IF : 3.62 | IC Value 70.36

Volume-5, Issue-5, May - 2016 • ISSN No 2277 - 8160

Research Paper

Microbiology

Insilico Interaction Study on Dispase Enzyme With Fibronectic Protein And Molecular Dynamic Study for Functional Stability

Geethanjali Murugesan | R & D Centre, Bharathiar University, Coimbatore – 641 046, TN

* Dr.N.Hema Shenpagam | Associate Professor, PG and Research Department of Microbiology Hindusthan College of Arts and Science, Avinashi Road, Coimbatore -641 028, TN * Corresponding Author

ABSTRACT

Aspergillus niger (A.niger) is a filamentous fungi and an essential microorganism used in the field of biotechnology. A.niger is cultured for industrial production of many substances. This study aimed at producing neutral protease (Dispase) from fungal strain, isolated from 50-75 agricultural soil sources obtained in and around Coimbatore district. The fungal strains were isolated by means of Sabouraud Dextrose Agar(SDA), Potato Dextrose Agar(PDA) and Czapek Dox Agar methods. Lactophenol Cotton Blue (LCB) staining method was used to identify the fungal colony as Aspergillus sp. Species name was identified as Aspergillus niger by using internal transcribed spacer nucleotide sequence alignment method. Groundnut seed oil cake, cotton seed oil cake, rice straw, wheat bran, rice bran from different varieties of rice like ponni, jaya, jyothi, MO5 and ASD 17 and sugarcane bagasse were used as substrates for dispase production and their effects on incubation time, pH of the medium and temperature were observed. Dispase was purified through ammonium sulfate precipitation and ion-exchange chromatography. Its molecular weight was estimated to be 48 kDa by sodium dodecyl sulphate polyacrylamide gel electrophoresis (SDS–PAGE) and further characterization was carried out on Fourier transform infrared spectroscopy (FTIR) and Nuclear magnetic resonance (NMR) spectroscopy. Docking and simulation techniques were carried out in silico to find interaction between fibronectin and dispase and to check the stability of the protein complex.

KEYWORDS : Aspergillus niger, Neutral Protease, Dispase, solid state fermentation, molecular docking, simulation, agro wastes

Introduction

Fungi: Fungi are found in all natural and artificial biotopes and can use highly diverse carbon source. They play a major role in the global carbon cycle by decomposing plant biomass and this biomass is the main carbon source for many fungi. The number of fungal species has been estimated at 1.5 million (Lutzoni *et al.,* 2004). The major groups which are recognized within fungi are Chytridiomycota, Zygomycota and Glomeromycota. Most species identified so far belong to the latter two phyla. Fungi can be unicellular (yeast) or multicellular (filamentous fungi) or dimorphic (both yeast and filamentous vegetative growth stage). Filamentous fungi grow by means of hyphae. These hyphae extend at their tips and branch subapically, forming an intricate network (Boyce and Andrianopoulos, 2006) Fungi use plant polysaccharides as their main carbon source but cannot take these polysaccharides up into the cell. They require a wide range of extracellular enzyme activities to degrade them into their monomeric components. The degraded monomers are taken up by the cell and are converted by a variety of carbon catabolic pathways to the compounds needed by the fungal cell for growth and reproduction (Countinho *et al.,* 2009).

Aspergillus sp: Aspergillus species are organisms of choice for enzyme production for pretreatment of plant material because they have high levels of protein secretion and they produce a wide range of enzymes for plant polysaccharide degradation. In nature, *Aspergillus* degrades the polysaccharides to obtain monomeric sugars that can serve as a carbon source (Srividya Sivakumar, 2012). Therefore, *Aspergillus* uses a variety of catabolic pathways to efficiently convert all the monomeric components (D-glucose, D-xylose, L-arabinose, D-galactose, D-mannose, L-rhamnose and D-galacturonic acid) of plant biomass. Several *Aspergillus* species are extensively used in industry because of their capability to produce a wide range of enzymes including those that degrade a variety of plant polysaccharides (Ferreira de Olivera *et al.,* 2010). Enzymes produced by *A.niger*, such as citric acid, amylases, lipases, cellulases and proteases are considered as Generally Recognized as Safe(GRAS) by United States Food and Drug Administration under the Federal Food, Drug and Cosmetic Act.

In silico **studies:** *In silico* studies are carried out to analyze protein-protein interactions. Protein-protein interactions play a key role in predicting the protein function of the target protein and drug ability of the molecules. While *in vitro* and *in vivo* experiments have their own limitations like cost, time and more false positives, the *in silico* experiments scale down the set of potential interactions to a subset of most likely interactions.

Retrieval of Protein structure from Protein Databank:
The Protein Data Bank (PDB) file format is a textual file format describing the three dimensional structures of molecules held in the Protein Data Bank. The PDB format accordingly provides for description and annotation of protein and nucleic acid structures including atomic coordinates, observed side chain rotamers, secondary structure assignments, as well as atomic connectivity. Structures are often deposited with other molecules such as water, ions, nucleic acids, ligands and so on, which can be described in the PDB format as well. The Protein Data Bank also keeps data on biological macromolecules in the newer mmCIF file format. The structures of the target proteins subjected to molecular docking study were obtained from the Protein Data Bank (PDB). The PDB is a repository for the 3-D structural data of large biological molecules submitted by scientists around the world. PDB sites such as the Research Collaboratory for Structural Bioinformatics (RCSB) are free accessible and contain information about experimentally-determined structures of proteins, nucleic acids, and complex assemblies.

Homology Modeling by Schrodinger Prime:
Prime 2.1 is a highly accurate protein structure prediction suite of programs that integrates comparative modeling and fold recognition into a single user-friendly, wizard-like interface. The comparative modeling path incorporates the complete protein structure prediction process from template identification, to alignment and to model building. Refinement can then be done from a separate panel, and involves side-chain prediction, loop prediction and minimization. The prime interface was designed to accommodate both novice and expert users, while the underlying programs were designed to produce superior results in a variety of applications, including high-resolution homology modeling, refinement of active sites, induced-fit optimization, fold recognition, structure-based functional annotation and generation of alternate loop conformations. Methods to predict *ab initio* modeling for queried sequence use Prime–SP(Standard Precision) module, Find Homologs, Comparative Modeling path, including the

Edit Alignment and Build Structure, Fold Recognition, Prime–Refinement module, Molecular Mechanics and Generalized Born model augmented with the hydrophobic solvent accessible surface area calculations.

Validation of Modeled Structure:
Modeled structure has to be validated against its property in the range and its structural similarity from sequence and structural way. Through the sequence analysis, we can get the residual mimic property through the mutant analysis but for conformation ally, there should be mutant in extended strand and loop.

Energy minimization of target protein:
MacroModel is a general purpose, force-field-based molecular modeling program with applicability to a wide range of chemical systems. MacroModel provides with multiple advanced methods to aid the understanding of chemical structure, energetics and dynamics. A large selection of force fields is available in MacroModel, including the latest technical advances introduced into OPLS_2010, a force field that Schrödinger is actively developing. Numerous minimization methods are available, enabling geometry optimizations for a broad selection of structural classes. A wide range of methods is available for conformational searching, allowing efficient sampling of the potential energy surface for low-energy structures, for systems ranging from small molecules to entire proteins. Solvation effects can be accounted for using the efficient continuum solvation models employed by MacroModel.

Molecular Docking of Target Protein and Interactable Protein:
Computational prediction of the 3D structures of molecular interactions is a challenging area, often requiring significant computational resources to produce structural predictions with atomic-level accuracy. This can be particularly burdensome when modelling large sets of interactions, macromolecular assemblies, or interactions between flexible proteins. Protein docking program, ZDOCK uses fast Fourier transform to perform a 3D search of the spatial degrees of freedom between two molecules. ZDOCK uses a fast Fourier transform to search all possible binding modes for the proteins, evaluating based on shape complementarily, desolvation energy and electrostatics. The top 2000 predictions from ZDOCK are then given to RDOCK where they are minimized by CHARMM to improve the energies and eliminate clashes, and then the electrostatic and desolvation energies are recomputed by RDOCK (in a more detailed fashion than the calculations performed by ZDOCK). The superior performance of ZDOCK and RDOCK has also been calculated in a community-wide protein docking blind test, CAPRI.

Molecular Dynamics Simulation
GROMACS is computational software to perform molecular dynamics simulations and energy minimization. These are two of the many techniques in computational chemistry and molecular modelling that provides the dynamical behaviour of the system in real time. Computational chemistry is a discipline that employs computational techniques in chemistry, ranging from quantum mechanics of small molecules to dynamics of large complex molecular aggregates and biological macromolecules like protein, peptides, and nucleic acids to obtain macroscopic observables. Molecular modelling is the process of describing complex chemical systems at the atomistic level. Molecular dynamics is the study of the system models by solving Newtonian equation of motion and atomic forces from initial atomic coordinates and velocities of the molecule. From the ensemble averages taken for the system, macroscopic properties are determined to understand the behaviour of system with time evolution. There are two methods for the generation of a representative equilibrium ensemble: (a) Monte Carlo (MC) simulations and (b) Molecular Dynamics (MD) simulations. Molecular dynamics is most appropriate and reliable method for the generation of non-equilibrium ensembles and for the analysis of dynamic process. Monte Carlo simulations do not require force computations; hence are simpler than MD but provide statistically inefficient results than MD for a given amount of computational time. For a system, if a starting configuration is away from equilibrium, the atomic forces may be excessively large and the MD simulation may fail. In those cases, a robust energy minimization is required. Energy minimisation is a molecular mechanics method of searching local minima of the system and prepare before performing dynamics simulation. Energy minimization performs non-dynamical calculation to predict potential energy at the stationary point and remove interatomic restraints.

In this study, protein-protein interactions were analyzed between fibronectin and dispase. (Fig 1). Protein-protein docking studies were done for the prediction of numerous possible Dispase enzyme complex structures of which only a few of them resemble the active structure. Choosing the near-native structures from the generated set is a key element in predicting correct binding mode of the protein complexes.

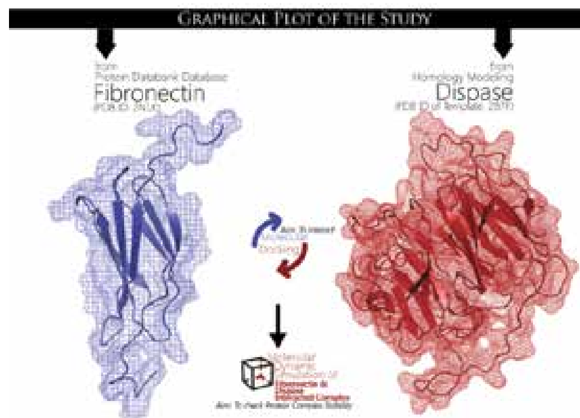


Fig 1 represents the graphical plot of the study:

Materials and methods
***In silico* tools used for the study:**
Preparation of Target Protein - Structure of Fibronectin
The domain structure of Human Fibronectin was retrieved from RCSB. The parameters were set to have less than 4 Å resolution calculations, monomeric protein and any bounded state. The PDB ID of the protein of interest is 2N1K. The resolution map calculations were from 33.48 Å to 1.90 Å as initial and ground state diffraction.

Preparation of Interactable Protein - Structure of Dispase
Homology modeling was used to predict the structure of Dispase. The following parameters were taken into consideration:
1. Limited number of amino acids for optimizing loop region to make it polarized one.
2. Third Party Access: Schrödinger makes available copies of the third-party databases that are used by Prime. However, these databases can also be updated directly from their respective sources.
3. Database against third part database, Maximum hits chosen 100 from the database, Lox complexity profile filter chose, Human repeats and specific nr sequence omission, word size of 28 , Maximum identity of Expected value 1e-1, Linear gaps inserted in the alignment profile and scoring match about 1-2 maximum.

Validation of Modeled Strcuture | Procheck [http://services.mbi.ucla.edu/SAVES]
The PROCHECK suite of programs provides a detailed check on the stereochemistry of a protein structure. Its outputs comprise a number of plots in PostScript format and a comprehensive residue-by-residue listing. These give an assessment of the overall quality of the structure as compared with well refined structures of the same resolution and also highlight regions that may need further investigation. The PROCHECK programs are useful for assessing the quality not only of protein structures in the process of being solved but also of existing structures and of those being modeled on known structures. Protein conformation was checked from most favored region, additional allowed regions, generously allowed region and disallowed region. In the residual property based on total of number of non glycine and non proline end residues (excl GLY and PRO), glycine residues, and proline residues were analyzed against total number of amino acids present in our modeled structure. Modeled Dispase structure was subjected to refinement and validation using Ramachandran plot with allowed region, disallowed region and quality of the amino acid positional conformation.

Volume-5, Issue-5, May - 2016 • ISSN No 2277 - 8160

Macromodel - Energy minimization of target protein:
Energy was minimized for the force field optimized domain structure of Human Fibronectin and Modeled Dispase enzyme structure. Polak-Ribiere Conjugate Gradient method was used to minimize the target protein with gradient of 2500 point iterations and threshold was set up to 0.05 Å. The potential parameters were kept within 8.0 Å of the vander Waal cut-off, 20.0 Å for electrostatic field, 4.0 Å for H-Bond interaction and 1.0 for dielectric constant.

Molecular Docking of Target Protein and Interactable Protein | Z Dock (http://zdock.umassmed.edu)
In this study, we incorporated a recently developed 3D convolution library into ZDOCK and modified ZDOCK to dynamically orient the input proteins for more efficient convolution. These modifications resulted in an average of over 8.5-fold improvement in running time when tested on 176 cases in a newly released protein docking benchmark, as well as substantial less memory usage, with no loss in docking accuracy. We also applied these improvements to a previous version of ZDOCK that uses a simpler non-pair wise atomic potential, yielding an average speed improvement of over 5-fold on the docking benchmark, while maintaining predictive success. This permits the utilization of ZDOCK for more intensive tasks such as docking flexible molecules and modelling of interatomics, and can be run more readily by those with limited computational resources.

Methodology and Parameterization for Molecular Dynamics Simulation for Docked Complex:
The molecular dynamic simulation was performed with GROMOS96 43a1 force fields using GROMACS 4.5.5-1 package. Before starting the simulations, all the models were solvated with the explicit Simple Point Charge (SPC216) water in a cubic box with the Periodic Boundary Conditions (PBC). The system was neutralized with 7 chlorine ions by solvent. The energy minimization was performed for the system concerned by using the steepest descent method. The topology files and charges for the ligand atoms were generated by the PRODRG2 Server. Then, 20 ns MD simulations were carried out with a time step of 2 fs. All simulations were run under periodic boundary condition with NVT ensemble by using Berenson's Temperature coupling algorithm with the temperature at 310 K. The Particle Mesh Ewald (PME) algorithm was used to calculate the electrostatic interactions and long-range electrostatics. The cut off for Vander Waals interactions was 10A˚. The NVT and NPT equilibration was done at 20 ns. All bonds were constrained by using the LINCS algorithm. The GROMACS 4.5.5-1 package was utilized to analyze the results.

Docked Fibronectin structure and Dispase Enzyme complex was simulated under water environment for 1 ns. The observation was under 1 pico second (ps) for 1 trajectory by Purified and Control peptide complex. 25000 iteration steps were maintained throughout the simulation period.

Results and Discussion
In silico **molecular docking and simulation results**

Validation of Modeled Dispase Structure:
The conformational profile from rotamer library was plotted against Ramachandran's library of conformations for each residue, which was derived from PDB template ID: 2B7F. Totally 92.3% amino acids are in core allowed region, 5.8 % amino acids are in additional allowed region, 1.2 % of amino acids are in generously allowed region and 0.7 % amino acids are in disallowed region out of 361 amino acids from modeled structure. Figure 14 shows the Ramachandran plot for Dispase structure.

Figure 14: Ramachandran plot and plot statistics against Dispase amino acids

Energy Minimization of Current and Post Minimized Modeled Dispase Enzyme

Energy Minimization of Current and Post Minimized Modeled Dispase Enzyme

Energy difference of Vanderwal, Coulombs and Electrostatic for modelled Dispase enzyme are $\int_{(-92.259)}^{(-88.908)} = -3.35\ KJ/Mol$, $\int_{(-349.560)}^{(-320.120)} = -29.44\ KJ/Mol$ and $\int_{(-221.382)}^{(-107.604)} = -113.718\ KJ/Mol$. The maximum limit is Control and minimum limit is Post minimized as observed in Fig 15. Reaction force energy rate is calculated on computed cavity energy surface which is around 42.169 E differences with Post minimized as 64.089 E and Control as 106.258 E. Total cavity energy was also computed on the reaction force energy i.e 17.249 E as Current energy and 17.232 E as Post minimized. Energy Plot states that post minimized structure has less energy threshold to take the structure for interaction and post interaction analysis. Energy minimized structures for modelled Dispase enzyme is shown in Fig 16. Modeled Dispase structure with current energy, energy minimized structure of Dispase and superimposed structures of current and minimized energy structures of Dispase can be seen from the figure.

Docking Complex of Fibronectin and Dispase Enzyme
Protein-protein docking studies were done for the prediction of numerous possible Dispase enzyme complex structures of which only a few of them resemble the active structure. Choosing the near-native structures from the generated set is a key element in predicting correct binding mode of the protein complexes. In the initial stage, the rigid body algorithm was used to generate protein-protein complex structures. This docking approach generated 3600 protein poses (hits) for the entire study as ZDOCK outputs 2000 protein poses per job. The reliable structure was then chosen for consideration based on higher ZDock Score after the ZDOCK has made its predictions. Table 4 shows best binding modes that were selected on the basis of volume and ZDOCK score of surface accessible.

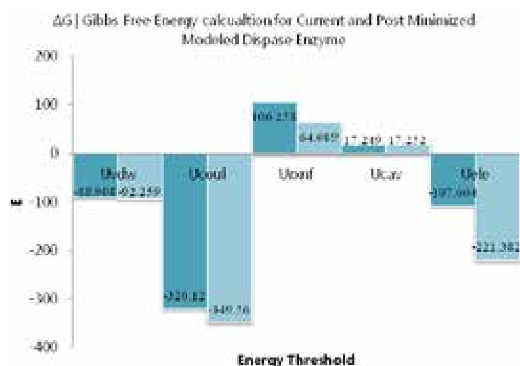


Figure 15: Energy calculations of Current and Post Minimized Modeled Dispase Enzyme. Free Energy calculations are Vanderwal energy (E_{vdw}), Coulombs energy (E_{coul}), Reaction force energy (Er_{xnf}), Cavity energy (E_{cav}) and Electrostatic energy (E_{ele}).

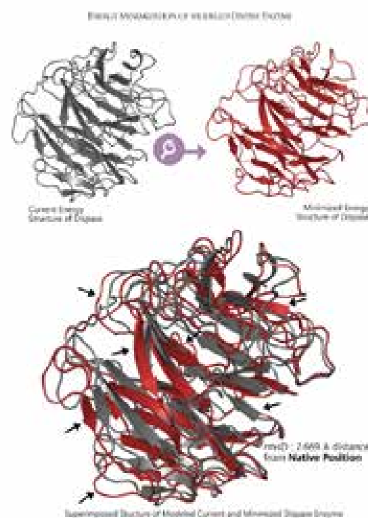


Figure 16: Energy minimization of modeled Dispase enzyme

Minimum Score (Last pose)	Maximum Score (First Pose)	Complex of Fibronectin and Dispase
Total 2000 Poses generated		
1081.338	1998.618 (Fig 1)	Conformer 1
940.985	1603.403	Conformer 2
793.426	1683.238	Conformer 3
819.886	1490.712	Conformer 4
937.74	1587.263	Conformer 5

Table 4: Docking Score of Fibronectin and Dispase Enzyme (Target Protein and Interactable Protein) from ZDock Server on basis of volume and Z score of the Surface Accessible

Target protein was interacted with receptor protein with rigid body docking algorithm using ZDock online server tool. The result was taken based on ZDock score along with Eulers Angle and Grid position in which target protein has taken most of the surface volume to interact. The criteria for selection of best binding mode of each complex were maximum score of protein - protein complex i.e 1998.618, 1603.403, 1683.238, 1490.712 and 1587.263 respectively. The total scores were validated and chosen out of 2000 poses from ZDock server. Complex that has higher ZDock score i.e, more than 1000 has higher interactive mechanism host (Target protein). So the complexes are more affinitive in the surface volume. Figure 17 shows the molecular docking pose of Fibronectin and Dispase.

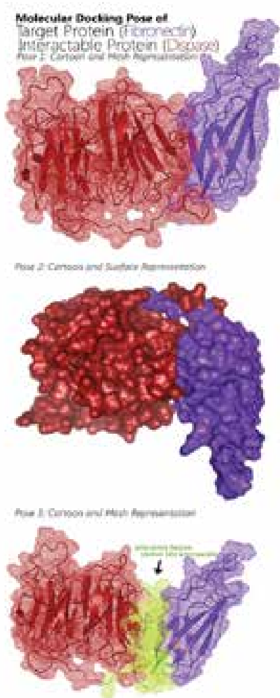
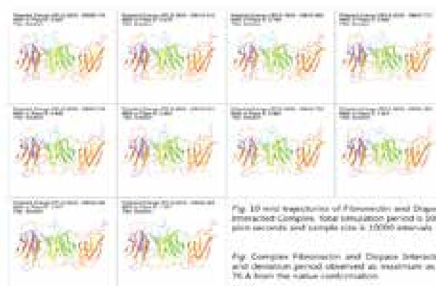


Figure 17: Molecular docking pose of Fibronectin and Dispase

Molecular Dynamic Simulation of Interacted Complex
The root mean square deviation (RMSD) analysis predicts the stability of protein and its structural variation while evolving with time. The RMSD of protein-protein complex for the trajectories written in the production run was analyzed to identify the stability of the system at each time interval. It was observed that the complex is stable within early 0.7 ns production run (Between 0.7 to 1.00 ns). The complex was observed to fluctuate from equilibrium with minor deviation of 0.1 nm, thereafter i.e., from 0.4 ns to 0.6 ns the system again attained equilibrium. On an average, system was observed to be stable with RMSD of 3.3 nm. The system achieves equilibrium in early 10-15 ps and remains stable thereafter for 1 ns simulation period (Fig 18). The complex of the Fibronectin stability shows that there is no major structural variation in Dispase after its binding. Since the Dispase complex is having more structural deviation between 0.25 Å till 0.45 Å distance, it shows instability within the active site and polar surface.

Figure 18: Graph: rmsD plot of Firbronectin and Dispase Complex for 1 nano second and 1000 trajectories interval



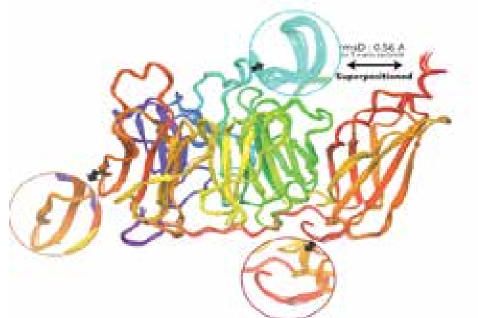


Figure 19 shows the results of 10 mid trajectories of Fibronectin and Dispase Interacted Complex and the deviation period observed by Fibronectin and Dispase Interacted complex.

Conclusion
From the wet lab results done on identifying the effects of Dispase enzyme production on pH, incubation ad temperature optimization towards various substrates from *Aspergillus niger*, we identified that on 15 days of incubation at pH 7 with 35°C on wheat bran substrate, there was higher Dispase enzyme production. The synthesized Dispase enzyme was partially purified from ammonium sulphate precipitation and ion exchange chromatography method. SDS-PAGE was used to evaluate the molecular weight and it was found to be 48kDa. FTIR and NMR spectroscopic techniques were employed to identify the presence of Dispase enzyme. Furthermore, *in silico* approaches (docking) were carried out to analyze protein-protein interactions of target protein Fibronection and Dispase interacted complex. Molecular simulation tool was used to check for the stability of the interacted complex and it was found that the complex displayed good stability

Volume-5, Issue-5, May - 2016 • ISSN No 2277 - 8160

within 0.7 ns production run. On an average, the complex was observed to be stable with RMSD of 3.3 nm. Thus the synthesized and purified Dispase enzyme has many potential applications in detergent pharmaceutical formulations, leather, laundry, food and waste processing industries and cancer biology field. Further work on this protease towards cancer cell line studies and molecular docking on oncoproteins is currently underway.

References

1. Bell DJ, Hoare M, Dunnill P. (1983). The formation of protein precipitates and their centrifugal recovery. *Adv. Biochem. Eng. Biotechnol.,* 26: 1-72.
2. Boyce KJ & Andrianopoulos A, (2006). *Morphogenesis: Control of cell types and shapes.* pp. 3-17
3. Countinho PM, Anderson MR, Kolenova K, vanKuyk PA, Benoit I, Gruben BS (2009). Post genomic insights into the plant polysaccharide degradation potential of *Aspergillus nidulans* and comparison to *Aspergillus niger* and *Aspergillus oryzae. Fungal Genetics and Biology:* FG & B, (Suppl.), S161-S169.
4. De Vries RP, Nayak V, van den Brink J, Vivas Duarte AD and Stalbrand H. (2012). Fungal degradation of oligo and poly saccharides. In P. Grunwald (Ed.,), Carbohydrate modifying biocatalysts (pp. 693-759)Singapore: Pan Stanford Publishing Pte, Ltd.
5. Ferreira de Olivera JM, van Passel MW, Schapp PJ, de Graaff LH (2010) Shotgun proteomics of *Aspergillus niger* microsomes upon D-xylose induction. *Appl Environ Microbiol* 76:4421-4429.
6. Hazem Aqel, Farouk Al-Quadan, Tahani K. YousefA novel neutral protease from thermophilic *Bacillus* strain HUTBS62 *J. BioSci. Biotech.* 2012, 1(2): 117-123.
7. Lowry OH, Rosebrough NJ, Farr AL, Randall RJ (1951). Protein measurement with the Folin phenol reagent. *J. Biol. Chem.,* 193(1): 265-275.
8. Lutzoni F, Kauff F, Cox CJ, McLaughlin D, Celio G, Dentinger B (2004). Assembling the fingal tree of life: progress, classification and evolution of subcellular traits. American Journal of Botany, 91, 1446-1480.
9. Srividya Shivakumar, (2012). Production and characterization of an acid Protease from a local *Aspergillus* Sp. by Solid substrate fermentation, Archives of Applied Science Research, 4 (1):188-199.
10. van den Brink J and de Vries RP, (2011). Fungal enzymes sets for plant polysaccharide degradation. Applied Microbiology and Biotechnology. 247, 229-239.