UNIVERSITY GRANTS COMMISSION
BAHADUR SHAH ZAFAR MARG
NEW DELHI – 110 002

Annual/Final Report of the work done on the Major Research Project.
(Report to be submitted within 6 weeks after completion of each year)

1. Project report No. 1st /2nd /3rd/Final: **Final**

2. UGC Reference No.F. : **41-547/2012(SR)**

3. Period of report: from : **05.12.2012** to **30.06.2015**

4. Title of research project: **Cloning and expression of a novel bacterial phytase in *Pichia pastoris* for extracellular secretion and optimization of production: An attempt to develop a commercial phytase production system**

5. (a) Name of the Principal Investigator: **Dr.Shilpi Ghosh**

   (b) Department: **Biotechnology**

   (c) University/College where work has progressed: **University of North Bengal**

6. Effective date of starting of the project: **05/12/2012**

7. Grant approved and expenditure incurred during the period of the report:

   a. Total amount approved: Rs. **12, 08, 235/- only**

   b. Total expenditure: Rs. **11, 61, 480/- only**

   c. Report of the work done: (Please attach a separate sheet)

i. Brief objective of the project:

- To isolate bacterial phytase ORF and analysis of gene sequence.
- To clone phytase gene in *Pichia* expression vector and analysis of expressed protein.
- To optimize production of phytase

ii. Work done so far and results achieved and publications, if any, resulting from the work (Give details of the papers and names of the journals inwhich it has been published or accepted for publication:

(A) Detail Report

Introduction

Phytic acid (myo-inositol 1,2,3,4,5,6-hexakis phosphate) and mixed cation salt of phytic acid, phytate, are a group of organic phosphorous (P) compounds found widely in nature. In terrestrial ecosystem, it is synthesized by plants and is the major storage form of P in seeds accounting for 80% of the total phosphorous of cereal and legumes. They are also strongly complexed in soils, representing an important class of organic P, which is only poorly available to plants. Monogastric animals (eg. swine, fish and poultry) are unable to degrade phytate in food. Phytic acid acts as an antinutritional agent by chelating various metal ions needed by the animals, such as calcium, copper and zinc. Due to unavailability of phytate phosphorous for plant and animal nutrition, it is a common practice to add inorganic P as plant fertilizer and as an animal feed supplement. The inefficient utilization of phytate consequently contributes to phosphorus pollution problems in the areas of intensive crop and livestock production. The phytic acid and phytate in soil are hydrolysed by phytate producing microorganism to release inorganic orthophosphate to water bodies contributing to eutrophication and algal bloom.

Phytic acid is hydrolysed by enzyme phytase (myo-inositol hexakisphosphate hydrolase) to less phosphorylated myo-inositol derivatives and inorganic phosphate. Most phytases belong to the family of histidine acid phosphatase and have been mainly classified as 3-phytase (EC 3.1.3.8), and 6-phytase (EC 3.1.3.26) based on the position of specificity of the initial hydrolysis of phytate. The enzyme is widespread in nature, occurring in microorganisms, plants, as well as in some animal tissues. Phytases from these sources exhibit variations in structure and catalytic mechanism and consequently, have been categorized into cystein phytases, histidine acid phosphatases, β-propeller phytases and purple acid phosphatases. Improved phosphorous nutrition is achievable by mobilization of phytate phosphorous by phytases. It has been applied as an additive in feed industry to improve the uptake of phosphorous and other nutrient and to reduce the phosphorous pollution in animal waste.

Pichia pastoris is a kind of methylotrophic yeast. It can grow with methanol as the sole carbon and energy source. *P. pastoris* grows to a very high cell density in simple defined media, and an extremely high yield of intracellular protein using the methanol-

controlled alcohol oxidase promoter. Using this system, many proteins have been produced with varying degrees of success (Primrose and Twyman, 2006).

Although the commercial production of phytase has focused on the fungus Aspergillus, studies have suggested bacterial phytase as more promising because of their higher substrate specificity, greater resistance to proteolysis and better catalytic efficiency. During the previous UGC major research project, PI's research group has isolated some novel phytase from rhizosphere. The purified enzyme has shown relatively high specific activity, substrate specificity, good pH profile, protease insensitivity and thermostability. The phytase encoding genes have been cloned and sequenced. The superior biochemical properties of the phytase suggest that it may be an attractive enzyme for various biotechnology applications. Hence, the present project was undertaken with the aim to clone and express bacterial phytase gene in *P.pastoris* for extracellular phytase secretion, which has great potential as a commercial phytase production system.

2.Methodology

2.1. Cloning of phytase gene from Shigella sp. CD2

2.1.1. Isolation of genomic DNA from *Shigella* sp.CD2

Genomic DNA was isolated from *Shigella* sp. by Murmer's method (1961). Bacterial strain was grown over-night in nutrient broth. The harvested bacterial cells were lysed in buffer containing 0.1M EDTA: 0.15M NaCl followed by treatment with lysozyme, SDS and proteinase-K. The DNA preparation was purified by phenol:chloroform extraction. DNA in the aqueous was precipitated by adding double volume of ethanol. The precipitated DNA was spooled out, washed in 70% ethanol, air dried and dissolved in Tris-EDTA (TE, pH8.0) buffer. The DNA was quantified spectrophotometrically and checked by agaroseTris acetate EDTA-gel electrophoresis.

2.1.2. PCR amplification of phytase ORF

The phytase ORF was amplified by PCR using genomic DNA as template. The forward primers, Phya 5'-GCC ATG AAA GCG ATC TTG ATC CCA-3' and reverse primer Phyb 5'-ATC CCA AAC TGC ACG CCG GTA TG -3' were used for amplification. PCR was performed for 30 cycles consisting of 94 ° C for 30 sec, 60 ^{o}C for 1min, 72 ^{o}C for 1min with a 7 min final extension at 72^{o}C. PCR reaction was analyzed by agarose gel electrophoresis. The PCR product was extracted from gel by using Gel extraction kit (Qiagen, Germany) following the method described by the manufacturer.

2.1.3. Agarose gel electrophoresis of DNA

1% agarose in 1X TAE (Tris-acetate EDTA) buffer was melted and then cooled to 50-60°C. It was then supplemented with 5µg ml^{-1} ethidium bromide (EtBr). The melted agarose was then poured in a casting tray fitted with a teflon comb forming wells. DNA sample was mixed with DNA loading dye (1X) prior to loading in the wells. Electrophoresis was performed in a horizontal electrophoresis tank using 1X TAE buffer. DNA bands were visualized on a UV-transilluminator (Genei, India).

2.1.4. Preparation of competent *E.coli* cells:

The E.coli strain JM109 (Promega) was made competent by chemical method using $CaCl_2$. The bacterial culture was grown in LB medium at 37^oC under shaking for overnight. The overnight grown culture was inoculated in fresh LB medium and grown till optical density of 0.4 to 0.6. The bacterial cells were pelleted by centrifugation at 6000 rpm for 10 min followed by washing with 100 mM $CaCl_2$. The cells were resuspended in the same solution and stored at 4^oC for further use.

2.1.5. Cloning of PCR product in pGEMT-easy vector and sequence analysis :

The PCR product was ligated to pGEMT-easy vector **(Figure 1)** by using T4 DNA ligase (Promega). The ligation reaction mix was transformed into competent E. coli JM-109 cell followed by selection of transformed bacterial colonies in presence of ampicillin, IPTG and X-gal. The recombinant plasmid was isolated from transformed white colonies and analysed for the presence of phytase gene by restriction digestion. The cloned gene was finally sequenced using T7 and SP6 universal primers.

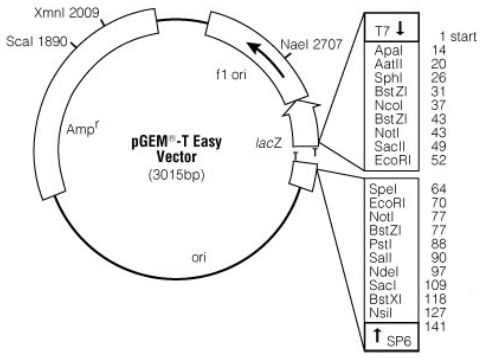

Figure 1: The map of pGEMT vector DNA: a double -stranded closed circular plasmid of 3015 base pairs; AmpR, ampicillin resistance gene; lacZ, enzyme beta-galactosidase gene; MCS (multiple cloning site) with a series of unique restriction sites; Ori, *E. coli* origin of replication and f1origin of replication.

Homology search in GenBank was done using the BLAST server (http://www.ncbi.nlm.nih .gov /BLAST (21). The amino acid sequence of the cloned gene was deduced and protein sequence was aligned by CLUSTAL W program (http:// www.ebi.ac.uk/clustal W). The phylogenetic analysis of the protein was performed by neighbour joining method using MEGA-4 (23). Bootstrap analysis was used to evaluate the tree topology of the neighbor joining data by performing 500 replicates. The tree was drawn to scale, with branch lengths in the same units as those of the evolutionary distances used to infer the phylogenetic tree. The recombinant pGEMT vector harbouring the full length phytase gene was named pGEMT-appA$_S$(f).

2.1.6. Expression of phytase AppA$_S$ in *Escherichia coli* and *Pichia pastoris* and characterization of the recombinant proteins

2.1.6.1. Clonimg of mature appAs and sequence analysis

Mature phytase gene without the signal sequence was amplified from pGEMT-appA$_S$ (f) by using internal primers, PhyF (5'-ATGAATTCGCTCAGAGTGAGCCGGAG-3´ with 5' *Eco*R1 restriction site) and PhyR (5´GATGCGGCCGCCAAACTGCACGCCGGTA TG -3' with 5' *Not*I site). PCR amplification was performed in 25µl reaction volume. The reaction mixture contained 5X reaction buffer, 5 µl; 25 mM $MgCl_2$, 2 µl; 10mM dNTP mix, 1µl; 50 pM forward and reverse primers, 1µl each; pGEMT-appAs(f), 2µl; Taq polymerase (5U μl^{-1}), 0.25 µl. PCR was performed for 30 cycles with initial denaturation at 94°C for 3min followed by a strand denaturation period at 94°C for 30 sec, primer annealing at 55°C for 1 min, strand extension at 72°C for 1 min for every individual cycle and a final extension period at 72°C for 7min. The PCR products were then loaded and separated on 1% agarose TAE gel. The PCR product was then cloned in pGEMT-Easy vector following manufacturer's instruction and sequenced using T7 and SP6 universal primers. The recombinant pGEMT vector harbouring the phytase gene was named pGEMT-appA$_S$.

2.1.6.2. Construction of *E.coli* and *P.pastoris* expression plasmids

Two different plasmids were constructed for heterologous expression in *E.coli* BL21(DE3) and *P.pastoris* GS115. For *E.coli* expression, the pGEMT-appA$_S$ plasmid was cut with *EcoR*1 and *Not*1. The resulting 1.2 kb DNA fragment was ligated into pET-20b (+) **(Figure 2)** to generate construct pET-20b (+) appA$_S$. The ligation reaction mixture in a total volume of 10µl contained: 10X ligation buffer, 1µl; prepared pET-20b(+) vector, 1µl (0.1 µg/µl); prepared insert, 2 µl; T4 DNA ligase (4 U/µl), 1 µl; 10

mM rATP (pH 7.0), 1 µl and 4 µl ddH_2O. The reaction mixture was incubated overnight at 4°C. The construct was transformed into *E.coli* BL21(DE3) by using heat-pulse in a 42°C water bath for 20 sec immediately followed by incubation on ice for 2 min. The transformed cells were revived in SOC medium (20.0 g of tryptone, 5.0 g of yeast extract, 0.5 g of NaCl in 970 ml distilled water, in which 10 ml of 1 M filter-sterilized $MgCl_2$ and $MgSO_4$ solutions, and 10 ml of 2 M filter-sterilized glucose solution added prior to use) for 1 h at 37°C with shaking at 250 rpm. The transformants were selected in LB- ampicillin (50 µg ml^{-1}) agar plates by incubating overnight at 37°C.

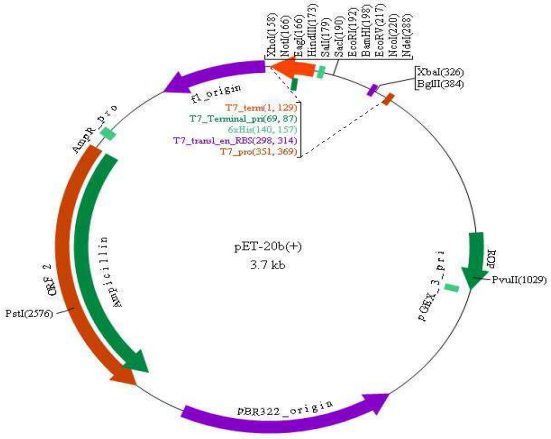

Figure 2: The map of pET-20b(+) plasmid vector: This cloning and expression vector includes f1 and pBR322 origin of replication; T7 promoter, T7 transcription start, T7 terminator sequences; AmpR, ampicillin resistance gene; Multiple cloning site (MCS); an N-terminal pelB signal sequence for potential periplasmic localization and an optional C-terminal His•Tag sequence.

For expression in *P.pastoris*, the 1.2 kb fragment released from the pGEMTappA_S plasmid was ligated into *P.pastoris* expression vector pPIC9 **(Figure 3)** digested with *EcoR*1 and *Not*I restriction enzymes. The resulting pPIC9appA_S was digested with *Bsp*E1 to linearize it and then transformed into *P.pastoris* GS115 by the spheroplasting protocol according to the manual of the *Pichia* Expression kit (Invitrogen, San Diego, CA). Transformants were selected for ability to grow on histidine-deficient medium. The His^+ transformants were screened for Mut^+ and Mut^S phenotypes. The integration of expression cassette into the genome of *P.pastoris* GS115 was ascertained by PCR using the 5' AOX1 and 3' AOX1 primers.

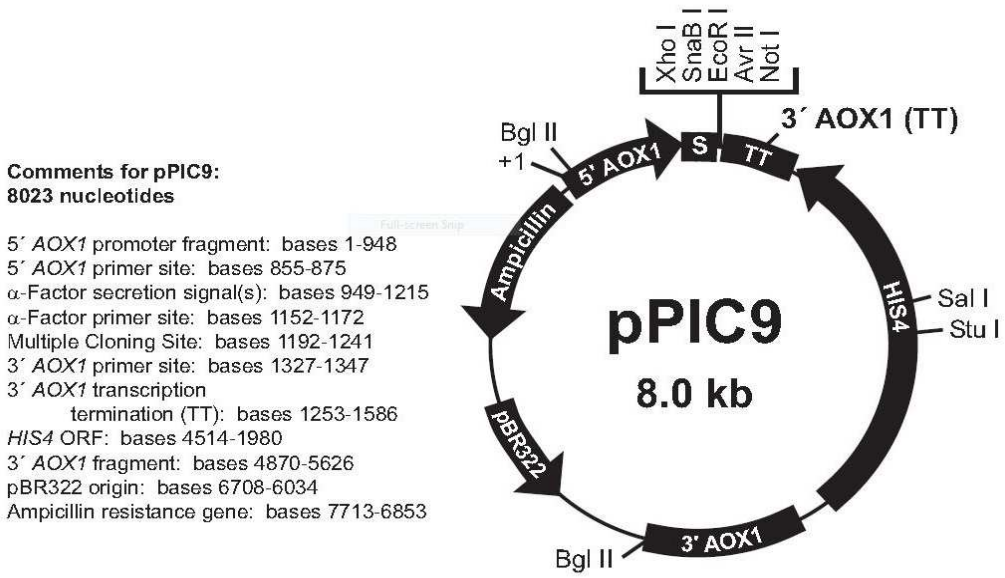

Figure 3: The map of pPIC9 plasmid vector for extracellular expression in *Pichia pastoris*. 5'*AOX1*, *AOX1* promoter; 3'*AOX1* TT, *AOX1* transcriptional terminator; 3'*AOX1*, *AOX1* downstream region; AmpR, ampicillin resistance gene; *HIS4*, *P. pastoris* wild-type gene coding for histidinol dehydrogenase; pBR322 origin of replication.

2.1.6.3. Expression of appA$_S$ in *E.coli* BL21(DE3)

The expression of appA$_S$ in *E.coli* BL21(DE3) was analysed by using MagicMedia^{TM} *E.coli* Expression Medium (Invitrogen, San Diego, CA) following manufacturer's instruction. *E. coli* BL21 (DE3) cells transformed with pET-20b(+)-appA$_S$ was grown overnight in LB medium at 37°C and 200 rpm shaking. The culture at 1% (v/v) was inoculated into the MagicMedia (19: 1, ready to use medium: IPTG solution) and grown overnight at 37°C and 300 rpm shaking. The cells were then harvested by centrifugation at 10,000 rpm for 10 min. The cell pellet was suspended in 50mM acetate buffer (pH 5.5), disrupted by sonication and centrifuged. The supernatant, and the pellet dissolved in 50 mM acetate buffer (pH 5.5) served as soluble and pellet fractions, respectively. Induction of appA$_S$ expression in both the fractions of was determined by 12% SDS-PAGE using protein molecular weight markers of 28, 36, 55, 72, 95 kDa. Both the fractions were also checked for phytase activity. *E. coli* BL21 (DE3) transformed with pET-20b(+) vector was used as control. Recombinant protein produced by appA$_S$ in *E.coli* BL21 (DE3) was named rAppA$_E$.

2.1.6.4. Expression of $appA_S$ in *P. pastoris* GS115

The Mut^+ *P. pastoris* GS115 cells transformed with pPIC9$appA_S$ was inoculated into 10 ml of YPD (1% yeast extract, 2 % peptone and 2 % dextrose) and incubated at 30°C overnight with vigorous shaking. One milliliter of starter culture was transferred to 100 ml of BMGY [1% yeast extract, 2% peptone, 100 mM potassium phosphate (pH 6.0), 1.34% YNB (yeast nitrogen base with ammonium sulfate without amino acids, 4×10^{-5}% biotin and 1% glycerol) and was grown at 30°C and 300 rpm shaking until cultures reached an OD_{600} of 1. Cells were subsequently harvested by centrifugation at 1500 rpm for 5 min and used to inoculate 100 ml of BMMY [1% yeast extract, 2% peptone, 100 mM potassium phosphate (pH 6.0), 1.34% YNB (yeast nitrogen base with ammonium sulfate without amino acids), 4×10^{-5}% biotin and 0.5% methanol] containing methanol as inducer and the culture was incubated with vigorous shaking at 30°C for 96 h. The induction was maintained by adding 0.5 % (v/v) methanol every 24 h. The culture was analyzed for extracellular and periplasmic phytase activity at 24 h intervals. For isolation of extracellular fraction, the culture was centrifuged at 1500 rpm for 5 min and the cell free medium was concentrated and diafiltered by Vivaspin-20 (30 kDa cutoff) sample concentrator (GE Healthcare, UK). For periplasmic fraction isolation, cell pellet suspended in 100 mM sodium acetate buffer (pH 5.5) containing zymolyase (3 mg/mL) was incubated at 37°C for 50 min and then centrifuged at 10,000 rpm for 10 min. The supernatant obtained after centrifugation served as periplasmic fraction. Induction of phytase overexpression was determined by 12% SDS-PAGE. The transformants with pPIC9 plasmid was used as control. Recombinant protein produced by $appA_S$ in *P.pastoris* was named $rAppA_P$

2.1.7. Purification of $rAppA_E$ and $rAppA_P$

For purification of $rAppA_E$, the IPTG induced culture of *E.coli* BL21(DE3) transformed with pET-20b (+) $appA_S$ was harvested by centrifugation at 10,000 rpm for 10 min. The cell pellet was suspended in 50 mM sodium acetate buffer (pH 5.5) and disrupted by sonication. The supernatant obtained after centrifugation of the sonicated cell lysate was loaded onto a Ni-Sepharose Fast Flow column (2 x 5 cm, GE Healthcare, UK) pre-equilibrated with 50 mM sodium acetate (pH 5.5) containing 10 mM imidazole. The column was washed with the same buffer and then bound proteins were eluted with an elution buffer containing 50 mM sodium acetate (pH 5.5) and 100 mM imidazole. Phytase activity was determined in each fraction and active fractions were pooled for subsequent studies. The $rAppA_P$ was purified from the cell free medium of pPIC9$appA_S$ transformed *P.pastoris* GS115 culture induced with methanol for 48 h. The concentrated and diafiltered cell free medium was loaded on to CM-cellulose column and bound protein was eluted by

acetate buffer (50 mM, pH 5.5) with linear gradient of 0-0.5 M NaCl. The active fractions were pooled and used for subsequent analysis.

2.1.8. Protein estimation and SDS-PAGE analysis

Total protein concentration was determined by the dye binding assay of Bradford using bovine serum albumin (BSA) as standard [21]. SDS-PAGE analysis was performed with 12% polyacrylamide gel according to the method of Laemmli [22]. After electrophoresis, the gel was stained with CBB R-250 reagent (0.1 % Coomassie Brilliant Blue R-250 in 10 % acetic acid and 40% methanol) and then destained. Broad range pre-stained protein standards were used as markers.

2.1.9. Phytase assay

Phytase activity was determined by the method described by Shimizu (13). The reaction mixture in a final volume of 2 ml contained 100 mM acetate buffer (pH-5.5), 2 mM sodium phytate, and 100µl enzyme preparation. The reaction was carried out at 37°C for 30 min followed by termination of reaction by adding 2 ml of 10 % trichloroacetic acid. The reaction mixture was centrifuged at 6,000 rpm for 10 min. The released inorganic P (Pi) was measured in the supernatant by adding 2ml of Pi-reagent containing ammonium molybdate (0.5%), sulphuric acid (5.0 N) and ascorbic acid (2.0 %), incubating the preparation for 20 min at 37°C and measuring absorbance at 610 nm. The amount of inorganic phosphate released from phytate was determined by using Pi Standard Curve. One unit (U) of phytase activity represents 1µmol of Pi released per min under assay conditions.

2.1.10. Characterization of phytase

The purified $rAppA_P$ and $rAppA_E$ were characterized for molecular weight, pH optima, temperature optima, K_m for phytate and thermostability.

2.1.10.1. Molecular weight determination

The purified phytase enzyme was fractionated by SDS-PAGE using 12% SDS-PAGE gel by the method of Laemmli. When protein sample and marker were loaded in the wells and electric field was applied, proteins were separated according to their molecular weight. The protein molecular weight markers used were: 210, 107, 75, 47, 32, 25, 15 kDa.

2.1.10.2. Effect of pH and temperature on phytase activity and thermostability of the enzyme

The pH optima of was determined by measuring the enzyme activity at pH 2.5-7.5 in the

following buffers: 100 mM glycine (pH 2.5 and 3.5), 100 mM sodium acetate (pH 4.5 and pH 5.5), 100 mM phosphate (pH 6.5) and Tris-HCl (pH 7.5). The normal assay procedure was followed only varying the buffer solution. The optimum temperature for phytase activity was determined at optimum pH and at temperature ranging from 20° to 80°C. The incubation temperature was varied for the standard assay procedure. Thermostability of the enzyme was determined by preincubating the enzyme at 20°-80°C for 30 min followed by measuring activity under standard conditions.

2.1.10.3. Determination of K_m

K_m value for phytate was determined by Lineweaver -Burk plot at optimum pH and at 37°C.

2.1.8.4. Effect of metal ions and chemical reagents on phytase activity

The effect of metal ions and chemical reagents on activity of the purified recombinant enzymes was determined by measuring phytase activity in presence of 2mM Ca^{2+}, Mn^{2+}, Mg^{2+}, Cu^{2+}, Zn^{2+}, Fe^{2+}, Co^{2+}, K^{+}, Na^{+}, SDS, EDTA, β-ME or Cystein-HCl. The test system without addition of ions or reagents was used as control. Data is shown as the mean of three independent experiments.

2.1.10.5. Substrate specificity of phytase

Substrate specificity of the enzyme was determined by replacing phytic acid in the standard assay mixture with an equal concentration (2 mM) of the following phosphorylated compounds: adenosine 5'-diphosphate (ADP), adenosine 5'-triphosphate (ATP), p-nitrophenyl phosphate (P-NPP) and glucose 6-phosphate (G6-P) and fructose 6-phosphate (F6-P).

2.1.10.6. Effect of protease on phytase activity

For determining the susceptibility to digestive protease, the 50 U of purified $rAppA_E$ or $rAppA_P$ was preincubated with pepsin and trypsin (30 U, Sigma) at 37°C and phytase activity was monitored 30 min later. itored 60 min later.

2.1.11. Deglycosylation

The deglycosylation of $rAppA_P$ was carried out using Endo H glycosidase (Endo H, New England Biolabs) following manufacturer's instruction. The reaction mix containing, 50 U of purified $rAppA_P$, 600 µL of 50 mM Tris buffer (pH 7.0) and 10 U of Endo H, was incubated at 37°C for 2 h. N-glycosylation was determined by assessing the migration shift of Endo H treated $rAppA_P$ in 12 % SDS-PAGE.

2.1.12. Western blot analysis

For immunoblot analysis, purified rAppA$_E$ and deglycosylated rAppA$_P$ proteins separated by 12% SDS-PAGE, were transferred to polyvinylidene difluoride (PVDF) membrane by semi-dry method using Electroblotting apparatus (Atto, Japan). A purified rabbit antibody raised against *E.coli* phytase, diluted 1:1000 prior to application, was the primary antibody. The reacted polypeptide was visualised with a secondary antibody, goat anti-rabbit IgG-alkaline phosphatase conjugate using colorimetric based nitroblue tetrazolium chloride and 5-bromo-4-chloro-3-indolylphosphate-p-toluidine (NBT/BCIP) detection kit (Invitrogen, USA). Broad range pre-stained standards were used as markers.

3. Results

3.1. Cloning and sequencing of phytase gene and phylogenetic analysis

The phytase gene from *Shigella* sp. CD2 was amplified by PCR using primers designed from the phytase sequences from related bacterial genus in the database **(Figure 4)**. The PCR product was cloned in pGEM-T Easy vector and sequenced using vector

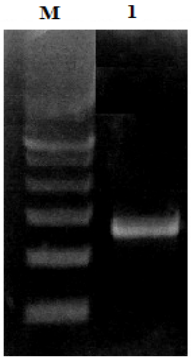

Figure 4: PCR amplification of phytase gene, Lane M: 500 bp ladder, Lane1:PCR product

specific T7 and SP6 universal primers. The sequence analysis of the insert indicated presence of an ORF of 1299 bp, encoding protein of 433 amino acids **(Figure 5)**. The nucleotide sequence was deposited in the GenBank database under accession number FR865899. Homology analysis of deduced amino acid sequence by BLAST program revealed 98 and 62 % similarity with AppA phytase of *E.coli* and *C. braakii*, respectively. Hence, *Shigella* sp. CD2 phytase ORF was named as appA$_S$ and encoded protein as AppA$_S$. An alignment of the amino acid sequence with enteric bacterial phytases from the GenBank using ClustalW program showed the presence of N-terminal RHGXRXP motif and the C-terminal HDTN motif at amino acid 38-44 and 325-326, respectively, which are

common in phytases belonging to HAP family (Yao et al. 2011). Moreover, five cysteine residues were also conserved among these phytases as shown in the Fig. 2.14. *E.coli*

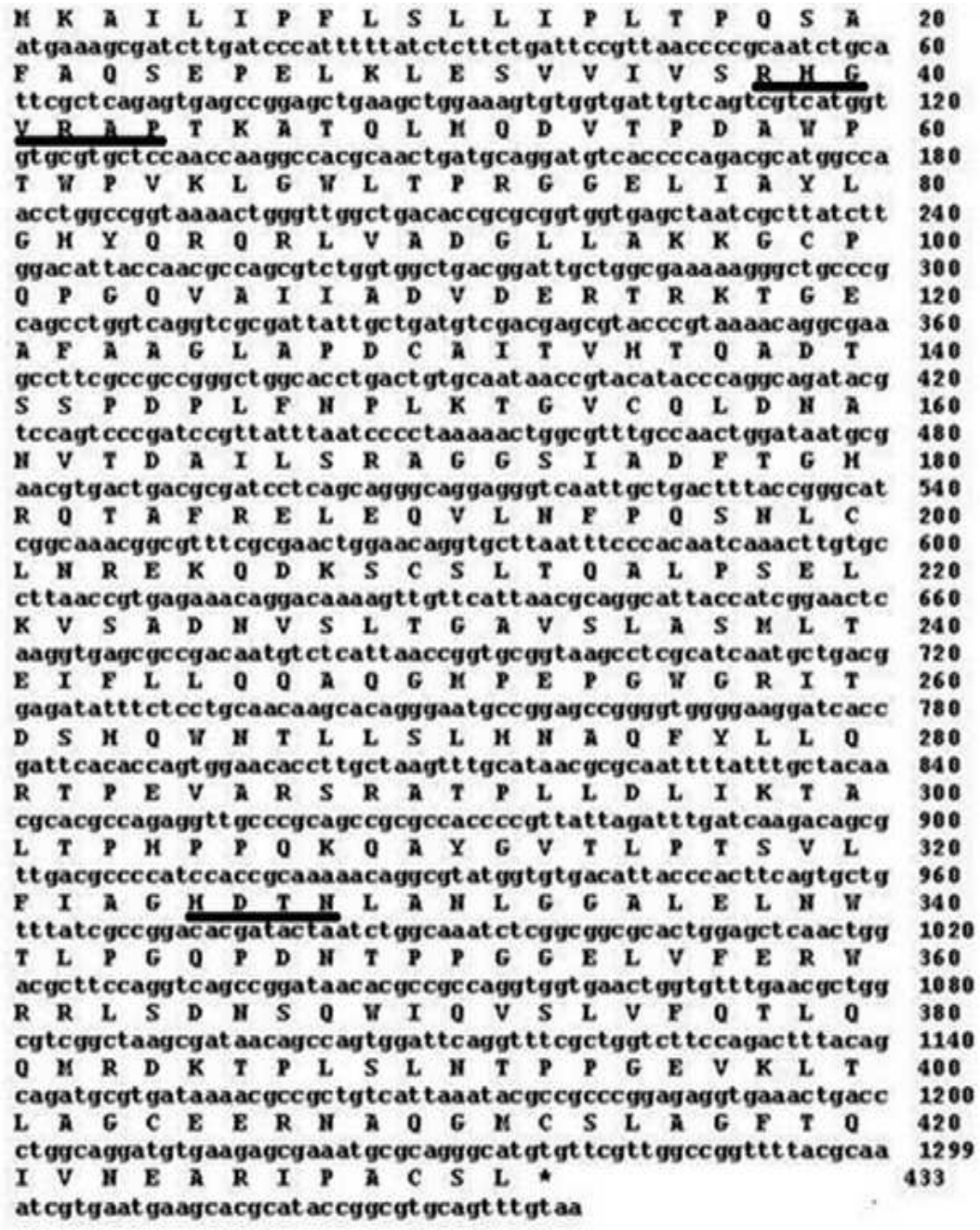

Figure 5: Nucleotide (1-1299) and deduced amino acid sequences (432) of the putative phytase gene appAs, from *Shigella* sp.CD2. The conserved histidine acid phosphatase family active site motifs are underlined. The stop codon is shown by asterisk.

AppA and *Shigella* sp CD2 AppAs differed in sequence at six positions. The amino acids S, R, K, E, M, A in *E.coli* AppA were substituted by P, Q, N, K, K, T in $AppA_S$ in the amino acid positions 102, 190, 202, 208, 298, 299, respectively **(Figure 6)**. *Shigella* sp.CD2 phytase ORF was named as $appA_S$ and encoded protein as $AppA_S$. A phylogenetic

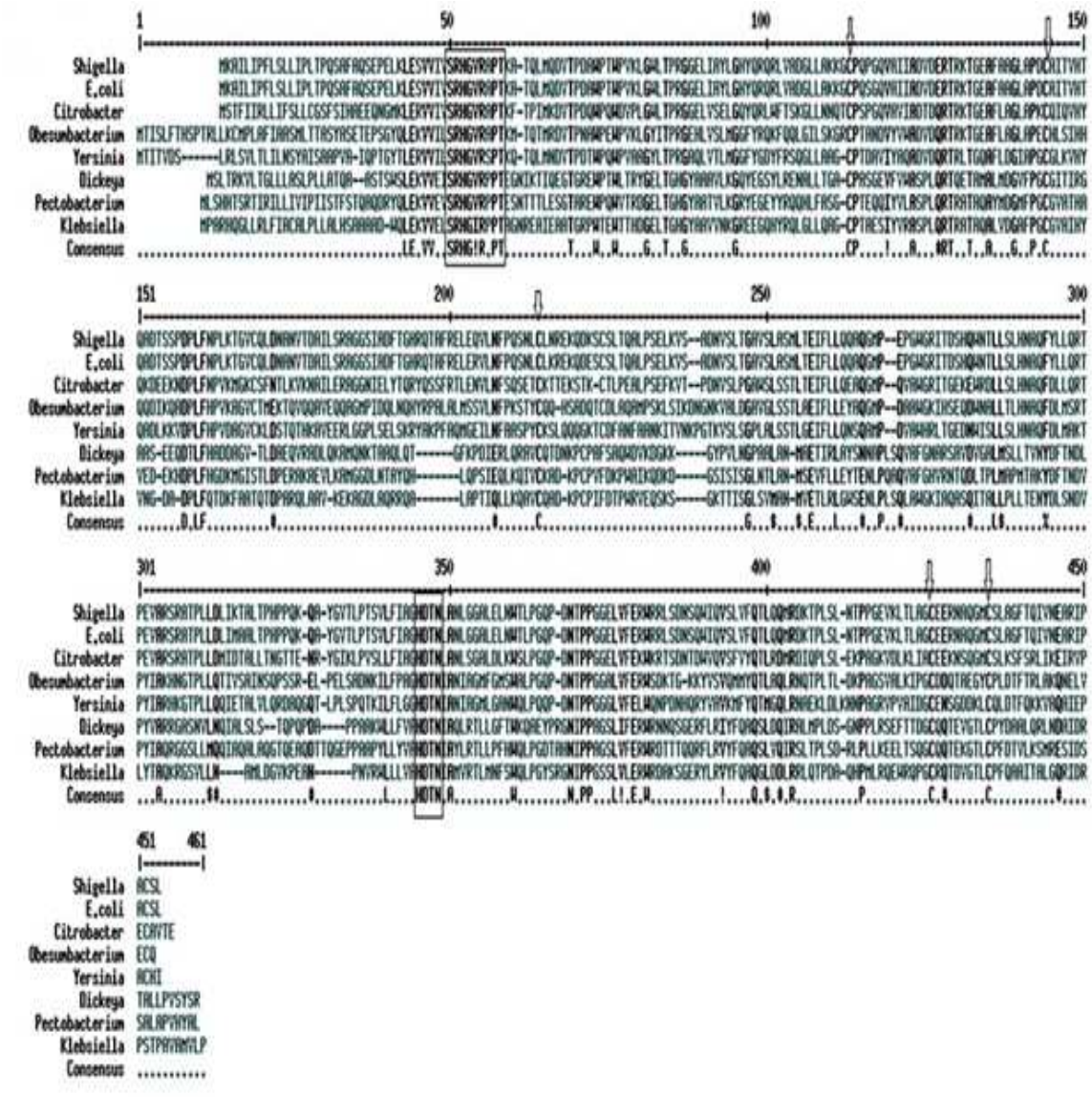

Figure 6: Multiple alignment of homologs of the *Shigella* sp. CD2 phytase $AppA_S$. Conserved active site motifs are boxed and conserved cystein residues are shown by arrows. The source and GenBank Accession Nos. of proteins are: *Shigella* sp.CD2, CCA94903; *E.coli* AppA, EDX38944; *Dickeya paradisiaca*, ABW76125; *Klebsiella pneumonia* ASR1 AAM23271; *Yersinia intermedia* ABI95370.1; *Citrobacter braakii* AAS45884; *Obesumbacterium proteus* AAQ90419; *Pectobacterium carotovorum* subsp. *carotovorum* ABY76184

tree then constructed based on the alignment using the neighbor joining method. The topology of the phylogram also confirmed that $AppA_s$ was closely related to AppA phytase of *E.coli* and *Citrobacter braakii* **(Figure 7).** The signal peptide cleavage site was identified using SignalP 4.1 Server (http://www.cbs.dtu.dk/services/SignalP/) (Petersen *et al.* 2011). The $AppA_S$ contained putative signal peptide of 22 amino acids, and three potential sites of N-glycosylation. The calculated molecular mass of the protein with and without signal sequence were about 47 and 45 kDa, respectively and a theoretical pI of 5.67.

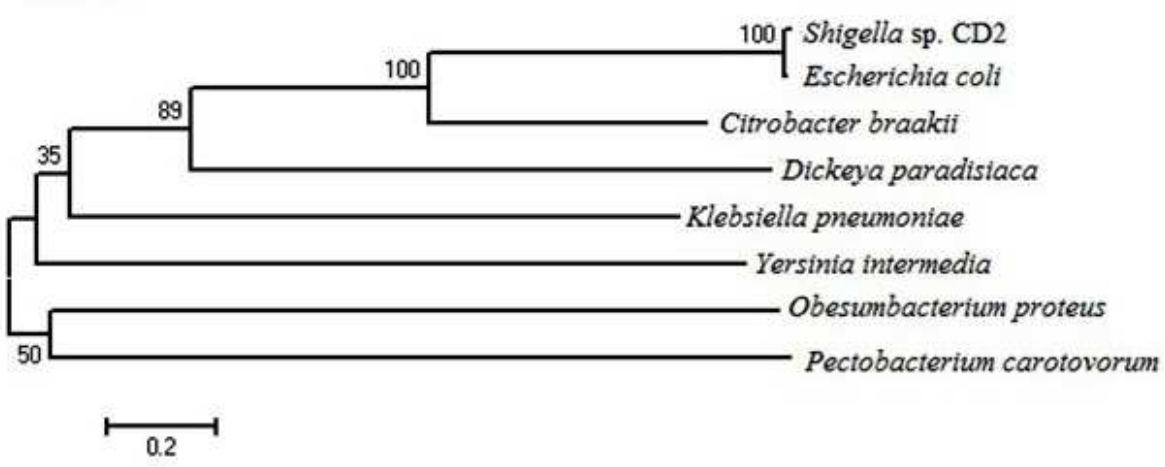

Figure 7: Phylogenetic tree of homologs of the *Shigella* sp. CD2 phtase $AppA_S$. Bootstrap values (%) from analysis of 500 bootstrap replicates are given at the respective nodes. The bar represents 2 substitutions per 10 amino acids. For GenBank Accession Nos. of proteins, see the Fig. 2 ligand.

3.2. Expression of *app*A_S in *P. pastoris* G115

The expression of enzymes as secreted protein is one of the useful and important characteristics for their economical production in industry. *P. pastoris* has been successfully used as host organism for extracellular production of recombinant proteins at high level, including phytases (Cereghino and Cregg 2000). The $appA_S$ was cloned in *P. pastoris* expression vector pPIC9. The recombinant plasmid pPIC9-$appA_S$ carried the $appA_S$-expression cassette consisting of 1.2 kb $appA_S$ gene in frame with *Sacharomyces cerevisiae* α-factor secretion signal, flanked by *AOX1* promoter and terminator sequences. Transformation of pPIC9-$appA_S$ into *P.pastoris* GS115 gave about 20 His^+ transformants. The integration of appAs-expression cassette into the genome of *P. pastoris* GS115 transformants was ascertained by PCR using 5' and 3' *AOX1* primers. PCR amplification products of about 0.5 kb and 1.7 kb in pPIC9 transformed and pPIC9-$appA_S$ transformed

P. pastoris GS115, respectively, indicated the integration of $appA_S$-expression cassette into the genome of the later **(Figure 8)**.

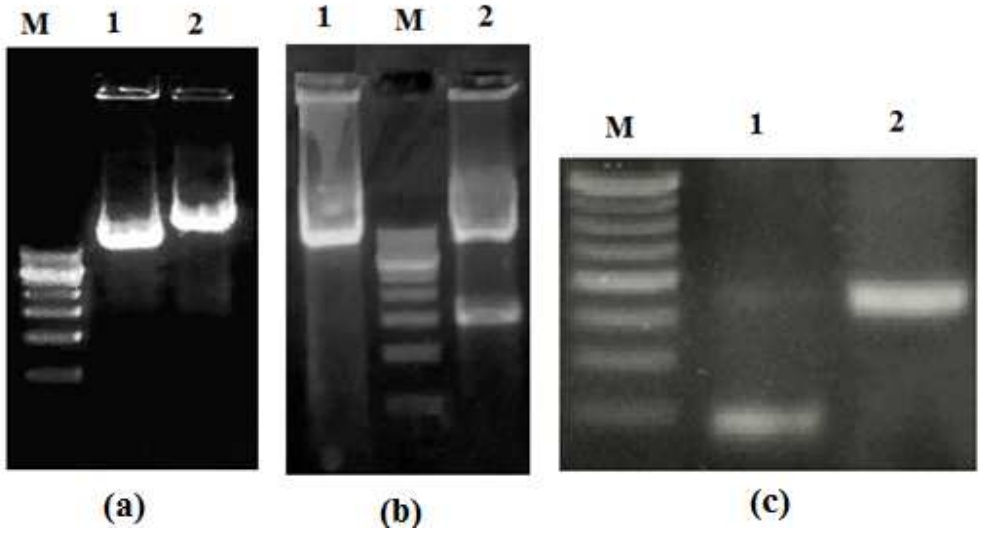

Figure 8: Cloning of $AppA_S$ in *P.pastoris* expression vector pPIC9 (a) Lane M-500bp ladder, Lane 1-pPIC9 plasmid, Lane2-pPIC9-$appA_S$ (b) Restriction analysis of recombinant pPIC9-$appA_S$ plasmid, Lane 1- pPIC9 plasmid, Lane 2- recombinant pPIC9-$appA_S$ plasmid pPIC9:phyS restricted with EcoRI and NotI, LaneM-500bp ladder, (c) PCR products of Lane 1- pPIC9 transformed and Lane2-pPIC9-$appA_S$ transformed *P. pastoris* GS115 using respective genomic DNA as template

The pPIC9-$appA_S$ transformed *P.pastoris* GS115 colonies were screened for Mut^+ and Mut^S phenotypes. Both Mut^+ and Mut^S transformants were analysed for extracellular and periplasmic phytase expression. The transformant was initially grown in BMGY medium until culture reached an OD_{600} value of 1 and then the culture was transferred to BMMY medium containing 0.5% methanol as inducer and grown for 96 h. Extracellular and periplasmic phytase activity and pH of the medium were determined at every 24 h intervals. A Mut^+ colony with highest extracellular phytase activity was selected for shake flask expression. The selected transformant showed maximum extracellular recombinant phytase ($rAppA_P$) activity of 62 U ml^{-1}, with specific activity 477 U mg^{-1}, at 60 h of methanol induction and an extracellular protein concentration of 0.13 mg ml^{-1}. The SDS-PAGE analysis of concentrated and diafiltered cell free extract showed two protein bands of approximate molecular mass 59 and 65 kDa **(Figure 9a)**. The enhanced

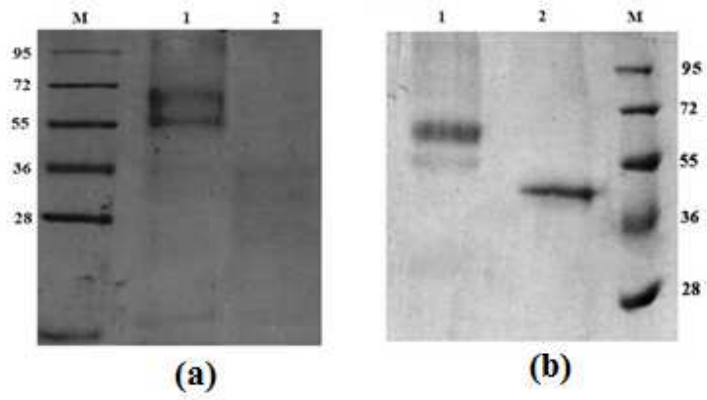

Figure 9: Expression of *Shigella* sp. CD2 phytase appA$_S$ in *P.pastoris* GS115 cells **(a)** SDS-PAGE analysis of recombinant phytase (rAppA$_P$). Lane, M- Molecular Weight markers in kDa, 1- extracellular fraction of *P.pastoris* GS115 transformed with pPIC9-appA$_S$, 2- extracellular fraction of *P.pastoris* GS115 transformed with pPIC9. Cultures were induced by 0.5 % methanol for 48 h and diafiltered culture supernatant containing 50 µg protein was loaded in each lane of a 12 % SDS-PAGE gel. **(b)** SDS-PAGE analysis of glycosylated and deglycosylated rAppA$_P$. Lane, M- Molecular Weight markers in kDa, 1- glycosylated rAppA$_P$, 2- deglycosylated rAppA$_P$. The rAppA$_P$ (50 U) was incubated with Endo H (10 U) for 2h. The protein was analyzed and compared with untreated enzymes.

molecular mass of rAppA$_P$ might be due to glycosylation that normally occurs for yeast secreted proteins. The deglycosylation treatment to rAppA$_P$ with Endo H resulted in single band with apparent molecular mass of 45 kDa, similar to that of *E. coli* expressed rAppA$_E$ **(Figure 9b)**.

3.3. Expression of *app*A$_S$ in *E.coli* BL21(DE3)

The mature appA$_S$ (without the signal sequence) was cloned into *E. coli* expression vector pET-20b (+) and the recombinant plasmid pET-20b(+)-appA$_S$ was transformed into *E. coli* BL21(DE3). The transformant was induced in Magic Media supplemented with IPTG and cells were disrupted by sonication after overnight induction. Recombinant phytase (rAppA$_E$) overexpression in the soluble and pellet fractions of sonicated cells was analyzed by SDS-PAGE. As shown in the results of **Figure 10**, the soluble fraction of the induced cell exhibited protein overexpression band of approximately 45 kDa, which agrees with the predicted molecular weight deduced from the amino acid sequence of AppA$_S$. The soluble and pellet fraction of IPTG induced *E. coli* BL21(DE3) transformant was also analyzed for phytase activity. Phytase activity level in the soluble fraction was 176 U ml^{-1} (specific activity 568 U mg^{-1}), whereas negligible activity was detected in the

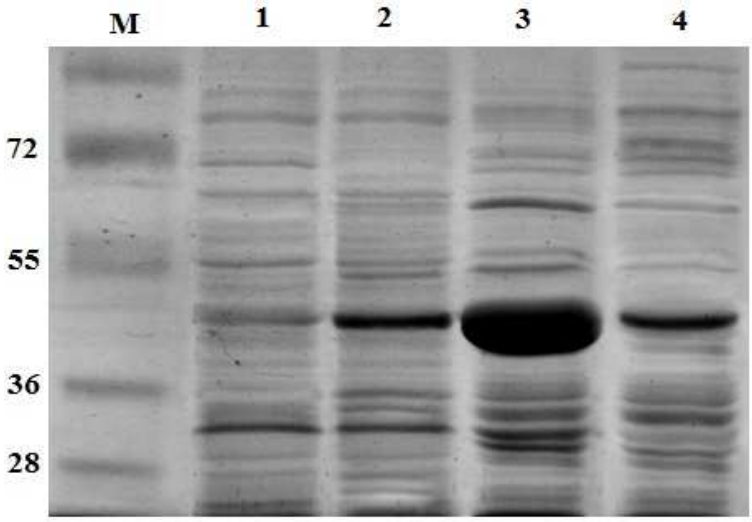

Figure 10: Bacterial expression of phytase $appA_S$ of *Shigella* sp. CD2 in *E.coli* BL21(DE3): SDS-PAGE analysis of recombinant phytase ($rAppA_E$). Lane, M- Molecular weight markers in kDa, **1**- soluble fraction of induced BL21 (DE3) transformed with pET20b (+) vector, **2**-pellet fraction of induced BL21 (DE3) transformed with pET20b (+) vector, **3**-soluble fraction of induced BL21 (DE3) transformed with pET20b(+) -$appA_S$, **4**- pellet fraction of induced BL21 (DE3) transformed with pET20b(+) -$appA_S$. Approx. 25 µg protein from the soluble fraction was loaded in each lane of 12 % SDS-PAGE gel.

pellet fraction. The recombinant phytase activity was about 101 folds greater than that of native phytase (5.60 U mg^{-1}) from *Shigella* sp CD2. The result thus indicates a correlation of $rAppA_E$ overexpression with phytase activity.

3.4. Purification and properties of $rAppA_E$ and $rAppA_S$

The $rAppA_E$ was purified from the soluble fraction of *E.coli* BL21 (DE3) using Ni-Sepharose Fast Flow affinity chromatography and $rAppA_P$ was purified by cation exchange chromatography of diafiltered extracellular fraction of pPIC9-$appA_S$ transformed *P. pastoris* GS115. The specific activities of purified $rAppA_P$ and $rAppA_E$ were 967 and 2982 U mg^{-1}, respectively. Western blot analysis of purified $rAppA_E$ and purified deglycosylated rAppAp using rabbit polyclonal antibody against *E.coli* phytase further demonstrated that the specific band with apparant molecular mass of 45 kDa was recombinant phytase **(Figure 11)**. Compared with the glycosylated $rAppA_P$, the non-glycosylated $rAppA_E$ was more active at pH 3.5- 7.5. Both the enzymes had more than 50 % activity in the pH range 3.5 to 6.5 with pH optima at 5.5 **(Figure 12a)**. Both $rAppA_E$ and $rAppA_p$ had temperature optima of 60°C. Compared to $rAppA_P$, $rAppA_E$ had 11 and 18 % greater relative activity at 37 and 50°C, respectively, whereas at higher incubation temperature $rAppA_P$ was more active than $rAppA_E$ **(Figure 12b)**. For determination of thermal stability, the purified

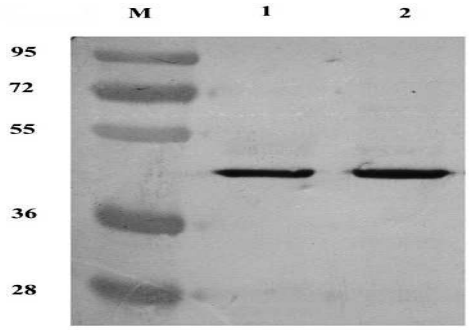

Figure 11: Western blot analysis. Lane, M- Molecular weight marker, 1-purified rAppA$_E$, 2-purified and deglycosylated rAppA$_P$.

rAppA$_E$ or rAppA$_p$ were pre-incubated at 10 to 80^{o}C for 30 min and then assayed for enzymatic activity. Although, the two enzymes didn't differ in their thermostability in the temperature range 10 to 50°C, rAppA$_p$ was more thermotolerant at higher temperature. Consequently at 60 and 70° C, phytase activity of rAppA$_p$ was 33 and 24 % higher than that of rAppA$_E$, respectively **(Figure 12c)**. K_m values for phytate as determined by Lineweaver-Burk plot were 0.18 and 0.22 mM for the rAppA$_E$ and rAppA$_P$, respectively **(Table 1)**.

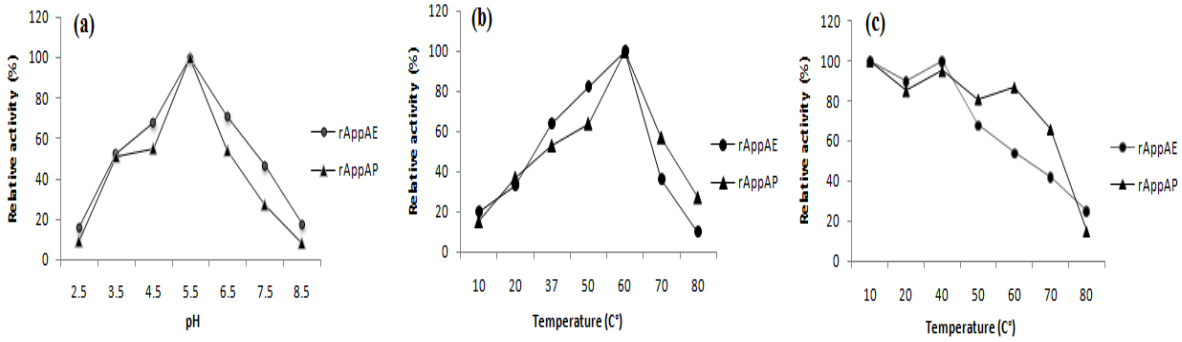

Figure 12: Characterization of purified recombinant phytases: **(a)** pH profile of rAppA$_E$ and rAppA$_P$. The effect of pH was determined in following buffers (100 mM): glycine-HCl (pH 2.5 and 3.5), sodium acetate (pH 4.5 and pH 5.5), and Tris-HCl (pH 6.5, 7.5 and 8.5). **(b)** Temperature profile of rAppA$_E$ and rAppA$_P$ at optimum pH and 10 to 80^{o}C. **(c)** Thermal stability of rAppA$_E$ and rAppA$_P$. Purified enzyme preparation was pre-incubated at indicated temperature for 30 min followed by determination of enzymatic activity as described in 'Materials and Methods' section.

Both the rAppA$_E$ and rAppA$_P$ showed high specificity to the substrate, exhibiting 100 % relative activity with sodium phytate. Activity with other phosphorylate substrates, such

as ATP, ADP, p-NPP, G6-P or F6-P was almost negligible. K_m values for phytate as determined by Lineweaver- Burk plot were 0.22 and 0.18 mM for the $rAppA_E$ and $rAppA_P$, respectively (Table 1). Analysis of the effect of various metal ions on phytase activities revealed that Ca^{2+}, Mg^{2+}, and Mn^{2+} had a stimulatory effect, as higher activity (100–130 % relative phytase activity) was detected from the enzymes incubated in the presence of these ions. On the other hand, Cu^{2+}, Fe^{2+}, Zn^{2+} or EDTA showed inhibitory effect toward the recombinant phytases. To determine the protease resistance the purified recombinant phytases (50 U) were pre-incubated separately with 30 U of either pepsin or trypsin at 37°C. The $rAppA_E$ and $rAppA_P$ retained 70 and 65% activity in presence of trypsin, and 55 and 50% of activity in presence of pepsin, respectively.

Table 1: Properties of phytase AppAs from *Shigella* sp.CD2 expressed in *E.coli* ($rAppA_E$) and *P. pastoris* ($rAppA_P$).

Properties	Results	
	$rAppA_E$	$rAppA_P$
*Substrate specificity (Sodium phytate)	100%	100%
K_m Phytate (mM)	0.22	0.18
Specific activity of purified enzyme (U mg^{-1} protein, 37°C)	2982	967
#Activity in presence of trypsin	70%	65%
#Activity in presence of pepsin	55%	50%
Activity in presence of metal ions (20 mM):		
Ca^{2+}	130%	105%
Mg^{2+}	125%	110%
Mn^{2+}	109%	102%

*Activity in presence of ATP, ADP, P-NPP, G6P, F6P was negligible.

Recombinant enzyme (50 U) was pre-incubated with pepsin or trypsin for 30 min followed by determination of phytase activity.

3.5. Optimization of phytase (rAppA$_P$) production by *P.pastoris*

Phytase production by *P.pastoris* was optimized with respect to carbon source, nitrogen source, medium pH, incubation temperature and time.

---

iii. Has the progress been according to original plan of work and towards achieving the objectives, if not, state reasons

iv. Please indicate the difficulties, if any, experienced in implementing the project__

v. If project has not been completed, please indicate the approximate time by which it is likely to be completed. A summary of the work done for the period (Annual basis) may please be sent to the Commission on a separate sheet.

vi. If the project has been completed, please enclose a summary of the findings of the study. One bound copy of the final report of work done may also be sent to University Grants Commission.

Summary and Conclusion

Phosphorus is predominately stored in mature seeds of cereals and legumes as organic-mineral complex known as phytate (Kumar et al. 2010). Phytate-phosphorus is poorly available to monogastric animals, though the primary constituents of diets for poultry, fishes and swine are plant-based ingredients which come primarily from the seeds (Lei et al. 2003). Phytase catalyses the hydrolysis of phytic acid to inorganic monophosphate and lower myo-inositol phosphate, and in some cases to free myo-inositol and thus its supplementation to the plant based animal feed can make phytate phosphorus available to the animals (Tan et al. 2014). Efficacy of phytase supplementation is dependent on several factors such as microbial source, temperature and pH optima and other kinetic properties of the enzyme, specificity to substrate, thermal tolerance during processing of feed, resistance to proteolytic break down by digestive proteinases encountered in the digestive tract. Hence, there is a constant search for a better commercial phytase with activity in acidic to neutral range. Bacterial phytases are an alternative to the fungal enzymes because of their higher specific activity, thermal stability, higher substrate specificity, greater resistance to proteolysis and better catalytic efficiency.

In the previous project from UGC (No. F.No.32-546/2006, dated: 1st March, 2007), a cell bound phytase was isolated from *Shigella* sp. CD2. The enzyme was purified with specific activity and purification fold of 780 U mg^{-1} and 133, respectively. The enzyme was active in the pH range 3.5 to 7.5 and temperature range 20 to 80°C with pH and temperature optima of 5.5 and 60°C, respectively. It showed high level of specificity to phytate and retained activity on treatment with proteolytic enzyme. Present project aimed at cloning, sequence analysis of phytase encoding gene and its expression of *P.pastoris*. The characteristic properties of phytase expressed by *Pichia* were compared to that expressed in *E.coli*. The gene and encoded protein $AppA_S$ showed highest homology with the $AppA_S$ phytase of *E.coli* and *C.braakii*. The $AppA_S$ amino acid sequence contained the conserved active site motifs, RHGXRXP and HDTN, and five conserved cystein residues, placing it in histidine acid phosphatase (HAP) family of phytase. Moreover, $AppA_S$ contained putative signal peptide of 22 amino acids indicating its periplasmic localization, and three potential sites of N-glycosylation. The calculated molecular mass of the protein with and without the signal sequence were about 47 and 45 kDa, respectively. The $appA_S$ ORF without the signal sequence was PCR amplified, cloned in pPIC9 vector for overexpressed in *P.pastoris* GS115. The recombinant phytase $AppA_P$ was expressed as extracellular protein with maximum activity level of 62 U ml^{-1} with specific activity of 477 U mg^{-1} protein at 60 h of methanol induction. The protein was glycosylated with molecular mass of 59 and 65 kDa and deglycosylation by Endo H glycosidase resulted in reduction in molecular mass of the enzyme to about 45 KDa. To examine the effect of glycosylation on enzymatic properties of $rAppA_P$, $appA_S$ was expressed in *E.coli* to produce $rAppA_E$. For this $appA_S$ ORF without the signal sequence was cloned in pET-20b(+) under the control of T7 promoter and recombinant plasmid was transformed into *E.coli* BL21(DE3) to produce $rAppA_E$. The $rAppA_E$ was overexpressed as intracellular protein with maximum activity of 176 U ml^{-1} and specific activity 568 U mg^{-1} protein at 24 h of IPTG induction. The recombinant glycosylated $rAppA_P$ and nonglycosylated $rAppA_E$ were purified with specific activity of 967 and 2982 U mg^{-1}protein, respectively. Thus, $rAppA_P$ represented 50% of the total secreted protein of recombinant *P.pastoris* GS115, whereas $rAppA_E$ was 19 % of the total intracellular protein of recombinant E.coli BL21(DE3). The biochemical properties of the purified phytases were compared. Both the enzymes had more than 50 % activity in the pH range 3.5 to 6.5 with pH optima at 5.5. Both $rAppA_E$ and $rAppA_p$ had temperature optima of 60°C. Compared to $rAppA_p$, $rAppA_E$ had 11 and 18 % greater relative activity at 37 and 50°C, respectively, whereas at higher incubation temperature $rAppA_p$ was more active than $rAppA_E$. Although, the two enzymes didn't differ in their thermostability in the temperature range 10 to 50°C, $rAppA_p$ was more thermotolerant at

higher temperature. Both the $rAppA_E$ and $rAppA_P$ were highly specific to sodium phytate as substrate and activity with either of phosphorylated substrates, such as ATP, ADP, pNPP, dSPP, G6P or F6P was negligible. The divalent cations, Ca^{2+}, Mg^{2+}, and Mn^{2+} had a stimulatory effect on phytase activity with 100-130 % relative activity was detected in presence of these ions. On the other hand, Cu^{2+}, Fe^{2+}, Zn^{2+} or EDTA showed inhibitory effect toward the recombinant phytases. The $rAppA_E$ and $rAppA_P$ retained 70 and 65 % activity in presence of trypsin, and 55 and 50% of activity in presence of pepsin, respectively, indicating greater resistance to trypsin.

In conclusion, phytase encoding gene $AppA_S$ was cloned and expressed in *P.pastoris* and *E.coli* to produce $rAppA_E$ and $rAppA_P$, respectively. The purified recombinant enzymes almost retained the biochemical properties of the native phytase, except for thermal stability. The $rAppA_P$ was more thermostable at higher temperature than $rAppA_E$ and native enzyme. Phytase AppA from *Shigella* sp.CD2 displayed 40-70 % activity in the pH range 3.5 to 6.5, which can facilitate phytate degradation in salivary gland (pH 5.0 – 7.0), stomach (fed state pH 6.5, reducing to 3.5 – 4.5 upon stimulation of acid secretion) and upper part of duodenum (pH 4.0 - 6.0). Hence, the enzyme can be used as feed additive for improving the utilization of phytate phosphorus by monogastric animals like, swine, poultry and farm animals. However, economical production of $rAPPA_P$ requires improving its expression by optimization of bioprocess and scaling up when the cells are grown in a fermenter.

vii. Any other information which would help in evaluation of work done on theproject. Atthe completion of the project, the first report should indicate the output, such as (a) Manpower trained (b) Ph. D. awarded (c) Publication of results (d) other impact, if any

(a). Manpower trained: The project fellow was trained on techniques related to Biochemistry and Molecular Biology research.

(b). PhD awarded: One thesis entitled “Isolation, characterization, molecular gene cloning and expression of novel bacterial phytase from environmental samples” has been submitted for award of PhD.

(c). Publication of results:

- Thesis: Isolation, characterization, molecular gene cloning and expression of novel bacterial phytase from environmental samples
- Moushree Pal Roy, Deepika Mazumdar, Subhabrata Dutta, Shyama Prasad Saha, Shilpi Ghosh* Cloning and expression of phytase appA gene from *Shigella* sp.CD2 in *Pichia pastoris* and comparison of properties of recombinant enzymes in *E.coli*

IGNATURE OF PRINCIPAL INVESTIGATOR

(Seal)

Dr Shilpi Ghosh
Dept. of Biotechnology
University of North Bengal

SIGNATURE OF REGISTRAR

15.01.19

Registrar
University of North Bengal

(Seal)

ı *Pichia*
mpt to

gal

pressed