Summary
Group A Streptococcus (GAS), or *Streptococcus pyogenes* is a gram positive bacterial pathogen causes a number of less severe suppurative infections to more severe nonsuppurative sequelae leading to high morbidity and mortality to human population. Due to its invasive nature GAS has also been termed as flesh eating bacteria. It is highly desirable to develop vaccine against GAS. The lack of information of distribution pattern of GAS is a great hindrance in development of universal vaccine against it. Developed countries have contributed a lot to add epidemiological information of GAS but little contribution is from developing countries like India.

Therefore we conducted epidemiological study of GAS in North India. We have collected a total of 160 clinical samples of throat swab, pus, blood and vaginal swab during 2004-2008. Out of these, 92 were identified as GAS. We found that all GAS isolates belonged to 17 different *emm* serotypes. This study not only added the information of epidemiology of GAS in India but also reported highly heterogenic distribution of GAS and found M1 as the most prevalent and M49 as second most prevalent serotype. Being the most and second most prevalent serotypes of GAS in our study, M1 and M49 were further selected to check and compare the invasiveness with human respiratory alveolar epithelial cells (A549) and hypo pharyngeal cells (HEp-2). We hypothesized that physiological conditions may regulate the invasiveness of GAS. For this purpose CFU count, growth pattern were checked and invasion assays were performed at different physiological conditions (25, 30, 41°C temperatures and salt stress 0.3 and 0.6 M NaCl) for both the serotypes. We observed that CFU counts were decreased as salt concentration was increased from 0.1 M to 0.9 M NaCl and decreased as the temperature was deviated from optimum (25° to 41°C). In case of M49 serotype, CFU counts were decreased as the salt concentration was increased from 0.1 M to 0.9 M NaCl and was increased as the temperatures were deviated. The viability assay of A549 and HEp-2 cells was done by using trypan blue exclusion method. The count of A549 cells (confluency > 95 %) per dish (60 mm) was 3.5 x 10^6 and for HEp-2 cells (confluency > 95 %), it was found 2.3 x 10^6 per dish (60 mm).
We have found that M1 was less invasive than M49 serotype. In case of M1 serotype maximum invasion i.e., 3.61 % was found at 0.6 M NaCl and in case of M49 it was 12.2 % at 25°C with A549 cells. We also checked the in vitro proinflammatory immune response of cytokines like TNF-α, IL-1β, IL-10, IL-6, IL-8, and IL-12P70 by A549 and HEp-2 cells during the 2 h interaction with M1 and M49 serotypes. HEp-2 cells produced significantly (P ≤ 0.001) higher amount of IL-6 i.e., 317.4 ± 12.7 and 305.45 ± 31 pg/ml when stimulated by M49 serotype grown at 25°C and 0.3 M NaCl conditions respectively as compared to normal conditions (0.1M NaCl, 37°C) (production was 184.05 ± 24.11 pg/ml). IL-6 and IL-8 production was higher in comparison to other cytokines suggesting the activation of humoral and cellular immune response by A549 and HEp-2 cells. We found that in vitro proinflammatory response by the host depends on cell type and serotype of GAS; and is inversely proportional to the invasiveness of GAS under different physiological conditions.

In the second phase of work, bioinformatics coupled with proteomics approach was done to identify proteins as vaccine candidates against GAS. By using in silico analysis, we have selected 147 most eligible vaccine candidates against GAS which were predicted to be expressed on GAS surface and having externally exposed conserved portion. These proteins mainly include transporters, proteases/peptidases, metabolic proteins, permeases, toxins and kinases etc.

Further to find exclusively or abundantly expressed proteins in highly invasive serotype i.e., M49 as compared to less invasive M1 serotype, 1D-ESI-LC-MS/MS analysis was performed. After proteomics analysis 919 proteins were detected in case of M49 serotype as compared to 407 proteins reported in case of M1. Among these proteins, total of 373, 307 and 239 were detected in cell wall/membrane, cytoplasmic and secretory fraction respectively. In case of M1 serotype a total of 128, 113 and 166 were detected in cell wall/membrane, cytoplasmic and secretory fraction respectively. Further we found that 583 proteins were detected exclusively in M49; 337 were detected in both serotypes and only 70 proteins were detected exclusively in M1. In case of M49, Pfam analysis of the detected proteins reveals that most of the proteins belong to metabolic enzymes, transporters, ribosomal proteins and protein synthesis. As these proteins are present in large numbers in case of M49 we conclude that these proteins are making M49 more
active metabolically, therefore making it more invasive in nature as compared to M1 in which a very less number of these proteins have been detected.

On comparing proteins detected in cell wall/membrane fraction with bioinformatically selected proteins, we found that 16 proteins were commonly present in cell wall/membrane fraction of both the serotypes that were also predicted by *in silico* analysis. While 23 predicted proteins were found common between *in silico* analysis and proteomics approach (cell wall and membrane fraction) of highly invasive M49 in comparison to only 1 in case of M1 GAS serotypes. We propose that these common proteins may be targeted as potential universal candidates for vaccine development against GAS in future.