# IPD Project Details

**Project ID:** IPD5439

**Project Title:** An integrated proteo-transcriptomics approach reveals novel drug targets against multidrug resistant Escherichia coli

**Description:** Infections due to multidrug-resistant (MDR) Escherichia coli are associated with severe morbidity and mortality, worldwide. Microbial drug resistance is a complex phenomenon which is conditioned by an interplay of several genomic, transcriptomic and proteomic factors. Here, we have conducted an integrated transcriptomics and proteomics analysis of MDR E. coli to identify genes which are differentially expressed at both mRNA and protein levels. Using RNA-Seq and SWATH-LC MS/MS it was discerned that 763 genes/proteins exhibited differential expression. Of these, 52 genes showed concordance in differential expression at both mRNA and protein levels with 41 genes exhibiting overexpression and 11 genes exhibiting under expression. Bioinformatic analysis using GO-terms, COG and KEGG functional annotations revealed that the concordantly overexpressed genes of MDR E. coli were involved primarily in biosynthesis of secondary metabolites, aminoacyl-tRNAs and ribosomes. Protein–protein interaction (PPI) network analysis of the concordantly overexpressed genes revealed 81 PPI networks and 10 hub proteins. The hub proteins (rpsI, aspS, valS, lysS, accC, topA, rpmG, rpsR, lysU, and spmB) were found to be involved in aminoacylation of tRNA and lysyl-tRNA and, translation. Further, it was discerned that three hub proteins - smpB, rpsR, and topA were non homologous to human proteins and were involved in several biological pathways directly and/or indirectly related to antibiotic stress. Also, absence of homology ensures a little cross-reactivity of their inhibitors/drugs with human proteins and undesirable side effects. Thus, these proteins might be explored as novel drug targets against both drug-resistant and -sensitive populations of E. coli.

**Principal Investigator:** Manish Kumar

**PI Affiliation:** Department of Biophysics, University of Delhi South Campus

**Sample Preparation:** Bacterial cells were washed with normal saline solution and dispersed in sonication buffer containing 50 mM Tris–HCl,10 mM MgCl2, 0.1% sodium azide, 1 mM phenylmethylsulphonyl fluoride and 1 mM; pH 7.4 (Singhal et al., 2012). Cells at a concentration of 1 g wet weight /5 mL of lysis buffer were broken down using a sonicator at 35% amplitude for 10 min at 4°C. The homogenate was centrifuged at

12,000 xg for 30 min at 4°C and the clear supernatant was overnight precipitated at -20? C with chilled acetone. The precipitated protein was collected by centrifugation (12,000 xg, 20 min), air dried and suspended in appropriate volume of protein dissolving buffer. Protein concentration was determined using the Bradford assay (Bradford, 1976).

**Peptide Separation:** Separation and identification of proteins by nanoLC AB Sciex Triple TOF 5600-MS Equal amount of protein samples was digested with trypsin and analyzed using a Triple TOF 5600 mass spectrometer (AB Sciex, USA) equipped with Eskigent MicroLC 200 system (Eskigent, Dublin, CA) with an Eskigent C18 - reverse phase column. 1 microgram of digested proteins was desalted online using the online C18 trap column with 98% water, 2% acetonitrile and 0.1% formic acid at flow rate of 5 ?L/min for 10 min. The desalted peptides were eluted on a C18 reverse phase analytical column for separation and analysis. A 120 min gradient in multiple steps (ranging from 5–50% acetonitrile in water containing 0.1% formic acid) was set for eluting the peptides from the ChromXP analytical column. The separated peptides were ionized and entered into the mass spectrometer and multiply charged molecules were fragmented using the IDA™ (information dependent data acquisition) criteria of the analyst software for library generation. In brief, 500 ng of all the samples were pooled together and run using IDA criteria of the mass spectrometer for library generation in triplicate. Mass spectrometric data for the first quadrupole was acquired in the range of 350 Da to 1,250 Da whereas 20 most abundant multiply charged peptides were fragmented in the mass range of 150 Da to 1,500 Da in the second quadrupole or collision induced dissociation cell. The accumulation time for each MS/MS experiment was 50 ms. The ionization potential for the turbo V ion source was kept at 4500 V and temperature for source was set 150°C, GS1 and GS2 were at 19 and 15 L/min, respectively. Declustering potential (DP) was set at 80 V. The resultant IDA data files were analyzed in ProteinPilot™ (Sciex software) for identification of peptides and proteins against E. coli proteome using Paragon algorithm and the pooled peptide list was used as the spectral library for SWATH analysis. Experiments were performed in technical replicates.

**Protein Characterization:** SWATH parameters for label free quantification In SWATH™ Sciex (Sequential window acquisition of all theoretical Spectra) DIA (Data independent acquisition) acquisition method, fixed value Q1 transmission window was kept at 12 Da for the mass range of 350–1,250 Da. Total 75 sequential windows were acquired independently with an accumulation time of 62 ms, along with three technical replicates for each of the sets. Total cycle time was kept constant at <5 s. For label free quantification, peak extraction and spectral alignment was performed using PeakView® 2.2 software with the following parameters: number of peptides selected for quantitation 2, confidence of peptide identification was set as more than 95%, number of fragment ions for each peptide was set as 5, extraction ion chromatogram (XIC) peak width was fixed 30 ppm for matching the RT whereas the XIC extraction window set for matching the peptide across the different samples was set at 5 min. For statistical interpretation, MarkerView software™ (ver. 1.3. 1; AB Sciex) was used. The SWATH acquisition data was processed using SWATH™ Acquisition MicroApp (ver. 2.0) in PeakView® software.

Data analysis Raw data was search processed with ProteinPilot™ using the Paragon and Progroup Algorithms for protein and peptide identification. Analysis was also done using the integrated tools in Protein Pilot at 1% false discovery rate (FDR). The identification file was used as spectral library in the quantitation experiment for SWATH/DIA. Retention time calibration was performed using the most abundant peptide across all the samples. Proteins whose log2 fold change (log2FC) values were > 0.5 or < ?0.5 (padj ?0.05) were considered as differentially expressed proteins (DEPs). Both DEGs and DEPs were visualized using ggVolcanoR.3

**Experiment Type:** Bottom-up

**Species:** Data in species_details  No Data

**Tissue:** Data in tissue_details  No Data

**Cell Type:** Data in cell_details  No Data

**Disease:** Data in disease_details  No Data

**Instrument Details:** Data in instrument_details  Data in instrument_details

**Protein Modifications:**

**PubMed ID:** 11893563