Chapter 3

Cloning, expression, purification, biophysical and biochemical characterization of Glyceraldehyde 3-phosphate dehydrogenase (GAPDH) from *Lactobacillus acidophilus*.

3.1. Introduction

Glyceraldehyde-3-Phosphate Dehydrogenase (GAPDH), EC 1.2.1.12, is a classical conserved intracellular glycolytic enzyme. It catalyses the oxidation and phosphorylation of G3P (glyceraldehyde 3-phosphate) to 1, 3-bisphosphoglycerate in the presence of NADH/ NADPH and inorganic phosphate (Harris & Waters, 1976). GAPDH is a ubiquitous enzyme found in eukaryotes, prokaryotes and viruses and primary function of GAPDH is its involvement in glycolysis process. Interestingly, GAPDH is considered as a principal "moonlighting" protein/multifunctional enzyme with additional roles in regulation of apoptosis (Hara *et al.*, 2005), iron homeostasis (Sheokand *et al.*, 2013), adherence to human colon (Kinoshita, Uchida, *et al.*, 2008), transcription activation (Zheng *et al.*, 2003) and host-pathogen interactions (Bergmann *et al.*, 2004) as summarized in figure 3.1. Proteins like GAPDH with multiple roles are known as "moonlighting proteins", a term first coined by C. Jeffery (Jeffery, 1999) and the multi-functions attributed to these proteins are mostly related to their cellular localization, oligomeric state, concentration of substrates, cofactor and ligands, type of cells or post-translational modifications (Amblee & Jeffery, 2015).

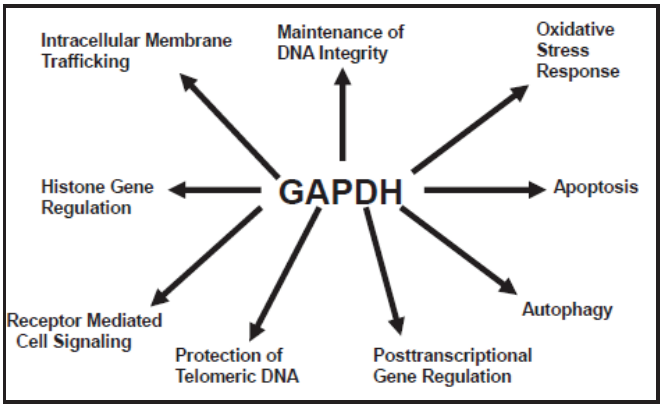


Figure 3.1: Functional diversity of GAPDH. Source: (Sirover, 2011).

GAPDH have been found on the cell surfaces, intracellularly and as secreted proteins in culture media in several bacterial species (Table 3.1). GAPDH is also commonly known as anchorless molecule on the cell surface which is involved in adhesion apart from its regular glycolytic function (Glenting *et al.*, 2013). It is also reported to bind to host epithelial mucus, extra-cellular matrix proteins like fibronectin and plasminogen or to epithelial cell lines like Caco-2 and HT29 thus enabling it to act as an essential component in host-microbe interactions (Kinoshita, Uchida, *et al.*, 2008; Pancholi & Fischetti, 1992; Ramiah *et al.*, 2008).

Table 3.1: Extracellular GAPDH reported in bacteria and their host targets.

Bacteria species	Host target	References
Streptococcus pyogenes	Fibronectin and Laminin.	(Boel *et al.*, 2005)
Streptococcus pneumoniae	Plasminogen	(Bergmann *et al.*, 2004)
Bacillus anthracis	Plasminogen	(Matta *et al.*, 2010)
Mycoplasma pneumoniae	Fibrinogen	(Dumke *et al.*, 2011)
Mycoplasma genitalium	Mucin	(Alvarez *et al.*, 2003)
Lysteria monocytogenes	Rab5a	(Alvarez-Dominguez *et al.*, 2008)
Escherichia coli (EHEC and EPEC)	Fibrinogen and plasminogen.	(Egea *et al.*, 2007)
Neisseria meningitides	Epithelial and endothelial cells	(Tunio *et al.*, 2010)
Lactobacillus plantarum	Mucin	(Kinoshita, Uchida, *et al.*, 2008)
Candida albicans	Plasminogen, fibrinogen and laminin.	(Crowe *et al.*, 2003)
Staphylococcus aureus & Staphylococcus epidermis	Transferrin and plasminogen.	(Ebner *et al.*, 2016)

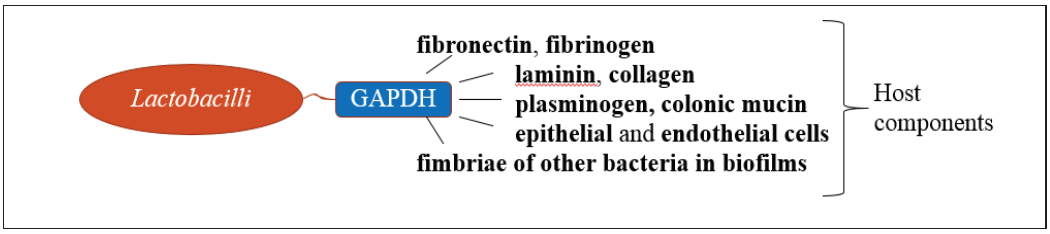


Figure 3.2: Schematic representation of GAPDH affinity to ECM components.

3.2. Materials

3.2.1. Chemicals and Reagents

The different chemicals and reagents used in the current study were of molecular biology grades. Enzymes used in cloning of genes like restriction endonucleases, TAQ polymerases and DNA ligases were purchased from Fermentas and Roche. Chemicals used in protein expression and purification like Ni-NTA agarose resins, IPTG, Triton X-100, PMSF protease inhibitor, SDS, glycine, Trizma base, etc. were purchased from Sigma-Aldrich. The protein molecular marker and DNA ladders were procured from Fermentas. Molecular biology grade chemicals like sodium chloride, imidazole, sodium phosphate, potassium chloride, etc. were purchased from HiMedia. The antibiotics ampicillin, kanamycin and chloramphenicol, were obtained from HiMedia. Media for culturing of bacteria viz. Luria-Bertani (LB), yeast extract, agar, tryptone, peptone, de Man Rogosa and Sharpe (MRS) media were purchased from HiMedia, India. Paraffin oil for anaerobic culturing was acquired from HiMedia, India. LB and MRS agar plates were prepared by supplementing media with 0.75-1 % agar. The sugars used for the studies were purchased from Sigma and HiMedia. The glycoproteins fetuin and type III mucin was purchased from Sigma. Bovine serum albumin and lysozyme was purchased from Sigma. Red blood cells (RBCs) for hemagglutination assay were isolated from rabbit and fresh human blood where as required. Rabbit blood was a kind gift from Dr Manish Nivsarkar, PERD Centre, India and Mr Bhavesh from Nirma University, India. The molecular weight markers used for molecular weight determination by gel filtration were composed of β-amylase (200kDa), alcohol dehydrogenase (150kDa), bovine albumin (66kDa), carbonic anhydrase (29kDa), cytochrome c (12.5kDa) obtained from Sigma-Aldrich.

3.2.2. *Lactobacillus* strain and genomic DNA isolation.

The *Lactobacillus acidophilus* strain was procured from National Dairy Research Institute (NDRI), Karnal, India. The genomic DNA was isolated using a standard procedure, and the strain was further confirmed by 16s rRNA amplification using standard universal primers – 27F (forward) and 1107R (reverse) and sequencing the amplicon. Chemicals used for genomic DNA isolation was purchased from HiMedia and oligos for 16s rRNA were synthesised from Xcelris Lab, India.

3.2.3. Cloning vectors and bacterial strains

The following cloning and expression vectors were used pET15b (a kind gift from Dr Hemanta Majumdar, IICB, Kolkata, India), pET28a (a kind gift from Dr Ashima Bhardwaj, IIAR) and pET30a (a kind gift from Dr C. Ratna Prabha, M.S. University of Baroda, India). *E. coli* cloning strains DH5α is a kind gift of Dr G. S. Khatri, IIAR and expression strain Rosetta (BL21-DE3) is a kind gift from Prof. M. Vijayan, IISc, India.

3.3. Methods

3.3.1. Genomic DNA isolation and strain confirmation

3.3.1.1. Genomic DNA isolation: *Lactobacillus acidophilus* was cultured from the stab in sterile MRS media and colonies obtained. The anaerobic conditions in the liquid culture media were maintained by creating a layer of sterile paraffin oil, and Glycerol stocks of cells were made by supplementing 15% glycerol in culture and subsequently stored at -80°C for long term storage. The cells were revived from a glycerol stock and re-inoculated into fresh MRS sterile media and incubated at 37°C for 24 hours. 1% pre-culture was used to inoculate 50ml MRS fresh media and incubated at 37°C for 12 hours.

For isolation of genomic DNA, 30ml of culture was centrifuged at 12000 x g for 5 minutes at 4°C, the supernatant discarded and pellet collected which was further washed with fresh MRS media. The pellet was re-suspended in buffer containing 6.7% w/v sucrose, 50mM Tris-HCl and 20mM EDTA at pH 8.0, with the addition of freshly prepared lysozyme (10mg/ml). The suspension was incubated at 37°C for 1-2 hour with intermittent shaking and then incubated for 1 hour at 55°C with the addition of lysis solution which consists of 500µl of 10% SDS and 100µl of Proteinase K (20mg/ml). After this 2ml of 5M NaCl, 6ml of chloroform-isoamyl alcohol mixture (24:1) was added and the mixture incubated at room temperature for 5 minutes. The mixture was subjected to centrifuge at 12000 x g for 10 minutes, and the upper aqueous layer was transferred to a fresh tube. To this solution add an equal volume of isopropanol alcohol and centrifuge at 12000 x g for 15 minutes. The DNA fragments were washed with 70% chilled ethanol and the pellet was air-dried. 5µl of RNase-A solution (100µg/ml) was added to remove RNA contamination and incubated for 1 hour at 37°C. To this mixture add an equal volume of TE-saturated phenol and centrifuge at 12000 x g for 10 minutes and further transfer the upper aqueous phase carefully to the fresh tube and repeat this step if necessary. To this mixture add an equal volume of phenol, chloroform and isoamyl alcohol mixture (25:24:1), centrifuged at 12000 x g for 5 minutes and transferred the upper layer into a fresh tube. DNA in the supernatant was precipitated out with two volumes of 100% ethanol and 3M sodium acetate, pH 5.2. The pellet thus obtained was dissolved in 50µl Tris-EDTA pH 7.5 and the concentration and purity was determined by recording absorbance at 260 nm (A_{260}) using Bio spectrophotometer (Eppendorf) and by resolving on 0.8% agarose gel electrophoresis

3.3.1.2. Strain confirmation: The 16s rRNA gene was amplified using PCR method from the genomic DNA using 27F (5’- GAGAGTTTGATCCTGGCTCCAG-3’) and 1107R (5’-GCTCGTTGCGGGACTAAACC-3’) as forward and reverse universal primers. The PCR reaction mixture contained 1X reaction buffer, 1mM of dNTP mix, 100ng of genomic DNA, 3mM $MgSO_4$, 10µM of primers each, 1U of Taq DNA polymerase enzyme in the 25µl reaction mixture. PCR amplification was performed using DNA thermal PCR cycler (Applied Biosystems, USA) with following parameters - initial denaturation for 5 minutes at 94ºC, followed by 35 cycles of denaturation for the 30s at 94ºC, annealing at 55ºC for 30s, extension at 72ºC for 90s and final denaturation at 72ºC for 10 minutes. The amplified gene was resolved on 0.8% agarose gel, isolated and purified and further subjected to sequencing for confirmation of strain.

3.3.2. Cloning and expression of recombinant LaGAPDH (r-LaGAPDH).

3.3.2.1. PCR amplification: The genomic DNA isolated from *L. acidophilus* strain was used as a template for the PCR amplification of the gene LBA0698 (UniProt: Q5FL51) encoding LaGAPDH using following forward and reverse primers (Table 3.2). The PCR reaction mixture (25µl total volume) consists of 1X Taq reaction buffer, 2mM $MgCl_2$, 50-100ng of genomic DNA, 0.5mM dNTPs mix, 10µM of primers each and 1U Taq DNA polymerase enzyme. The amplification was performed using DNA thermal PCR cycler (Applied Biosystems, USA) programmed with following parameters - initial denaturation for 2 minutes at 94ºC, followed by 35 cycles of denaturation for the 30s at 94ºC, annealing at 50ºC for 35s, extension at 72ºC for 90s and final denaturation at 72ºC for 10 minutes. The amplified product was resolved on 0.8% agarose gel electrophoresis followed by purification of amplified gene using PCR gel extraction kit (Qiagen) as per manufacturer's protocol.

Table 3.2: Description of vectors and respective primers used for LaGAPDH cloning.

Plasmid/Vector	Antibiotic marker	Promoter	Tag	Primers (restriction sites are underlined)	
pET15b	Ampicillin	T7 *lac*	HIS	F	GTA CATATG ATGGCATTTGACGGATTA
				R	GTA GGATCC TTAAGCCATAATTTCGTCTTC
pET28a, pET30a	Kanamycin	T7 *lac*	HIS	F	GTA CATATG ATGGCATTTGACGGATTA
				R R	GTA GGATCC TTAAGCCATAATTTCGTCTTC GTA AGATCTG TTAAGCCATAATTTCGTCTTC

CATATG – Nde*I; GGATCC* – BamH*I; AGATCTG* - Bgl*II*

3.3.2.2. Restriction digestion and ligation reaction: The plasmid vector and the PCR amplified product of LaGAPDH (1014 bp) were double digested with restriction endonucleases and incubated at 37ºC for 16 hours. The digested product was resolved on 0.8% agarose gel and purified plasmid and amplicon were extracted using PCR gel extraction kit as per manufacturer's protocol, and the final ligation reaction of amplicon and plasmid (1:3 molar ratio) was done using T4 DNA ligase in 1X T4 DNA ligase buffer at 16ºC overnight. The purified cloned plasmid was resolved in 0.8% agarose gel electrophoresis and a shift in plasmid band viz. control plasmid without ligation was a positive indication of a clone which was further confirmed by colony PCR and restriction digestion to visualise insert release, and the clone was finally confirmed through sequencing. The vector maps with the cloned genes are shown in Figure 3.3.

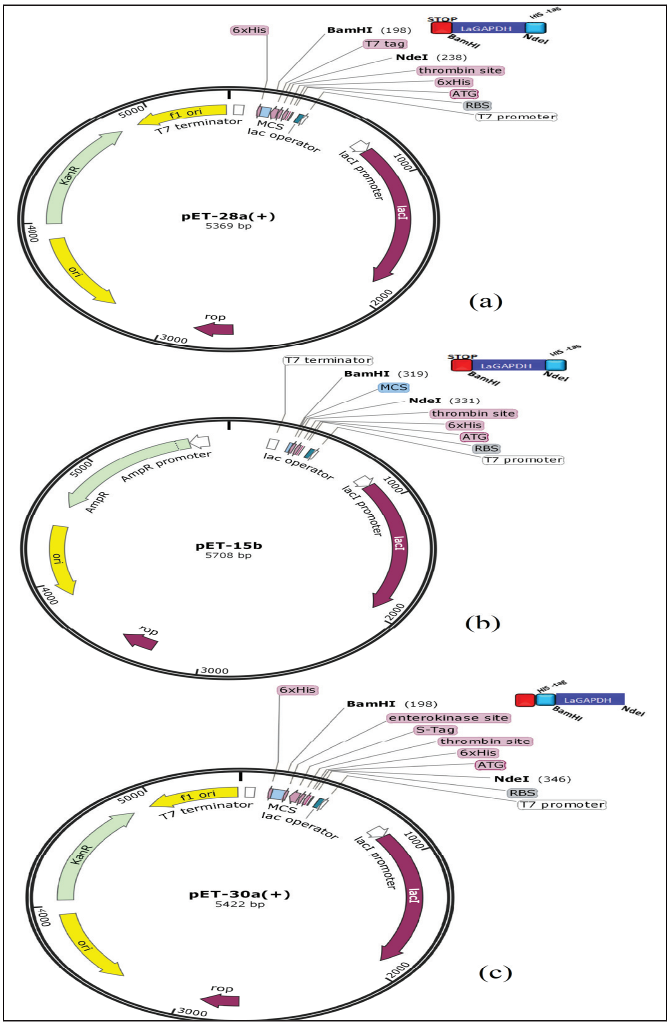


Figure 3.3: The vector map used in cloning of LaGAPDH: (a) The vector map of pET28a showing the restriction sites and directional cloning strategy used for cloning LaGAPDH gene; (b) The vector map of pET15b showing the restriction sites and directional cloning strategy used for cloning LaGAPDH gene; (c) The vector map of pET30a showing the restriction sites and directional cloning strategy used for cloning LaGAPDH gene. Source: Image of vector map generated using SnapGene tool (www.snapgene.com/resources).

3.3.2.3. Chemical competent cells preparation and clone transformation into *E.coli*:

The *E. coli* DH5α cells were maintained on LB agar plates or as 15% glycerol stock stored and used as required. 50ml LB media was inoculated with *E. coli* DH5α cells and cultured at 37ºC until the absorbance O.D. (A600) reading was 0.5. The cells were then chilled on ice for 30 minutes before subjecting to centrifuge at 4ºC and 3000 X g. The pellet was resuspended in 0.1M$CaCl_2$ and kept on ice for 15 minutes and subjected to centrifuge at 4ºC at 3000 X g. The pellet thus obtained was resuspended in a solution of 0.1M $CaCl_2$ and 15% glycerol and stored at -80ºC in aliquots of 50µl each.

After the ligation reaction, vectors were transfected into competent *E.coli* DH5α cells by heat shock method. The cell suspension with the recombinant vectors was cooled on ice for 30 minutes and then subjected to heat shock at 55ºC for the 60s followed by incubation on ice for a further 10 minutes. 1ml of fresh LB media was added to the mixture and cells were allowed to culture for 1 hour at 37ºC and then the harvested cells were plated on LB agar plates containing the antibiotics as selection marker and incubated at 37ºC overnight. The growing colonies on the plate were screened manually for positive clones carrying the desired insert into the plasmid. Plasmid were isolated from the individual colony using Alkaline-Lysis method mentioned in Molecular Cloning (Sambrook & Russell, 2012).

3.3.2.4. Heterologous expression of r-LaGAPDH: The confirmed recombinant plasmid as selected in colonies mentioned above was isolated by plasmid isolation kits and transformed into chemically competent *E. coli* Rosetta BL21 (DE3) expression strains. The transfected cells were cultured on LB agar media plate supplemented with 50µg/ml of antibiotics like ampicillin, kanamycin and chloramphenicol used as per requirement of strains and plasmid and a single colony was inoculated into 5ml of LB media supplemented with required antibiotics and culture overnight at 37°C. Further 500ml of LB media was inoculated with 1% of this primary culture and cells cultured until the A_{600} absorbance O.D. reached 0.6. The protein expression was then induced by the addition of IPTG (0.1 – 2mM) and further incubation of transformed cells with different temperature from 16 to 37°C with incubation time from 4 hours to overnight to get the optimum conditions.

The cells were harvested by subjecting to centrifuge at 10000 x g for 20 minutes at 4°C, and the obtained pellet was suspended in lysis buffer consisting of 50mM Tris-HCl pH 7.4, 150mM NaCl, 10% glycerol, 0.1% Triton X-100, 20mM Imidazole and 1mM PMSF. The cell suspension in this buffer was subjected to sonication using sonicator (Vibra-cell Sonics VCX750, USA) with 10s pulse on and 20s pulse off for 5 minutes. The lysate was further centrifuged at 10000 x g for 20 minutes at 4°C to separate soluble and insoluble fraction. The protein samples were analysed by SDS-PAGE gel electrophoresis using 15% polyacrylamide gel and resolved on a Mini-Protean II system (Bio-Rad, Hercules, USA) following Laemmli method (Laemmli, 1970).

3.3.2.5. Purification of r-LaGAPDH using affinity chromatography: The soluble fraction of lysate was further subjected to affinity chromatography for purification of r-LaGAPDH using the Ni-NTA affinity resins. The lysate was passed through Ni-NTA column pre-equilibrated with 10 column bed volumes of equilibration buffer (50mM Tris-HCl pH 7.4, 150mM NaCl and 20mM Imidazole). Unbound non-specific protein was removed by washing the column with 20 column bed volumes of washing buffer (50mM Tris-HCl pH 7.4, 150mM NaCl and 30mM Imidazole). Finally, the bound r-LaGAPDH was eluted from the column using elution buffer whose composition is same as above buffers with 250mM Imidazole and the eluted purified protein samples were analysed by running on 15% SDS-PAGE electrophoresis.

3.3.3. Biochemical characterization of r-LaGAPDH.

3.3.3.1. Hemagglutination and carbohydrate specificity: The hemagglutination assay was performed using rabbit RBCs suspension. Briefly, 50µl of purified protein in PBS solution (1mg/ml), 50µl of 2% rabbit RBCs suspension in PBS and 50µl of PBS was added to wells making up the total reaction of 150µl in U-shaped 96 well microtiter plate and incubated the mixture at room temperature for one hour. Visual examination of RBC sedimentation (button formation) was observed after one-hour incubation at room temperature where the button formation / sedimentation indicated no agglutination and vice versa. Hemagglutination is expressed in hemagglutination units (HU) which are defined as the reciprocal of the highest dilution showing detectable hemagglutination and the specific activity was calculated by dividing the activity in HU by the protein concentration (mg ml^{-1}) i.e. as the number of hemagglutination unit per mg of protein.

For inhibition assay, sugars in the concentration range of 1–150mM in PBS were added to the wells instead of blank. The following sugars were tested: D-mannose, D-galactose, Lactose, N-acetyl-D-glucosamine, N-acetyl-D-galactosamine, sialic-acid, fucose, D-glucose along with complex glycoprotein-like 1% fetuin and type-III mucin. The minimum concentration of sugar showing detectable button formation in the assay was taken as minimal inhibitory concentration (MIC).

3.3.3.2. Effect of temperature: The effect of temperature on the hemagglutination activity of the purified r-LaGAPDH was determined by incubating different aliquots of purified protein in PBS at a temperature ranging from 30 to 90 °C for 30 minutes and then cooling the tubes on ice for half an hour followed by the hemagglutination assay as described above to evaluate its hemagglutination activity.

3.3.3.3. Mucin binding assay: The wells of a 96-well untreated microtiter plate were coated with 200µl of 3mg/ml of mucin prepared in PBS and incubating the plates for 1 hour at 37°C followed by overnight incubation at 4°C. The wells were then washed with PBS to remove unbound mucin and then non-specific blocking was done by addition of 1% BSA in PBS to the wells and incubating for 1-2 hours at 37°C. Subsequently, the wells were washed with PBS twice to remove unbound BSA and further 20µl of 5mg/ml r-LaGAPDH protein in PBS was added to each well and incubated for 2 hours at 37°C for the binding reactions, and the unbound proteins were removed by washing twice with PBS. To extract the bound proteins 60µl of 1% (w/v), SDS was added with incubation for 1 hour at 37°C, and the resulting suspension was analysed by SDS-PAGE with commasive blue staining or silver staining if required. As a negative control r-LaGAPDH alone was used in the wells which were not coated with mucin and treated similarly.

3.3.3.4. Enzyme kinetic studies of r-LaGAPDH: The enzymatic assay of LaGAPDH activity was done according to Pancholi and Fischetti (Pancholi & Fischetti, 1992). Since GAPDH catalyses the oxidative phosphorylation of G3P (glyceraldehyde-3-phosphate) to 1, 3-diphosphoglycerate with the presence of inorganic phosphate and NAD (B-nicotinamide adenine dinucleotide), the activity was measured as an increase in absorbance at 340 nm due to NADH formation. The assay buffer consists of 40mM triethanolamine, 50 mM Na_2HPO_4 and five mM EDTA and the final assay reaction contained 2mM G3P, 1mM NAD, 180µM of r-LaGAPDH protein in the assay buffer to make the final volume of 200µl at pH 8.6. Enzyme kinetic activity of r-LaGAPDH (110 µM) was determined with varying concentrations of NAD and with a fixed concentration of G3P and vice versa respectively to determine the Km and Vmax for NAD and G3P. The increase in the absorption of NADH formation was measured at 30s interval for 3 minutes at 340 nm. The pre-determined molar extinction coefficient of NADH (6.22×10^3 $M^{-1}cm^{-1}$) was used for the calculations (Horecker & Kornberg, 1948). The kinetic coefficients K_m and V_{max} were calculated using the Michaelis-Menten plot on each substrate using Graph Pad Prism software's enzyme kinetics equation. All experiments were performed in replicates, and final values were represented with standard error mean. The GAPDH activity was also measured in the cell-free media supernatant to detect whether GAPDH is secreted by *L. acidophilus* native cells cultured in MRS media. The enzyme reaction consists of the total proteins isolated from extracellular media, and the total enzyme activity was estimated with presence and absence of G3P which serves as a control.

3.3.4. Biophysical characterization

3.3.4.1. Peptide mass fingerprinting (MALDI MS/MS): Mass spectrometry is a highly sensitive technique, and the mass spectrometer consists of an ionisation source, which ionises the sample, an analyser, which separates the ions according to their mass to charge ratios (m/Z), and a detector, which detects the ions. The purified r-LaGAPDH was analysed by SDS-PAGE gel electrophoresis and the band corresponding to respective molecular weight size was excised from the destained gel and further in-gel digested with trypsin. The trypsinized peptides fragments were subsequently subjected to MALDI-TOF using AB SCIEX LC-MS/MS system (AB-SCIEX, USA). The time-of-flight (TOF) analyser separates the ions according to their m/z ratios, by measuring the time taken for the ions to pass through a field-free region with 99.9% accuracy. The peptides peak list generated by 4000 Series Explorer software was searched against the MASCOT search engine (http://www.matrixscience.com) to identify them and compare with theoretically digested peptides from known NR protein set.

3.3.4.2. Circular dichroism spectroscopy: The circular dichroism (CD) spectra can be used to estimate the secondary structural elements of protein since which are based on the differential absorption of the left and right circularly polarised components of pLane polarised light. The CD spectra of r-LaGAPDH were recorded using Jasco J-815 spectrophotometer (Japan Spectroscopic, Japan). The protein was dialysed against 20mM sodium phosphate buffer at pH 7.8 and the spectra for far-UV in the wavelength range of 190-250nm was recorded using 1mm quartz cuvette at 20°C, 1nm bandwidth, data integration time of 1s and scanning rate of 100nm/min. The baseline was recorded using a blank buffer without protein and was corrected automatically from the recorded spectra. Each spectrum was an average accumulation of three readings. The machine recorded units of millidegrees ellipticity (mdeg) were used to convert to mean residue molar ellipticity (θ) using the following equation, where n is number of peptide bonds in protein:

$$\theta\ (deg.cm^2.dmoL^{-1}) = \frac{Ellipticity(mdeg).10^6}{Pathlength(mm).Protein\ conc.(\mu M).n}$$

3.3.4.3. Molecular mass identification of r-LaGAPDH: The molecular weight and homogeneity of the purified r-LaGAPDH were analysed on 15% SDS-PAGE gel electrophoresis using standard Laemmli method (Laemmli, 1970) under reducing and non-reducing conditions. For molecular weight estimation, a standard protein molecular weight marker from 14 to 118kDa in size was used as a reference. For exact molecular weight estimation in native condition and to identify its oligomeric state r-LaGAPDH protein (2ml of 5mg/ml in PBS, pH 7.4) was resolved on Superose-6 column (GE Life Sciences) using ÅKTA prime purification system (GE Life Sciences). The total void volume (V_o) was determined by measuring the elution of blue dextran, and elution volume (V_e) was determined by measuring the absorption of the peak at 280nm. The known molecular weight standards (β-amylase 200kDa; alcohol dehydrogenase 150kDa; albumin 66kDa; carbonic anhydrase 29.2kDa) were used as a reference marker. The molecular mass of r-LaGAPDH was estimated from the plot of V_e/V_o against the logarithm of the molecular mass (kDa).

3.3.4.4. Fluorescence spectroscopy: The interaction of different carbohydrates with r-LaGAPDH was investigated by monitoring changes in the fluorescence intensity and ʎmax of the protein induced by binding of carbohydrates. The fluorescence spectrum was recorded on a Hitachi F-7000 FL (Hitachi, Japan) with excitation and emission monochromators slits set by 5nm, the response time and scan speed were 0.5s and 1200nm / minute, respectively. The protein 25µM in PBS, pH 7.5 was excited at 280nm, and the emission spectra were recorded between ʎ 300–400nm. Titrations were carried out by adding 1µl aliquots of carbohydrates stock of 1000mM prepared in PBS to protein solution and mixing thoroughly with autopipettes and avoiding bubble formation. All spectra were recorded at room temperature, and each spectrum was an average of three successive scans. All spectra of titration with carbohydrates were corrected to compensate the dilution effect to the addition of carbohydrates in each titration. The carbohydrates were added, and each spectrum was recorded until the fluorescence emission gets saturated and the change in volume by the addition of carbohydrates remains less than 5% of the solution. The dissociation constant (Kd) of the interaction of each carbohydrate with purified protein was calculated by a known method (Chipman *et* *al.*, 1967). The dissociation constant (Kd) values for different sugars were calculated from the plot of log (Fo − Fc)/(Fc − F∞) against log [S], where Fo, Fc, and F∞ are the fluorescence intensity values of purified protein alone, purified protein in the presence of a concentration [S] of sugar and purified protein saturated with carbohydrate respectively. The relative fluorescence value of r-LaGAPDH saturated with carbohydrate, F∞ was extrapolated from the experimental values by the plot of Fo/(Fo − Fc) against 1/[S]. The final value of Kd is the calculated antilog of the value of log [S] at log (Fo − Fc)/(Fc − F∞) = 0. The final value of Kd was calculated from three independent readings and represented as an average of final Kd values along with the standard error mean.

3.4. Results

3.4.1. Genomic DNA isolation and strain identification

The *L. acidophilus* strain was cultured in MRS media under anaerobic conditions at 37°C for 48-72h, and the purity of the genomic DNA was checked. Protein concentration is estimated by measuring absorbance at 260nm and 280nm respectively. The absorbance ratio of A_{260}/A_{280} was 1.78, and A_{260}/A_{280} ratio in between 1.7-2.0 indicates good quality DNA. The final concentration of genomic DNA used for further experiments was 65.6ng/µl, and the purity of genomic DNA was also checked by resolving it on 0.8% agarose gel electrophoresis (Figure 3.4a). For 16s rRNA sequencing an amplicon of ~1000 bp was amplified from *L. acidophilus* genome (Figure 3.4b) and confirmed through nucleotide sequencing (Figure 3.5a) and subsequently searching in the Ribosomal Database Project (RDP) (https://rdp.cme.msu.edu/) (Figure 3.5b) and BLAST search against NR data reveals the sequence of *L. acidophilus* NCFM strain.

3.4.2. Cloning, expression and purification of r-LaGAPDH from *L. acidophilus*

3.4.2.1. PCR amplification and clone confirmation: The GAPDH gene (LBA0698) was amplified using gene-specific primers (Table 3.2) with *L. acidophilus* genomic DNA as the template. A 1014 bp long LaGAPDH gene was successfully amplified and resolved on 0.8% agarose gel (Figure 3.6a). The PCR amplified product was accordingly digested (using specific restriction enzymes mentioned in Table 3.2) for cloning in to pet15b, pET28a and pET30a vectors. After ligation, the recombinant plasmids were transformed into *E.coli* DH5α cells and the presence of LaGAPDH insert in the selected positive clones were confirmed through plasmid shift (Figure 3.6b), double digestion of clones with the restriction enzymes to release the insert (Figure 3.6c) and finally through nucleotide sequencing confirming the cloning of LaGAPDH into the desired vector (Figure 3.7). The results show the successful ligation of LaGAPDH gene in the vectors.

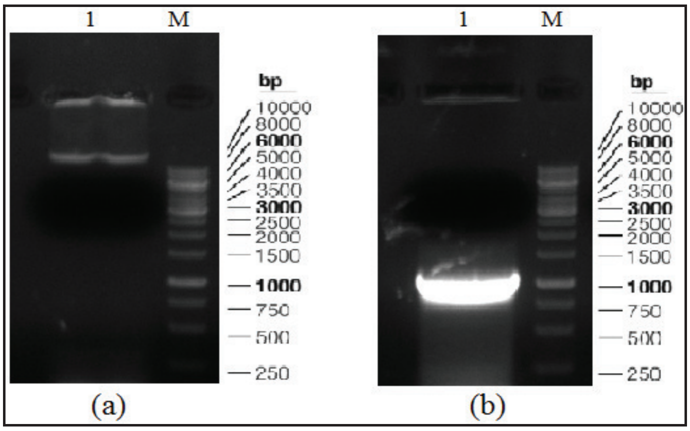
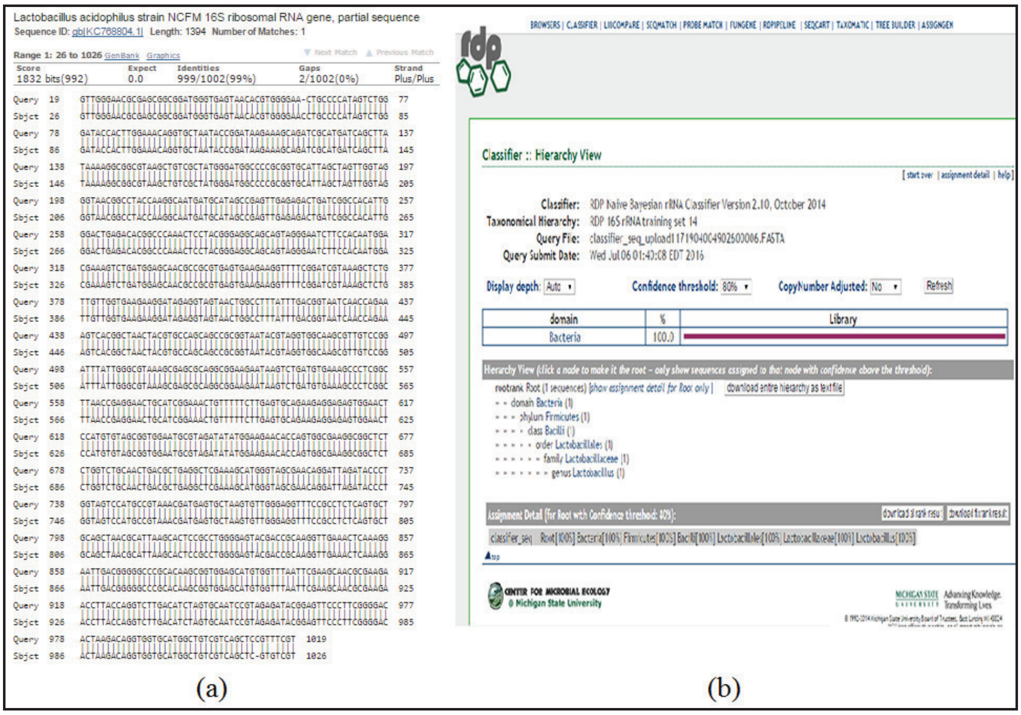


Figure 3.4: Agarose electrophoresis gel for genomic DNA and 16s rRNA amplicon: (a) Genomic DNA of *L. acidophilus*, Lane M: 1kb ladder, Lane 1: genomic DNA sample; (b) PCR amplification of *L. acidophilus* 16s rRNA gene (~1000 bp), Lane M: 1kb ladder, Lane 1: 16s rRNA amplicon.



Lactobacillus acidophilus strain NCFM 16S ribosomal RNA gene, partial sequence
Sequence ID: gb|KC768804.1| Length: 1394 Number of Matches: 1

Range 1: 26 to 1026 GenBank Graphics

Score	Expect	Identities	Gaps	Strand
1832 bits(992)	0.0	999/1002(99%)	2/1002(0%)	Plus/Plus

```
Query  19    GTTGGGAACGCGAGCGGCGGATGGGTGAGTAACACGTGGGGAA-CTGCCCCATAGTCTGG  77
             ||||||||||||||||||||||||||||||||||||||||||| ||||||||||||||||
Sbjct  26    GTTGGGAACGCGAGCGGCGGATGGGTGAGTAACACGTGGGGAACCTGCCCCATAGTCTGG  85

Query  78    GATACCACTTGGAAACAGGTGCTAATACCGGATAAGAAAGCAGATCGCATGATCAGCTTA  137
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Sbjct  86    GATACCACTTGGAAACAGGTGCTAATACCGGATAAGAAAGCAGATCGCATGATCAGCTTA  145

Query  138   TAAAAGGCGGCGTAAGCTGTCGCTATGGGATGGCCCCGCGGTGCATTAGCTAGTTGGTAG  197
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Sbjct  146   TAAAAGGCGGCGTAAGCTGTCGCTATGGGATGGCCCCGCGGTGCATTAGCTAGTTGGTAG  205

Query  198   GGTAACGGCCTACCAAGGCAATGATGCATAGCCGAGTTGAGAGACTGATCGGCCACATTG  257
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Sbjct  206   GGTAACGGCCTACCAAGGCAATGATGCATAGCCGAGTTGAGAGACTGATCGGCCACATTG  265

Query  258   GGACTGAGACACGGCCCAAACTCCTACGGGAGGCAGCAGTAGGGAATCTTCCACAATGGA  317
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Sbjct  266   GGACTGAGACACGGCCCAAACTCCTACGGGAGGCAGCAGTAGGGAATCTTCCACAATGGA  325

Query  318   CGAAAGTCTGATGGAGCAACGCCGCGTGAGTGAAGAAGGTTTTCGGATCGTAAAGCTCTG  377
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Sbjct  326   CGAAAGTCTGATGGAGCAACGCCGCGTGAGTGAAGAAGGTTTTCGGATCGTAAAGCTCTG  385

Query  378   TTGTTGGTGAAGAAGGATAGAGGTAGTAACTGGCCTTTATTTGACGGTAATCAACCAGAA  437
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Sbjct  386   TTGTTGGTGAAGAAGGATAGAGGTAGTAACTGGCCTTTATTTGACGGTAATCAACCAGAA  445

Query  438   AGTCACGGCTAACTACGTGCCAGCAGCCGCGGTAATACGTAGGTGGCAAGCGTTGTCCGG  497
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Sbjct  446   AGTCACGGCTAACTACGTGCCAGCAGCCGCGGTAATACGTAGGTGGCAAGCGTTGTCCGG  505

Query  498   ATTTATTGGGCGTAAAGCGAGCGCAGGCGGAAGAATAAGTCTGATGTGAAAGCCCTCGGC  557
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Sbjct  506   ATTTATTGGGCGTAAAGCGAGCGCAGGCGGAAGAATAAGTCTGATGTGAAAGCCCTCGGC  565

Query  558   TTAACCGAGGAACTGCATCGGAAACTGTTTTTCTTGAGTGCAGAAGAGGAGAGTGGAACT  617
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Sbjct  566   TTAACCGAGGAACTGCATCGGAAACTGTTTTTCTTGAGTGCAGAAGAGGAGAGTGGAACT  625

Query  618   CCATGTGTAGCGGTGGAATGCGTAGATATATGGAAGAACACCAGTGGCGAAGGCGGCTCT  677
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Sbjct  626   CCATGTGTAGCGGTGGAATGCGTAGATATATGGAAGAACACCAGTGGCGAAGGCGGCTCT  685

Query  678   CTGGTCTGCAACTGACGCTGAGGCTCGAAAGCATGGGTAGCGAACAGGATTAGATACCCT  737
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Sbjct  686   CTGGTCTGCAACTGACGCTGAGGCTCGAAAGCATGGGTAGCGAACAGGATTAGATACCCT  745

Query  738   GGTAGTCCATGCCGTAAACGATGAGTGCTAAGTGTTGGGAGGTTTCCGCCTCTCAGTGCT  797
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Sbjct  746   GGTAGTCCATGCCGTAAACGATGAGTGCTAAGTGTTGGGAGGTTTCCGCCTCTCAGTGCT  805

Query  798   GCAGCTAACGCATTAAGCACTCCGCCTGGGGAGTACGACCGCAAGGTTGAAACTCAAAGG  857
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Sbjct  806   GCAGCTAACGCATTAAGCACTCCGCCTGGGGAGTACGACCGCAAGGTTGAAACTCAAAGG  865

Query  858   AATTGACGGGGGCCCGCACAAGCGGTGGAGCATGTGGTTTAATTCGAAGCAACGCGAAGA  917
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Sbjct  866   AATTGACGGGGGCCCGCACAAGCGGTGGAGCATGTGGTTTAATTCGAAGCAACGCGAAGA  925

Query  918   ACCTTACCAGGTCTTGACATCTAGTGCAATCCGTAGAGATACGGAGTTCCCTTCGGGGAC  977
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Sbjct  926   ACCTTACCAGGTCTTGACATCTAGTGCAATCCGTAGAGATACGGAGTTCCCTTCGGGGAC  985

Query  978   ACTAAGACAGGTGGTGCATGGCTGTCGTCAGCTCCGTTTCGT  1019
             |||||||||||||||||||||||||||||||||| || ||||
Sbjct  986   ACTAAGACAGGTGGTGCATGGCTGTCGTCAGCTC-GTGTCGT  1026
```

(a)

BROWSERS | CLASSIFIER | LIBCOMPARE | SEQMATCH | PROBE MATCH | FUNGENE | RDPIPELINE | SEQCART | TAXOMATIC | TREE BUILDER | ASSIGNGEN

Classifier :: Hierarchy View

[start over | assignment detail | help]

Classifier: RDP Naive Bayesian rRNA Classifier Version 2.10, October 2014
Taxonomical Hierarchy: RDP 16S rRNA training set 14
Query File: classifier_seq_upload1171904004902600006.FASTA
Query Submit Date: Wed Jul 06 01:40:08 EDT 2016

Display depth: Auto
Confidence threshold: 80%
CopyNumber Adjusted: No

domain	%	Library
Bacteria	100.0	▬

Hierarchy View (click a node to make it the root – only show sequences assigned to that node with confidence above the threshold):

rootrank Root (1 sequences) [show assignment detail for Root only]
• • domain Bacteria (1)
• • • phylum Firmicutes (1)
• • • • class Bacilli (1)
• • • • • order Lactobacillales (1)
• • • • • • family Lactobacillaceae (1)
• • • • • • • genus Lactobacillus (1)

Assignment Detail (for Root with Confidence threshold: 80%):

classifier_seq Root[100%] Bacteria[100%] Firmicutes[100%] Bacilli[100%] Lactobacillales[100%] Lactobacillaceae[100%] Lactobacillus[100%]

CENTER FOR MICROBIAL ECOLOGY @ Michigan State University

MICHIGAN STATE UNIVERSITY Advancing Knowledge. Transforming Lives

© 1990-2014 Michigan State University Board of Trustees. East Lansing MI 48824

(b)

Figure 3.5: 16s rRNA analysis of strain: (a) Pairwise alignment of the amplicon and 16s rRNA gene sequence available with NCBI; (b) Search result of the sequence of amplicon from RDP classifier.

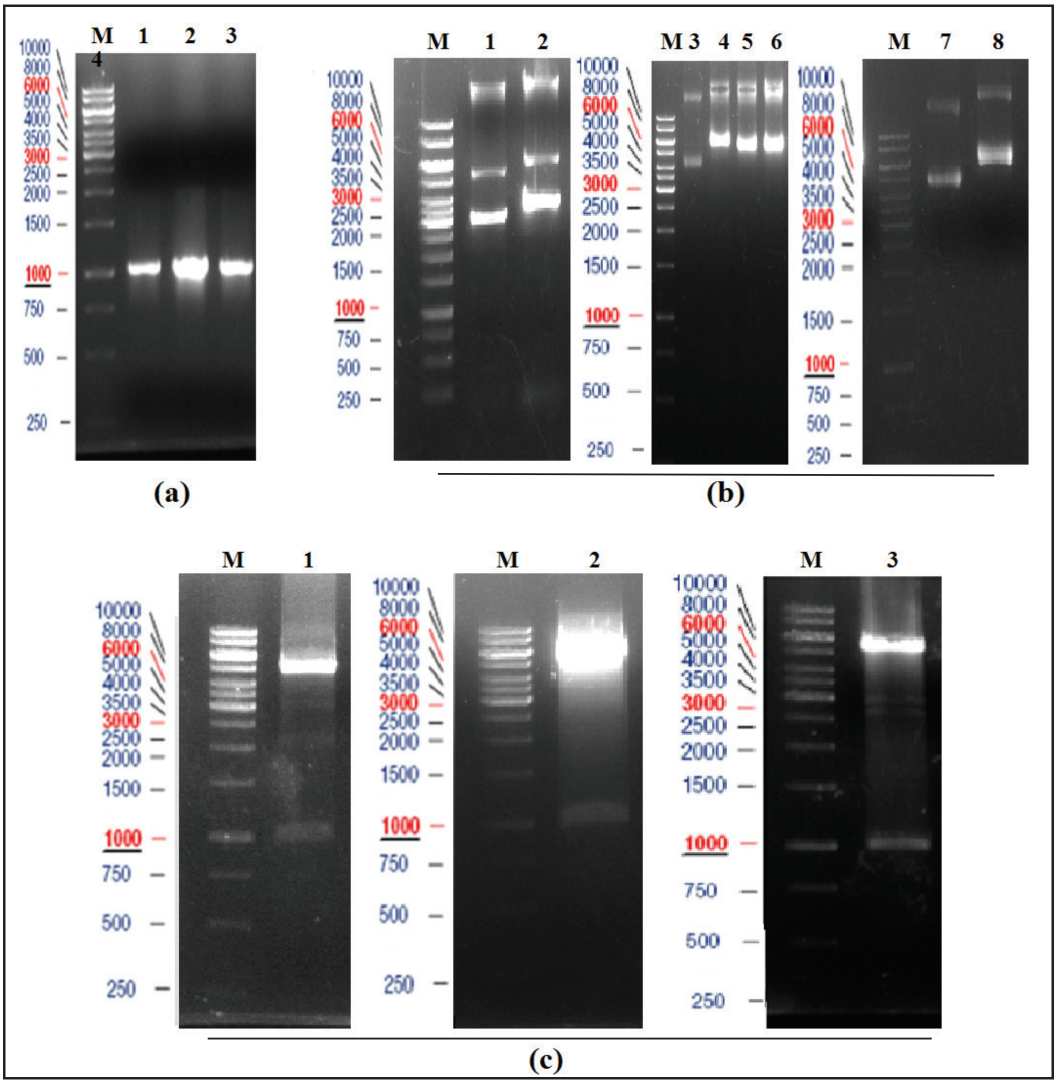
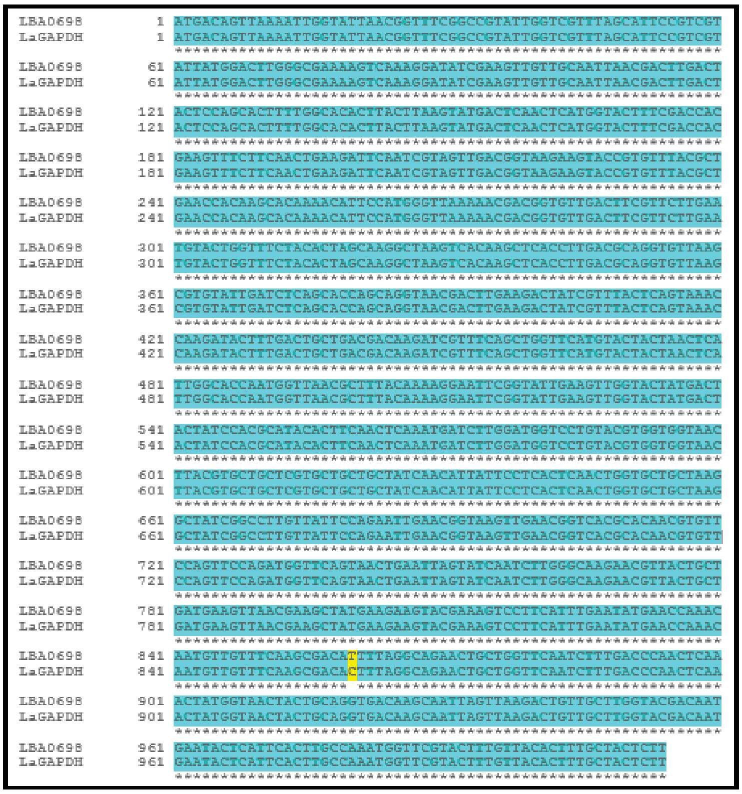


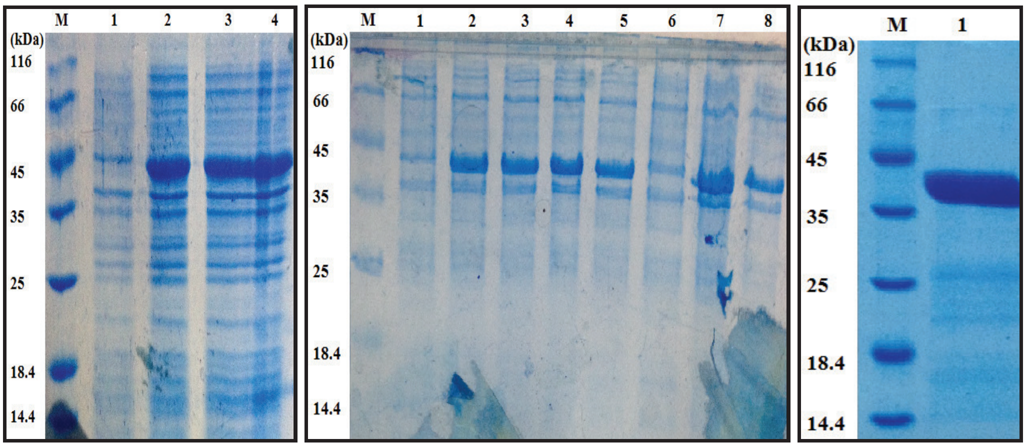
Figure 3.6: Cloning results of LaGAPDH: (a) PCR amplification of LaGAPDH gene (1014 bp), Lane M: 1kb ladder, Lane 1-3 LaGAPDH amplicon; (b) Plasmid shift to assess positive clones, Lane M: 1kb ladder, Lane 1: pET30a control vector, Lane 2: positive clone with pET30a and LaGAPDH insert showing shift, Lane 3: pET15b control vector, Lane 4-6: positive clone with pET15b and LaGAPDH insert showing shift, Lane 7: pET28a control vector, Lane 2: positive clone with pET28a and LaGAPDH insert showing shift; (c) Lane M: 1kb ladder, Lane 1: digestion of positive clone in pET30a with *Nde*I and *Bam*HI restriction enzymes showing LaGAPDH insert release, Lane 2: digestion of positive clone in pET28a with *Nde*I and *Bam*HI restriction enzymes showing LaGAPDH insert release, Lane 3: digestion of positive clone in pET15b with *Nde*I and *Bam*HI restriction enzymes showing LaGAPDH insert release.



```
LBA0698        1 ATGACAGTTAAAATTGGTATTAACGGTTTCGGCCGTATTGGTCGTTTAGCATTCCGTCGT
LaGAPDH        1 ATGACAGTTAAAATTGGTATTAACGGTTTCGGCCGTATTGGTCGTTTAGCATTCCGTCGT
                 ************************************************************

LBA0698       61 ATTATGGACTTGGGCGAAAAGTCAAAGGATATCGAAGTTGTTGCAATTAACGACTTGACT
LaGAPDH       61 ATTATGGACTTGGGCGAAAAGTCAAAGGATATCGAAGTTGTTGCAATTAACGACTTGACT
                 ************************************************************

LBA0698      121 ACTCCAGCACTTTTGGCACACTTACTTAAGTATGACTCAACTCATGGTACTTTCGACCAC
LaGAPDH      121 ACTCCAGCACTTTTGGCACACTTACTTAAGTATGACTCAACTCATGGTACTTTCGACCAC
                 ************************************************************

LBA0698      181 GAAGTTTCTTCAACTGAAGATTCAATCGTAGTTGACGGTAAGAAGTACCGTGTTTACGCT
LaGAPDH      181 GAAGTTTCTTCAACTGAAGATTCAATCGTAGTTGACGGTAAGAAGTACCGTGTTTACGCT
                 ************************************************************

LBA0698      241 GAACCACAAGCACAAAACATTCCATGGGTTAAAAACGACGGTGTTGACTTCGTTCTTGAA
LaGAPDH      241 GAACCACAAGCACAAAACATTCCATGGGTTAAAAACGACGGTGTTGACTTCGTTCTTGAA
                 ************************************************************

LBA0698      301 TGTACTGGTTTCTACACTAGCAAGGCTAAGTCACAAGCTCACCTTGACGCAGGTGTTAAG
LaGAPDH      301 TGTACTGGTTTCTACACTAGCAAGGCTAAGTCACAAGCTCACCTTGACGCAGGTGTTAAG
                 ************************************************************

LBA0698      361 CGTGTATTGATCTCAGCACCAGCAGGTAACGACTTGAAGACTATCGTTTACTCAGTAAAC
LaGAPDH      361 CGTGTATTGATCTCAGCACCAGCAGGTAACGACTTGAAGACTATCGTTTACTCAGTAAAC
                 ************************************************************

LBA0698      421 CAAGATACTTTGACTGCTGACGACAAGATCGTTTCAGCTGGTTCATGTACTACTAACTCA
LaGAPDH      421 CAAGATACTTTGACTGCTGACGACAAGATCGTTTCAGCTGGTTCATGTACTACTAACTCA
                 ************************************************************

LBA0698      481 TTGGCACCAATGGTTAACGCTTTACAAAAGGAATTCGGTATTGAAGTTGGTACTATGACT
LaGAPDH      481 TTGGCACCAATGGTTAACGCTTTACAAAAGGAATTCGGTATTGAAGTTGGTACTATGACT
                 ************************************************************

LBA0698      541 ACTATCCACGCATACACTTCAACTCAAATGATCTTGGATGGTCCTGTACGTGGTGGTAAC
LaGAPDH      541 ACTATCCACGCATACACTTCAACTCAAATGATCTTGGATGGTCCTGTACGTGGTGGTAAC
                 ************************************************************

LBA0698      601 TTACGTGCTGCTCGTGCTGCTGCTATCAACATTATTCCTCACTCAACTGGTGCTGCTAAG
LaGAPDH      601 TTACGTGCTGCTCGTGCTGCTGCTATCAACATTATTCCTCACTCAACTGGTGCTGCTAAG
                 ************************************************************

LBA0698      661 GCTATCGGCCTTGTTATTCCAGAATTGAACGGTAAGTTGAACGGTCACGCACAACGTGTT
LaGAPDH      661 GCTATCGGCCTTGTTATTCCAGAATTGAACGGTAAGTTGAACGGTCACGCACAACGTGTT
                 ************************************************************

LBA0698      721 CCAGTTCCAGATGGTTCAGTAACTGAATTAGTATCAATCTTGGGCAAGAACGTTACTGCT
LaGAPDH      721 CCAGTTCCAGATGGTTCAGTAACTGAATTAGTATCAATCTTGGGCAAGAACGTTACTGCT
                 ************************************************************

LBA0698      781 GATGAAGTTAACGAAGCTATGAAGAAGTACGAAAGTCCTTCATTTGAATATGAACCAAAC
LaGAPDH      781 GATGAAGTTAACGAAGCTATGAAGAAGTACGAAAGTCCTTCATTTGAATATGAACCAAAC
                 ************************************************************

LBA0698      841 AATGTTGTTTCAAGCGACATTTTAGGCAGAACTGCTGGTTCAATCTTTGACCCAACTCAA
LaGAPDH      841 AATGTTGTTTCAAGCGACACTTTAGGCAGAACTGCTGGTTCAATCTTTGACCCAACTCAA
                 ******************* ****************************************

LBA0698      901 ACTATGGTAACTACTGCAGGTGACAAGCAATTAGTTAAGACTGTTGCTTGGTACGACAAT
LaGAPDH      901 ACTATGGTAACTACTGCAGGTGACAAGCAATTAGTTAAGACTGTTGCTTGGTACGACAAT
                 ************************************************************

LBA0698      961 GAATACTCATTCACTTGCCAAATGGTTCGTACTTTGTTACACTTTGCTACTCTT
LaGAPDH      961 GAATACTCATTCACTTGCCAAATGGTTCGTACTTTGTTACACTTTGCTACTCTT
                 ******************************************************
```

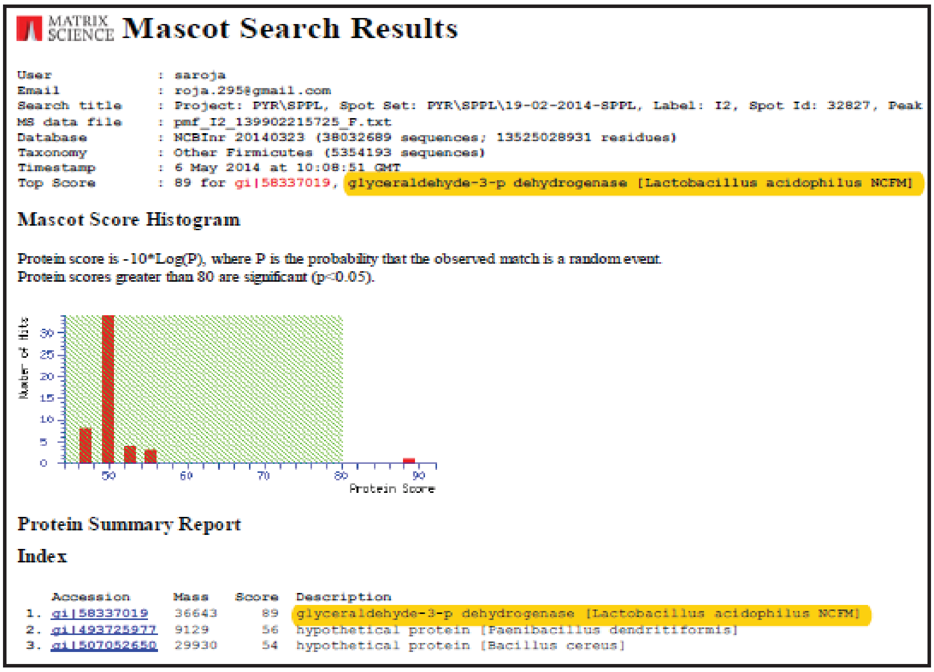
Figure 3.7: Sequence comparison of LaGAPDH nucleotide sequencing result and LBA0698 GAPDH entry from NCBI.

3.4.2.2. Heterologous expression and purification: After confirming the successful cloning of LaGAPDH in pET30a, pET28a and pET15b vectors, the protein expression studies were carried out using IPTG as inducers in *E. coli* Rosetta (DE3) strain. The protein expression trials are summarised in Table 3.3, and significant expression of r-LaGAPDH was detected as resolved on 15% SDS-PAGE gel (Figure 3.8a, b). The expression was detected at around ~40 kDa which is nearer to its expected theoretical molecular mass of 36.4 kDa. In pET30a-LaGAPDH construct, the purification was not achieved after passing lysate through Ni-NTA affinity column chromatography. Sequencing of the insert later identified that due to a frameshift in nucleotide a proline tag was inserted in place of HIS tag. Hence, for further structural and functional analysis of the r-LaGAPDH protein was expressed using pET28a-LaGAPDH construct at 0.25-0.5mM IPTG concentration and 18°C induction temperature which was further purified by Ni-NTA affinity column chromatography (Figure 3.8c). The r-LaGAPDH protein was purified as a single band and further its identity was confirmed by MALDI-TOF/TOF analysis which confirmed that the expressed protein is GAPDH from *L. acidophilus* (Figure 3.9).



(a) (b) (c)

Figure 3.8: SDS-PAGE of protein expression and purification results of r-LaGAPDH in *E. coli* Rosetta (DE3): (a) Protein expression results of pET30a-LaGAPDH construct, Lane M: protein marker, Lane 1: 0.5mM IPTG-induced pET30a control vector, Lane 2: 0.5mM IPTG-induced pET30a-LaGAPDH construct, Lane 3: 1mM IPTG-induced pET30a-LaGAPDH construct, Lane 4: 1.5mM IPTG-induced pET30a-LaGAPDH construct; (b) Protein expression results of pET15b-LaGAPDH and pET28a-LaGAPDH constructs, Lane M: protein marker, Lane 1: 0.5mM IPTG-induced pET15b control vector, Lane 2: 2mM IPTG-induced pET15b-LaGAPDH construct, Lane 3: 1.5mM IPTG-induced pET15b-LaGAPDH construct, Lane 4: 1mM IPTG-induced pET15b-LaGAPDH construct, Lane 5: 0.5mM IPTG-induced pET15b-LaGAPDH construct, Lane 6: 0.5mM IPTG-induced pET28a control vector, Lane 7: 2mM IPTG-induced pET28a-LaGAPDH construct, Lane 8: 1.5mM IPTG-induced pET28a-LaGAPDH construct; (c) Lane M: protein marker, Lane 1: purified recombinant LaGAPDH protein through Ni-NTA affinity column used for further structural and functional characterization.



Mascot Search Results

```
User            : saroja
Email           : roja.295@gmail.com
Search title    : Project: PYR\SPPL, Spot Set: PYR\SPPL\19-02-2014-SPPL, Label: I2, Spot Id: 32827, Peak
MS data file    : pmf_I2_139902215725_F.txt
Database        : NCBInr 20140323 (38032689 sequences; 13525028931 residues)
Taxonomy        : Other Firmicutes (5354193 sequences)
Timestamp       : 6 May 2014 at 10:08:51 GMT
Top Score       : 89 for gi|58337019, glyceraldehyde-3-p dehydrogenase [Lactobacillus acidophilus NCFM]
```

Mascot Score Histogram

Protein score is -10*Log(P), where P is the probability that the observed match is a random event.
Protein scores greater than 80 are significant (p<0.05).

Protein Summary Report

Index

```
   Accession        Mass    Score  Description
1. gi|58337019      36643     89   glyceraldehyde-3-p dehydrogenase [Lactobacillus acidophilus NCFM]
2. gi|493725977     9129      56   hypothetical protein [Paenibacillus dendritiformis]
3. gi|507052650     29930     54   hypothetical protein [Bacillus cereus]
```

Figure 3.9: MALDI-TOF MS/MS confirmation of r-LaGAPDH.

Table 3.3: Summary of expression trials of r-LaGAPDH detected on SDS-PAGE gel in different vectors.

E. coli	**Construct**	**IPTG induction**						**Remark**
		≤0.5 mM		**0.5-1.0 mM**		**≥1 mM**		
		37° C	**18° C**	**37° C**	**18° C**	**37° C**	**18° C**	
Rosetta (DE3)	pET30a-LaGAPDH	+/-	+	+/-	+/-	+/-	+/-	No HIS tag in final product
	pET15b-LaGAPDH	+/-	+	+/-	+/-	+/-	+/-	Less product in soluble fraction than pET28a.
	pET28a-LaGAPDH	+/-	+	+/-	+/-	+/-	+/-	Soluble expression at lower IPTG & temperature.

(+) indicates soluble expression in supernatant, (-) indicates insoluble expression and (+/-) indicates expression in partially soluble and insoluble fractions.

3.4.3. Biochemical and functional characterization of LaGAPDH.

3.4.3.1. Hemagglutination assay and carbohydrate specificity: r-LaGADPH showed hemagglutination activity against 2% rabbit RBCs with MIC of 45µg (Figure 3.10a). The hemagglutination activity was inhibited only by complex glycoprotein fetuin at 0.1% (w/v) (Figure 3.10b). Glycoprotein mucin was also tested for inhibition but due to its viscosity, the results of hemagglutination assay was not conclusive. None of the carbohydrates was able to inhibit the hemagglutination. The hemagglutination process was also monitored under the microscope with a positive control protein from *Solanum tuberosum* (PotHg) previously known to agglutinate rabbit RBCs (Figure 3.10c) (Shah *et al.*, 2016). The specific activity was 42.5 HU/mg (Figure 3.10d). The hemagglutination activity of r-LaGAPDH was stable up to 60°C of which it was the activity was not regained (Figure 3.11a, b).

3.4.3.2. Mucin-binding assay: r-LaGAPDH was tested for binding to mucin immobilised on 96-well microtiter plates, and it bound to mucin in a dose-dependent manner (Figure 3.12a). There was a band of ~40 kDa protein on SDS-PAGE along with a band of 60 kDa which corresponds to LaGAPDH and BSA protein which was used for blocking wells. The absence of protein band in SDS-PAGE in control indicated that the binding of r-LaGAPDH to immobilised mucin was specific. Additionally to confirm the specificity another protein adenylate kinase (LdAK) from *Leishmania donovani* was used as a negative control which also did not show any binding with mucin (Figure 3.12b).

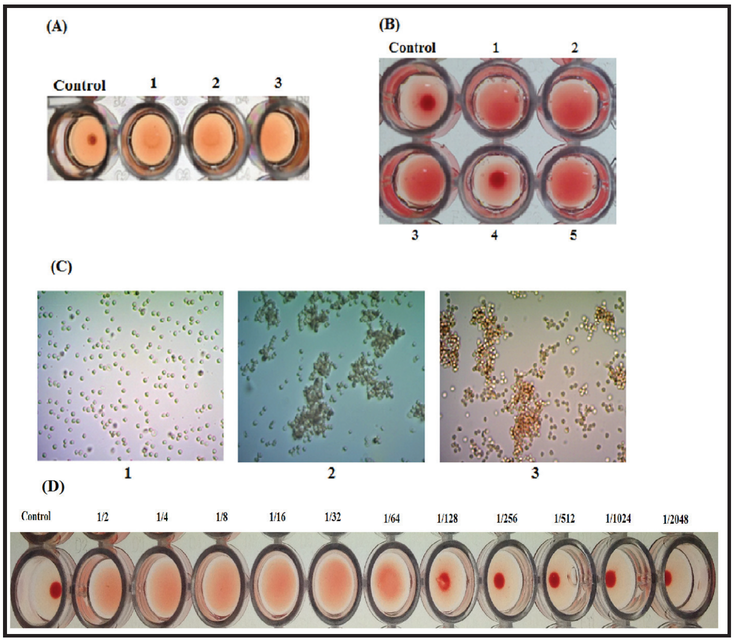


Figure 3.10: Results of hemagglutination assay of r-LaGAPDH towards rabbit erythrocytes: (A) r-LaGAPDH agglutinating 2% rabbit RBCs, C: Control having only rabbit RBCs, 1: Purified r-LaGAPDH (45µg), 2: Purified r-LaGAPDH (60µg), 3: Purified r-LaGAPDH (75µg); (B) Hemagglutination inhibition assay with carbohydrates, C: Control having only rabbit RBCs, 1: Hemagglutination inhibition using r-LaGAPDH and mannose, 2: Hemagglutination inhibition using r-LaGAPDH and glucose, 3: Hemagglutination inhibition using r-LaGAPDH and galactose, 4: Hemagglutination inhibition using r-LaGAPDH and fetuin, 5: Positive control (PotHg); (C) Microscopic observation of hemagglutination by r-LaGAPDH, 1: Control having only rabbit RBCs, 2: Positive control-PotHg hemagglutination with rabbit RBCs, 3: r-LaGAPDH hemagglutination with rabbit RBCs; (D) r-LaGAPDH hemagglutination titer analysis.

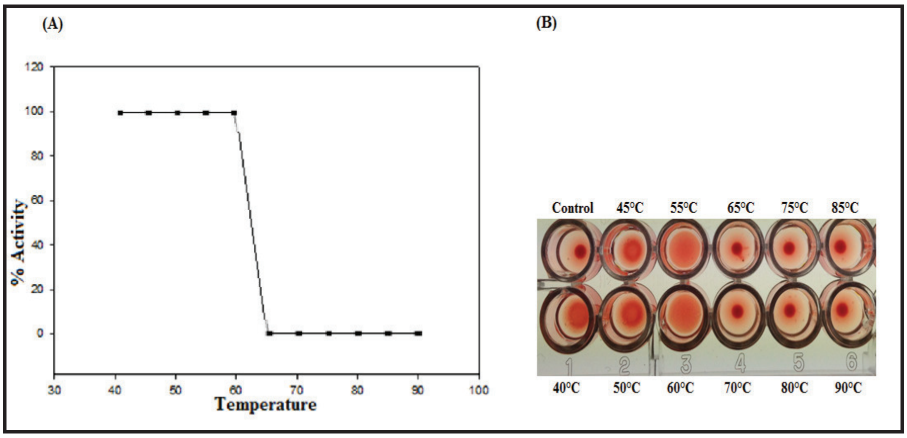


Figure 3.11: The effect of temperature on the hemagglutination of r-LaGAPDH in the hemagglutination assay: (A) the value of temperature is plotted on the x-axis, and the activity was plotted on the y-axis. The dots on the line represent the temperature tested in the assay and activity was plotted either 100% or 0% judged by the absence and presence of button formation respectively; (B) Image of 96-well microtiter plate of hemagglutination assay of r-LaGAPDH at different temperature mentioned above each well.

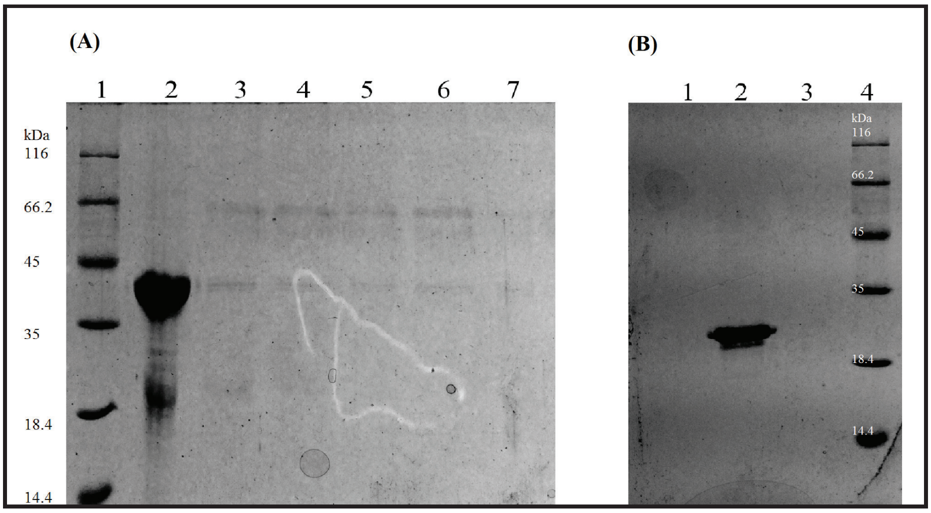


Figure 3.12: Representative SDS-PAGE gel showing r-LaGAPDH protein recovered from the coated mucin after the mucin binding assay: (A) Lane 1: Protein molecular weight marker, Lane 2: r-LaGAPDH as a reference, Lane 3: Purified r-LaGAPDH (125µg), Lane 4: Purified r-LaGAPDH (100µg), Lane 5: Purified r-LaGAPDH (75µg), Lane 6: Purified r-LaGAPDH (50µg), Lane 7: Purified r-LaGAPDH (25µg); (B) Lane 1: Purified r-LaGAPDH recovered from the well without mucin, Lane 2: Purified Leishmania donovani Adenylate kinase (LdAK) as a reference, Lane 3: Purified LdAK recovered from the well, Lane 4: Protein molecular weight marker.

3.4.3.3. Enzyme kinetic studies: The enzyme kinetics assay was performed to assess whether the r-LaGAPDH retains its enzymatic activity. In the presence of G3P and triethanolamine buffer, it showed a dose-dependent conversion of NAD to NADH which was observed by monitoring the NADH formed at 340nm.

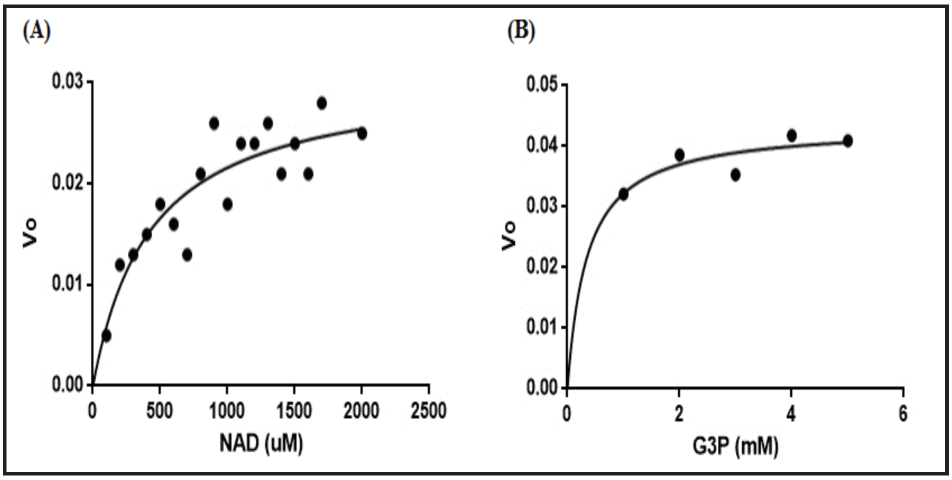


Figure 3.13: Enzyme kinetics plot of r-LaGAPDH using Michaelis-Menten equation: (A) 110µM of r-LaGAPDH protein was assayed as a variable function of NAD (100-2000 µM) in the presence of G3P (1 mM) with an estimated K_m of 333 µm ± 16; (B) 110µM of r-LaGAPDH protein was assayed as a variable function of G3P (1-5 mM) in the presence of NAD (250 µM) with an estimated K_m of 333 µm ± 16.

The enzyme reaction rate was determined using 110µM of r-LaGAPDH and with different concentration of G3P and NAD respectively. The results were analysed with Michaelis-Menten equation using Graph Pad Prism software. The K_m of NAD (100-2000µM) in the presence of G3P (1mM) was estimated to be 459µM ± 29.1 (Figure 3.13a) and K_m of G3P (1-5mM) in the presence of NAD (250µM) was estimated to 333µm ± 16 (Figure 3.13b). The enzyme activity was also detected in the cell-free media supernatant obtained from a culture of *L. acidophilus* cells thus confirming the presence of GAPDH as a secretory protein and the enzyme activity was estimated to be 11 x 10^{-3} ± 0.0005µmoles per min in comparison to -0.4 x 10^{-3} ± 0.0003µmoles per min for control reaction in which G3P was absent.

3.4.4. Biophysical studies

3.4.4.1. Molecular mass: r-LaGAPDH was detected as a single band on SDS-PAGE corresponding to ~40 kDa, which is the theoretical expected molecular weight (36.4 kDa) (Figure 3.8c). The identity of the protein was confirmed as GAPDH from *L. acidophilus* (UniProt: LBA0698) by MALDI-TOF MS/MS analysis (Figure 3.10). The molecular mass of r-LaGAPDH in solution was estimated by gel filtration chromatography using Superose-6 column. It eluted as a single peak with an apparent molecular weight of about 149 kDa indicating that the r-LaGAPDH exists as a tetramer in solution (Figure 3.14). The r-LaGAPDH was also detected as a single band in SDS-PAGE when resolved both in presence and absence of reducing agent like DTT and βME indicating that the oligomerization is independent of disulphide linkages. All this result indicates that r-LaGAPDH is present as an oligomer in solution probably as a homotetramer.

3.4.4.2. CD spectroscopy analysis: CD spectra for r-LaGAPDH showed minima at 208 and 215nm (Figure 3.15a) which indicates the presence of alpha helix and beta sheets as secondary structural elements which are corroborated with findings from the in-silico secondary structure prediction. The CD thermal melt analysis indicates that the protein is stable up till 65°C and beyond that the protein starts unfolding which is indicated by a dip in the molar ellipticity values (Figure 3.15b). This finding is by the temperature studies using hemagglutination assay.

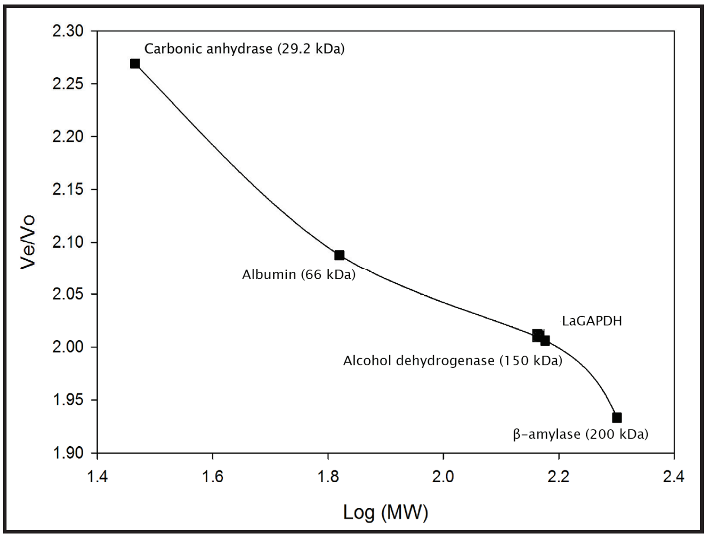


Figure 3.14: r-LaGAPDH molecular mass identification through gel filtration column.

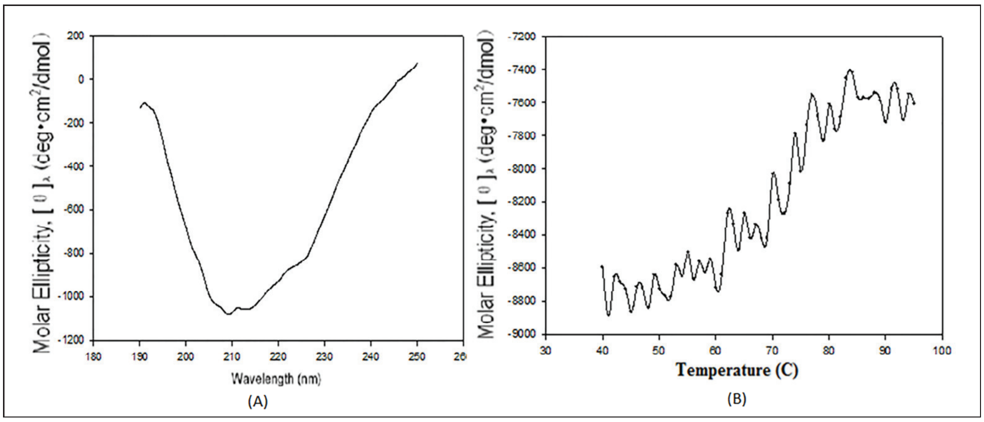


Figure 3.15: Circular dichroism spectra (CD) of r-LaGAPDH: (A) CD spectra of r-LaGAPDH (far UV range) molar ellipticity versus wavelength; (B) CD thermal melt showing stable secondary structure up to 65°C.

3.4.4.3. Fluorescence spectroscopy: The fluorescence emission spectra of r-LaGAPDH when excited at 280 nm, showed a ʎmax of 323–326 nm. When excited at 295 nm no emission spectra were observed indicating that the tryptophan residues are not exposed to the environment. The solution containing r-LaGAPDH was titrated with carbohydrates to monitor the change in the microenvironment of aromatic amino acids which is either by enhancing or by quenching of the fluorescence emission spectra. Titration of r-LaGAPDH with mannose, galactose, N-Acetyl-D-galactosamine and N-Acetyl-D-glucosamine resulted in a decrease in emission spectra without altering the ʎmax. These particular sugars were selected because they are the primary composition in glycoprotein mucin which forms the large and basic molecular site for recognition in the gut. A spectrum corresponding to a representative fluorescence titration of LaGAPDH with different carbohydrates is given in Figure 3.16 and Figure 3.17. The plots in Figure 3.16 and 3.17 represent an average plot of three independent readings. The calculated K_d values are represented along with standard error mean in Table 3.4.

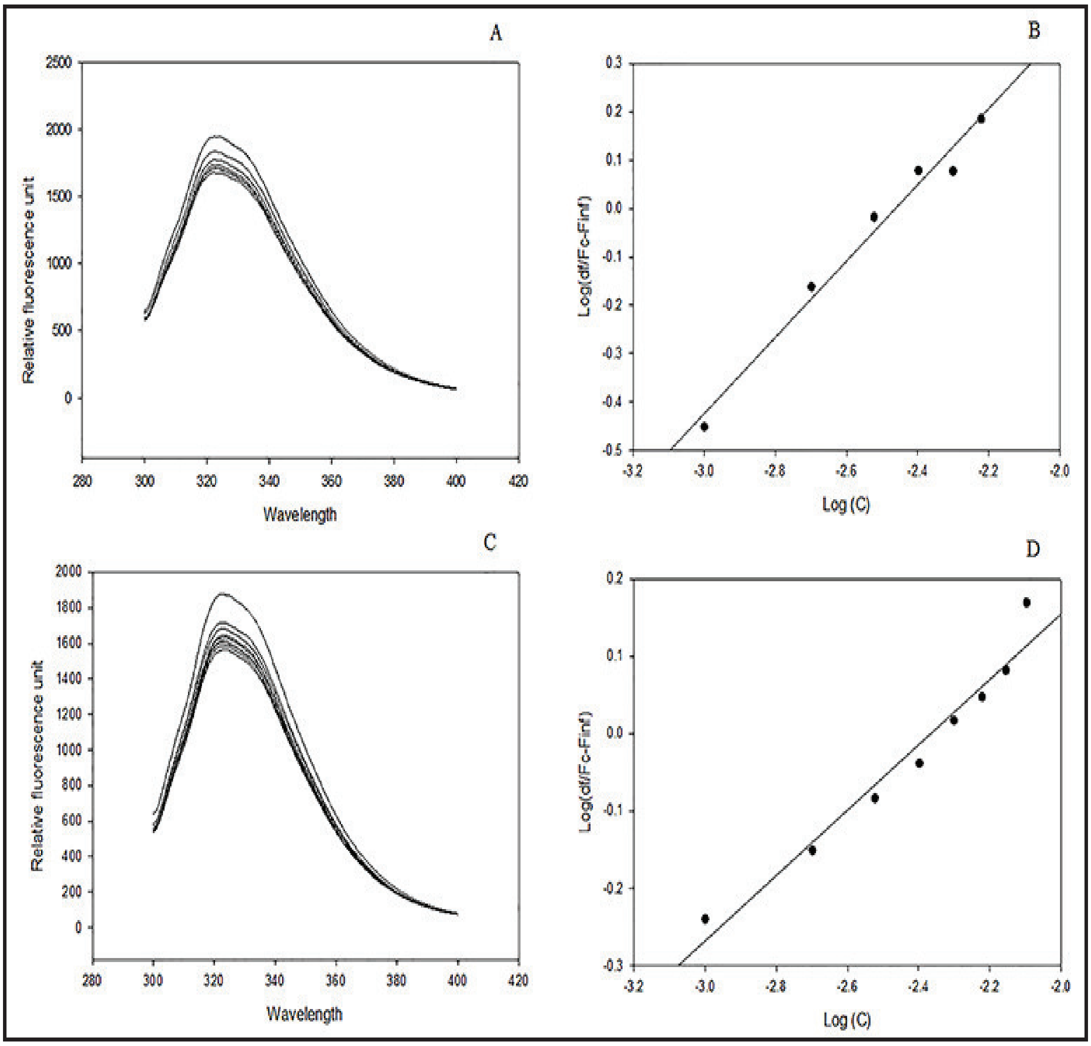


Figure 3.16: Fluorescence spectroscopy studies of r-LaGAPDH with mannose and galactose: (A) Fluorescence emission spectra of r-LaGAPDH along with increasing concentration of mannose; (B) The plot of log [dF/(F_c − F_{inf})] versus log C for the interaction of mannose with r-LaGAPDH. Antilog value of the X-intercept of the plot gives the K_d value of the interaction; (C) Fluorescence emission spectra of r-LaGAPDH along with increasing concentration of galactose; (D) The plot of log [dF/(F_c − F_{inf}))] versus log C for the interaction of galactose with r-LaGAPDH. Antilog value of the X-intercept of the plot gives the K_d value of the interaction.

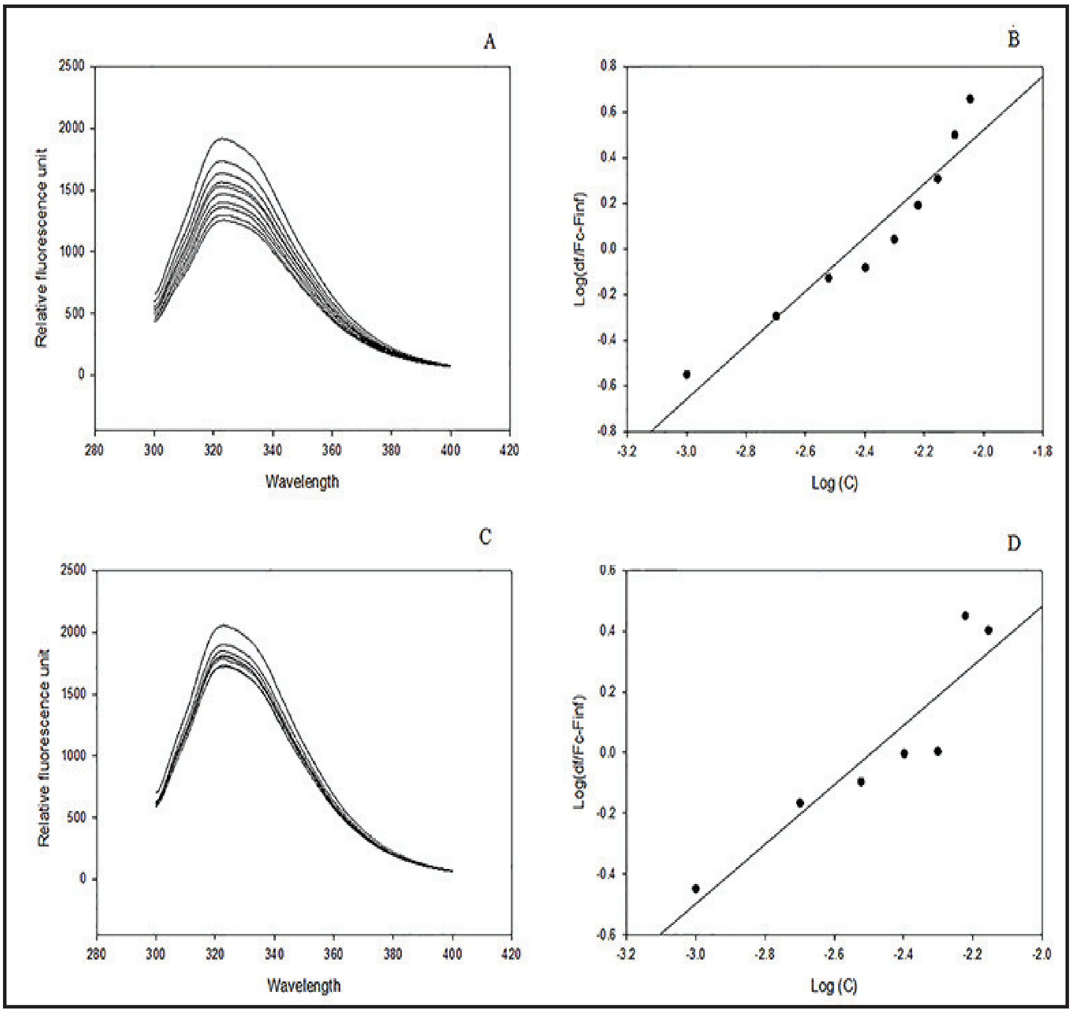


Figure 3.17: Fluorescence spectroscopy studies of r-LaGAPDH with N-Acetyl-D-galactosamine and N-Acetyl-D-glucosamine: (A) Fluorescence emission spectra of r-LaGAPDH along with increasing concentration of N-Acetyl-D-galactosamine; (B) The plot of log $[dF/(F_c - F_{inf})]$ versus log C for the interaction of N-Acetyl-D-galactosamine with r-LaGAPDH. Antilog value of the X-intercept of the plot gives the K_d value of the interaction; (C) Fluorescence emission spectra of r-LaGAPDH along with increasing concentration of N-Acetyl-D-glucosamine; (D) The plot of log $[dF/(F_c - F_{inf}))]$ versus log C for the interaction of N-Acetyl-D-glucosamine with r-LaGAPDH. Antilog value of the X-intercept of the plot gives the K_d value of the interaction.

Table 3.4: Dissociation constant of different carbohydrates with r-LaGAPDH determined by fluorescence spectroscopy.

Carbohydrate	**Dissociation Constant (K_d)**
Mannose	$3.6 \pm 0.7 \times 10^{-3}$ M
Galactose	$4.34 \pm 0.09 \times 10^{-3}$ M
N-Acetyl-D-galactosamine	$4.0 \pm 0.87 \times 10^{-3}$ M
N-Acetyl-D-glucosamine	$3.7 \pm 0.28 \times 10^{-3}$ M

(±) indicates values represented with standard error mean.

3.5. Discussion

It is evident that the human body lives in mutualism with a complex gut microbial population which is a home for thousands of bacterial species. This health-promoting bacteria has formed the basis of probiotic microorganism. Probiotics which are usually found in the healthy human gastrointestinal tract are the class of beneficial organisms, most of which have different but overlapping benefits. At molecular levels, the probiotic organisms work through several interrelated mechanisms to harness benefits to human host (Thomas & Versalovic, 2010). Probiotics reduces the risk of infection or toxin-mediated diseases caused by pathogenic organisms, regulates immune responses to infection, suppress excessive inflammation along with providing a tight epithelial junctions and additional barrier to the entry of dangerous microorganisms and toxins (Girardin & Seidman, 2011; Ly *et al.*, 2011; Vanderpool *et al.*, 2008). To impart such a huge health promoting effects of the human host, adhesion and colonisation to the human gut are a primary criterion for lactobacillus.

Adhesion of lactic acid bacteria to host tissues, epithelial and mucosa is recognised as a necessary first step of bacterial colonisation. It is a complicated process involving multiple different factors. One of the prime factors of adhesion is mediated by cell surfaces proteins also referred to as adhesins which are involved in specific recognition of host components, a central process in bacterial colonisation and adhesion. These adhesins can be broadly categorised as S-layer proteins, LPXTG anchoring proteins, transporter proteins, and N- or C-terminal membrane protein along with multifunctional house-keeping proteins which are also involved in physiological processes in the bacteria.

Glyceraldehyde-3-phosphate dehydrogenase (GAPDH) is a classical conserved intracellular glycolytic enzyme that catalysis the oxidative phosphorylation of G3P (glyceraldehyde 3-phosphate) to 1, 3-bisphosphoglycerate in the presence of NADH/NADPH and inorganic phosphate (Harris & Waters, 1976). Apart from a well-known housekeeping protein, GAPDH is also considered as a principal moonlighting protein with additional roles in apoptosis initiation (Hara *et al.*, 2005), transcription activation (Zheng *et al.*, 2003) and host-pathogen interactions (Bergmann *et al.*, 2004). GAPDH is considered the prototype of moonlighting proteins because this is a first such housekeeping protein identified for its multiple roles and initial characterization was done as a virulence factor in gram-positive pathogen Streptococci (Pancholi & Fischetti, 1992). GAPDHs also acts as adhesin which binds to host epithelial mucin, or ECM components (fibronectin, mucin, plasminogen, etc.) and thus make it an essential mediator of host-microbe interactions (Kinoshita, Uchida*, et al.*, 2008; Pancholi & Fischetti, 1992; Ramiah *et al.*, 2008).

Since then, the majority of study is done to identify the bacterial species having a repertoire of such moonlighting proteins, but the molecular and mechanistic aspects are not clear. There exists only a few preliminary reports of adhesins from *L. acidophilus* group which binds to ECM through lectin-like recognition (Mukai *et al.*, 1992; Yamada *et al.*, 1994) but more efforts are needed to understand the mechanistic and structural basis of these host-microbe interactions. Although the role of GAPDH in host-microbe interactions is well known, only a past known study have been undertaken to evaluate its lectin-like recognition of *Lactobacillus plantarum* cell surface GAPDH with human blood group antigens (Kinoshita, Uchida*, et al.*, 2008).

In this study, LaGAPDH was successfully cloned in pET28a vector and overexpressed as a recombinant protein in *E. coli* Rosetta (DE3) strain. The SDS-PAGE resolved the HIS-tagged r-LaGAPDH protein at ~40 kDa which is in agreement with calculated molecular weight of 6X(His) tag recombinant LaGAPDH (38.9 kDa). The correct in-frame gene orientation was confirmed by nucleotide sequencing, and the identity of the purified recombinant protein was confirmed by MALDI-TOF analysis of trypsinized digested band as GAPDH from *L. acidophilus* NCFM (UniProt ID: LBA0698). Molecular weight determination using gel filtration studies revealed that LaGAPDH exists as a homotetramer and not as a monomer in solution. Under the reducing and non-reducing condition, the relative molecular mass of r-LaGAPDH in SDS-PAGE revealed a single band at ~40kDa indicating that the oligomerization of r-LaGAPDH was not due to disulphide linkages between cysteine residues despite the unusual presence of three cysteine residues (position 101, 156, and 326). The oligomer state of a protein is critical from moonlighting property (Jeffery, 1999) and a previous study showed that the human GAPDH in the tetrameric state functions as glycolytic enzyme but when in a monomeric state it functions as nuclear uracil-DNA glycosylase (Meyer-Siegler *et al.*, 1991).

Microbial adhesins recognise the oligosaccharide moieties of glycoproteins either in mucus layer or mucin and thus resembles lectin-like recognition. To explore and understand the protein-glycan mediated host-microbe interaction in more detail we performed adhesion assay and hemagglutination assay on r-LaGAPDH in the present study. r-LaGAPDH had a positive binding interaction with mucin as found in our studies. Mucins are rich in Ser/Thr residues which are sites for O-glycosylation primary with N-acetyl-D-galactosamine (GalNAc), upon which long and more complex oligosaccharides chains are generated. These diverse glycan structures are believed to be the binding sites for microbial adhesins along the GI tract (Etzold *et al.*, 2014). To further investigate the lectin-like recognition in the GAPDH-mucin interaction hemagglutination assay was done and it indicated that r-LaGAPDH agglutinated rabbit RBCs, and this interaction was only inhibited by fetuin, a complex glycoprotein. The results of temperature stability study in hemagglutination assay and the CD thermal melt studies were in agreement. To further exploit the molecular interaction a fluorescence-based spectroscopy study was performed for r-LaGAPDH interaction with four carbohydrates mannose, galactose, N-acetyl-D-galactosamine (GalNAc) and N-acetyl-D-glucosamine (GlcNAc) primarily due to their significant presence in mucin. r-LaGAPDH showed the highest affinity with mannose sugar followed by GlcNAc amongst the tested carbohydrates. The affinity for GlcNAc was little higher than GalNAc which may be due to the difference in orientation of C4 hydroxyl group. Also, the enzyme assay showed that the recombinant r-LaGAPDH is enzymatically active, and the native GAPDH from *L. acidophilus* cells is secreted into media supernatant.

Several reports are indicating the presence and involvement of GAPDH on cell surface mediating adhesion to ECM or cell lines, but very few reports exploring the mechanistic/ molecular information are reported. Apart from the existence of GAPDH on the cell surface of Lactobacillus, the GAPDH is found as secreted in enterohemorrhagic and enteropathogenic *E.coli* (Egea *et al.*, 2007), Streptococcus (Pancholi & Fischetti, 1992) and even in Mycoplasma (Alvarez *et al.*, 2003). Earlier reports suggested that the extracellular GAPDH which is either secreted or cell-surface associated is enzymatically active which is extremely relevant for a multifunctional protein (Pancholi & Fischetti, 1992) which was in agreement with our results of enzymatically active LaGAPDH. This ability indicates maintenance of the protein structure and folds throughout the export process. Plasminogen binding GAPDH from *L. crispatus* was found to be localised on the cell surface at pH 5 while in alkaline pH it gets released in to the environment (Antikainen *et al.*, 2007). A single study exists which indicated that MSG1, a surface localised protein exhibiting GAPDH activity is involved in binding to erythrocytes

(Hoelzle *et al.*, 2007).

Only two studies are reported which suggested that GAPDH interaction with mucin or ECM components is lectin-like recognition. The surface bound GAPDH protein from *Mycoplasma genitalium*, when pretreated with mucin-associated sugars, inhibits the *M. genitalium* adherence to mucin. The most extensive reduction in adhesion is by N-acetyl-D-glucosamine and galactose. These two sugar molecules bridge individual sugar side chains to the mucin and are more abundant than other mucin-associated sugars (Alvarez *et al.*, 2003). Another study from *L. plantarum* indicates that the GAPDH binds with trisaccharide i.e. human ABO type blood group antigens present on human colonic mucin. *L. plantarum* GAPDH binding to mucin is decreased when the protein is incubated with NAD before adhesion. These suggest that the NAD binding and carbohydrate binding sites are same in GAPDH (Kinoshita, Wakahara*, et al.*, 2008). While another report suggests that in the presence of G3P, heparin a glycosaminoglycan cannot bind with GAPDH suggesting a competition between heparin and G3P for enzyme binding (Torres-Bugeau *et al.*, 2012). The only way to the molecular determination of GAPDH binding with carbohydrates is the structural studies of GAPDH in complex with sugars. Thus in this chapter, the preliminary characterization of r-LaGAPDH was achieved indicating a lectin-like recognition mechanism towards host mucin, which may be due to the interaction between r-LaGAPDH and carbohydrates present on the glycoprotein mucin. Further, the crystallisation of LaGAPDH is carried out to address the adherence mechanism to the host ECM.