their corresponding residues, side chain, Hydropathy index, Percentage of occurence in chitinase protein structure was predicted.

BMRB homepage can be found in (http://www.bmrb.wisc.edu)

3.7 IDENTIFYING MOTIF REGION:

3.7.1 MOTIFSCAN

Motif scanning means finding all known motifs that occur in a sequence. This database lets us to paste a protein sequence, select the collections of motifs to scan for, and launch the search. A document deals with the interpretation of the match scores. The best motif region present in the protein structure with the score was predicted using Motif scanning.

3.7.2 MotifSearch

Motif Search allows to specify several different criteria when searching for sites, such as a p-value or an approximate false-positive rate given a set of negative examples. Motif Search is free and is licensed under the GNU General Public License (GPL); The motif search gives the detailed information about the motif region present in the protein structure with the score and matches the similar motif present in the protein sequence.

3.7.3 Motif 3D

Motif3D is a web-based protein structure viewer designed to allow sequence motifs, and in particular those contained in the fingerprints of the PRINTS database, to be visualised on three-dimensional (3D) structures. It is also used to view the motif present in the amino acid sequence. The Protein Data Bank Id is given as input to find the finger motifs present in the structure. Using this the finger motif region and secondary structure was also visualized. Motif3D can be used via the web interface available at: (http://www.bioinf.man.ac.uk/dbbrowser/motif3d/motif3d.htm)

3.7.4 Improbizer:

Improbizer searches for motifs in DNA , RNA and protein sequences that occur with improbable frequency (to be just chance) using a variation of the expectation maximization (EM) algorithm. works in an iterative fashion. It starts with one pattern. It scans the data set for matches and near matches to that pattern. It then collects the matches and near matches, and averages them together to create a new pattern one which reflects the sequence data as well as the initial pattern. Improbizer uses all initial-motif-sized subsequences of the first sequences as initial patterns. It runs each of these subsequences over the entire data set for one iteration and then takes the most promising looking subsequences through repeated iterations until the pattern no longer changes to fit the data. Though the final pattern does depend on the initial pattern somewhat, many different initial patterns will converge into the same final pattern. Because of this, and since screening the initial patterns is the slowest part of the process, Using this best motif region with the amino acid residue and position with motif score was found.

3.8 FINDING CYSTEINE SIDE CHAIN IN THE STRUCTURE WITH THEIR POSITION:

3.8.1 DIANNA 1.1

DIANNA 1.1 (DiAminoacid Neural Network Application) is a web server that

provides two services:

- Cysteine classification prediction
- Disulfide connectivity prediction

DiANNA 1.1 determines the cysteine species (free cysteine, half-cystine or ligand-bound) by using a support vector machine (SVM) with degree 2 polynomial kernel for the

spectrum representation, if a cysteine is predicted to be ligand-bound, then the most likely of the four most common ligands (iron, zinc, cadmium, carbon) is proposed. The cysteine side chain of the chitinase protein, their position and score were analyzed using DIANNA 1.1.

3.9 PROTEIN FUNCTIONAL SITE PREDICTION:

3.9.1 eFSEEK

Molecular function of proteins are determined by their three dimensional structures, thus the similarity of protein structure can give some clues to infer their functions. In many cases, the molecular function are begun with the molecular interaction with small molecules (ligands). eF-seek is a web server to search for the similar ligand binding sites for the uploaded coordinate file with PDB format. The representative binding sites in eF-site database are search by algorithm based on the clique search algorithm. eFsite (electrostatic-surface of Functional site) is a database for molecular surfaces of proteins' functional sites.

3.10 FINDING PEST REGION IN THE PROTEIN:

3.10.1 PESTFIND

epestfind allows rapid and objective identification of PEST motifs in protein target sequences. PEST motifs were defined as hydrophilic stretches of at least 12 amino acids length with a high local concentration of critical amino acids. Remarkably, negatively charged amino acids are clustered within these motifs while positively charged amino acids, arginine (R), histidine (H) and lysine (K) are generally forbidden. The quality of 'valid' PEST motifs is refined by means of a scoring parameter based on the local enrichment of critical amino acids as well as the motif's hydrophobicity. The output from **epestfind** is a simple text one. It reports poor and potential PEST motifs together with their PEST score, mass percent of DEPST and their hydrophobicity index. 'Valid' PEST motifs below the threshold score (5.0) are considered as 'poor', while PEST scores above the threshold score are of real biological interest. The higher the PEST score, the more likely is degradation of proteins.

Itis available at http://emboss.bioinformatics.nl/cgi-bin/emboss/epestfind

3.11 FOLD RECOGNITION:

3.11.1 Desc – Fold

DescFold(Descriptor-based Fold Recognition System) is a web server for protein fold recognition, which can predict a protein's fold type from its amino acid sequence. The server combines with a six effective descriptors a profile-sequence-alignment-based descriptor using Psi-blast e-values and bit scores, a sequence-profile-alignment-based descriptor using Rps-blast e-values and bit scores, a descriptor based on secondary structure element alignment (SSEA), a descriptor based on the occurrence of PROSITE functional motifs, a descriptor based on profile-profile-alignment(PPA) and a descriptor based on Profile-structural-profile-alignment (PSPA) .

3.11.2 Foldindex

Foldindex is the web based server used to find the number of folded region the aminoacid residues involved with the position. Protein folding type and classification with the identity score and the structure was also predicted for the chitinase protein using Fold index server database.

3.11.3 Foldnpro

FOLDnpro Protein Fold Recognition and Template-Based 3D Structure Prediction. FOLDpro is a web server to predict protein 3D structure using a machine learning fold recognition approach. It makes predictions in three steps.

- Step 1: Use a machine learning information retrieval approach to rank template proteins for the query protein, integrating a variety of similarity features including sequence/family information, sequence-sequence alignment, sequence-profile (profile-sequence) alignment, profile-profile alignment, and structural features
- Step 2: Generate profile-profile alignments between the query protein and the top ranked template proteins. Multiple templates are used to improve both the alignment and the structure modeling if necessary.
- Step 3: Based on the query-template alignments and 3D structures of the templates, Modeller (Sali and Blundell, 1993) is used to generate structure models for the query protein. Five models are generated for the query. The models ranked higher are generated based on the templates ranked higher. So they are presumably, but not always, better than the models ranked lower (e.g., fold1.pdb is likely better than fold2.pdb. fold2.pdb is likely better than fold3.pdb).

3.11.4 MEMSAT

MEMSAT program implements a new method for the prediction of the secondary structure and topology of all-helix integral membrane proteins based on the recognition of topological models. The method employs a set of statistical tables (log likelihood ratios) compiled from well-characterized membrane protein data, and a novel

dynamic programming algorithm to recognize membrane topology models by expectation maximization. The statistical tables encode definite biases towards certain amino acid species being on the inside, middle and outside of a cellular membrane.

3.11.5 GENThreader

GenTHREADER an efficient and reliable protein fold recognition method for genomic sequences is a new protein fold recognition method which is both fast and reliable. The method uses a traditional sequence alignment algorithm to generate alignments which are then evaluated by a method derived from threading techniques. As a final step, each threaded model is evaluated by a neural network in order to produce a single measure of confidence in the proposed prediction. The speed of the method, along with its sensitivity and very low false-positive rate makes it ideal for automatically predicting the structure of all the proteins in a translated bacterial genome (proteome) This method can be accessed from the web page (http://globin.bio.warwick.ac.uk/psipred).

3.12FINDING THE TYPE OF FOLD IN PROTEIN:

3.12.1 PSCS

Protein Structure Classification server allow to predict the class of protein from its amino acid sequence.It predict weather protein belong to class Alpha or Beta or Alpha+Beta or Alpha/Beta. This method is based on a statistical method. Using this server the class of the chitinase protein was predicted as Alpha + Beta.

This server can be accessed from the web page(http://www.imtech.res.in/raghava/proclass).

3.12.2 CSSP

CSSP is an interactive protein secondary structure prediction Internet server. The server allows a single sequence or multiple alignment to be submitted and returns predictions from six secondary of the structure prediction algorithms that exploit evolutionary information from multiple sequences. The server can be accessed from the web page http://barton.ebi.ac.uk/servers/jpred.html. A consensus prediction is also returned which improves the average Q3 accuracy of prediction by 1% to 72.9%.The server simplifies the use of current prediction algorithms and allows conservation patterns important to structure and function to be identified.

3.12.3 PSIPRED

PSIPRED protein structure prediction server allows users to submit a protein sequence, perform a prediction of their choice and receive the results of the prediction both textually via e-mail and graphically via the web. The user may select one of three prediction methods to apply to their sequence: PSIPRED, a highly accurate secondary structure prediction method; MEMSAT 2, a new version of a widely used transmembrane topology prediction method; or GenTHREADER, a sequence profile based fold recognition method. Using PSIPRED the fold was recognized with its folding type for the chitinase protein.

Freely available to users at (http://globin.bio.warwick.ac.uk/psipred)

3.13 ANALYSING HELIX STABILITY AND MUTATION:

3.13.1 MUpro

MUpro: Prediction of Protein Stability Changes for Single-Site Mutations from Sequences. It is under SCRATCH server for predicting protein tertiary structure and structural features. The SCRATCH software suite includes predictors for secondary structure, relative solvent accessibility, disordered regions, domains, disulfide bridges, single mutation stability, residue contacts versus average, individual residue contacts and tertiary structure. The user simply provides an amino acid sequence and selects the desired predictions, then submits to the server. Results are emailed to the user. The helix stability of the chitinase protein structure and the confidential score was also analyzed using mupro.

The server is available at (http://www.igb.uci.edu/servers/psss.html)

3.13. IDENTIFICATION OF STABILIZING RESIDUES

The **identification of stabilizing residues** in protein structures using the parameters, surrounding hydrophobicity, long-range order, stabilization center and conservation score. Proteins with known three dimensional structures can be submitted to the server. The server takes a four letter PDB code or an uploaded coordinate file in PDB format as input.

It can be accessed from the web page (http://sride.enzim.hu/)

3.13.3 I- MUTANT

The **I-MUTANT** is the software used to find whether the mutation affects the native protein structure and it is represented in the graphical form. Using this software the single mutation change in the amino acid residues affects

the stability of the chitinase protein structure and the normal fold of the protein is also being deviated.

3.14 CALCULATING FORCE FIELD WITH RESPONSIBLE AMINOACID RESIDUES INVOLVED:

3.14.1 SPBDV

Swiss-Pdb Viewer (Deep View) has been developed since 1994 by Nicolas Guex. Swiss-Pdb Viewer is tightly linked to SWISS-MODEL, an automated homology modeling server developed within the Swiss Institute of Bioinformatics (SIB) at the Structural Bioinformatics Group at the Bio zentrum in Basel. Swiss-Pdb Viewer can also read electron density maps, and provides various tools to build into the density. In addition, various modeling tools are integrated and command files for popular energy minimization packages can be generated.

3.14.2 ANNOLEA

ANNOLEA calculate the force field of the chitinase protein swiss-model was used. plots of Annolea(Melo et al), GROMOS empirical force field energy (van Gunsteren *et al*) were used to calculate the force field. The atomic empirical mean force potential **ANOLEA** (Melo *et al.*) is used to assess packing quality of the models. The program performs energy calculations on a protein chain, evaluating the "Non-Local Environment" (NLE) of each heavy atom in the molecule. The y-axis of the plot represents the energy for each amino acid of the protein chain. Negative energy values (in green) represent favorable energy environment whereas positive values (in red) un favorable energy environment for a given amino acid.

3.14.3 GROMOS

In **GROMOS** y-axis of the plot represents the (van Gunsteren *et al.*) empirical force field energy for each amino acid of the protein chain. Negative energy values (in green) represent favorable energy environment whereas positive values (in red) unfavorable energy environment for a given amino acid.

3.15 PHYSICO CHEMICAL PROPERTIES:

3.15.1 ProtParam

ProtParam is a tool, which allows the computation of various physical and chemical parameters for, a given protein stored in SWISSPROT OR TrEMBL for a user entered sequence. The computed parameters are the molecular weight, theoretical isoelectric point (pI), amino acid composition, extinction coefficient ,estimated half – life, instability index, aliphatic index and grand average of hydropathicity (GRAVY).The protein can be specified as swissprot or tremble accession number or Id, or in the form of raw sequence. If the accession number of a SWISSPROT OR TrEMBL entry provided, then we prompted with an intermediary page that allows us to select the portion of the sequence on which analysis is to be performed. The choice includes a selection of mature chains or peptides and domains from the SWISSPROT/TrEMBL table , as well as the possibility to enter the start and end position. By default the complete sequence will be analyzed. ProtParam tool has been used for analyzing physicochemical properties of chitinase protein, the amino acid sequence of chitinase protein in FASTA format is given as input in ProtParam tool available at (www.expasy.org).The physicochemical properties of chitinase protein were analyzed.

4. RESULTS AND DISCUSSION

The result and discussion pertaining to study “Determination of Helix stability and Hydrophobicity in Folding mechanism of chitinase protein in the plant Brassica juncea” is listed below. The results are discussed under the following topics.

4.1 Identification of chitinase protein

4.1.1 Amino acid content in chitinase protein

4.1.2 Physicochemical properties of protein

4.1.3 Visualizing Helix,Sheet and Coil present in the protein structure

4.2 Prediction of protein structure and finding motif region in the chitinase protein

4.2.1 Prediction of protein structure

4.2.2 Helix wheel for the predicted protein

4.2.3 Amino acid composition in protein sequence

4.3 Identifying motif region

4.3.1 Finding motif region in the chitinase protein structure

4.3.2 Visualizing the Motif region

4.3.3 Motif score

4.4 Analysing different type of motif region present in the chitinase protein

4.4.1 Type of motif in A - Chain

4.4.2 Type of motif in B - Chain

ISSN 2229-5518

4.4.3 Type of motif in C - Chain

4.4.4 Type of motif in D- Chain

4.5 Finding Hydrophobic and Hydrophilic region in the chitinase protein

4.5.1Hydrophobic aminoacids involved in chitinase protein structure

4.5.2 Prediction of polar and non – polar residues present in protein helical region

4.5.3 Hydrophobic amino acid residues and their corresponding residue with percentage

4.5.4 Hydrophilic amino acid residues and their corresponding residue with percentage

4.5.5 Hydrophobicity plot

4.5.6 Hydrophobicity analysis

4.5.7 Kyte – Doolitle and Hoop and Woods algorithm

4.5.8 Hydropathy score

4.5.9 Hydrophobic cluster analysis

4.6 Hydrophobic and cysteine amino acid involved in folding mechanism of chitinase protein

4.6.1Finding hydrophobic amino acid with their side chain

4.6.2 Finding cysteine amino acid involved in fold

4.7 Protein folding type and classification

4.7.1 protein folding and classification

4.7.2 Protein folding types

4.7.3 Disulphide bonds

4.7.4 Fold class recognition and deviation in fold

4.7.5 Identifying PEST region and finding functional site in chitinase protein

4.8 Helix stability and mutation

4.8.1 Identifying amino acid residues which makes the helix structure stable

4.8.2 Mutation changing the protein structure stability

4.9 Interpretation of force field

The present study "Determination of Helix stability and Hydrophobicity in Folding mechanism of chitinase protein in the plant Brassica juncea " has been undertaken to find the role of chitinase protein in the plant brassica juncea.

4.1 Identification of chitinase protein

The chitinase protein is involved in growth and development of the plant and it is expressed under stress condition of both biotic and abiotic by phytohormones such as ethylene, jasmonic acid and salicylicacid. Chitinase protein was taken from the plant BRASSICA JUNCEA and 3D structure of protein was identified from Protein data bank. PDB ID of the protein is 2z37. The Brasica juncea catalytic module with bound chloride ions (2z38) and Brassica juncea catalytic module Glu234ala mutant with bound chloride ions (2z39) are the two related PDB entries for chitinase protein. In Table1 The complete overview of the chitinase protein structure was shown with the Identity, Molecular weight, Sequence length, and Dihedral angle for the protein was analyzed. The Table2 represents the amino acid content with the percentage.Table3 shows the physicochemical properties of the protein. The 3D Visualization software ANTHEPROT 3D is mainly used to visualize the secondary structure of the protein. In Figure2 The Helix, sheet and coil present in the chitinase protein structure was visualized.

Table2: An Overview of chitinase protein

IJSER

ISSN 2229-5518

PROTEIN NAME	CHITINASE PROTEIN
ORGANISM	BRASSICAJUNCEA
PDB-ID	2Z37
STRUCTURE WEIGHT	1082.17
NUMBER OF CHAINS	A,B,C,D(4 CHAINS)
NUMBER OF HELICES	12 HELICES
NUMBER OF STRANDS	3 STRANDS
IDENTITY	80 – 90%
SEQUENCE LENGTH	244
DOMAINS	4 DOMAINS
DIHEDRAL ANGLE	3 ANGLE (PSI-34),(PHI-87),(CHI-12)
FIRST RESIDUE	DLSGGHI
LAST RESIDUE	SYTSSRRH

In the above Table2 shows an complete view about the protein structure their identity, molecular weight, number of dihedral angle present in the protein structure, prediction of functional site domains, number of helices and number of strands present in the structure, First and last amino acid residue present in the chitinase protein sequence was also analyzed.

4.1.1 Amino acid content in chitinase protein

Table3:Amino acid content in the protein with percentage

AMINOACIDS		CONTENT	PERCENTAGE
ALA	(A)	19	7.8%
ARG	(R)	9	3.7%
ASN	(N)	14	5.7%
ASP	(D)	21	8.6%
CYS	(C)	9	3.7%
GLN	(Q)	6	2.5%
GLU	(E)	8	3.3%
GLY	(G)	25	10.2%
HIS	(H)	5	2.0%
ILE	(I)	14	5.7%
LEU	(L)	11	4.5%
LYS	(K)	12	4.9%
MET	(M)	7	2.9%
PHE	(F)	14	5.7%
PRO	(P)	16	6.6%
SER	(S)	16	6.6%
THR	(T)	12	4.9%
TRP	(W)	7	2.9%
TYR	(Y)	11	4.5%
VAL	(V)	8	3.3%
PYL	(O)	0	0.0%
SEC	(U)	0	0.0%

The Table3 determines the single and triple letter code of the amino acid residues with content and percentage present in the chitinase protein sequence was analyzed.

4.1.2 Physicochemical Properties Of Protein

IJSER

Table4: properties of chitinase protein

PROPERTY	CHITINASE HYDROLASE PROTEIN
MOLECULAR WEIGHT	27028.2
THEORETICAL PI	5.05
TOTAL NUMBER OF AMINOACIDS	244
NEGATIVELY CHARGED RESIDUES (ASP + GLU)	29
POSITIVELY CHARGED RESIDUES (ARG + LYS)	21
NUMBER OF CARBON ATOMS	1203
NUMBE ROF HYDROGEN ATOM	1778
NUMBER OF OXYGEN ATOMS	320
NUMBER OF NITROGEN ATOM	362
NUMBER OF SULPHUR ATOM	16
TOTAL NUMBER OF ATOM	3679
INSTABILITY INDEX	31.57
PROTEIN	STABLE
ALIPHATIC INDEX	57.25
GRAND AVERAGE OF HYDROPATHICITY	0.429

The Table4 represents the properties of protein with the number of amino acids, negatively and positively charged amino acid residues, number of carbon, hydrogen, oxygen, nitrogen and sulphur atom, the stability of protein, and hydrophobicity percentage for the chitinase protein was predicted.

4.1.3 Visualizing helix,sheet and coil in the protein structure

The **ANTHEPROT 3D** visualization software mainly used for visualizing the secondary structure of protein. The Figure2 shows the helix, sheet and coil region plays an one of the important role for analyzing the protein which is mainly used to determine the stability of the chitinase protein present in the plant Brassica juncea.

Figure2: Showing Helix,Sheet and Coil region in the protein structure

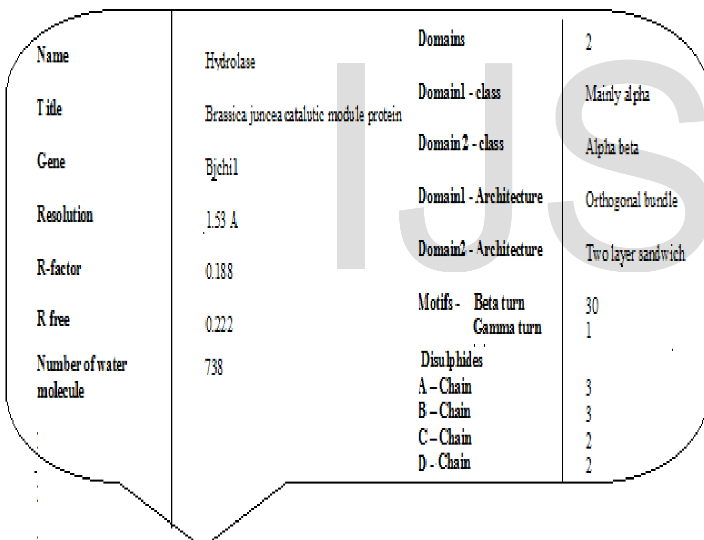
4.2 PREDICTION OF PROTEIN STRUCTURE AND FINDING MOTIF REGION IN THE CHITINASE PROTEIN.

The protein structure for the chitinase protein was predicted using the database **PDB SUM**. This database provides a detailed information about the protein with their structure and function. The Table5 represents the function of protein, type of fold, number of folds, number of water molecule, domains and motifs in the chitinase protein. These database is used in the prediction of protein secondary structure . Figure3 shows the length of the sequence and structure with the resolution and r-factor. The Helix, sheet, domains, motifs and disulphide bonds with their position in the chitinase protein sequence was shown in Figure4.Helix wheel and amino acid composition for the protein sequence was being identified for the predicted protein structure using **BIOEDIT** software. Figure5 represents the helix wheel for the predicted secondary structure. The amino acid composition with the position in sequence Figure6 and the amount present in the chitinase protein structure is also shown in Figure6 in graphical representation.

Motif region for the chitinase protein was identified using **Motif 3D** and **Motif search**. The motif 3D was used to identify motif region, polar and non polar group. Motif 3D score for the predicted structure of chitinase protein was also analyzed. Motif search was used to find number of motif present in the sequence, motif position and to identify the specific prosite pattern. Motif region with related sequence and structure is shown in Figure7. The best motif region was predicted in the protein sequence was shown in Figure8. The amino acid residues present in the chitinase protein structure was visualized in different colour in Figure9. The predicted motif region, Helix, Hydrophilic region, and amino acid involved in chitinase protein structure was visualized using MOTIF 3D in the Figure10 and Figure11 .The motif score was shown in Figure12 for the ten best motif region present in the protein sequence.

4.2.1 Prediction of protein structure.

Table 5: The following parameters was identified for chitinase protein.



Name	Hydrolase	Domains	2
Title	Brassica juncea catalutic module protein	Domain1 - class	Mainly alpha
Gene	Bjchil	Domain2 - class	Alpha beta
Resolution	1.53 A	Domain1 - Architecture	Orthogonal bundle
R-factor	0.188	Domain2 - Architecture	Two layer sandwich
R free	0.222	Motifs - Beta turn Gamma turn	30 1
Number of water molecule	738	Disulphides A – Chain B – Chain C – Chain D - Chain	 3 3 2 2

The Table 5 represents the resolution, r- factor, domain, fold, disulphide's, water molecules ,motifs present in the chitinase protein structure was analyzed.

In Figure3 the sequence length with the position and the total number of amino acid present in the protein with the secondary structure was analyzed for the chitinase protein.

The Figure4 shows the secondary structure prediction. Helix region present in the sequence with the position is labeled as H1 and H2. Disulphide bonds are shown in yellow colour and Sheets present in the sequence are labeled in red colour. The domains with type and class was also being predicted for the chitinase protein.

4.2.2 Helix wheel for the predicted protein.

The helix wheel was predicted for the motif region present in the chitinase protein structure using BIOEDIT software was shown in the Figure 5.

4.2.3 Amino acid composition in protein sequence.

The Figure 6 represents the total number of amino acid present and occurence in graphical form with the amino acid position, molecular weight and percentage.

4.3 IDENTIFYING MOTIF REGION.

The number of motif region present in the chitinase protein structure was predicted with the score from that the best motif region in the sequence was analyzed using Motif search and visualized using Motif 3D.

4.3.1 Finding motif region in the chitinase protein structure.

In the Figure7 the position where the motif is present in the chitinase protein with the related sequence and structure was predicted. The motif region was analyzed for the protein to find highly conserved region with the position and sequence can be predicted using the Motif search database.

The Figure8 shows the best motif region as **IAFKTAIWFWM** in the position **150………160** was predicted.

4.3.2 Visualizing the Motif region.

The Figure9 shows the chitinase protein structure and different colour represents different type of amino acid present in the protein sequence.

Residue colours also allows colouring of residues by amino acid properties:

- **RED = ACIDIC (ASP + GLU),**
- **BLUE = BASIC (HIS, LYS + ARG),**
- **GREEN = POLAR NEUTRAL (ASN, GLN, SER + THR),**
- **GREY = HYDROPHOBIC/ALIPHATIC (ALA, ILE, LEU, MET + VAL),**
- **PURPLE = HYDROPHOBIC/AROMATIC (PHE, TRP + TYR),**
- **YELLOW = DISULPHIDE BONDING (CYS),**
- **BROWN = SPECIAL STRUCTURAL PROPERTIES (GLY + PRO).**

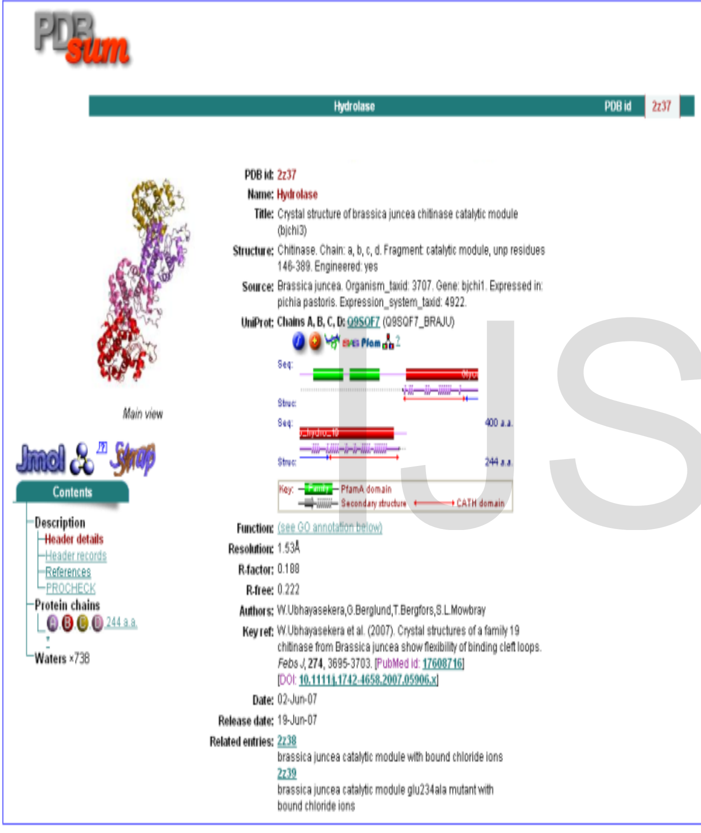
The Figure10 and Figure11 shows the Helix, Motif and hydrophilic region present in the chitinase protein structure was visualized and predicted using MOTIF 3D.

BLUE = HELIX, RED = MOTIF

GREEN = HYDROPHILIC REGION

4.3.3 Motif score.

The motif 3D score shown in Figure12 represents all the motif present in sequence with length and position these also shows the best motif as the same which was predicted



earlier as **IAFKTAIWFWM** in the position **150………160** was predicted.

Figure3: Sequence and structural representation with amino acid for chitinase protein

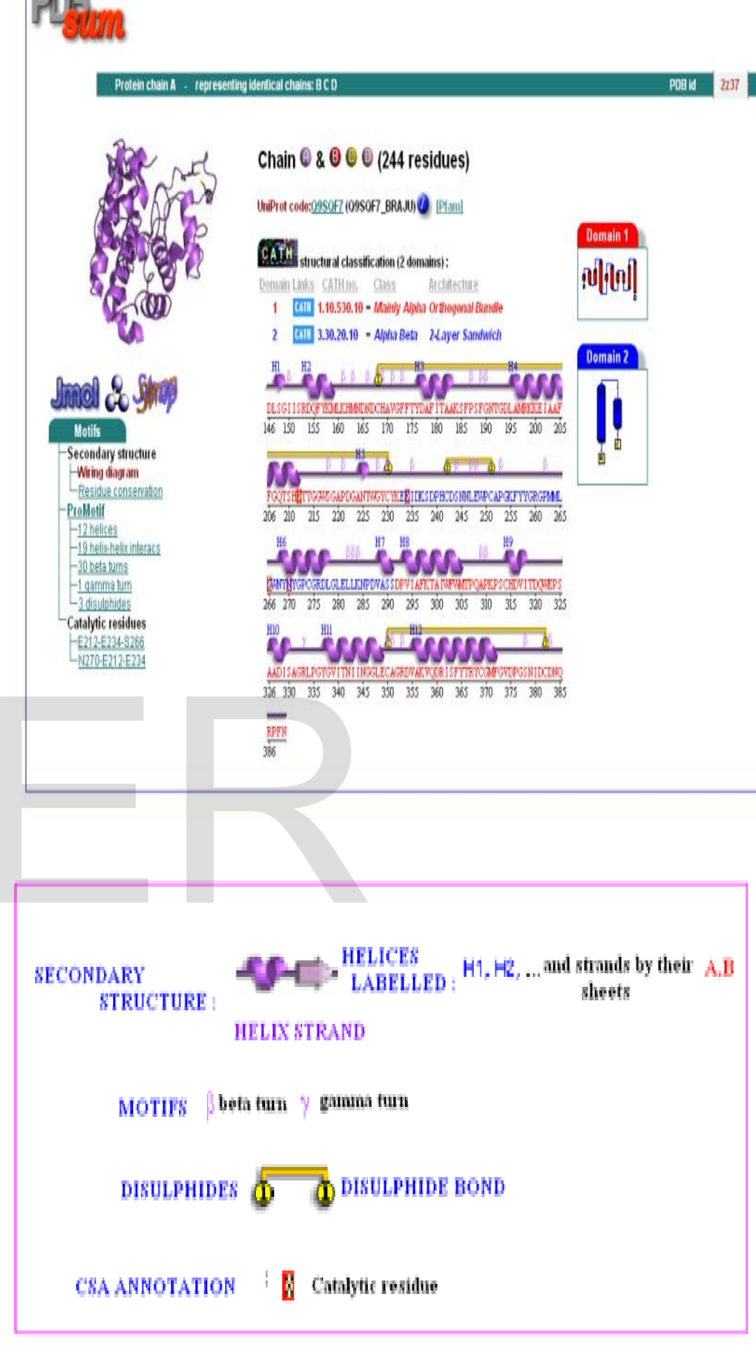


Figure4: Prediction of the secondary structure for chitiinase protein

IJSER

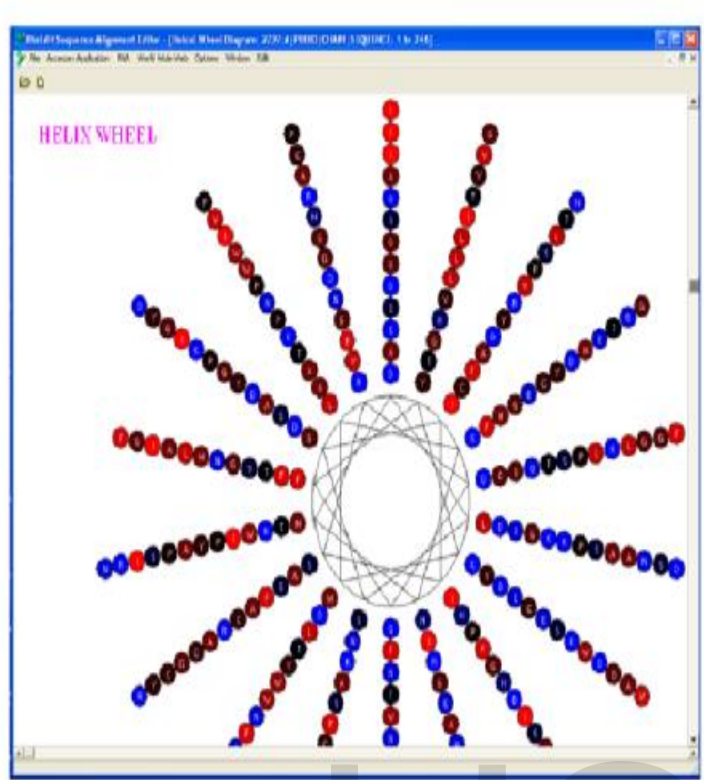
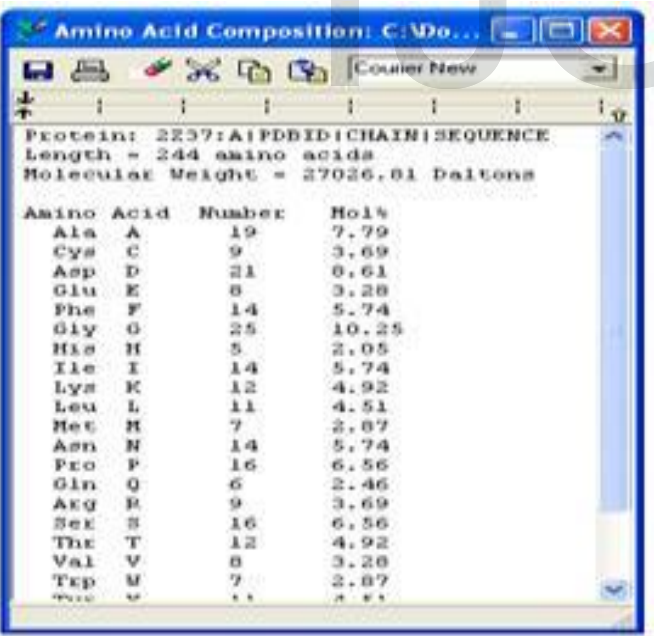


Figure5:Helix wheel for the motifs present in the protein

Figure7: Identifying motif position

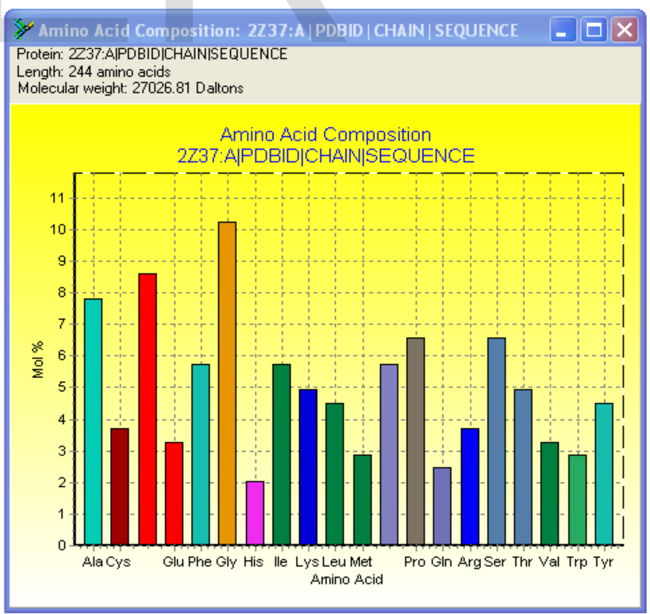


Amino Acid Composition: C:\Do...

Protein: 2Z37:A|PDBID|CHAIN|SEQUENCE
Length = 244 amino acids
Molecular Weight = 27026.81 Daltons

Amino Acid		Number	Mol%
Ala	A	19	7.79
Cys	C	9	3.69
Asp	D	21	8.61
Glu	E	8	3.28
Phe	F	14	5.74
Gly	G	25	10.25
His	H	5	2.05
Ile	I	14	5.74
Lys	K	12	4.92
Leu	L	11	4.51
Met	M	7	2.87
Asn	N	14	5.74
Pro	P	16	6.56
Gln	Q	6	2.46
Arg	R	9	3.69
Ser	S	16	6.56
Thr	T	12	4.92
Val	V	8	3.28
Trp	W	7	2.87

Figure6 : Amino acid composition with position and percentage



ISSN 2229-5518

Motif CHITINASE_19_2 in your sequence

Prosite ID:

CHITINASE_19_2 (PS00774)

Description:

Chitinases family 19 signature 2.

Pattern:

[LIVM]-[GSA]-F-x-[STAG](2)-[LIVMFY]-W-[FY]-W-[LIVM].

Appearance:

Position	Found Motif
150..160	IAFKTAIWFWM

Sequence:

DLSGIISRDQFYKMLKHMNDNDCHAVGFFTYDAFITAAKSFPSFGNTGDLAMRK
KEIAAF
FGQTSHETTGGWSGAPDGANTWGYCYKEEIDKSDPHCDSNNLEWPCAPGKFYY
GRGPMML
SWNYNYGPCGRDLGLELLKNPDVASSDPVIAFKTAIWFWMTPQAPKPSCHDVI
TDQWEPS
AADISAGRLPGYGVITNIINGGLECAGRDVAKVQDRISFYTRYCGMFGVDPGSNI
DCDNQRPFN

Figure8: Best motif region identification

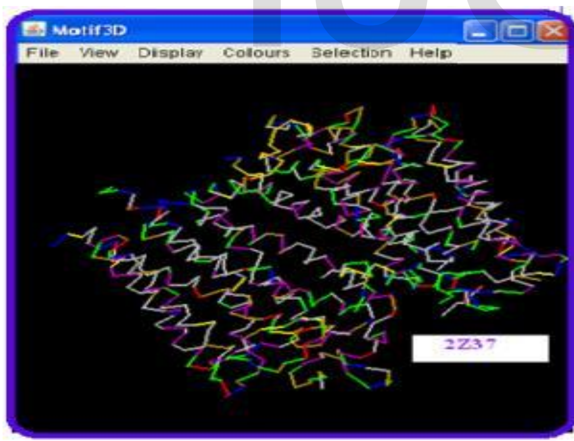


Figure9:Chitinase protein sequence and structure

>CHITINASE PROTEIN SEQUENCE
DLSGIISRDQFYKMLKHMNDNDCHAVGFFTYDAFITAAKS
FPSFGNTGDLAMRKKEIAAFFGQTSHETTGGWSGAPDGAN
TWGYCYKEEIDKSDPHCDSNNLEWPCAPGKFYYGRGPMML
SWNYNYGPCGRDLGLELLKNPDVASSDPVIAFKTAIWFWM
TPQAPKPSCHDVITDQWEPSAADISAGRLPGYGVITNIIN
GGLECAGRDVAKVQDRISFYTRYCGMFGVDPGSNIDCDNQ
RPFN

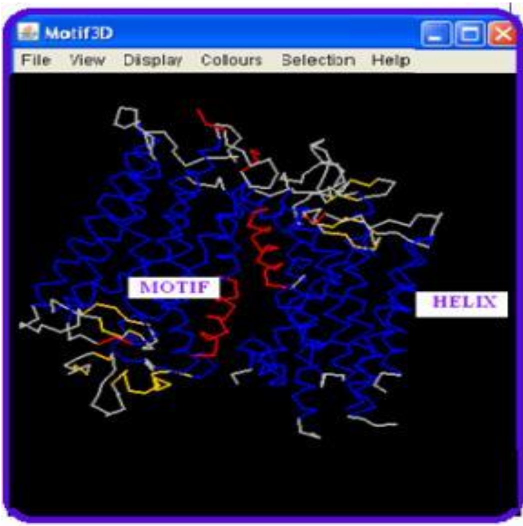


Figure10: Helix and Motif region

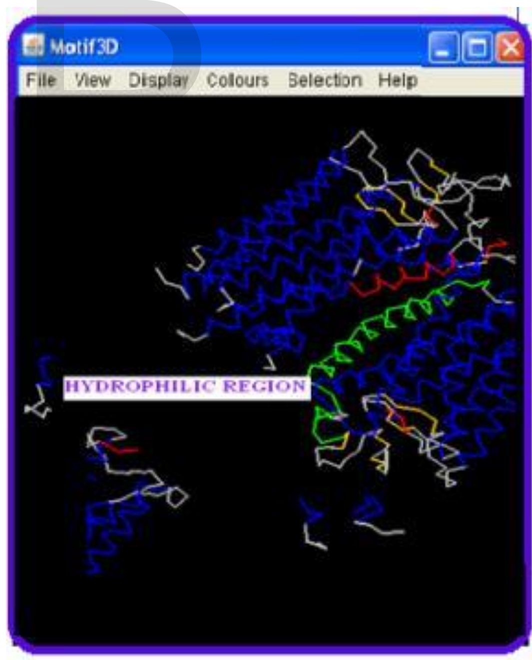


Figure11:Hydrophilic region

IJSER

MOTIF 3D SCORE

Ten top scoring fingerprints for your query

Ancestry	Fingerprint	No. of Motifs	SumId	AveId	PfScore	Pvalue	Evalue	GRAPHScan	
DPTHRIATOXIN	DPTHRIATOXIN	2 of 10	57	28	680	2.2e-05	41	.....I.i..	Graphic
CHAPERONIN60	CHAPERONIN60	2 of 5	47.62	23.81	510	2.9e-05	53	..1.1	Graphic
FLGPRINGFLGI	FLGPRINGFLGI	2 of 9	40.47	20.24	397	0.00018	3.8e+02	...i..i..	Graphic
7TM-->GPCRCLAN-->GPCRRHODOPSN-->MUSCARINICR	MUSCARINICR	2 of 9	68.25	34.12	481	0.00027	9.8e+02	..i...I..	Graphic
ALYTICPTASE	ALYTICPTASE	2 of 5	43.46	21.73	369	0.00031	7.9e+02	..i.i	Graphic
THRMBOMODULN	THRMBOMODULN	2 of 9	54.91	27.46	689	0.00062	1.3e+03	i....I...	Graphic
RNADNAPOLMS	RNADNAPOLMS	2 of 9	44.30	22.15	398	0.0017	4.3e+03	...i..i..	Graphic
HOMSERKINASE	HOMSERKINASE	2 of 5	55.90	27.95	329	0.0018	4.7e+03	Ii...	Graphic
7TM-->GPCRCLAN-->GPCRRHODOPSN-->LTBRECEPTOR-->LTB1RECEPTOR	LTB1RECEPTOR	2 of 6	57.33	28.67	388	0.0018	5.3e+03	.I.i..	Graphic
GENE66	GENE66	2 of 8	74.03	37.01	453	0.0021	5.6e+03	..i....I	Graphic

Ten top scoring fingerprints for your query. Detailed by motif

FingerPrint Name	Motif Number	IdScore	PfScore	Pval	Sequence	Length	low	Pos	high
DPTHRIATOXIN	6 of 10	30.43	330	1.42e-02	SFPSFGNTGDLAMRKKEIAAFFG	23	0	40	0
	8 of 10	26.09	350	1.57e-03	DISAGRLPGYGVITNIINGGLEC	23	0	183	0
CHAPERONIN60	3 of 5	19.45	215	2.71e-02	YDAFITAAKSFPSFGNTGDLAMRK	24	0	31	0
	5 of 5	28.17	295	1.07e-03	DQWEPSAADISAGRLPGYGVIT	22	0	175	0
FLGPRINGFLGI	4 of 9	21.00	220	1.02e-02	TYDAFITAAKSFPSFGNTGD	20	0	30	0
	7 of 9	19.47	177	1.74e-02	SAGRLPGYGVITNIINGGL	19	0	185	0
MUSCARINICR	3 of 9	23.32	191	9.28e-02	GFFTYDAFITA	11	0	27	0
	7 of 9	44.93	290	2.87e-03	ASSDPVIAFKTA	12	0	144	0
ALYTICPTASE	3 of 5	19.75	178	2.74e-02	ISAGRLPGYGVITNIING	18	0	184	0
	5 of 5	23.70	191	1.12e-02	GMFGVDPGSNIDCDN	15	0	225	0
THRMBOMODULN	1 of 9	23.33	329	3.52e-02	GQTSHETTGGWSGAPDGANT	20	0	62	0
	6 of 9	31.58	360	1.75e-02	TRYCGMFGVDPGSNIDCDN	19	0	221	0
RNADNAPOLMS	4 of 9	19.30	191	9.77e-02	GLELLKNPDVASSDPVIAF	19	0	134	0
	7 of 9	25.00	207	1.70e-02	KVQDRISFYTRYCG	14	0	212	0
HOMSERKINASE	1 of 5	31.94	187	4.64e-02	WNYNYGPCGRDLGLEL	16	0	122	0
	2 of 5	23.96	142	3.83e-02	GVITNIINGGLECAGR	16	0	193	0
LTB1RECEPTOR	2 of 6	30.77	188	9.50e-02	KMLKHMNDNDCHA	13	0	13	0
	4 of 6	26.56	200	1.90e-02	NIINGGLECAGRDVAK	16	0	197	0
GENE66	3 of 8	28.57	217	8.75e-02	KHMNDNDCHAVGFFTYDAFIT	21	0	16	0
	8 of 8	45.45	236	2.43e-02	GRDLGLELLKN	11	0	130	0

Figure12: Score for predicted motif with the position and length

COLOUR TEXT VIEW OF MOTIF

Different colour represent different motif.darker colour represent strong motif.

1 gnncagtataagntgaaagtttgggagan
2 tgccnncagggnngcggnaata-------
3 tactaaagggtnnggcagatcgggncn--
4 n----------------------------

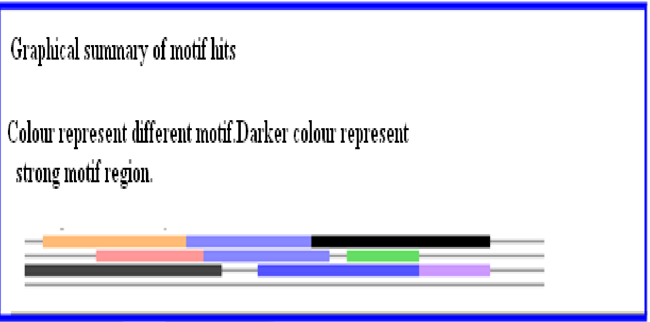
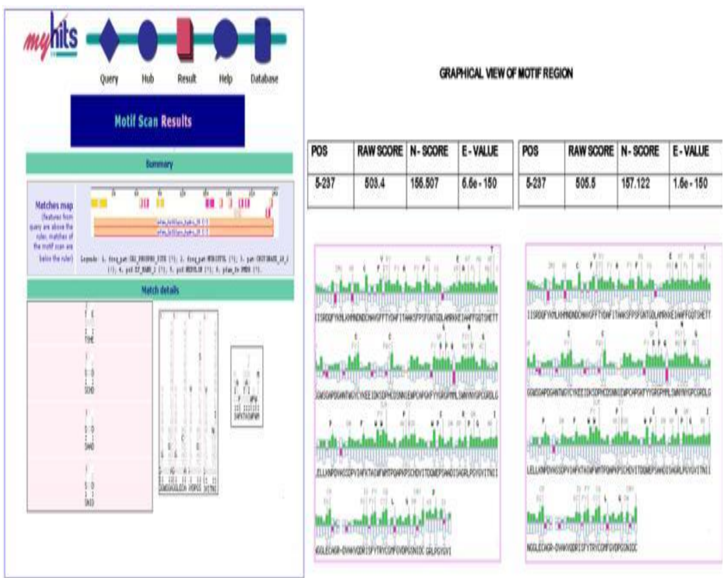


Figure13:Different type of motif and stronger motif in chitinase protein



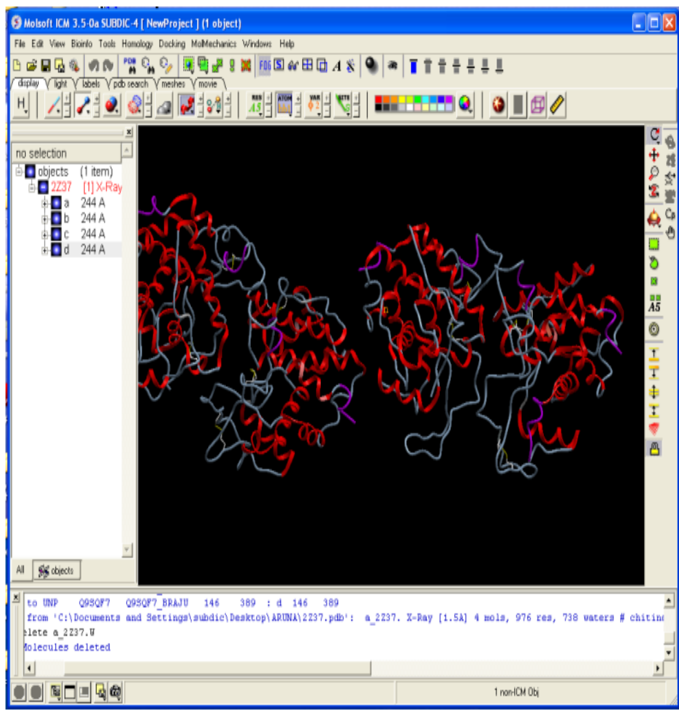
POS	RAW SCORE	N - SCORE	E - VALUE
1-21	88	7.428	0.79

POS	RAW SCORE	N - SCORE	E - VALUE
86-95	104	4.053	1.9E + 03

POS	RAW SCORE	N - SCORE	E - VALUE
187-195	0.2	7.180	1.4

Figure14: Graphical view of motif region with similarity

IJSER



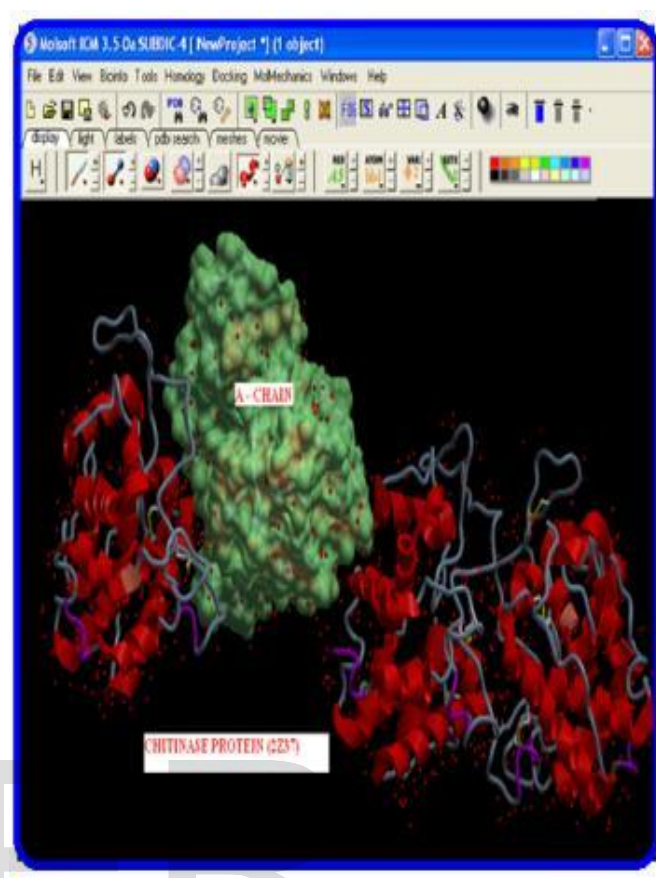


Figure15: Chitinase protein structure

IJSER

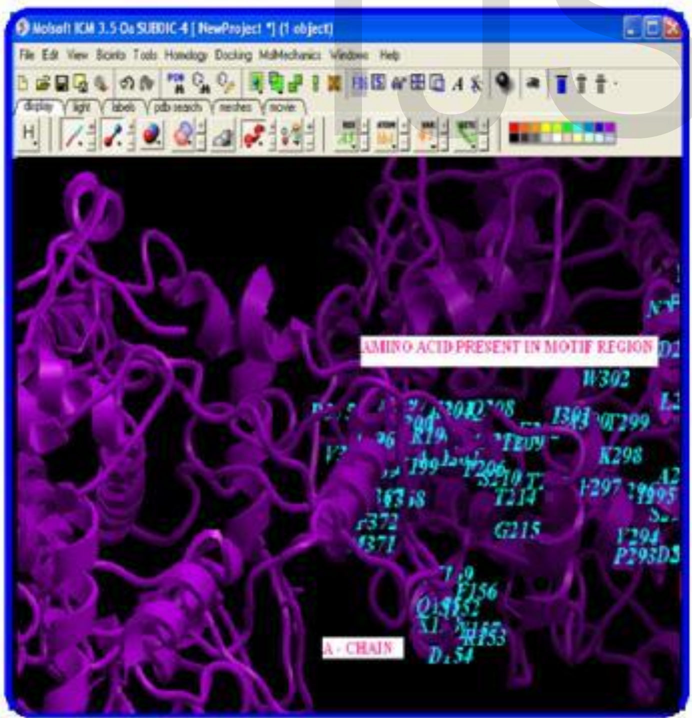
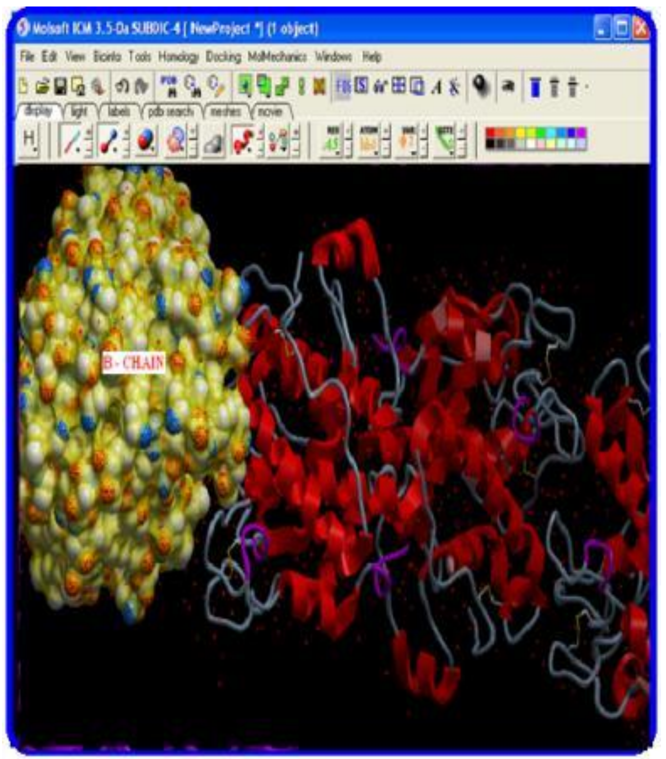


Figure16: A – chain



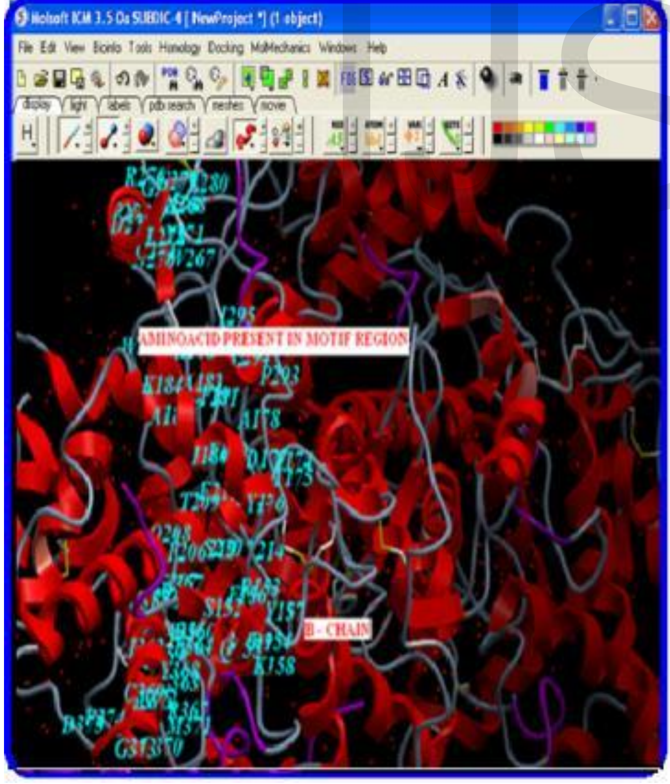
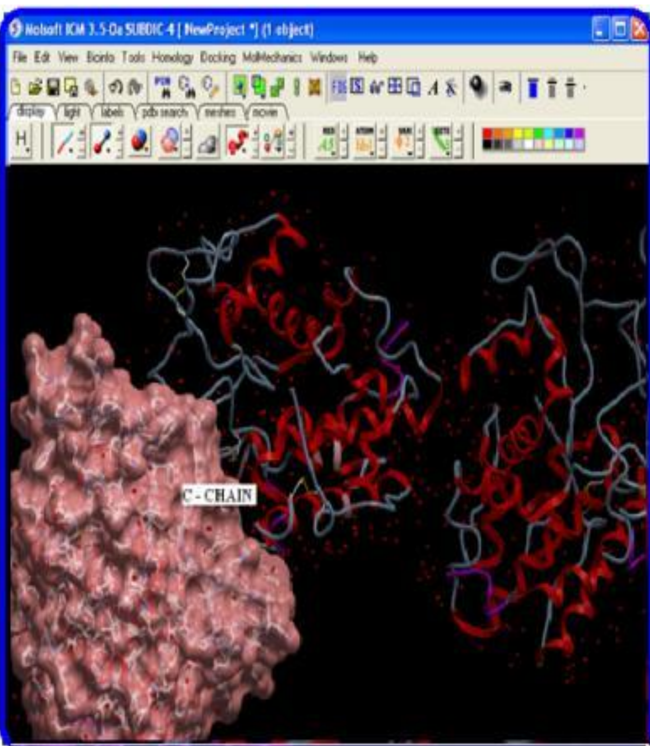


Figure18: B– chain

Figure19: Motif region in B- Chain



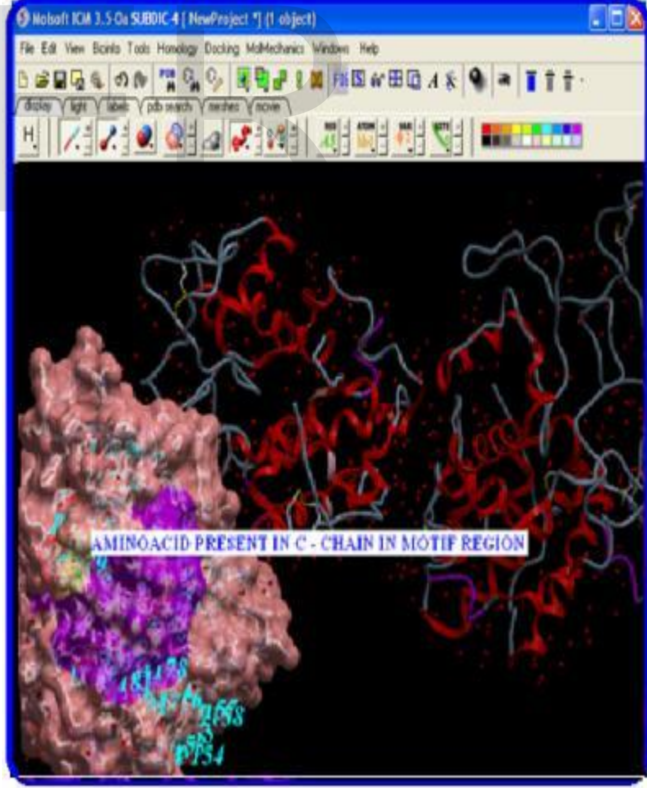


Figure20: C– chain \

Figure21: Motif region in C - Chain

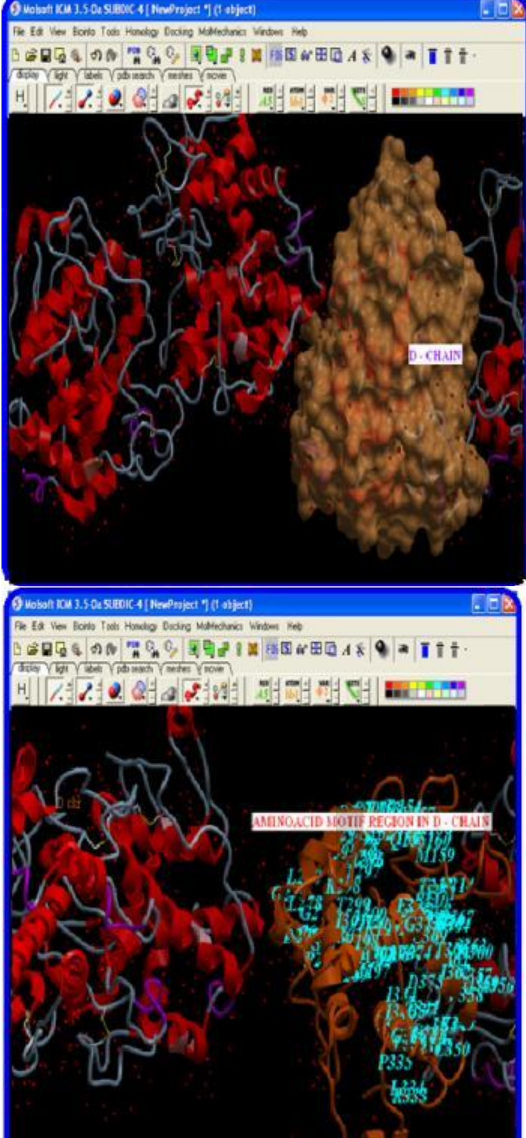


Figure22: D– chain

Figure23: Motif region in D – Chain

4.5 FINDING HYDROPHOBIC AND HYDROPHILIC REGION IN THE CHITINASE PROTEIN.

The Hydrophobic and Hydrophilic residues region are present inside the core region of the protein and plays an one of the important role for the structure stability. The Hydrophobic amino acid which are mainly used for the formation of the protein structure can be mainly predicted based on the hydrophobicity score. The Figure 24 represents the hydrophobic amino acid involved. Figure25 represents both hydrophobic, hydrophilic region and amino acid present in the protein was analyzed. The Table10 shows the frequency and percentage of amino acid involved in hydrophobicity. The helix wheel representation of hydrophobic amino acid was shown in Figure26.The hydrophobic and hydrophilic region present in the chitinase protein structure was visualized using Chimera is represented in Figure27 and Figure28.The Table11 and Table12 shows amino acid involved and corresponding residue with the percentage.Figure29 shows the Hydrophobic moment. The hydrophobicity plot with the hydrophobicity percentage is analyzed in the Figure30,31and Figure32. Kyte and hoop algorithm was predicted to find the best amino acid involved in hydrophobicity is shown in Figure33 and Figure34.The hydrophobic amino acid involved in core region of the protein can be predicted based on the hydrophobicity score is shown in the Table13.The hydrophobic cluster analysis was also done for the chitinase protein to find the hydrophobic and hydrophilic residues present in the protein Figure35.

4.5.1Hydrophobic aminoacids involved in chitinase protein structure.

The GENTHREADER is used to find the hydrophobic region present in the chitinase protein . It is viewed in three different colour in which red colour represents the residues which are highly hydrophobic. BLUE colour represents the residues which are hydrophilic is represented in the Figure24.The polar and non – polar residues are represented in the form of histogram plot. In which red color denotes hydrophobic region and GREEN color represents hydrophilic region in shown in the Figure25.

The Table10 shows hydrophobic amino acid with single letter and triple letter code. Frequency and percentage was predicted using GENEIOUS software. The Leucine and Iso leucine are some of the hydrophobic amino acid but they are not involved all the time in the fold only some of the amino acid are involved and they are packed tightly in the core of the protein which helps in the stability of the protein.

The hydrophobic amino acid which occurs in the core can be predicted using the algorithm and hydrophobicity score these plays an one of the important role in the chitinase protein taken from the plant Brassica juncea.

Table10: Hydrophobicity frequency statistics table

AMINO ACID	THREE LETTER	SINGLE LETTER	FREQUENCY	PERCEN
Alanine	Ala	A	32	(05
Cysteine	Cys	C	0	(00
Aspartic acid	Asp	D	2	(22
Glutamic acid	Glu	E	0	(00
Phenyl alanine	Phe	F	1	(17.
Glycine	Gly	G	1	(00
Histidine	His	H	103	(00.
Iso leucine	Ile	I	2	(00.
Lysine	Lys	K	0	(26
Leucine	Leu	L	1	(00
Methionine	Met	M	0	(23
Asparagine	Asp	N	0	(00
Proline	Pro	P	0	(00.
Glutamine	Glu	Q	1	(05
Arginine	Arg	R	128	(15
Serine	Ser	S	140	(19
Threonine	Thr	T	0	(5.
Valine	Val	V	0	(00.
Tryptophan	Try	W	0	(13
Tyrosine	Tyr	Y	128	(26
Selenocysteine	Sec	U	0	(20.
Pyrrolysine	Pyl	O	0	(2.

4.5.2 Prediction of polar and non – polar residues present in protein helical region.

The polar and non – polar residues present in the helical region of the chitinase protein was predicted and shown in the Figure26.These figure also represents the residues in the helix wheel model.

4.5.3 Hydrophobic amino acid residues and their corresponding residue with percentage

The chimera software is used to find the hydrophobic region present in A,B,C and D chain in the chitinase protein structure was visualized and shown in Figure 27 The hydrophobic amino acid residue ,their corresponding residue, position, Hydrophobic moment and Hydrophobicity percentage for the chitinase protein structure was shown in Table11.

Table11 Amino acid residue involved in hydrophobicity

HYDROPHOBICITY					
NAME OF RESIDUE	CORRESPONDING RESIDUE	POSITION	PERCENTAGE	HYDROPHOBIC MOMENT	HYDROPHOBICITY
Y,I,Y,L,F,M	T,D,K,A,R,D,S,Q	1 - 14	50%	0.153	0.372
Y,F,Y,L,F,M	K,S,D,D,R,A,Q	2 - 15	50%	0.114	0.371
L,F,M,I,Y,Y, F	T,D,K,R,Q,D	3 - 16	57.14%	0.172	0.503
Y,F,Y,L,F,M, I,Y	K,D,T,A	7 - 20	64.29%	0.077	0.510
I,M,F,L	G,D,K,A,T	8 - 21	57.14%	0.081	0.441
L,M,Y,F,I	G,A,A,K,T	9 - 22	57.14%	0.122	0.457
L,M,I,Y,F,A	G,A,T,D,K,A	10 - 23	57%	0.198	0.491
L,L,I,Y,F	A,G,A,D,A	11 – 24	57.14%	0.303	0.391
I,L,Y,F	T,A,G,A,D,K	12 – 25	64.29%	0.257	0.461
L,I,Y,F,M	T,A,K,D,R,G	13 – 26	57.14%	0.126	0.320
L,I,F,M	K,K,A,T,G,A,R,D	14 – 27	57%	0.098	0.304
I,L,M	K,K,K,T,D,G,A,R	16 – 29	42.86%	0.111	0.308
M,I,L	A,A,Q,K,K,S,R,G,A,E,T	24 – 37	42.79%	0.170	0.151
M,I,F	A,A,Q,K,K,S,R,G,A,E,T	25 – 38	35.71%	0.184	0.139
I,F	A,A,Q,K,K,S,R,G,A.	26 – 39	28.57%	0.226	0.005
I,F,W	Q,D,E,E,A,S,T,G	29 – 42	35.71%	0.174	0.416
Y,W,Y	E,N,G,T,H,S,P	35 – 48	28.57%	0.090	0.264
Y,W,Y,C	G,T,R,N,D,E,G,N,Q	39 – 52	35.7%	0.240	0.238
V,L,C,L	A,A,R,C,G,P,Q,D,S	47 - 60	42.86%	0.151	0.326

4.5.4 Hydrophilic amino acid residues and their corresponding residue with percentage

The chimera software is used to find the hydrophilic region present in A,B,C and D chain in the chitinase protein structure was visualized and shown in Figure 28.

The hydrophilic amino acid residue , their corresponding residue, position, Hydrophobic moment and Hydrophilicity percentage for the chitinase protein structure was shown in Table12.

Table12: Amino acid residue involved in hydrophilicity

HYDROPHILICTY					
NAME OF RESIDUE	CORRESPONDING RESIDUE	POSITION	PERCENTAGE	HYDROPHILIC MOMENT	HYDROPHILICITY
S,T	T,D,K,Q,I,R,S,A,E	1 - 14	50.00%	0.153	0.372
S,T,D,A	F,K,T,R,S,Q,M,L	2 - 15	50.00%	0.114	0.371
T	Y,D,F,A,T,K,L,F,I	3 - 16	57.14%	0.172	0.503
T,T,S	D,F,K,Y,L,A,M,Q,C	4 -17	57.14%	0.193	0.594
S,T,T,L,A	Q,K,E,A,D,G,H,I,S	27 - 40	64.29%	0.127	0.671
S,T,S,T	K,E,A,G,H,Q,A,S	28 - 41	74.13%	0.111	0.709
T,T,S,T	E,A,Q,W,A,H,I,K	29 - 42	42.26%	0.077	0.510
T,S,T,P,T	G,N,Y,Q,W,H,I	35 - 48	35.12%	0.081	0.441
T,T,P	R,G,N,E,Q,H,N,A	38 - 51	32.00%	0.122	0.457
T,P	G,C,N,G,D,N,R,Q	41 - 54	38%	0.198	0.491
S	R,A,G,G,D,E,Q	46 - 59	57%	0.303	0.391
S,S	R,L,A,G,D,C,L	47 - 60	50.00%	0.257	0.461
S,S,T	K,A,A,D	55 - 68	28.32%	0.126	0.320
T,T	H,N,D,E,G	75 - 88	4.71%	0.098	0.308

The Figure29 shows the hydrophobicity and hydrophobic moment plot for the residues present in the protein .The hydrophobic moment is a measure of hydrophobicity of the resisues when rotated in different angle.

IJSER

4.5.5 Hydrophobicity plot

The Hydrophobicity plot for the chitinase protein with the position was shown in Figure30 by which the position of amino acid are taken along X – axis and hydrophobicity along Y – axis. Circles in the plot represents the position of amino acid residues.

4.5.6 Hydrophobicity analysis

The hydrophobicity analysis for the chitinase protein was predicted and shown in Figure31 and Figure32 using MPex software. The polar and non – polar residues are shown in different shape the best hydrophobic region with the position was analyzed.

4.5.7 Kyte – Doolitle and Hoop Woods algorithm.

The Kyte – Doolitttle plot gives the information about the possible amino acid residues present in the protein structure is shown in Figure33.

The Hoop woods algorithm was developed to find the hydrophobicity residues in the protein. The amino acid residues are predicted based on the hydrophobicity score was shown in Figure34.

4.5.8 Hydropathy score

The hydropathy score Table13 shows the hydropathicity values. If the value is positive then the protein have hydrophobic residues. If the value is negative with less score shows that the protein have more hydrophobic amino acid region in the core region of chitinase protein.

Table13: Hydrophobicity score table

Hydrophobicity Scales		
	Kyte-Doolittle	Hopp-Woods
Alanine	1.8	-0.5
Arginine	-4.5	3.0
Asparagine	-3.5	0.2
Aspartic acid	-3.5	3.0
Cysteine	2.5	-1.0
Glutamine	-3.5	0.2
Glutamic acid	-3.5	3.0
Glycine	-0.4	0.0
Histidine	-3.2	-0.5
Isoleucine	4.5	-1.8
Leucine	3.8	-1.8
Lysine	-3.9	3.0
Methionine	1.9	-1.3
Phenylalanine	2.8	-2.5
Proline	-1.6	0.0
Serine	-0.8	0.3
Threonine	-0.7	-0.4
Tryptophan	-0.9	-3.4
Tyrosine	-1.3	-2.3
Valine	4.2	-1.5

Amino Acid Hydropathy Scores

Amino Acid Name	One Letter Code	Hydropathy Score
Isoleucine	I	4.5
Valine	V	4.2
Leucine	L	3.8
Phenylalanine	F	2.8
Cysteine	C	2.5
Methionine	M	1.9
Alanine	A	1.8
Glycine	G	-0.4
Threonine	T	-0.7
Tryptophan	W	-0.9
Serine	S	-0.8
Tyrosine	Y	-1.3
Proline	P	-1.6
Histidine	H	-3.2
Glutamic acid	E	-3.5
Glutamine	Q	-3.5
Aspartic acid	D	-3.5
Asparagine	N	-3.5
Lysine	K	-3.9
Arginine	R	-4.5

4.5.9 Hydrophobic cluster analysis.

The hydrophobic cluster analysis is a method for comparing and aligning protein sequence. The hydrophobic clusters are determined in the 3D pattern and then used for sequence comparison is analyzed for the chitinase protein and represented in the Figure35.

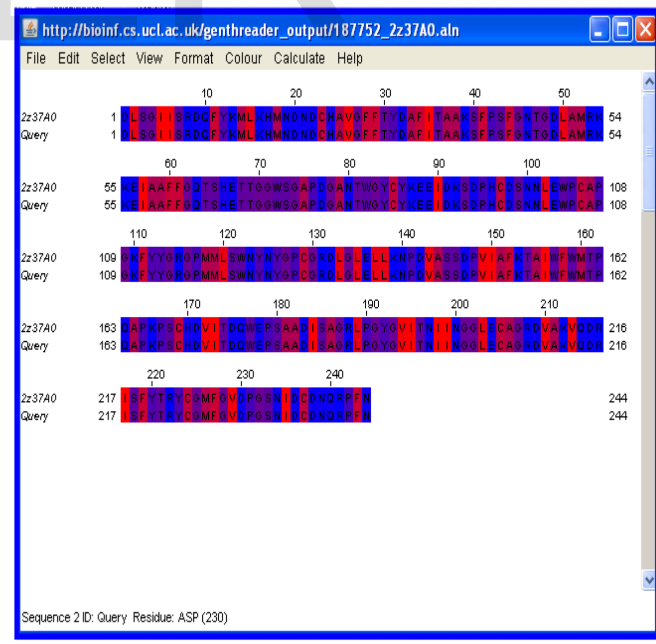


Figure24 Hydrophobic amino acid residues are shown in red colour

IJSER

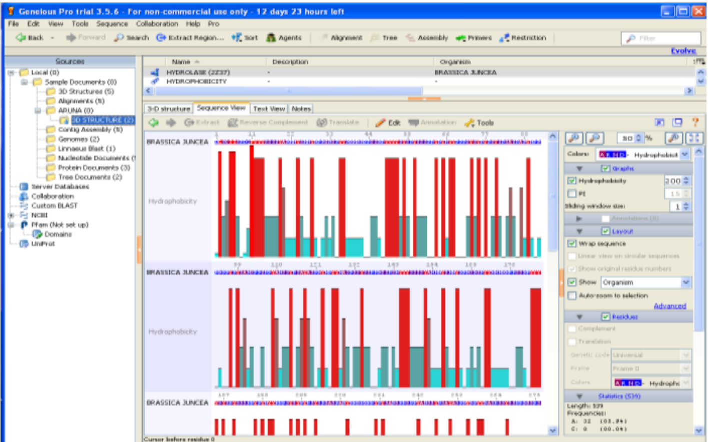


Figure25: Histogram showing hydrophobic amino acid in red colour

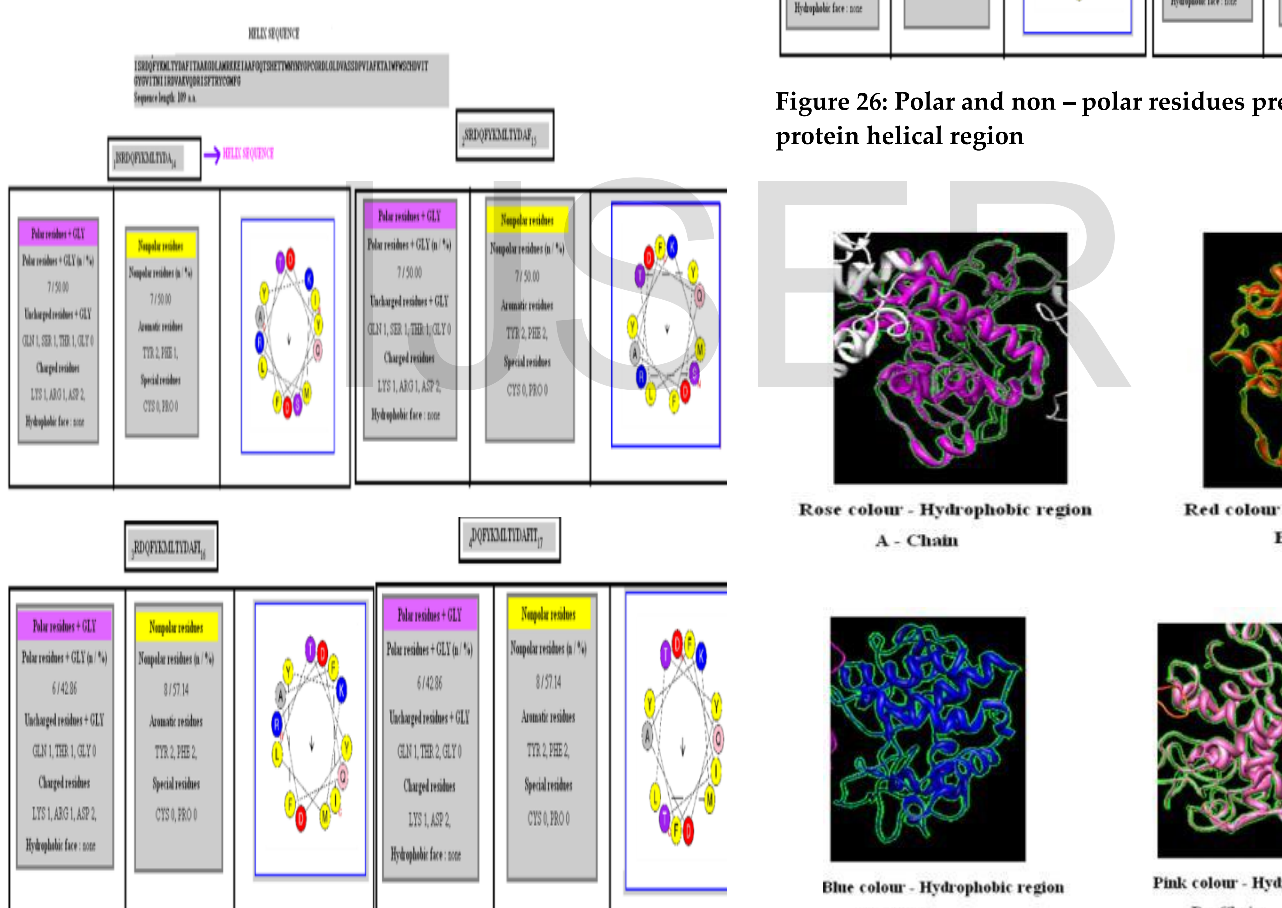


Figure 26: Polar and non – polar residues present in protein helical region

Figure 27: Visualizing hydrophobic region

IJSER

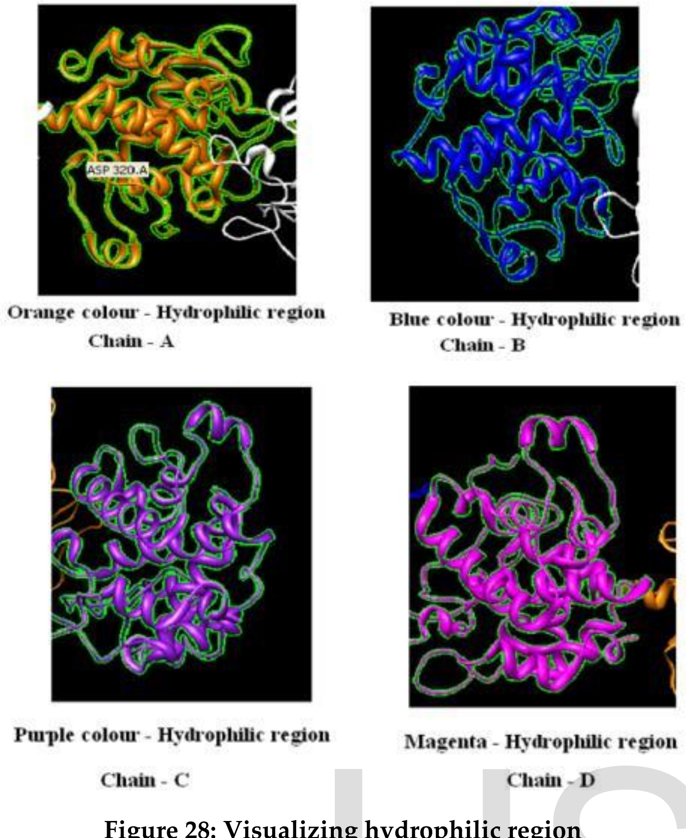


Figure 28: Visualizing hydrophilic region

Figure 29: Hydrophobic moment plot

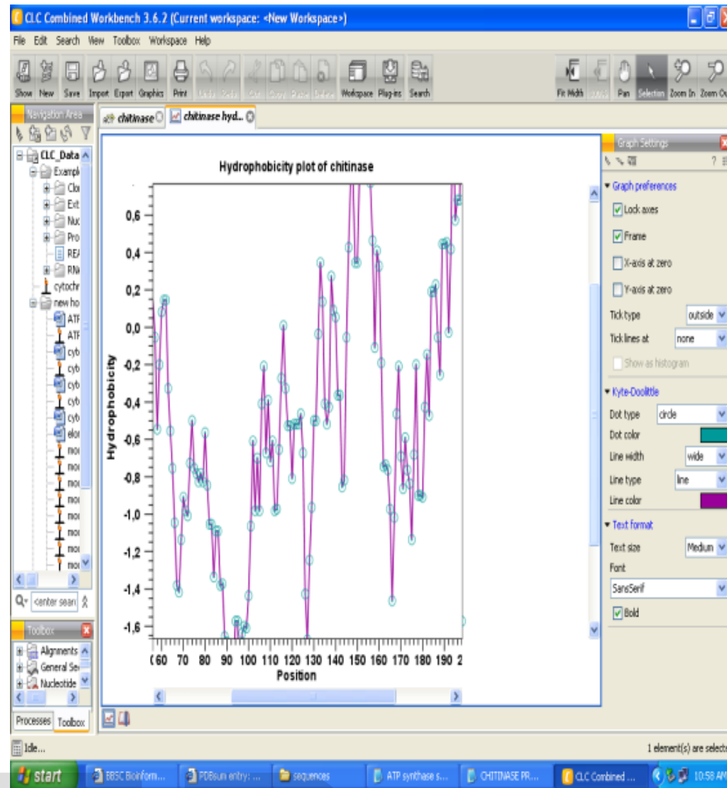


Figure 30: Hydrophobicity plot

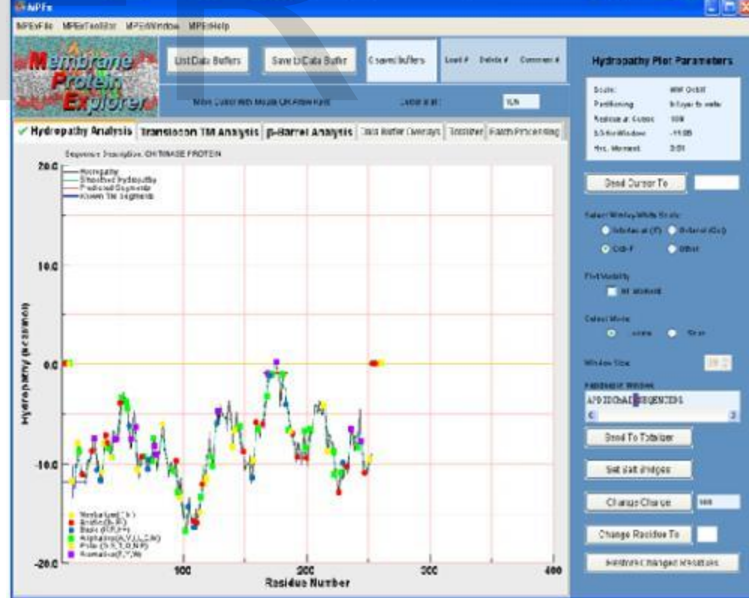


Figure 31: Hydropathy analysis

IJSER

Figure 32: Finding strong Hydrophobic region

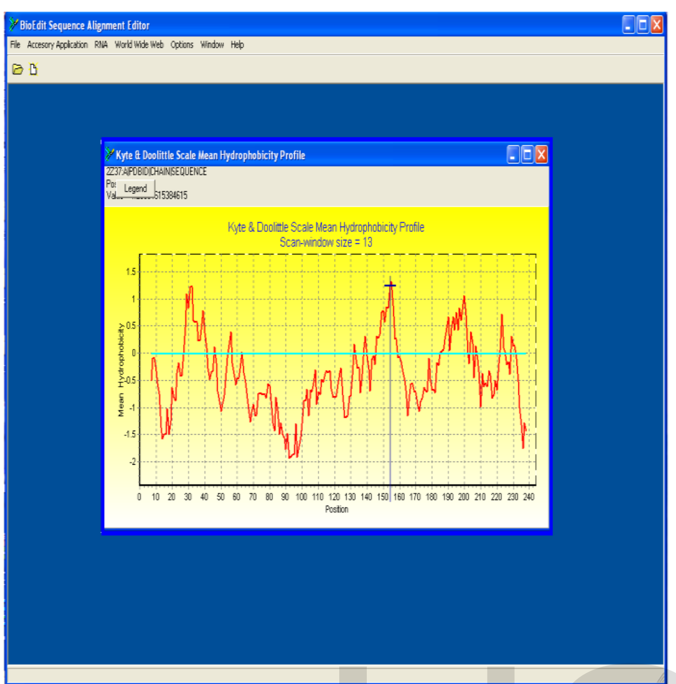


Figure 33: Kyte – Doolittle hydrophobicity plot

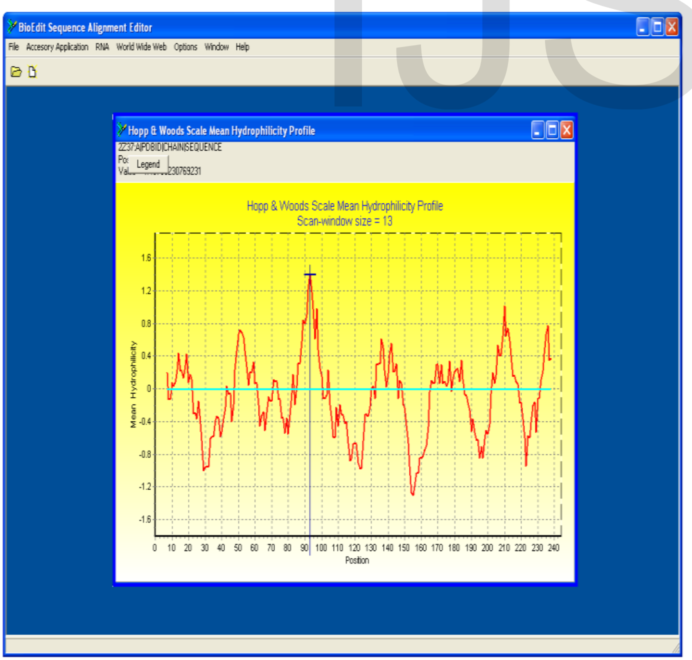


Figure 34:Hoop and Woods hydrophobicity plot

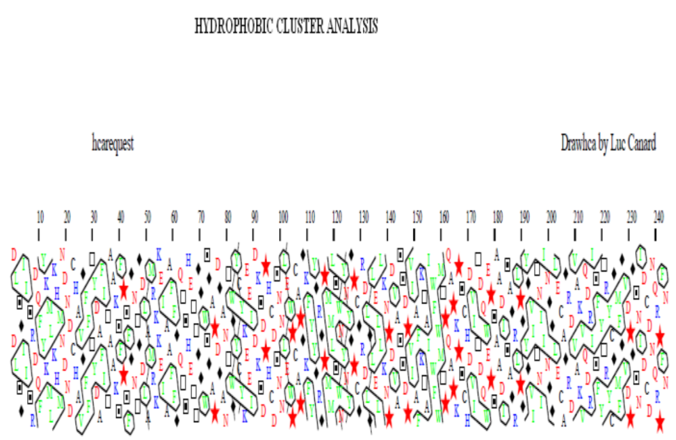


Figure 35:Hydrophobic cluster analysis

4.6 HYDROPHOBIC AND CYSTEINE AMINO ACID INVOLVED IN FOLDING MECHANISM OF CHITINASE PROTEIN.

The amino acids in a protein form associations with each other that try to keep hydrophobic amino acids together and separate them from the water and hydrophilic amino acids. If a hydrophobic amino acid such as valine were surrounded by water, this would be a very unstable structure. So the valine amino acids try to stay away from the water by associating with other hydrophobic amino acids such as phenylalanine and leucine. The hydrophobic amino acid involved in chitinase protein are shown in the Figure36 in which the amino acid which makes the protein structure stable are predicted based on hydrophobicity score. The Table14 represents the non polar amino acids.

IJSER

ISSN 2229-5518

Table14:Hydrophobic amino acid and their side chain.

NON POLAR AMINOACID RESIDUE	NON POLAR SIDE CHAIN (R-GROUP)	POSITION	HYDROPATHY INDEX	OCCURRENCE IN PROTEIN PERCENTAGE
CYSTEINE	CH2 - SH	23 - 35	2.5	1.9
PHENYL ALANINE	CH2 (benzene ring)	28 - 40	2.8	3.9
ISOLEUCINE	CH3 - CH - CH2 - CH3	90 - 100	4.5	6.3
LEUCINE	CH2 - CH - CH3 - CH3	102 - 110	3.8	9.1
METHIONINE	CH2 - CH2 - S - CH3	118 - 130	1.9	2.3
VALINE	CH (CH3, CH3)	143 - 151	4.2	6.6
TRYPTOPHAN	CH2 - C - CH - N (O)	72 - 85	- 0.9	1.4

The amino acid with their side chain,hydropathy index,amino acid position and their percentage of occurrence in the chitinase protein is represented in Table14.

4.6.1Finding hydrophobic amino acid with their side chain.

The hydrophobic amino acid was predicted using EMBL and BMRD. The residues which make the protein structure stable was analyzed based on the score. The side chain structure of all the non polar amino acids with their molecular formula, reference table, percentage of occurrence in the chitinase protein was determined. This was shown in the Figure36 and Figure37.The Hydrophobic amino acid in A,B,C, and D chain is shown in the Figure38 using Chimera software. The violet colour shown in the figure represents the Non – polar region present in the chitinase protein structure.

4.6.2 Finding cysteine amino acid involved in fold.

The Cysteine amino acid play an major role in protein because these amino acid mainly occurs near the disulphide bonds. The fold in the protein will be present near the cysteine residues Figure39 shows the total number of cysteine residues present in the protein with their position and bond index. The free and ligand bound cysteine with their score was predicted using DIANNA 1.1 server for the chitinase protein.

Table15:Cysteine residues involved with their position

Predicted disulfide bonds(cysteine pairs) ordered by probability in descending order:

Bond_Index	Cys1_Position	Cys2 Position
1	205	237
2	97	106
3	23	85
4	129	169

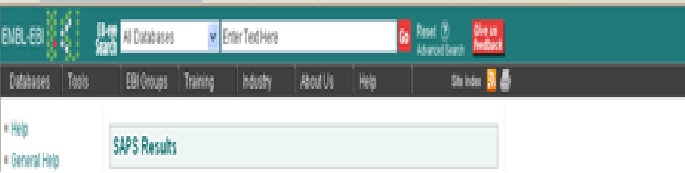
The Table15 shows the total number of cysteine,predicted number of bonds with their position was analysed to find the number of folds in the protein structure.

Total number of cysteines: 9

Predicted number of bonds: 4

Cysteines at the following positions are predicted to form the disulfide bond: 23,85,97,106,129,169,205,237

IJSER

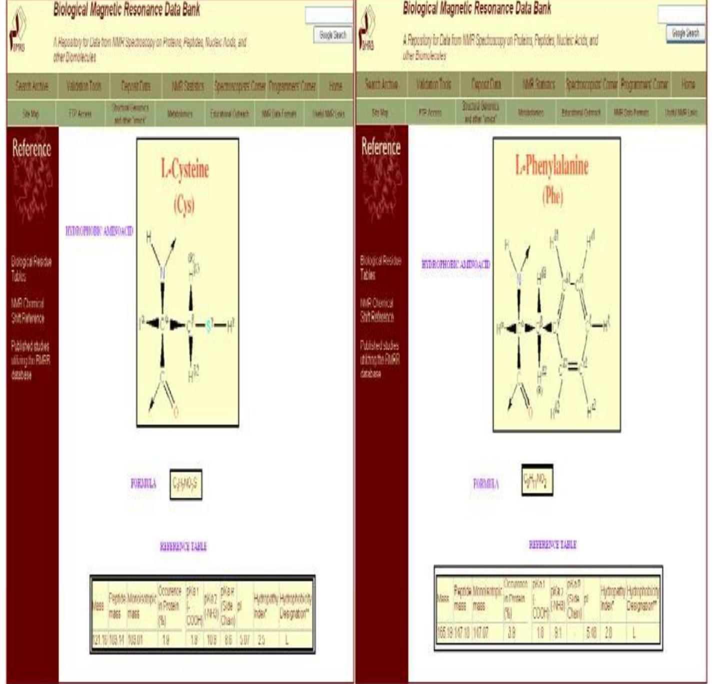


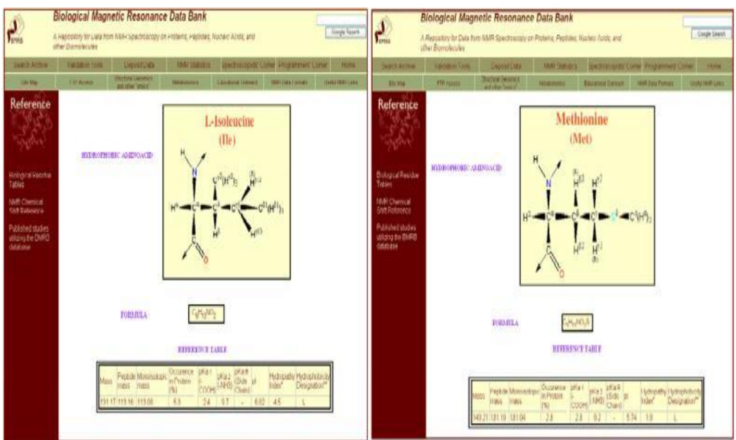
Hydrophobic aminoacids involved in Chitinase protein

High scoring aminoacid segments in protein

(LVIFM)	2.00
(AGYCW)	1.00
(PH)	-2.00
(STNQ)	-4.00
(KEDR)	-8.00

Figure 36: Hydrophobic amino acid involved in chitinase protein





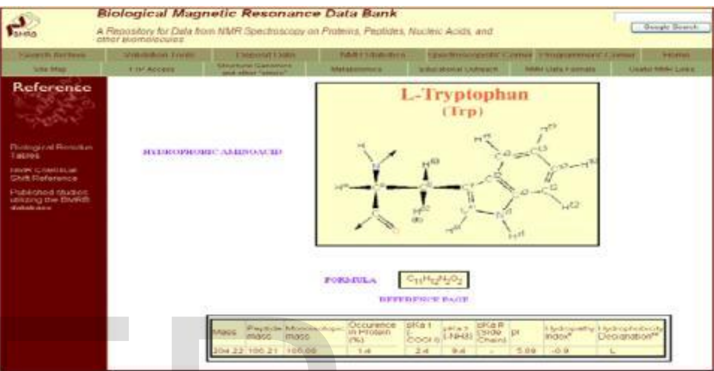


Figure 37: Side chain for non – polar aminoacid

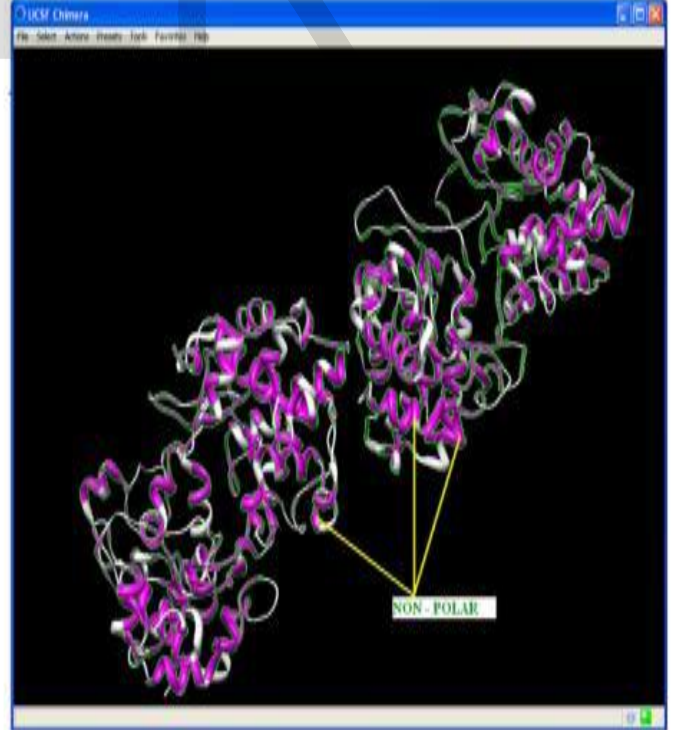


Figure 38:Non – polar region

IJSER

DiANNA 1.1 web server

Sequence **inputSeq** Length **249** residues				
Cysteines in this sequence: 9				
Cysteine Class prediction				
Cysteine	Half-cystine	Free cysteine	Ligand-bound	Ligand
23	0.136920	0.715272	0.147808	-
85	0.576235	0.382538	0.041227	-
97	0.723460	0.213258	0.063282	-
106	0.475319	0.394707	0.129974	-
129	0.240000	0.493511	0.266489	-
169	0.276237	0.512887	0.210876	-
205	0.144755	0.708736	0.146509	-
224	0.291337	0.653765	0.054897	-
237	0.267057	0.639773	0.093170	-

Introduction Instructions References About Help! Contacts

© Boston College

Figure 39:Cysteine residues with position

4.7 PROTEIN FOLDING TYPE AND CLASSIFICATION.

The secondary structure of chitinase protein was predicted.The folded region in the protein was analyzed using FOLD pro Recognition tool. It was visualised using Molsoft software. The number of folds present in the Chitinase protein structure was visualized and shown in the Figure40 and Figure41.The folded,un folded,hydrophobic and charged region were differentiated using different colurs and represented in Figure42.Ten folds were predicted using Desc fold database in Figure43 and the folds are seen using MolSoft software was shown in Figure44.Figure45 determines the total number of disulphide donds in the chitinase protein structure.The Figure46 represents the structure class of the chitinase protein.The ten fold was predicted and each fold belongs to different type was also analyzed for chitinase protein which was shown in the Figure47. Fold recognisation and the deviation in the predicted fold was analyzed.Figure48 shows the type of fold recognised with structure.The fold deviation predicted was shown in Figure49. Figure50 and Figure51 shows the Half life and functional site of chitinase protein structure.

4.7.1 protein folding and classification

The protein folding mechanism plays an one of the important role in structure stability. Total number of folds present in the chitinase protein structure was visualized using Molsoft and Yasara is shown in the Figure40 and Figue41.

The Figure42 shows the graphical representation of fold in the protein with the position and minimum energy. The Desc fold database is used to predict the total number of fold and their structure in PDB format was analyzed and shown in Figure43.The fold structure for the protein was seen in CN3D.The A chain(3 fold),B chain(3 fold),C chain(2 fold) and D chain(2 fold) was also seen in the Figure44.

In the Figure46 chitinase protein structure class was predicted using protein structure classification server. The query protein sequence is given as input and the class of the protein was found to be Alpha + Beta.

4.7.2 Protein folding types

The protein secondary structure classification server is used to classify the fold region of the protein which is based on the Support Vector Machine. The Figure47 shows type of fold predicted for the chitinase protein with the rank and structure. It was predicted that the fold present in the chitinase protein mainly belong to the type All alpha, All beta and Alpha + Beta.

4.7.3 Disulphide bonds

The ten disulphide bridges are predicted in the folded region of the chitinase protein was visualized using the Visualization software was shown in Figure45.The region shown in the form of capsules denote the folded region(Disulphide bridge formed between the cysteine residues) in the protein structure of A,B,C and chain.

Command used in rasmol: Rasmol>ssbonds on
Disulphide bridges……….10
Rasmol> ssbonds 100
Rasmol>color ssbonds red

4.7.4 Fold class recognition and deviation in fold.

The type of the fold was predicted with the architecture and the structure were also analyzed for fold type of chitinase protein their net score, Identity, p – value, alignment score, alignment length for the protein is shown in Figure48.

The predicted fold in the structure may undergo deviation due to some mutation. If the fold region present in the chitinase protein undergoes deviation the place where fold is deviated is shown by structure in graphical form was also shown in Figure49.

Table16: Fold deviation score

Individual Values	
Residue	Value
ASP1	1.27
LEU2	0.48
SER3	0.40
GLY4	0.79
ILE5	0.58
ILE6	1.50
SER7	1.00
ARG8	0.17
ARG8	0.17
ASP9	0.50
GLN10	0.32
PHE11	0.41
TYR12	0.69
LYS13	0.53
LYS13	0.53
MET14	0.39
LEU15	1.23
LYS16	0.64
HIS17	0.04
MET18	0.57
ASN19	0.85
ASP20	1.47
ASN21	0.28
ASP22	0.49
LYS23	1.91
HIS24	0.79
ALA25	1.06
VAL26	1.22
GLY27	1.01
ILE57	0.37
ALA58	0.26
ALA59	0.36
PHE60	0.29
PHE61	0.67
GLY62	0.37
PHE28	0.41
PHE29	1.02
THR30	1.65
TYR31	0.42
ASP32	0.37
ALA33	0.19
PHE34	0.21
ILE35	0.31
THR36	0.28
ALA37	0.31
ALA38	0.40
LYS39	0.53
SER40	0.40
PHE41	1.42
PRO42	0.98
SER43	0.70
PHE44	1.18
GLY45	1.70
ASN46	0.86
THR47	1.44
GLY48	2.39
ASP49	1.60
LEU50	0.37
ALA51	0.40
MET52	0.32
ARG53	0.29
LYS54	0.27
LYS55	0.31
GLU56	0.40
ILE57	0.37
ALA58	0.26
ALA59	0.36
PHE60	0.29
PHE61	0.67
GLY62	0.37
GLN63	0.36
THR64	0.46
SER65	0.46
HIS66	0.40
GLU67	0.30
THR68	1.07
THR69	1.63
GLY70	0.98
GLY71	2.03
TRP72	2.29
SER73	0.58
GLY74	1.00
ALA75	1.32
PRO76	1.32
ASP77	1.52
GLY78	2.76
ALA79	0.36
ASN80	0.01
THR81	0.80
TRP82	1.06
GLY83	1.01
TYR84	1.47
ILE90	0.28
ASP91	1.10
LYS92	0.70
SER93	0.47
ASP94	1.53
PRO95	0.70
HIS96	1.49
CYS97	1.67
ASP98	1.32
SER99	0.52
ASN100	0.88
ASN101	1.27
LEU102	0.45
GLU103	0.33
TRP104	1.32
PRO105	1.58
CYS106	1.20
ALA107	1.46
PRO108	1.57
GLY109	1.72
LYS110	1.60
PHE111	1.53
THR112	1.02
TYR113	1.57
GLY114	1.70
ARG115	1.70
GLY116	2.01
PRO117	0.56
MET118	0.83
MET119	1.66
LEU120	1.29
SER121	1.92
SER121	2.43
TRP122	0.43
ASN123	0.76
TYR124	0.63
ASN125	0.44
TYR126	0.65
GLY127	0.54
PRO128	0.43
CYS129	0.79
GLY130	0.43
ARG131	0.33
ASP132	0.64
LEU133	1.92
GLY134	1.64
LEU135	0.00
GLU136	0.44
LEU137	0.47
LEU138	0.30
LYS139	1.31
ASN140	0.57
PRO141	0.40
ASP142	0.64
VAL143	0.57
ALA144	0.96
SER145	0.68
SER146	1.07
ASP147	0.76
PRO148	0.57
VAL149	1.92
ILE150	0.46
ALA151	0.49
PHE152	0.39
LYS153	0.51
THR154	0.38
ALA155	0.23
ILE156	0.51
TRP157	0.44
PHE158	0.32
TRP159	0.45
MET160	0.82
THR161	1.54
PRO162	1.53
GLN163	1.42
ALA164	0.50
PRO165	0.94
LYS166	1.65
PRO167	1.86
SER168	1.50
CYS169	0.30
HIS170	0.25
ASP171	0.32
VAL172	0.37
ILE173	0.57
THR174	0.00
ASP175	1.48
GLN176	0.67
TRP177	1.07
GLU178	1.43
PRO179	1.65
SER180	1.49
ALA181	0.70
ALA182	0.26
ASP183	0.29
ILE184	0.34
SER185	0.32
ALA186	0.39
GLY187	1.67
ARG188	0.99
THR189	1.62
PRO190	1.56
GLY191	2.00
TYR192	0.26
GLY193	0.32
VAL194	0.60
VAL194	0.60
ILE195	0.54
THR196	0.40
ASN197	0.47
ILE198	0.38
ILE199	0.55
ASN200	1.27
GLY201	0.40
GLY202	0.77
LEU203	0.78
GLU204	0.86
CYS205	1.70
ALA206	1.34
GLY207	1.71
ARG208	0.36
ASP209	1.41
VAL210	0.33
ALA211	0.30
LYS212	0.27
VAL213	0.30
GLN214	0.39
ASP215	0.37
ARG216	0.32
ILE217	0.32
SER218	0.49
PHE219	0.51
TYR220	0.43
THR221	0.45
THR221	0.45
ARG222	0.44
TYR223	0.31
CYS224	0.29
GLY225	0.18
MET226	0.28
PHE227	0.55
GLY228	1.86
VAL229	1.78
ASP230	1.59
PRO231	0.79
GLY232	2.34
SER233	1.56
ASN234	1.42
ILE235	0.45
ILE235	0.45
ASP236	1.00

The Table16 shows the deviation score for all the amino acids present in the chitinase protein. The deviation in the structure occurs due to mutation which leads to change in the structure stability. So there will be a deviation in the chitinase protein fold mechanism.

4.7.5 Identifying PEST region and finding functional site in chitinase protein

The PESTFIND is a tool used to find the half- life of the protein and the amino acid involved .proline (P), Glutamic acid(E), Serine(S), Threonine(T) are the amino acids used to determine the half – life of the protein. Lower the number of pest region(Less than 3) increases the half – life, The number of pest region in the chitinase protein is more than 3 which indicates that the half – life of the protein is low. which can withstand not more than one day is shown in Figure50.

The pest score for the protein sequence was plotted in the graph. It was found that all the pest region lie below the threshold level which indicates that the protein has decreased half – life.

In the above graph the scores in

Maroon colour represents – invalid region
Red colour represents - poor region

The functional site was shown in Figure51.The chitinase protein have same functional site and position. In A,B,C and D chain these shows that the protein have best functional site in the particular region. The amino acid residues present in the functional site of the chitinase protein is (LAFKTAIW FW) their Description, position (295–305),chain and source was present same in all the chain of chitinase protein.

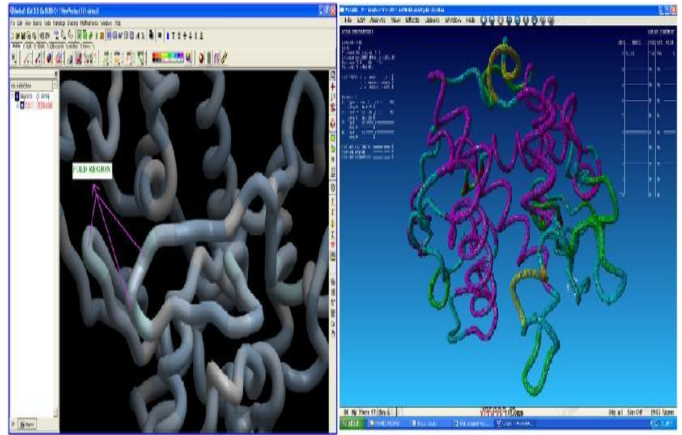
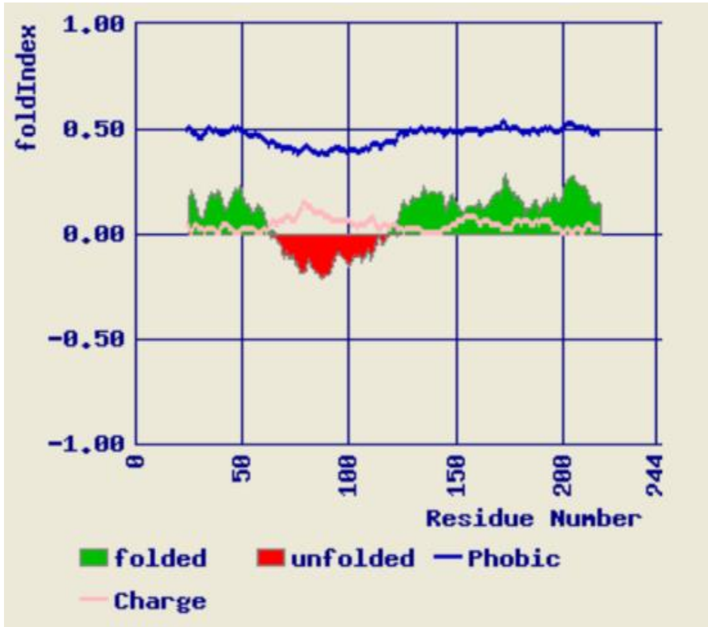


Figure 40 **Figure 41**

The image shows fold region



244 residues, unfoldability 0.076 (Charge: 0.033, Phobic: 0.

Figure 42:Graph showingFold,un folded and hydrophobic region.

#	SCOP Name	SCOP Class	Z-Score	PDB	SEQ	Motif	Psi-blast Alignment	Rps-blast Alignment	SSEA Alignment	PPA Alignment	PSPA Alignment	RSA Alignment
1	d1cnsa_	d.2.1.1	34.52	pdb	seq	motif	aln	aln	aln	aln	aln	aln
2	d1dxja_	d.2.1.1	33.95	pdb	seq	motif	aln	aln	aln	aln	aln	aln
3	d1qgia_	d.2.1.7	2.67	pdb	seq	motif	aln	aln	aln	aln	aln	aln
4	d2eqla_	d.2.1.2	2.63	pdb	seq	motif	aln	aln	aln	aln	aln	aln
5	d1el1a_	d.2.1.2	2.63	pdb	seq	motif	aln	aln	aln	aln	aln	aln
6	d1gd6a_	d.2.1.2	2.63	pdb	seq	motif	aln	aln	aln	aln	aln	aln
7	d1qsaa2	d.2.1.6	2.62	pdb	seq	motif	aln	aln	aln	aln	aln	aln
8	d1lb6a_	d.2.1.2	2.62	pdb	seq	motif	aln	aln	aln	aln	aln	aln
9	d1ivma_	d.2.1.2	2.61	pdb	seq	motif	aln	aln	aln	aln	aln	aln
10	d1hfba_	d.2.1.2	2.61	pdb	seq	motif	aln	aln	aln	aln	aln	aln

Figure 43:Identification of fold

Figure 44: Fold region in A,B,C and D chain

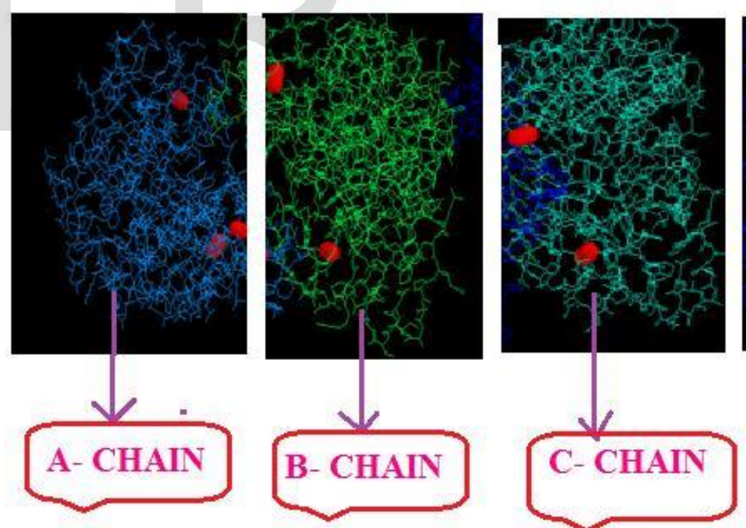


Figure 45:Number of disulphide bonds in protein

IJSER

PROTEIN STRUCTURE CLASSIFICATION SERVER

Query Sequence:

CHITINASE PROTEIN(2Z37)

DLSGIISRDQFYKMLKHMNDNDCHAVGFFTYDAFITAAKSFPSFGNTGDLAMRKKEIAA
FFGQTSHETTGGWSGAPDGANTWGYCYKEEIDKSDPHCDSNNLEWPCAPGKFYYGRGP
MMLSWNYNYGPCGRDLGLELLKNPDVASSDPVIAFKTAIWFWMTPQAPKPSCHDVITD
QWEPSAADISAGRLPGYGVITNIINGGLECAGRDVAKVQDRISFYTRYCGMFGVDPGSNI
DCDNQRPFN

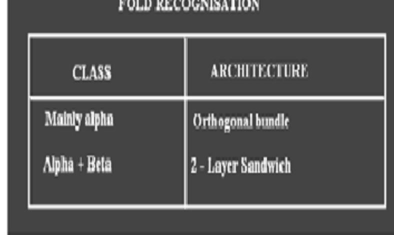
Result

Class of your protein: Alpha+Beta

Figure 46: Class of chitinase protein

Genthreader Scores

Conf.	Net Score	p-value	PairE	SolvE	Aln Score	Aln Len	Str Len	Seq Len	Alignment	SCOP Codes
CERT	154.251	1e-14	-157.4	-13.3	887.4	244	244	244	2z37A0	-

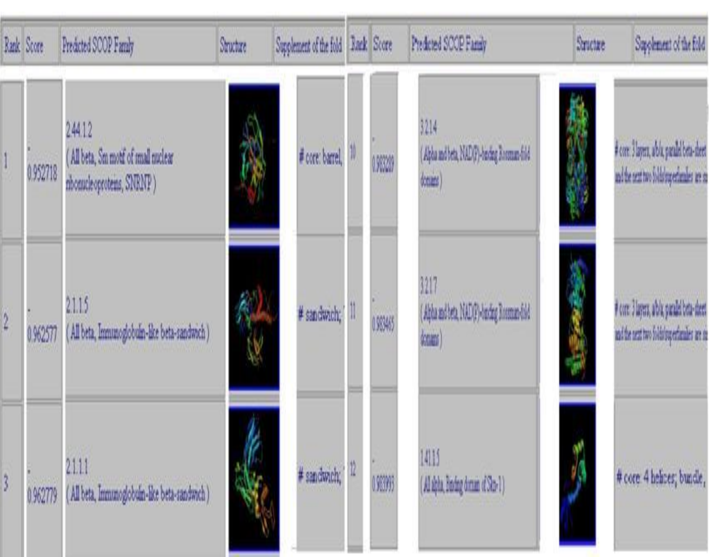


Structure

ORTHOGONAL BUNDLE

2 - LAYER SANDWICH

Figure 48:Fold recognition with type



Rank	Score	Predicted SCOP Family	Structure	Supplement of the fold	Rank	Score	Predicted SCOP Family	Structure	Supplement of the fold
1	0.952718	2.44.1.2 (All beta, Sm motif of small nuclear ribonucleoproteins, SNRNP)		# core: barrel,	10	0.983209	3.2.1.4 (Alpha and beta, NAD(P)-binding Rossmann-fold domains)		# core: 3 layers, a/b/a, parallel beta-sheet and the next two folds/superfamilies are …
2	0.962577	2.1.1.5 (All beta, Immunoglobulin-like beta-sandwich)		# sandwich;	11	0.983465	3.2.1.7 (Alpha and beta, NAD(P)-binding Rossmann-fold domains)		# core: 3 layers, a/b/a, parallel beta-sheet and the next two folds/superfamilies are …
3	0.962779	2.1.1.1 (All beta, Immunoglobulin-like beta-sandwich)		# sandwich;	12	0.983993	1.41.1.5 (All alpha, Binding domain of Skn-1)		# core: 4 helices; bundle,

Figure 47:Chitinase protein folding type

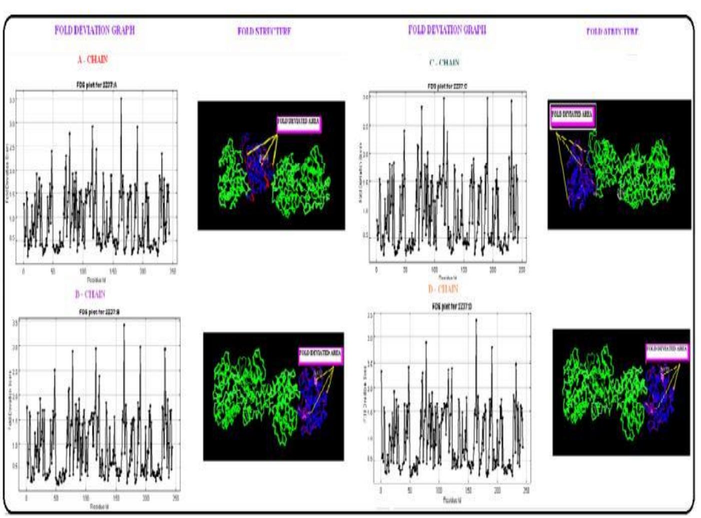


Figure 49: Fold deviation

OUTPUT FILE :

CHITINASE PROTEIN 2Z37

PEST-find: Finds PEST motifs as potential proteolytic cleavage sites.

11 PEST motifs were identified in
from positions 1 to 244 and sorted by length.

Poor PEST motif with 20 amino acids between position 66 and 87.
66 HETTGGWSGAPDGANTWGYCYK 87
PEST score: -8.86

Poor PEST motif with 19 amino acids between position 188 and 208.
188 RLPGYGVITNIINGGLECAGR 208
PEST score: -27.35

Poor PEST motif with 18 amino acids between position 222 and 241.
222 RYCGMFGVDPGSNIDCDNQR 241
PEST score: -15.51

Poor PEST motif with 17 amino acids between position 170 and 188.
170 HDVITDQWEPSAADISAGR 188
PEST score: -3.78

Invalid PEST motif with 15 amino acids between position 115 and 131.
115 RGPMMLSWNYNYGPCGR 131

Invalid PEST motif with 14 amino acids between position 24 and 39.
24 HAVGFFTYDAFITAAK 39

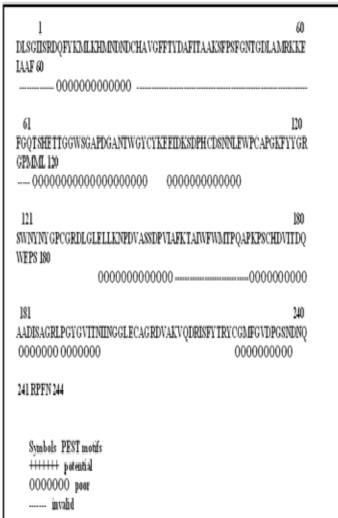
Poor PEST motif with 13 amino acids between position 39 and 53.
39 KSFPSFGNTGDLAMR 53
PEST score: -17.21

Poor PEST motif with 13 amino acids between position 96 and 110.
96 HCDSNNLEWPCAPGK 110
PEST score: -10.43

Poor PEST motif with 13 amino acids between position 139 and 153.
139 KNPDVASSDPVIAFK 153
PEST score: -12.37

Invalid PEST motif with 12 amino acids between position 153 and 166.
153 KTAIWFWMTPQAPK 166

Invalid PEST motif with 10 amino acids between position 55 and 66.
55 KEIAAFFGQTSH 66



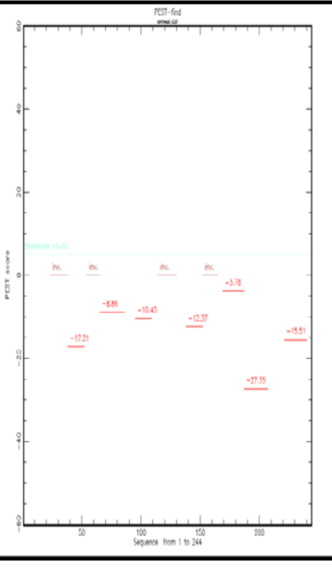


Figure 50:PEST region

PDBj eF-seek

eF-site ID	2z37-ABCD
PDB Code	2z37
Chain	A, B, C, D
Title	Crystal structure of Brassica juncea chitinase catalytic module (Bjchi3)
Classification	HYDROLASE
Compound	Chitinase
Source	Brassica juncea (Leaf mustard) (Indian mustard) (Q9SQF7_BRAJU)
Description	(1) Chitinase (E.C.3.2.1.14)

Sequence

A: DLSGIISRDQFYKMLKHMNDNDCHAVGFFTYDAFITAAKS
FPSFGNTGDLAMRKKEIAAFFGQTSHETTGGWSGAPDGAN
TWGYCYKEEIDKSDPHCDSNNLEWPCAPGKFYYGRGPMML
SWNYNYGPCGRDLGLELLKNPDVASSDPVIAFKTAIWFWM
TPQAPKPSCHDVITDQWEPSAADISAGRLPGYGVITNIIN
GGLECAGRDVAKVQDRISFYTRYCGMFGVDPGSNIDCDNQ
RPFN

B: DLSGIISRDQFYKMLKHMNDNDCHAVGFFTYDAFITAAKS
FPSFGNTGDLAMRKKEIAAFFGQTSHETTGGWSGAPDGAN
TWGYCYKEEIDKSDPHCDSNNLEWPCAPGKFYYGRGPMML
SWNYNYGPCGRDLGLELLKNPDVASSDPVIAFKTAIWFWM
TPQAPKPSCHDVITDQWEPSAADISAGRLPGYGVITNIIN
GGLECAGRDVAKVQDRISFYTRYCGMFGVDPGSNIDCDNQ
RPFN

C: DLSGIISRDQFYKMLKHMNDNDCHAVGFFTYDAFITAAKS
FPSFGNTGDLAMRKKEIAAFFGQTSHETTGGWSGAPDGAN
TWGYCYKEEIDKSDPHCDSNNLEWPCAPGKFYYGRGPMML
SWNYNYGPCGRDLGLELLKNPDVASSDPVIAFKTAIWFWM
TPQAPKPSCHDVITDQWEPSAADISAGRLPGYGVITNIIN
GGLECAGRDVAKVQDRISFYTRYCGMFGVDPGSNIDCDNQ
RPFN

D: DLSGIISRDQFYKMLKHMNDNDCHAVGFFTYDAFITAAKS
FPSFGNTGDLAMRKKEIAAFFGQTSHETTGGWSGAPDGAN
TWGYCYKEEIDKSDPHCDSNNLEWPCAPGKFYYGRGPMML
SWNYNYGPCGRDLGLELLKNPDVASSDPVIAFKTAIWFWM
TPQAPKPSCHDVITDQWEPSAADISAGRLPGYGVITNIIN
GGLECAGRDVAKVQDRISFYTRYCGMFGVDPGSNIDCDNQ
RPFN

Functional site

1)	chain	A
	residue	295-305
	type	prosite
	ligand	
	sequence	IAFKTAIWFWM
	description	Chitinases family 19 signature 2. [LIVM]-[GSA]-F-x-[STAG](2)-[LIVMFY]-W-[FY]-W-[LIVM]
	source	prosite : PS00774
2)	chain	B
	residue	295-305
	type	prosite
	ligand	
	sequence	IAFKTAIWFWM
	description	Chitinases family 19 signature 2. [LIVM]-[GSA]-F-x-[STAG](2)-[LIVMFY]-W-[FY]-W-[LIVM]
	source	prosite : PS00774
3)	chain	C
	residue	295-305
	type	prosite
	ligand	
	sequence	IAFKTAIWFWM
	description	Chitinases family 19 signature 2. [LIVM]-[GSA]-F-x-[STAG](2)-[LIVMFY]-W-[FY]-W-[LIVM]
	source	prosite : PS00774
4)	chain	D
	residue	295-305
	type	prosite
	ligand	
	sequence	IAFKTAIWFWM
	description	Chitinases family 19 signature 2. [LIVM]-[GSA]-F-x-[STAG](2)-[LIVMFY]-W-[FY]-W-[LIVM]
	source	prosite : PS00774

Figure 51:Functional site of chitinase protein

4.8 HELIX STABILITY AND MUTATION

IJSER

Geneious Software is used to predict the entire helical region in the protein.From the entire helical region a single helix was viewed with their aminoacid residues and it's hydrophobicity was analyzed . The helical region present in the chitinase protein was predicted and shown in Figure52.The total number of alpha helix was predicted with position,length and percentage for the protein was shown in Table17 and Figure53.The amino acid residues which makes the helix structure stable was analyzed and shown in Figure54.The mutation in the sequence affects the protein structure stability.The single residue which affects the structure was shown in Figure55 and Figure56 for the chitinase protein.

4.8.1 Identifying amino acid residues which makes the helix structure stable

Identification of stabilizing residues in protein is the server which is used to find the amino acid residues responsible for the stability of the secondary structure of protein was shown in Figure54.

Alanine, Glycine, Tyrosine, Asparagine, Phenylalanine are the amino acids involved to make the protein helix structure stable in the chitinase protein. Even a single change in the above amino acid residue**(A,G,T, F,N)** affects the **stability of helical region in the chitinase protein**.

Table17: Predicted Alpha helix

The following alpha helices were predicted:

Helix No	Alpha Helix	Position	Length
1	RDOFYKMLKH	8 to 17	10
2	YDAFITAA	31 to 38	8
3	DLAMRKKEIAAF	49 to 60	12
4	VAKVQ	210 to 214	5

The following beta strands were predicted:

Strand no.	Beta Strand	Position	Length
1	AVGFF	25 to 29	5
2	GYCYK	83 to 87	5
3	MMLSW	118 to 122	5
4	IAFK	150 to 153	4
5	AIWFW	155 to 159	5
6	VITNII	194 to 199	6
7	ISFYT	217 to 221	5

The Consensus secondary structure prediction server was used to predict the percentage of helix, strand, their position, amino acid residue involved and length. It is also used to predict the best helix and strand region present in the chitinase protein structure was represented in Table17.

4.8.2 Mutation changing the protein structure stability.

MUpro is the server used to analyze the stability of protein .It was found that a single change in the amino acid residue affects the structure and decreases the stability of protein .Hence the energy gets increased. Using MuPro server instead of G aminoacid A is substituted which decreases the stability of the protein. From this even a single mutation change in the structure affects the stability of helical region in the chitinase protein structure was shown in Figure55.

The Figure56 shows the protein stability and their changes in Helix structure during mutation.Using this server, when there is a change in aminoacid residue and stability in the secondary structure.The stability of helix region is decreased and increse in energy leads to change in the pH,Reliability index and temperature. By these even a single mutation affects helix stability in chitinase protein structure .

C

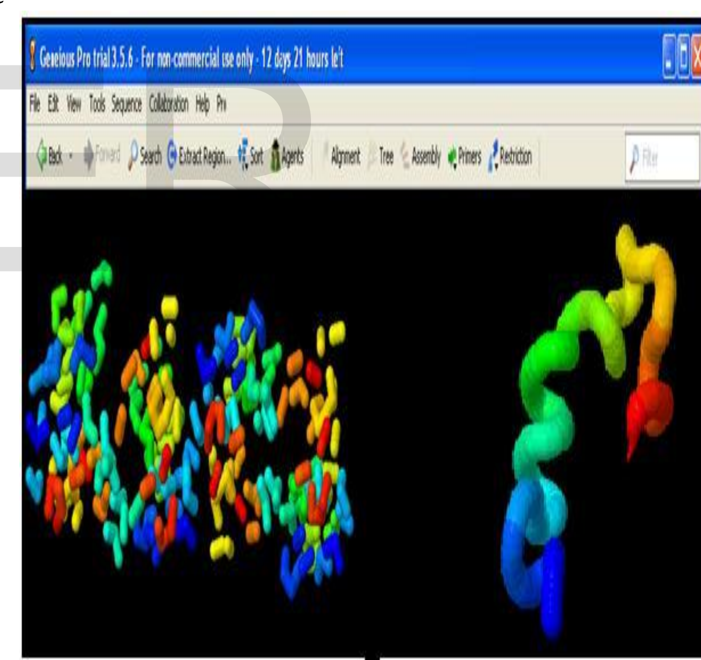


Figure 52:Helical region in chitinase protein

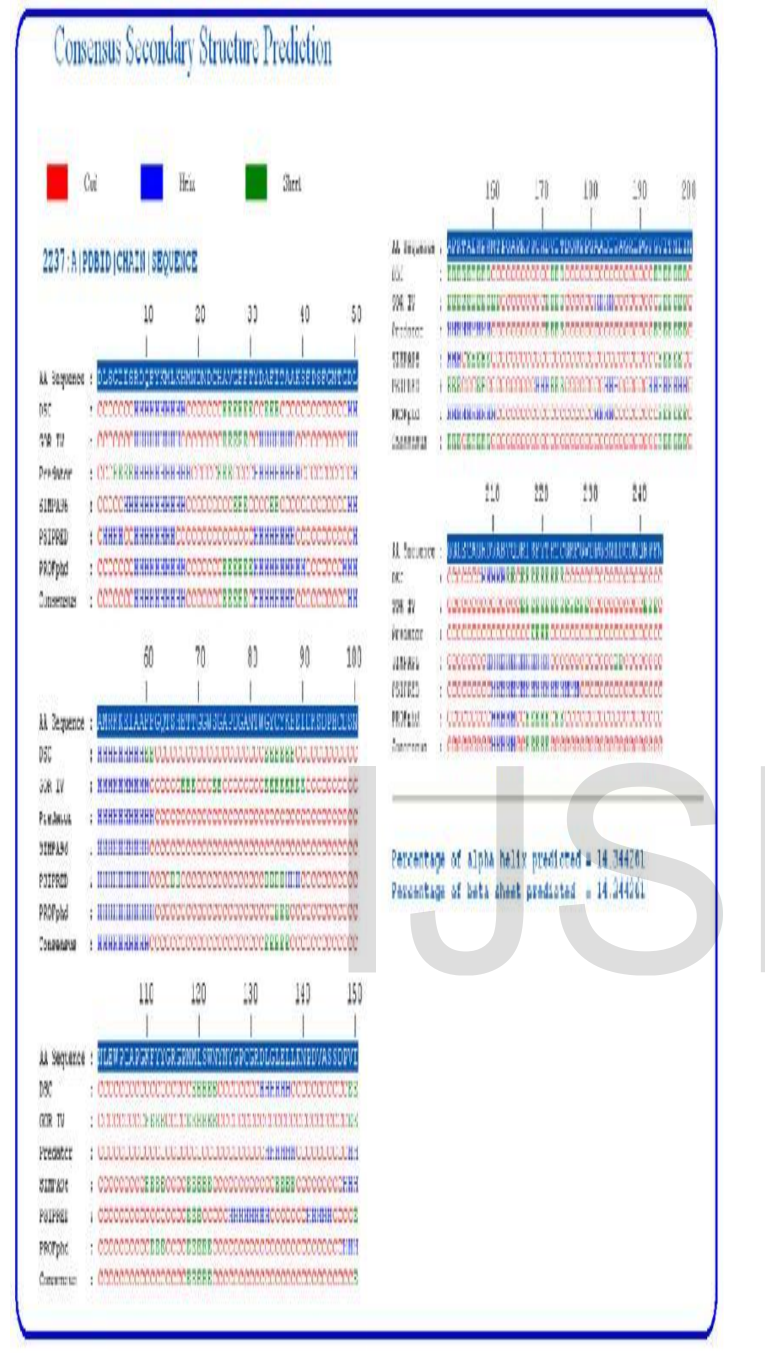


Figure 53:Identifing total number of helix

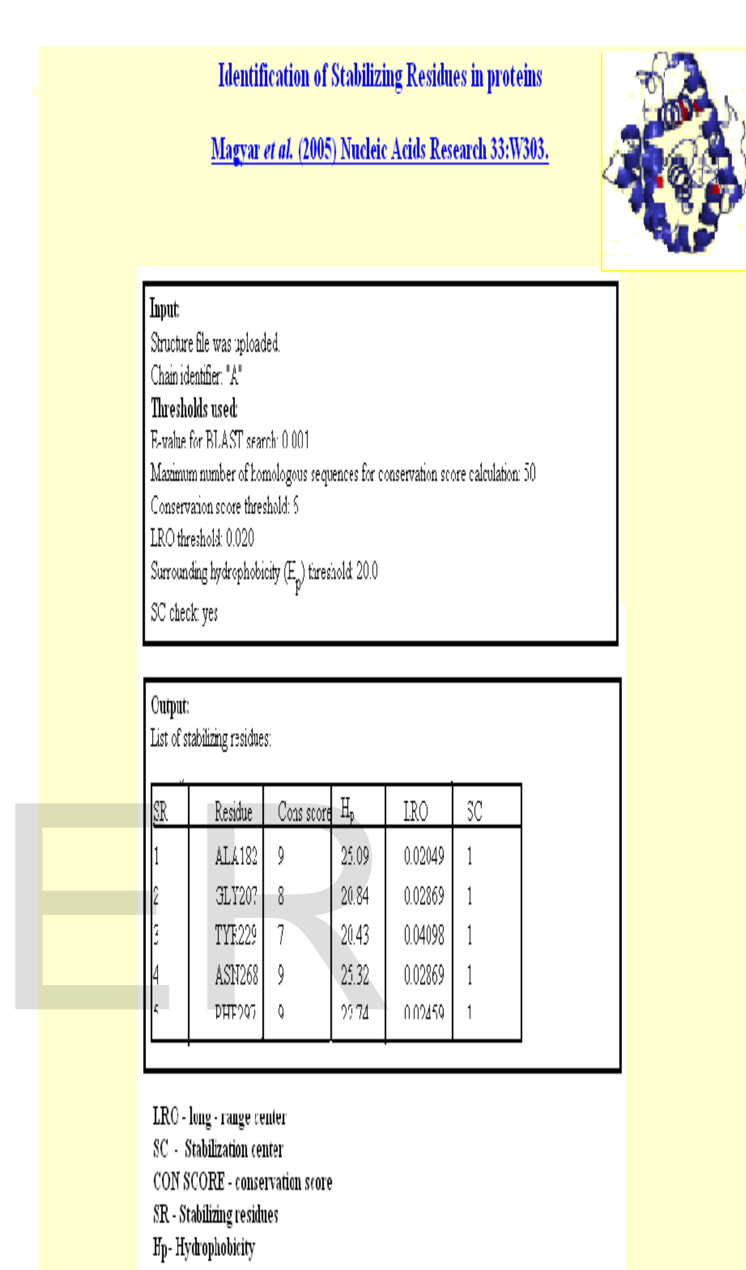
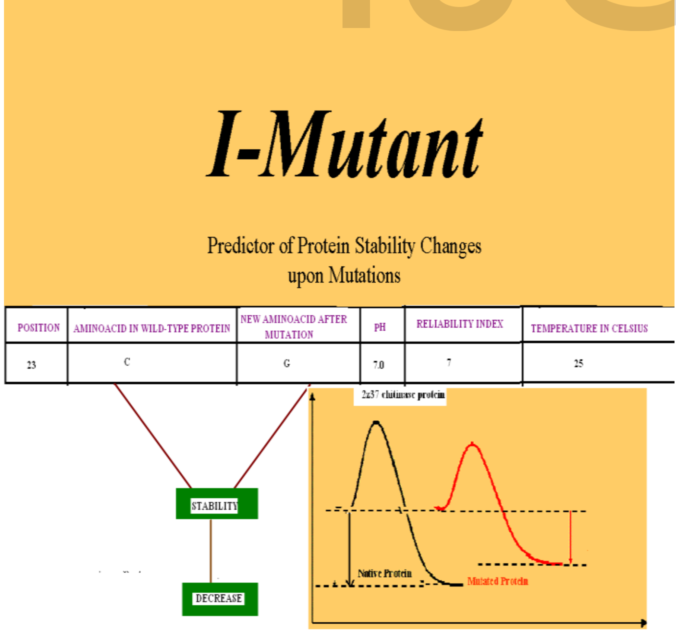


Figure 54:Amino acid which makes the structure stable

IJSER

Figure 55: Mutation



POSITION	AMINOACID IN WILD-TYPE PROTEIN	NEW AMINOACID AFTER MUTATION	PH	RELIABILITY INDEX	TEMPERATURE IN CELSIUS
23	C	G	7.0	7	25

Figure 56: Mutation and change in native protein

4.9 INTERPRETATION OF FORCE FIELD

The force field was found using ANOLEA and GROMOS present in the SWISS – MODEL work space.In the above plot favourable and unfavourable regions are pedicted in two different colours.The force field with the score was shown in the Figure57.

GREEN- Favourable region energy values which makes structure stable)

RED - Un favourable region(Positive values protein unstable)

Force field can also be calculated using Swiss – Pdb Viewer.The PDB structure for the chitinase protein is loaded in the workspace.The force field option is selected from the tools icon.Force field settings will be displayed and the energy to be calculated is selected from the settings toolbox.The predicted force field energy was represented in Figure58.

The DBCP server was used to find the disulphide bridges present in the folded region of chitinase protein with their aminacid residue,position,distance,Probability and Metal binding site was found.The force field was used to find energy and stability of protein secondary structure based on fold.**The best predicted disulphidebond is at the position [23 – 85] [97 – 106] [205 – 237].As the maximum probability score (0.98227,0.99541,0.60965)and distance (5.83,10.81,4.30) for the bonds are at the position [23 – 85] [97 – 106] [205 – 237]** .Hence the structure is more stable at these position because of protein folding.The helix structure is stable at these position.These was shown in the Figure59.

The Table18 shows the aminoacid residues involved in folding of Chitinase protein.The disulphide bond present in fold region with distance and forcefield energy was also calculated using SPBDV.It was found that the protein has the best fold region in the cysteine residue at the **position 230** in all the chains.It was predicted that the energy was least in **A – Chain** than in other chain.

The plant Brassica juncea is grown commonly in the Andhaman and nicobar island.The growth and development of the plant is affected by fungal disease.During this attack chitinase protein is expressed to prevent the plant from disease. The brassica plant is protected by protein only for two days because the half life of the protein is decreased and energy increased.Due to less half life, the chitinase protein can withstand in the disease affected plant not more than two days.So the fungus present in the plant spreads slowly and affects the growth(Third day).To prevent these,the Stability,hydrophobic amino acids involved in the core of the protein,mutation,folding mechanism of the chitinase protein was predicted.This work was done mainly to manufacture pesticide and bio fertilizers from the protein chitinase which helps the plant in their growth and development.

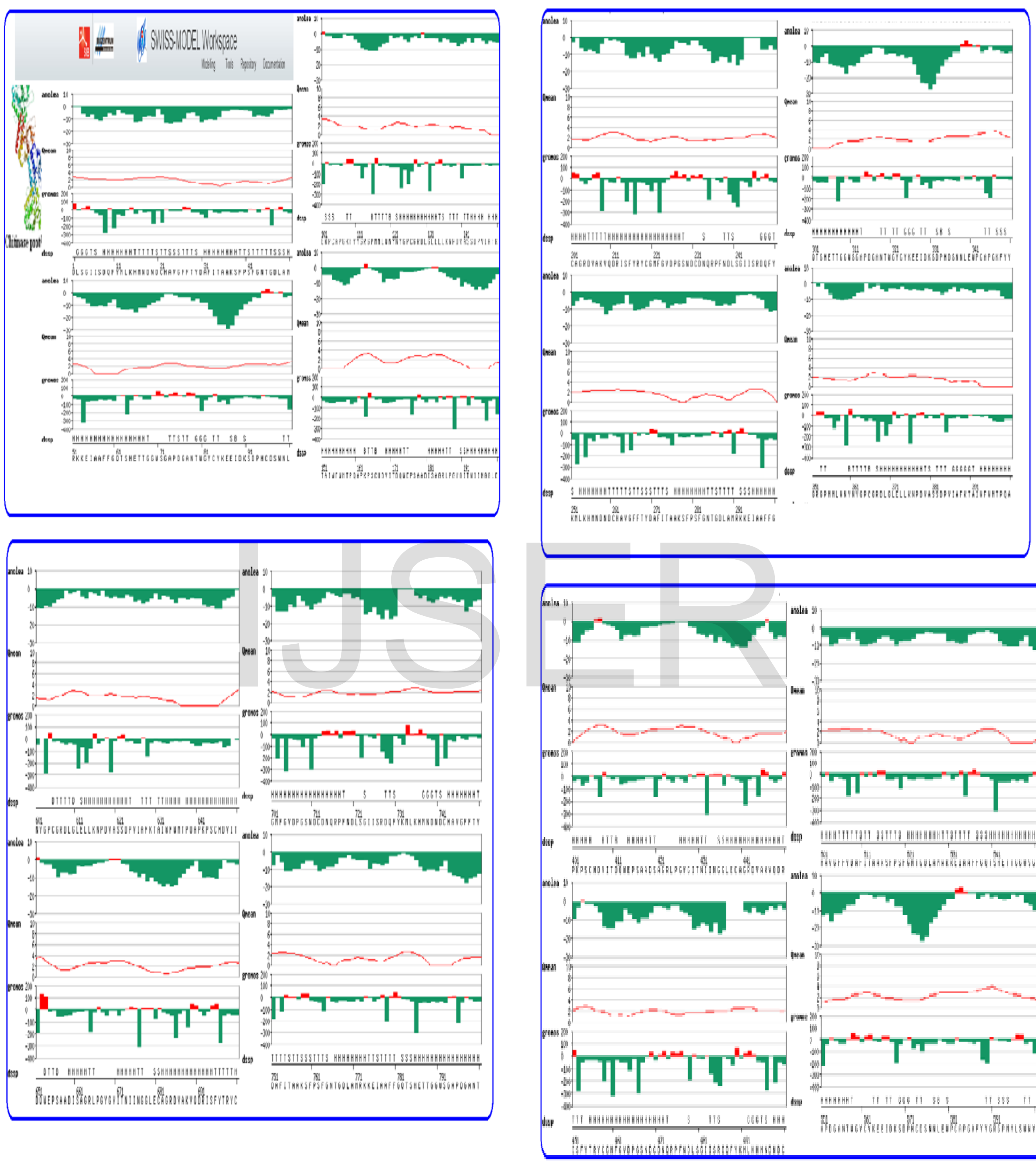


Figure 57:Force field calculation

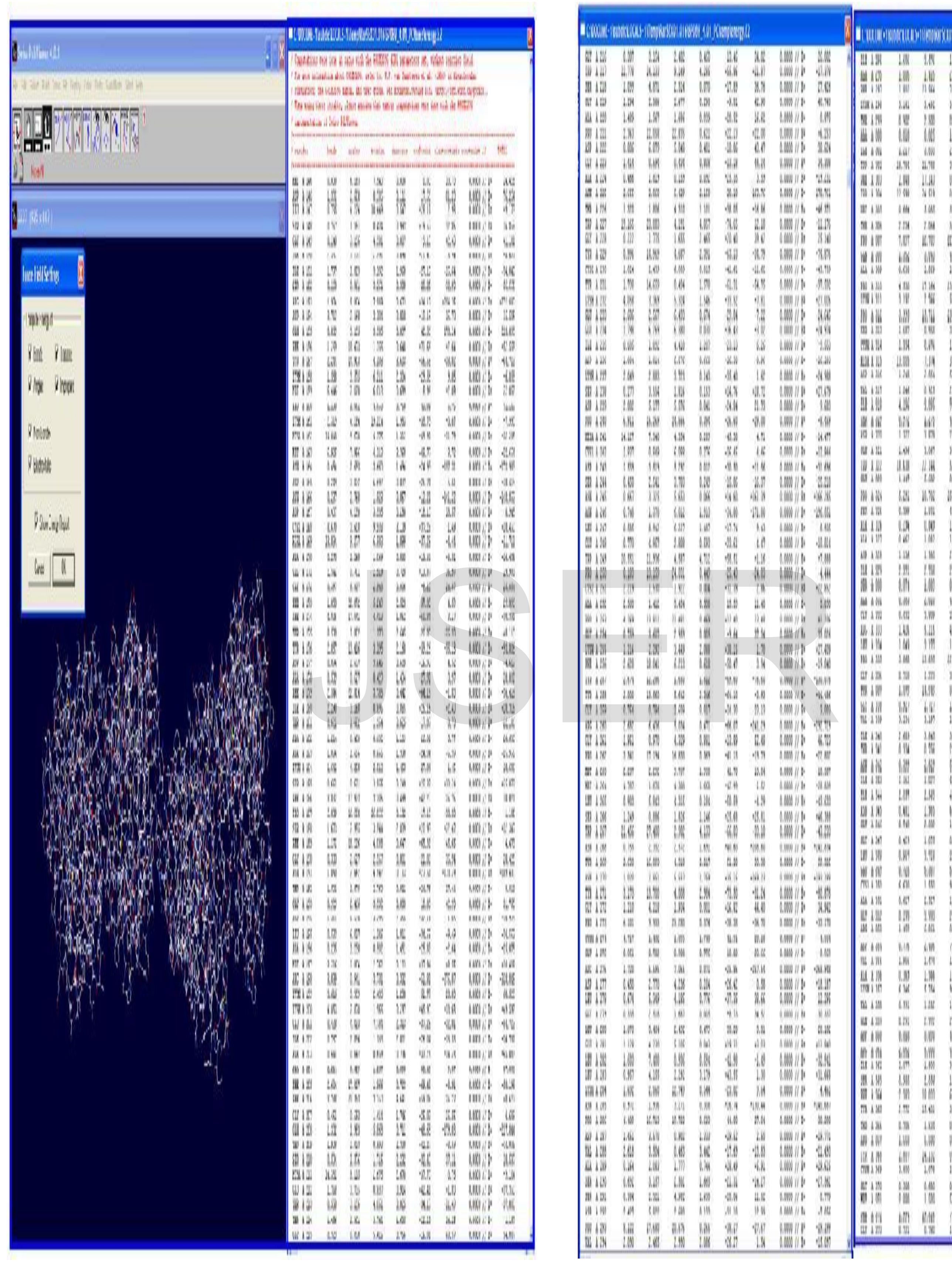


Figure 58: Force field energy and score

DBCP: A web server for Disulfide Bonding Connectivity Pattern prediction without the prior knowledge of the bonding state of cysteine

Prediction Result

Job ID	79844562	Query Name	
Template PDB ID	2Z37A		
SEQUENCE	DLSGIISRDQFYKMLKHMNDNDCHAVGFFTYDAFITAAKSFPSFGNTGDLAMRKKEIAAF FGQTSHETTGGWSGAPDGANTWGYCYKEEIDKSDPHCDSNNLEWPCAPGKFYYGRGPMML SWNYNYGPCGRDLGLELLKNPDVASSDPVIAFKTAIWFWMTPQAPKPSCHDVITDQWEPS AADISAGRLPGYGVITNIINGGLECAGRDVAKVQDRISFYTRYCGMFGVDPGSNIDCDNQ RPFN Sequence length: **244** residues		
Positions of cysteines	23 85 97 106 129 169 205 224 237		
e-value	7e-146	Sequence identity	100.00%

Disulfide bond connectivity prediction score

Position	Cysteines pair Euclidean distance	Separation	Probability	Metal-binding site score
23 - 85	5.43	62	0.98227	0.12911
23 - 97	28.49	74	0.00010	0.00010
23 - 106	27.11	83	0.00010	0.00010
23 - 129	25.36	106	0.00010	0.00010
23 - 169	28.69	146	0.00010	0.00010
23 - 205	31.82	182	0.00010	0.00010
23 - 224	26.72	201	0.00010	0.00010
23 - 237	31.54	214	0.00010	0.00010
85 - 97	24.61	12	0.00010	0.00010
85 - 106	24.54	21	0.00010	0.00010
85 - 129	21.65	44	0.00010	0.00010
85 - 169	23.75	84	0.00010	0.00010
85 - 205	27.01	120	0.00010	0.00010
85 - 224	23.94	139	0.00010	0.00010
85 - 237	26.86	152	0.00010	0.00010
97 - 106	10.81	9	0.99541	0.18415
97 - 129	16.57	32	0.00010	0.00010
97 - 169	27.18	72	0.00010	0.00010
97 - 205	26.40	108	0.00010	0.00010
97 - 224	40.58	127	0.00010	0.00010
97 - 237	29.53	140	0.00010	0.00010
106 - 129	13.32	23	0.00474	0.21989
106 - 169	30.44	63	0.00010	0.00010
106 - 205	35.19	99	0.00010	0.00010
106 - 224	43.59	118	0.00010	0.00010
106 - 237	37.63	131	0.00010	0.00010
129 - 169	19.34	40	0.00010	0.00010
129 - 205	31.70	76	0.00010	0.00010
129 - 224	34.62	95	0.00010	0.00010
129 - 237	33.00	108	0.00010	0.00010
169 - 205	21.93	36	0.00010	0.00010
169 - 224	20.15	55	0.00010	0.00010
169 - 237	21.07	68	0.04944	0.21990
205 - 224	26.13	19	0.00010	0.00010
205 - 237	4.30	32	0.60965	0.17598
224 - 237	22.62	13	0.00010	0.00010

Positions of oxidized cysteines	23 85 97 106 205 237
Predicted disulfide bonds	[23-85] [97-106] [205-237]
Predicted positions of cysteines in metal binding site	None
Note	The color for Probability in cysteines pair indicates the pair is predicted to have a disulfide bond. The color for Metal-binding site score in cysteines pair indicates that the pair is probable in metal-binding site.

IJSER

Figure 59:Finding disulphide bond with distance and Energy.

5. SUMMARY AND CONCLUSION

Brassica juncea chitinase is an endo-acting, pathogenesis-related protein that is classified into glycoside hydrolase family 19, with highest homology (50–60%) in its catalytic domain to class I plant chitinases. the active-site residues are identified as Glu212 and Glu234. Glu212 is believed to be the catalytic acid in the reaction, whereas Glu234 is thought to have a dual role both activating a water molecule in its attack on the anomeric carbon, and stabilizing the charged intermediate, Chitin, an insoluble homopolymer of β-(1→4)-linked units of N-acetylglucosamine, is one of the most abundant polysaccharides on earth. It is found, for example, in the exoskeletons of insects and in fungal cell walls. Chitinases (EC 3.2.1.14) catalyze the random cleavage of internal glycosidic linkages in chitin chains. In insects, the primary roles of chitinases seem to lie in growth and morphogenesis, because they can break down and so enable rebuilding of chitin-containing structures, as reviewed earlier . In other organisms such as fungi, bacteria and animals, chitinases appear to have their most important functions in catabolism and development. In plants, these

enzymes represent the single largest group of pathogenesis-related proteins; they are strongly upregulated under pathogen (insect and fungal) attack and in response to other stresses (heat, cold, etc.), and so are thought to represent a vital part of plant defense mechanisms .

The Hydrophobic,Hydrophilic aminoacid residues their charged side chain with position , Motifs,Type of motif and aminoacid residues involved,Protein folding type and classification, Fold deviation,Half – life,Functional site of protein,Mutation and Helix stability was analyzed for chitinase protein taken from the plant Brassica juncea using different tools and softwares.During these studies the following results were found.The best motif region present in the chitinase protein was predicted as (IAKFTAIWFWM).The type of motif with their position was also analyzed.The Hydrophobic and Hydrophilic aminoacid residues plays an important role in helix stability.Ten folded region were found in protein structure and the type of fold is predicted as Alpha+Beta class.The fold predicted is based on the disulphidebond present in the structure.The fold deviation also occurs in the chitinase protein and scores for the deviated fold was analyzed for all the aminoacid residues present in the protein sequence.The disadvantage of chitinase protein is that it has very low half – life because of more than three pest sites in the structure.The half – life of Chitinase protein is decreased and it can not withstand not more than two days.But the chains A,B,C,and D in the Chitinase protein structure have same functional site.The mutation also affects the stability of the structure .The mutation mainly occurs by the change in single aminoacid residue in the protein sequence.The Ala,Gly,Asn,Tyr,Phe are the aminoacids involved in the chitinase protein and makes the protein secondary (helix) structure stable .Force field energy is calculated to find the energy for both the favourable and unfavorable region in the chitinase protein structure.The Cysteine residue at the position 230 has maximum energy in A,B,C and D chain which occurs in favourable region of the chitinase protein this shows that the structure have best fold region at that position.The stability is also more in that region of protein structure.

FUTURE RECOMMENDATIONS

◎ The brassica juncea is a plant which grows large amount in Andhaman and nicobar Islands and most of the varities of Brassica Juncea grown here are not found in other parts of the country.

◎ The disease which affects this plant is mainly due to the fungus and bacteria and main symptoms of the disease affected plant is holes in the leaves and black colour formation leads to blackleg disease caused by the fungus pythium species.

◎ Black Leg can attack at any stage of plant and affects development.When the symptoms appear it is usually too late and any healthy roots and leaves in the plant are dead.

◎ Chemical applications of Subdue MAXX, Banrot are effective in the prevention of blackleg disease but they affect the growth of the plant.The role of chitinase protein in the plant Brassica Juncea is that it helps in the growth and development of the plant and prevents the plant from fungal attack.

◎ The results from the above studies can be used in the future to find the pesticide to prevent the fungal attack and the Bio – fertilizers can be manufactured using chitinase protein.

BIBLIOGRAPHY

- Duke J A and Wain K K (1981). Medicinal plants of the world, *“Journal of Human Genetics*” 3 : 234- 237.
- Knowles P F, Kearney T E and Cohen (1981). Species of rapeseed and mustard as oil crops in California , “*AOCS Monograph American Oil Chemists' Society Champaign*” 7: 255–268.
- Leung A Y(1980). Encyclopedia of common natural ingredients used in food, drugs, “*Journal of Pharmacological uses in plants*” New York. 2: 23 – 27.
- Maity P K, Sengupta A K , and Jana P K (1980). Response of mustard variety varuna (*Brassica juncea*) to levels of irrigation and nitrogen *Indian Agriculturist FCAB*”. 24(1):43–47.
- Patel J R Parmar and Patel J C (1980). Effect of different sowing dates, spacings, and plant populations on yield of mustard,”*Journal of microbial technology*” 25(3): 526–527.
- Perry L M (1989). Medicinal plants of east and southeast Asia,*“MIT Press Cambridge”* 4:20 –22.
- Elizabeth Jeffery (2005). Maximizing The Anti-Cancer Power Of Broccoli,”*Science Daily, University of Cambridge” 22(4): 56-59.*
- Anamika S S and M H Fulekar (2009). Phytoremediation of cadmium, lead and zinc by *Brassica juncea,” Journal for Application in Bioscience*” 13:726–736.
- Begum F *et al* (1995). Somatic hybrids between *Brassica juncea* (L) Czern and *Diplotaxis harra* (Forsk) Boiss and the generation of backcross progenies,”*Theoretical. Applications in Genetics”* 91:1167–1172.

IJSER

- Choudhary B R. *et al* (2008). Cytogenetics of *Brassica juncea* × *Brassica rapa* hybrids and patterns of variation in the hybrids derivative, "*Journal for medicinal uses in plant (New York)"* 121:292–296.
- Chrungu B *et al* (1999). Production and characterization of interspecific hybrids between *Brassica maurorum* and crop brassicas. "*Theoretical Applications in Genetics*" 98:608–613.
- Duke J A *et al* (2002). Medicinal uses of plants,"*Handbook of medicinal herbs*" (CRC MedHerbs) 2: 23 – 28.
- Erickson L R *et al* (1983). Restriction patterns reveal origins of chloroplast genomes in *Brassica* amphiploids,"*Journal for Medicinal plant and uses*" 65:201–206.
- FitzJohn R G *et al* (2007). Hybridisation with in *Brassica* and allied genera: evaluation of potential for transgene escape, "*Euphytica*" 158:209–230.
- Gladis T K and Hammer (1992). Die Gaterslebener *Brassica*-Kollektion - *Brassica juncea*, *B. napus*, *B. nigra* und *B. rapa*," *Feddes Repert*" 103:469–507.
- Gleason H A and A Cronquist (1991). Manual of vascular plants, "*United States and adjacent Canada*"(Glea Cron ed2) 34 – 38.
- Halldén, C *et al* (1987). Distribution and evolution of a tandemly repeated DNA sequence in the family Brassicaceae, "*Journal of Molecular Evolution*" 25:318–323.
- Kunakh V A *et al* (2008). Mixoploidy in wild and cultivated species of Cruciferae capable of hybridizing with rapeseed *Brassica napus*, "*Cytology & Genetics*" 42:204–209.
- Lefol E *et al* (1997). Sexual hybridisation in crosses of cultivated *Brassica* species with the crucifers Erucastrum gallicum and Raphanus raphanistrum,"*potential for gene introgression Euphytica*" 95:127–139.
- Moffat A S (1995). Plants proving their worth in toxic metal cleanup,"*Journal of plant Science*" 4:269-300.
- Plieske J and Struss (2001). STS markers linked to Phoma resistance genes of the Brassica B-genome revealed sequence homology between Brassica nigra and Brassica napus , "*Theoretical Application in Genetics*" 102:483–488.
- Velasco L *et al*(1998).Variability for the fatty acid composition of the seed oil in a germplasm collection of the genus *Brassica*,"*Genetical Resources of plant Crop*" 45:371–382.
- Alberts Bruce and Alexander Johnson (2002). The Shape and Structure of Proteins," *Molecular Biology of the Cell*" Garland Science. 3: 23 – 26.
- Robson B and Vaithilingham A (2008). Protein Folding Revisited. Progress in Molecular Biology and Translational Science, "*Molecular Biology of Protein Folding*"84:161-202.
- Chiti F and Dobson C (2006). Protein misfolding, functional amyloid, and human disease, " *Annual review of biochemistry*" 75: 333–366.
- Berg B Ellis and J Dobson (1999). Effects of macromolecular crowding on protein folding and aggregation,"*The EMBO Journal*"18: 6927-6933 (1999).
- Lienqueo M E and Mahn A V (2002). Mathematical Correlations for Predicting Protein Retention Time in Hydrophobic Interaction Chromatography, "*Journal of Chromatograph*" 978:71 – 79.
- J Kyte and R Doolite (1982). A Simple Method for Displaying the Hydropathic Character of a Protein, "*Journal of Molecular Biology*" 157 : 105-132.