



IJIRSET

International Journal of Innovative Research in
SCIENCE | ENGINEERING | TECHNOLOGY

e-ISSN: 2319-8753 | p-ISSN: 2320-6710



INTERNATIONAL JOURNAL OF INNOVATIVE RESEARCH

IN SCIENCE | ENGINEERING | TECHNOLOGY

Volume 10, Issue 7, July 2021

ISSN INTERNATIONAL
STANDARD
SERIAL
NUMBER
INDIA

Impact Factor: 7.569



9940 572 462



6381 907 438



ijirset@gmail.com



www.ijirset.com

Escherichia coli: A Model Organism in Life Science Research

Archana Moon,^{*1} Virendra Meshram², Deeba Khan³, Ankita Pise⁴, Arshiya Mirza⁵, Himanshu Ukey⁶,
Kaustubh Dandekar⁷, Megha Kewalramani⁸, Nitu Payasi⁹, Priya Khatri¹⁰, Samiksha Meshram¹¹,
Shweta Mishra¹²

Post Graduate Teaching Department of Biochemistry, Rashtrasant Tukadoji Maharaj Nagpur University,

Nagpur, India¹⁻¹²

*Corresponding Author

ABSTRACT: *E. coli* is one of the most widely studied micro-organism. *E. coli* has several advantages like its simple nutritional requirements, fast growth rate, steady genetics, and its easy manipulation which has ideally made it a model organism in life science researches. An attempt has been made to introduce the morphology, biochemistry, diversity, genomics, proteomics, interactions and its role in diseases etc. *E. coli* is a host of choice for metabolic engineering to enhance the production of desired chemicals, since it is easily manipulated due to availability of many genetic tools. It is a versatile host for production of recombinant proteins, which are commonly used to produce pharmaceutical products and protein-based polymers for drug delivery. Moreover, this review presents the various roles of *E. coli* in life science research.

KEYWORDS: *Escherichia coli*, morphology, biochemistry, diversity, role in disease, model organism in life science.

I. INTRODUCTION

Escherichia coli

Escherichia coli, originally known as *Bacterium coli* commune, was discovered by the German pediatrician, Theodor Escherichia in 1885. It is widely distributed in the intestine of humans and warm-blooded animals and is the predominant facultative anaerobe in the bowel of the essential intestinal flora that maintains the physiology of the healthy host. In 1892, Sharding proposed the use of *E. coli* as an indicator of fecal contamination because it is abundant in human and animal feces and not usually found in other niches. *E. coli* ferments lactose to produce acid and gas within 48 hrs at 35°C.



Fig.1 : *Escherichia coli*. © 2021 Cleveland Clinic

I. SCIENTIFIC CLASSIFICATION OF *E. COLI*

Classification is necessary to arrange any organism or bacteria into groups so that on the basis of their mutual similarity or evolutionary relatedness it can assist researchers.

DOMAIN: Bacteria

PHYLUM: Proteobacteria

CLASS: Gammaproteobacteria

ORDER: Enterobacterales

FAMILY: Enterobacteriaceae

GENUS: *Escherichia*

SPECIES: *E.coli*

II. BASIC CHARACTERISTICS SHOWING THE MORPHOLOGY OF *E. COLI*

Escherichia coli include in the family of enterobacteriaceae. It is Gram negative facultatively anaerobic rods i.e. possessing both a fermentative and respiratory metabolism. *E.coli* can be either non motile or motile by peritrichous flagella. The cells of *E.coli* are typically 1.1-1.5 micrometer wide by 2-6 micrometer long and occur as straight rods. *E.coli* possess flagella in addition to that some strains also produce fimbriae or Pilli which are proteinaceous structures that extend outwards from the bacterial surface and play role in attachment of the cells to other cells or to host tissues. They are non-spore forming bacteria. *E.coli* are chemoorganotrophic, catalase positive, oxidase negative, and mesophilic and producing acid from glucose. *Escherichia coli* is arranged singly or in pairs. Capsules are present in some *E.coli* strains which can be easily demonstrated using India ink preparation, appear as a clear halo in a dark background(3). Some *E.coli* strains are harmful therefore depending upon the pathogenicity of *E.coli* they are divided into two subgroups (2).

Below is the tree chart showing *E.coli* types:

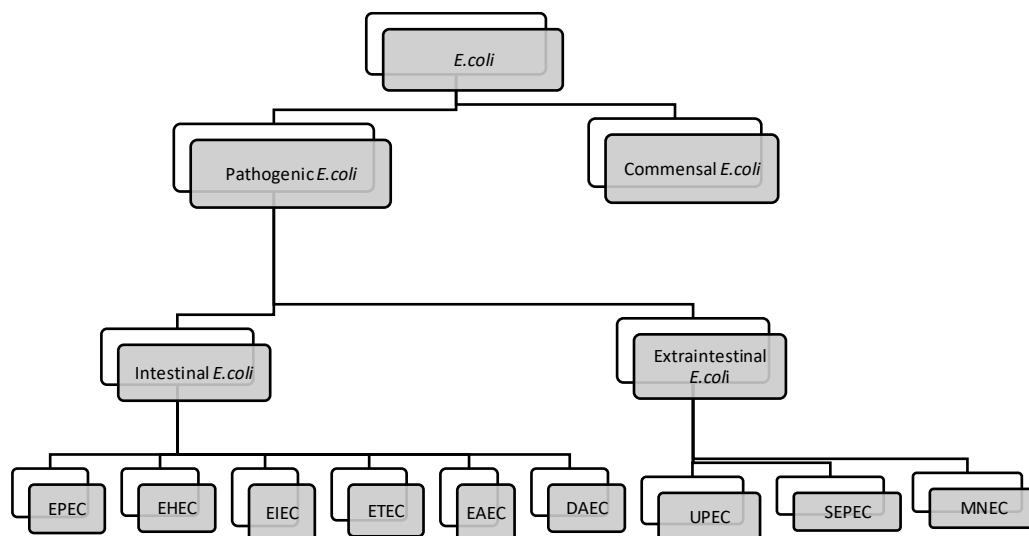


Fig. 2 : Classification of *E.coli* where EPEC is Enteropathogenic *E.coli*, EHEC is Enterohemorrhagic *E.coli*, EIEC is Enteroinvasive *E.coli*, ETEC is Enterotoxigenic *E.coli*, EAEC is Enteraggregative *E.coli*, DAEC is Diffusely adherent *E.coli*, UPEC is Uropathogenic *E.coli*, SEPEC is Sepsis-associated *E.coli*, MNEC is Meningitis-associated *E.coli*.

III. CULTURE REQUIREMENTS OF *Escherichia coli* (*E.coli*)

Special requirements- Normally *Escherichia coli* can readily grow in an ordinary media like nutrient agar medium commonly the (NAM) and MacConkey agar medium is used for the cultivation of *Escherichia coli* in the laboratory. *Escherichia coli* have no complex nutritional requirements.

Optimum temperature- *Escherichia coli* ranges from 10-40°C but the optimum temperature for the most of the strains is 37°C and usually cultivated at this temperature in laboratories.

Optimum pH- *E.coli* can survive at 4.5-9.5 pH but the maximum growth observed at 7.0 i.e. neutral pH. Also, the pH requirements vary as per the strain of *Escherichia coli*.

Oxygen requirements- *Escherichia coli* is an aerobic bacterium i.e. grow best in the presence of oxygen and it is also a facultative anaerobic organism i.e. can grow in a low oxygen environment.

There are various culture media used for the cultivation of *Escherichia coli* in the laboratory and most commonly the Nutrient agar medium and MacConkey agar medium is used, the other media are as follows-

Eosin Methylene Blue Agar (EMB Agar) medium (**selective medium**)

Sorbitol MacConkey Agar medium (**selective medium**)

Modified Tryptic Soy Broth (**selective medium**)

IV. METABOLISM OF *Escherichia coli*

Metabolism is a term that is used to describe all chemical reactions involved in maintaining the living state of cells and the organism. Nutrition is the key to metabolism. The pathways of metabolism rely upon nutrients that they breakdown in order to produce energy. By manipulating the metabolism of *Escherichia coli* we can produce our desired product. The advantages that *E.coli* have like rapid growth, complete genetic background, and the reliable genetic operation technique, *Escherichia coli* has become the preferred host cell for microbial biochemicals production. And it is not surprising that *E.coli* has been widely used by applied microbiologist to try to steer the metabolism of this organism toward the production of molecules with biotechnological value. In most studies the use of *E.coli* for applied processes, sugars are used as carbon and electron source. And their import can be achieved via different uptake machineries. For example, glucose is mainly imported and simultaneously phosphorylated to glucose-6-phosphate using the phosphotransferase system (1).

V. GROWTH AND CELL CYCLE

The *E. coli* Cell Cycle is divided into three Stages B, C and D during which DNA is replicated, chromosomes are segregated and cells grow and divide. In B Phase Cell Cycle events are absent thus called as Gap Phase. During C period DNA is replicated. Bacterial DNAs segregated during D period producing two daughter cells (34). Chromosome Segregation occur simultaneously with DNA replication with certain delay meaning that it starts during C period and extend into D period (35). Cell division is initiated when the divisome protein FtsZ forms a ring structure at midcell, which often occurs in the C period before replication is completed. Cell occur throughout the cycle (8). The time between two consecutive rounds of cell division can be 20 min. During multifork replication, a new round of replication is initiated while previous rounds are still ongoing. So, replication initiation for a new cell cycle take place before cell division of the previous cell cycle has been completed.

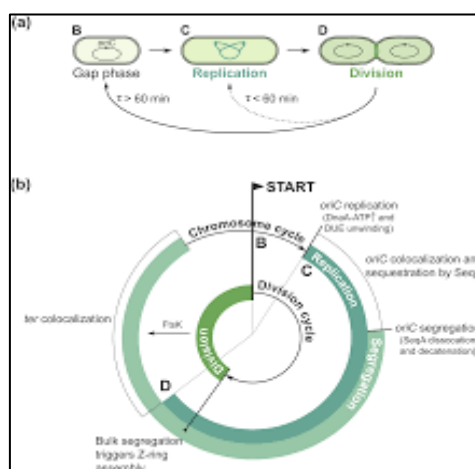


Fig. 3 : The bacterial cell cycle.
Copyright © 2021 Oxford University Press

E. coli can grow with both linear and overlapping chromosome/division cycles, with wide variations in the durations of component processes in different situations among different individual cells.

VI. GENETIC ADAPTION

The Calvin Benson Basshan (CBB) cycle produce organic biomass inorganic carbon. A combination of rational metabolic engineering and laboratory evolution can achieve the successful incorporation of non -native CBB cycle into *E. coli*. The evolved strain is able to utilize CO₂ as an only carbon source for the synthesis of all sugar derived metabolic building blocks. The successful evolution of hemi-autotrophic strain suggests that the non- native carbon fixation pathway incorporated into the metabolic network of heterotrophic host to allow autocatalytic conversion of CO₂ to biomass. Prs and pps genes are introduced into the ancestor strain. The modified strain that contain two additional mutations was still unable to grow hemi-autotrophically, evolving it in a lab resulted in a hemi-autotrophic strain with fewer required mutation. Five mutations are sufficient for the adaption of rationally design ancestor *E. coli* to produce sugar from CO₂.

VII. SEROTYPES

A serotype is defined as the distinct variation within a species of bacteria, virus or among immune cells of different individuals. These organisms and cells are classified on the basis of their cell surface antigens. Kauffmann laid the foundation for *E. coli* serotyping more than 40 years ago. Taking in consideration *E. coli* serotypes there have been traditional methods of identifying *E. coli* which uses antibodies to test for surface antigens: The O- polysaccharide antigens, flagellar H-antigens, and capsular K-antigens. Each *E. coli* has been primarily subdivided into a number of O groups which have been subdivided into subgroups with different K antigens. Each subgroup includes strains with different H antigens; separate numbers were previously allotted for B antigens, but afterwards these numbers have been included in the same consecutive series as the rest of K antigens. There are currently approximately 186 different *E. coli* O-groups and 53 H-types, so serotyping is highly complex. Based on the antigenic diversity among the different O-antigens, they have been targeted as biomarkers for classification of *E. coli* since the 1940s. Serotyping of *Escherichia coli* is useful, but complex, with 173 O antigens, 80 K antigens, and 56 H antigens, which can all be subdivided into partial antigens. The final number of *E. coli* serotypes is very high, 50,000 – 100,000 or more. The two groups of pathogenic *E. coli* are both composed of a limited number of clones for which the O:K:H serotypes are excellent, although not faultless, markers.

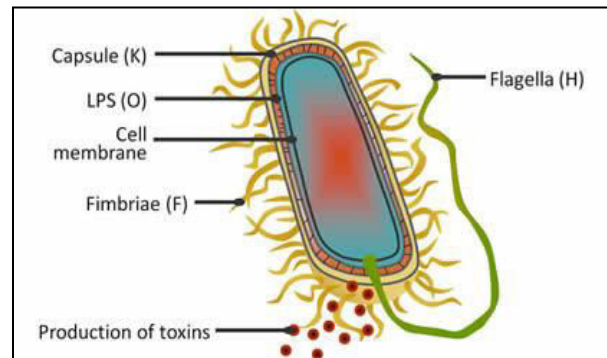


Fig. 4: *E.coli* Surface Antigens
Copyright © 2004 EcL.

I. Genome Plasticity and Evolution:

Living organisms are defined by the genes they possess. Control of expression of this gene set, both temporally and in response to the environment, determines whether an organism can survive changing conditions and can compete for the resources it needs to reproduce. Bacteria are no exception; changes to the genome will, in general, threaten the ability of the microbe to survive, but acquisition of new genes may enhance its chances of survival by allowing growth in a previously hostile environment (4). Genome plasticity refers to the alterable condition of prokaryotic genomes that enables the fluid exchange of DNA from one microorganism to another and allows prokaryotes to adapt their genomes rapidly so they can survive changes in the environmental conditions. *Escherichia coli* is a good example of the interdependency of genome plasticity and microbial traits. *E. coli* pangenome studies indicate that the species *E. coli* is characterized by an open pangenome which indicates that the species still evolves by gene acquisition and diversification. Evolution of the *E. coli* variants as we know them today results from ordered gene acquisition events and loss of genetic information together with DNA rearrangements and point mutations (5).

II. *E. coli* Phylogeny :

The history of the evolution of species or group, especially in reference to lives of descent and relationship among broad groups of organisms (10). Actually, the phylogenetic grouping can be done by multi-locus enzymes electrophoresis or ribotyping. Phylogenetic analysis has shown that *E. coli* strains fall into four main phylogenetic group i.e. A₁, B₁, B₂, and D which can be subdivided into seven groups – A₀, A₁, B₁, B₂, B₂, D₁ & D₂(10). *Escherichia coli* isolates from the non- pathogenic phylogenetic groups A and B₁ rarely cause extraintestinal infections (11). *Escherichia coli* is one of the most versatile bacterial species. It alternates between its primary habitat, the gut of vertebrates, where it lives as a commensal and its secondary habitat, water and sediment. It may also function as an intra- and extraintestinal pathogen in human and many other animal species (11).

III. Guide to *E. coli* Genotype and Genetic Marker Nomenclature

Choosing the right competent cell line for your assay and achieving a successful transformation are dependent on understanding genotypes and genetic markers. Here, we describe the nomenclature of *E. coli* genotypes to help guide you in choosing your competent cells.

Genotype

The genotype (genes) of an organism, including *E. coli*, is expressed resulting in a phenotype, which is an observable and measurable characteristic. The genotype will contain information for genes, mutation types, allele numbers, antibiotic resistance, insertions, fusions, inversions, plasmids and deletions carried by a specific bacterial strain.

Genes: In *E. coli*, the genotype only includes the genes that carry a loss of function mutation. The gene name is listed as three letters in lower case and italics (DNA methylase is written as *dam*). Different genes affecting the same function/pathway are identified with different uppercase italic letters. *GalE* and *galK* are examples of genes encoding proteins that function in the same pathway.

Prophages: A prophage, also known as a “phage,” can infect bacterial cells resulting in its genome integrating into the bacterial DNA chromosome (lysogenization) or becoming an extrachromosomal plasmid (lytic growth). $\phi 80$ = Strain carries lambdoid prophage $\phi 80$. Recognizes the FhuA receptor on *E. coli* allowing binding to *E. coli* membrane.

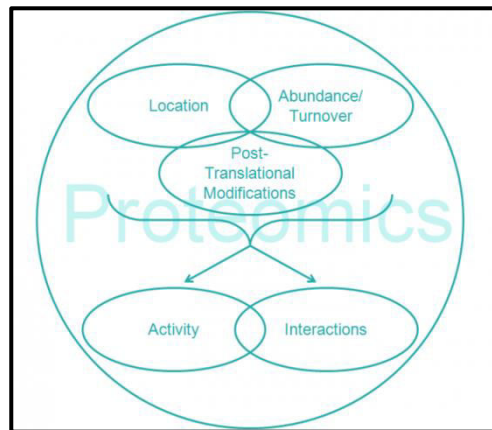
Restriction-modification system: *E. coli* have restriction (cleavage of DNA) and modification (by methylation) systems that confer protection from foreign DNA.

Other Phenotypic Traits: Relevant phenotypic characteristics, including antibiotic resistance and sensitivity to cold, can also be described in the genotype using the following specific nomenclature. Antibiotic resistance = Resistance to a specific antibiotic is indicated by two, three or four letters for the antibiotic and a superscript uppercase R (R).

StrRSmR = streptomycin resistance

CarbR = carbenicillin resistance

Sensitivity: Conditional alleles with different sensitivities are indicated in the genotype. Sensitivity to cold and temperature are indicated as lowercase superscript cs (cs) and ts(ts), respectively. leuAts = Temperature sensitive allele of leuA.



Other conditional alleles: Other mutations can be indicated with two lowercase superscript letters. Examples are the amber mutation, the opal mutation and the ochre mutation, which introduce a stop codon and results in a truncated protein product.

argEam = Amber mutation in argE introducing a nonsense codon (UAG).

argEop = Opal mutation in argE introducing a nonsense codon (UGA).

argEoc = Opal mutation in argE introducing a nonsense codon (UAA) (12).

IV. Proteomics :

Proteomics is the large-scale study of the structure and function of proteins in complex biological sample. A proteome is a set of proteins produced in an organism, system, or biological context. We may refer to, for instance, the proteome of a species (for example, *Homo sapiens*) or an organ (for example, the liver). The proteome is not constant; it differs from cell to cell and changes over time. To some degree, the proteome reflects the underlying transcriptome. However, protein activity (often assessed by the reaction rate of the processes in which the protein is involved) is also modulated by many factors in addition to the expression level of the relevant gene. Such an approach has the potential value to understand the complex nature of the organism. Current proteomic tools allow large-scale, high-throughput analyses for the detection, identification, and functional investigation of proteome.

Normal Microbiota

Normal microbiotas are the microorganisms that live on another living organism (human or animal) or inanimate object without causing disease. The human body is not sterile; we become colonised by bacteria from the moment we are born. We are covered with, and contain within our intestines, approximately one hundred trillion bacteria that form the normal flora of our bodies. This normal flora helps to prevent us becoming colonized with more dangerous bacteria, which might lead to infection.

Normal Microbiota of *Escherichia coli*: *Escherichia coli* are very common bacteria in gastrointestinal tract, and part of the normal bacterial flora. However, some *E. coli* strains are able to produce a toxin that could produce serious infection (13).

Therapeutics uses of *Escherichia coli*

E. coli is most beneficial microorganism to express heterologous proteins for therapeutic uses as around 30% of approved beneficial proteins are currently being produced using it as a host due to its rapid growth high yield of product, cost effectiveness and easy scale up process, *E. coli* is an expression host of choice in the biotechnology industries for large scale production of protein, specially non glycosylated proteins for therapeutic use.

Escherichia coli used as vector for drug delivery

The development of save advanced tool for gene therapy and other novel therapies will allow for targeted drug delivery of pharmaceutically active agents and carries promise to improve therapies both through increase efficiency as well as low toxicity.

II. *E.coli* as a model organism

Because of its easy manipulation this organism has various advantages in the fields of life science researches.

In biotechnology, *E.coli* has led to the development of strains capable of synthesizing several novel compounds. *E.coli* is used in genetic engineering and synthetic biology, because it is a suitable platform for the useful production of molecules for the pharmaceutical industry, metabolites and molecular relevant for food additives, pigments and more recently complex aliphatic molecules. (14).

E.coli is a very versatile host for the production of heterologous or recombinant proteins. Heterologous expression refers to the expression of a gene or part of a gene in a host organism which does not naturally have this gene or gene fragment. As various proteins expression systems have been developed which allows the production of recombinant proteins in *E.coli*. Researches can introduce genes into the *E.coli* using plasmid which permit high level expression of protein, and such protein can be mass produced in industrial fermentation processes. One of the first useful applications of recombinant DNA technology was the manipulation of *E.coli* to produce insulin for diabetic patients (15).

E.coli is a primary host in metabolic engineering for the production of industrially important biochemicals such as biofuel, organic acids, and amino acids. *E.coli* are examined largely for their capability to produce biofuels due to the presence of well- established tools for genetic modification, well studied growth metabolism, and their successful use in other industrial applications. Especially, *E.coli* has the unique advantage of being the best studied model organism in terms of gene regulation and expression and also as an organism with the largest molecular tools available for genetic engineering. *E.coli* strains can naturally utilize a variety of carbon sources under both aerobic and anaerobic conditions. Metabolic engineering of *E.coli* for the enhanced biochemical production has been significantly promoted by the integrated use of recent developments in system biology, synthetic biology and evolutionary engineering (16).

Genome modifications to understand *E.coli*

In *E. coli*, several tools for genome modification have been developed. Some of the most important methods involve either the generation of deletion mutants by removing specific genes; one outstanding case is the use of the lambda Red system for inhibiting linear DNA degradation and by homologous recombination, the deletion of specific genes using PCR-derived selection marker cassettes with homologous sequences with target gene. By using

Cre/lox recombination, considerable genomic fragments can be deleted or sequentially removed, rendering the nonessential regions (regardless the genes present) from the genome. In the case of *E. coli*, there are several pieces of evidence that points out that at least 23% of the genome can be eliminated gaining genomic stability and normal growth. Also, eliminating insertion sequences can enhance the capacity of *E. coli* to synthesize proteins due to the decrease or insertions on plasmids and strains exhibit normal growth plus increased genome stability.

Literature Review :

Metabolic Engineering of *E.coli*

BIOFUELS PRODUCTION

The production of fuel substitutes such as bio-alcohols and biodiesel from renewable resources has gained significant attention because of the rising energy price and environmental concerns. Generally, biofuels include the bioethanol and other advances fuels like butanol, propanol, isopropanol.

Bioethanol Production

Bioethanol is the predominant renewable liquid energy source capturing 90% of the current world biofuel market even though it has certain limitations such as corrosiveness and low energy. The native *E.coli* is capable of producing ethanol by an endogenous process under anaerobic conditions in which one mole of glucose is metabolized into two moles of formate, two moles of acetate, and one mole of ethanol (Fig. 37A). In the last step there is a reduction of acetyl-coA into ethanol by AdhE (aldehyde alcohol dehydrogenase *E.coli*). The reduction reaction consumes two NADH molecules, while the initial step in which glucose is converted to pyruvate produces only 1NADH leading to redox imbalance. This native fermentation process leads to the sub optimal level of production of ethanol, which is estimated to be 0.26 g ethanol/g of glucose. To overcome the problems existing in the endogenous ethanol production process, Ingram et al have made successful attempts to produce high yields of ethanol by genetic engineering of *E.coli* by inserting genes such as *pdc* and *adhB* from *Z.mobilis*. These genes were expressed in operon from a plasmid under a constitutively expressed artificial *pet*(production of ethanol) promote to produce pyruvate decarboxylase and alcohol dehydrogenase II. This heterologues fermentation pathway produces 95% of the final products as ethanol without creating any redox imbalance (Fig. 37B). In order to further stabilize the *E.coli* to continuously generate ethanol production, Ohta et al constructed an *E.coli* ATCC 11303 strain KO3 through chromosomal integration of *pdc* (pyruvate decarboxylase), *adhB* (alcohol dehydrogenase II) genes along with a selective chloramphenicol resistance gene (16).

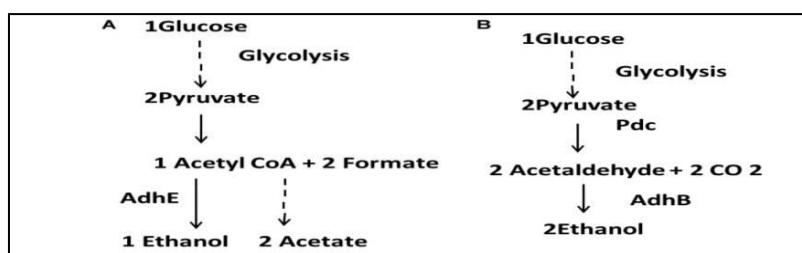


Fig. 6: Strategies for the production of bioethanol from *E.coli* (A) Endogeneous ethanol production pathway for ethanol production in *E.coli*. (B) Metabolic engineering pathway for ethanol production in which endogenous *E.coli* ethanol production pathway was engineered by expressing *pdc* and *adhB* genes of *Zymomonasmobilis*).

Production of *n*-butanol

Higher carbon atoms, such as *n*-butanol and isobutanol as fuels are much better than ethanol due to their less corrosiveness, high energy, high blending capability and their use in conventional combustion engines without modifications. *Clostridium* has a native butanol- production pathway and it produces a carbon yield of 0.41 g butanol/g of glucose. However, it also produces several byproducts and cause NADH redox imbalance and thus lowering the

productivity yield. And also, the genetic tools are not available to modify clostridium to overcome this redox imbalance genetically engineered *E.coli* is the best choice for production of *E.coli* (18).

Glycerol is a byproduct generated in the production process of biodiesel. As compared to glucose, glycerol has highly reduced carbon atoms and enables generation of more reducing equivalents. However, glycerol metabolism is less effective than glucose metabolism in *E.coli*. In recent study, researchers used engineered *E.coli* for the production of n-butanol from glycerol. They carried out this by glycerol catabolism and by manipulations of central metabolism. *E.coli* strains applied in their studies are BuT-8, BuT-12A, BuT-12-2, BuT-12-3, BuT-16. In *E.coli*, two catabolic pathways of glycerol existing include the *glpK-glpD* and the *gldA-dhaKLM*. The *gldA-dhaKLM* pathway mediated the conversion of glycerol to dihydroxyacetone phosphate (DHAP) generates extra NADH and this may increase the n-butanol production. Strain BuT 12-2 was thus obtained by genomic fusion of a strong promoter with *gldA* and the *dhaKLM* operon to enhance their expression levels in strain BuT-12A. At the end of fermentation, the strain utilized 10.5 g/L glycerol and produced 3 g/L n-butanol at 24 hr (Fig. 7). The synthesis for one molecule of n-butanol requires 4 molecules of NADH and 2 molecules of acetyl CoA (19).

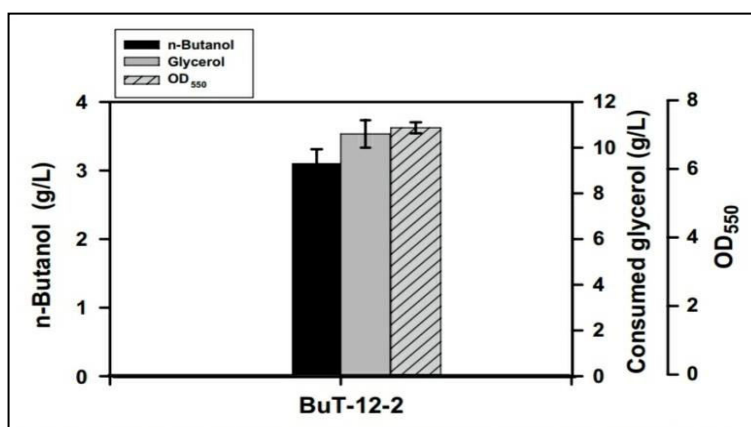


Fig. 7 : Microaerobic production of n-butanol in the strain with the enhanced glycerol catabolism. *E.coli* strain BuT-12-2 was engineered by enhancement of the *gldA-dhaKLM* catabolic route. The strain was grown in M9Y medium containing 20 g/L pure glycerol and the fermentations were carried out for 24 hr (19).

Production Of Organic Acids

Organic acids are an important component of the building –block chemicals that can be produced by microbial processes. Most microbial organic acids are end products, or at least natural intermediated in major metabolic pathways. Because of their functional groups, organic acids are extremely useful as starting materials for the chemical and food industries. Recent advances in strain development for organic acid production by metabolic engineering of *E.coli* are discussed (20).

Production Of Succinic Acid

Succinic acid was identified as one of the top 12 building block chemicals by the U.S. department of energy. It is not only used in many fields such as agriculture, chemical, food, and pharmaceutical industries, but also can offer a huge potential alternative as a C₄-dicarboxylic acid platform that can be converted to other chemicals such as 1,4-butanediol, tetrahydrofuran, N-methyl pyrrolidinone.

E.coli can utilize multiple pathways to form succinic acid. *E.coli* can naturally produce succinic acid but at a small amount as an intermediate of the central metabolism or as a fermentation end product via the reductive tricarboxylic acid (TCA) cycle (Fig.40) There are three alternative paths to form succinic acid: the PEP-pyruvate-oxaloacetate node, the oxidative TCA branch and the glyoxylate shunt (Fig.8) under anaerobic conditions (21). In order to improve succinic acid production, genetically modified metabolic pathways are first constructed to increase the key enzyme activity. For example, succinic acid production from glucose by *E.coli* was significantly increased by over-expression of endogenous phosphoenolpyruvate carboxylase (PEPPc). Using *E.coli* AFP111 as biocatalyst in which the

pyc gene (pyruvate carboxylase) was over expressed a final succinic acid concentration and productivity reached 99.2 g/l and 1.3 g/l/h. As fumarate reductase (*frdABCD*) is also a key enzyme in this metabolic pathway, two recombinant *E.coli* strains with amplified *frdABCD* activity have been constructed for the conversion of fumarate to succinic acid. The other metabolic engineering strategy to increase the production of succinic acid was to knock out or inhibit the enzyme in succinic acid competition pathways. Sanchez et al. constructed *E.coli* strain, SBS550MG that inhibited the central metabolic pathway by deactivating *adhE* (encoding aldehyde dehydrogenase), *ldhA* (encoding lactate dehydrogenase) and *ack-pta* (encoding acetate –phosphate acetyltransferase) from the central metabolic pathway and activating the glyoxylate pathway through the inactivation of *iclR* (encoding isocitrate lyase). Because anaerobic fermentation of succinic acid has a number of disadvantages, Lin et al. formed a mutant *E.coli* strain by deletion of *sdh* (encoding succinic acid dehydrogenase), *poxB* (encoding pyruvate oxidase), *pta-ack*, *iclR*, and *ptsG* genes, and overexpression of *ppc* gene. Fed-batch fermentation of Lin's strain resulted in a concentration of up to 58.3 g/l of succinic acid with an overall yield of 0.85 mol/mol glucose (22).

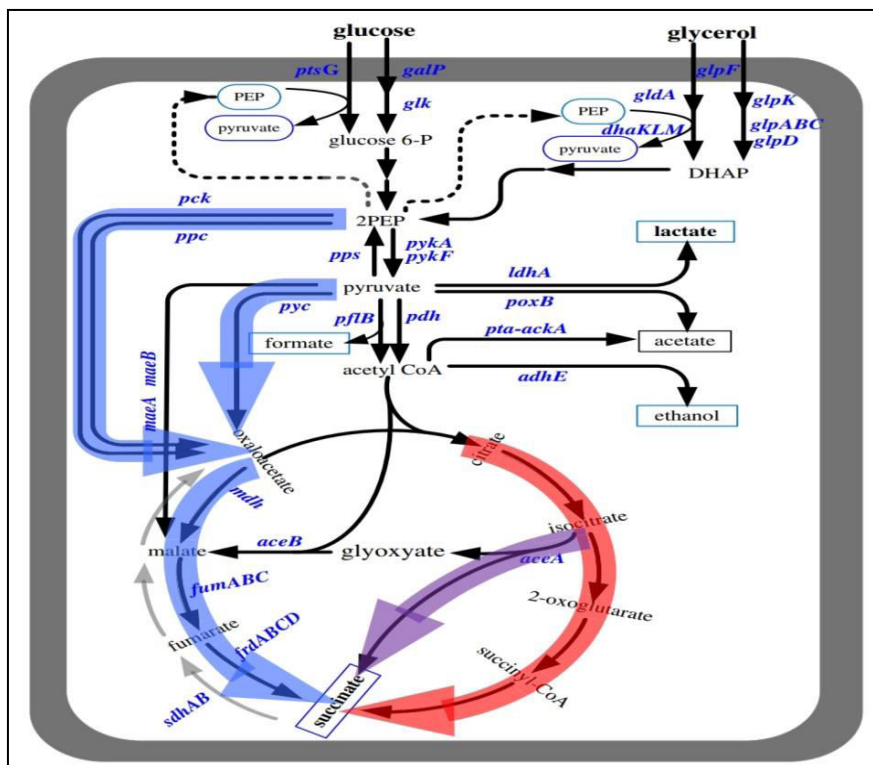


Fig. 8 : Lactic acid and succinic acid pathways from glucose and glycerol in *E. coli* and enzymes involved in metabolic regulation. The three pathways for succinic acid production are indicated by the thick blue, red, and purple arrows, respectively. Relevant biochemical reactions are represented by the names of the gene(s) coding for the enzymes: *aceA*: isocitrate lyase; *aceB*: malate synthase; *ackA*, acetate kinase; *adhE*, acetaldehyde/alcohol dehydrogenase; *dhaKLM*, dihydroxyacetonekinase; *frdABCD*, fumarate reductase; *fumABC*, three isoenzymes of fumarases; *gldA*, glycerol dehydrogenase; *glk*, glucose kinase; *glpABC*, anaerobic glycerol-3-phosphate dehydrogenase; *glpF*, glycerol facilitator; *glpK*, glycerol kinase; *ldhA*, D-lactate dehydrogenase; *maeA/maeB*, NADH/NADPH malice enzymes, respectively; *mdh*, malate dehydrogenase; *pck*, phosphoenolpyruvate carboxykinase (*E. coli* or *A. succinogenes*); *pdh*, pyruvate dehydrogenase complex; *pflB*, pyruvate formate-lyase; *poxB*, pyruvate oxidase; *ppc*, phosphoenolpyruvate carboxylase; *pps*, PEP synthase; *pta*, phosphoacetyl transferase; *ptsG*, phosphotransferase system; *poxB*, pyruvate oxidase; *pyc*, L lactis

Lactic Acid Production

Lactic acid is an important organic acid. It is a bulk chemical increasingly used in the food and pharmaceutical industries products such as food, beverages and biodegradable polymers. Lactic acid is produced in two racemic isomers forms such as D or L lactic acid. The production of lactic acid is of biotechnological interest mostly due to the production of polylactic acid (PLA) based plastic materials. Previously, lactic acid bacteria were used for the production of lactic acid but due to its high nutritional requirements *E.coli* seems to be better choice for the production of lactic acid.

In *E.coli*, D-lactate is produced by homofermentative or heterofermentative pathways together with NADH consumption in order to achieve a redox balance under anaerobic culture conditions (Fig.8) however; *E.coli* does not naturally produce L-lactate in noticeable amounts, but can catabolize it via a dehydrogenation reaction mediated by an FMN- dependent L- lactate dehydrogenase. In order to expand the use of lactic acid especially in biopolymers, it is necessary to improve the optical purity of the acid required for polyactide use in bio-based plastics, and to reduce production costs. Therefore, *E.coli* strains have been developed for effective production of D- and L- lactic acid using a metabolic engineering approach (17).

In a first successful example of D- or L- lactate production by engineered *E.coli*, Chang et.al created two strains. One strain produced D-lactate with a titre of 62.2 g/L and one produced L-lactate with a concentration of 45 g/L. Two genes *pta* and *ldhA* were deleted and L- lactate dehydrogenase gene from *Lactobacillus casei* was overexpressed in the L-lactate producing strain (193). In aims to make the *E.coli* strain to produce L- lactate, Niu et.al constructed three L- lactate dehydrogenase genes from different bacteria. The L-lactate producing strains 090B1, 090B2, 090B3 were developed from a previously developed D- lactate over producing strain which is B0013-070 by three different ways 1) deleting *IdhA* to block D-lactate formation. 2) deleting *IldD* to block the conversion of L-lactate to pyruvate and 3) expressing an L- lactate dehydrogenase to convert pyruvate to L-lactate under the control of the *IdhA* promoter. Fermentation tests were carried out in a shaking flask and in a 25-l bioreactor. Strains 090B1, 090B2 or 090B3 were shown to metabolize glucose to L- lactate instead of D- lactate. *E.coli* 090B3 was able to produce 142.2 g/l of L- lactate with no more than 1.2 g/l of byproducts (mainly acetate, pyruvate and succinate) accumulated (23).

Zhou et.al manipulated a wild type *E.coli* B0013 by deletion of eight genes (*ackA-pta*, *pps*, *pflB*, *dld*, *poxB*, *adhE*, and *frdA*) to over produce D-lactate. Coupling lactate fermentation with cell growth was employed and the cells in the bioreactor yielded an overall volumetric productivity of 5.5 g/L/h and a yield of 86 g lactic acid/100 g glucose which were 66% higher and at the same level compared to that of the aerobic and oxygen limited two phase fermentation. Mazumdar et al. created a recombinant *E.coli* strain by overexpressing pathways involved in the conversion of glycerol to D-lactate and the elimination those leading to the synthesis of competing by-products. Thus, an engineered homofermentative *E.coli* strain was developed that converted 40 g/L glycerol to 32 g/L D-lactate with a yield of 85% of the theoretical maximum (24).

Strategies For The Production Of Recombinant Protein In *E.coli*

In the recent past years, a large number of proteins have been expressed in *Escherichia coli* with high productivity due to fast development of genetic engineering technologies. There are many hosts used for the production of recombinant protein but the preferred choice is *E. coli*. *E.coli* facilitates protein expression by its relative simplicity, its inexpensive and fast high-density cultivation the well-known genetics. But there may occur a problem in the expression of foreign genes in *E. coli*. Researchers frequently face problems producing soluble recombinant proteins for over-expression. Here we describe the few methods that can improve the level of expression, non-expression, or less expression of the gene of interest in *E. coli*. 1- changing the vector, 2- changing the host strain 3- changing the culture parameters of the recombinant host strain, 4- co-expression of other genes.

Changing The Vector

Changing the vector means either changing the promoter with which the gene of interest is to be cloned or changing the fusion 'tag' which control solubilization of recombinant protein in *E. coli* host (Table 1). The pET vector with 6-Histidine tag (His-tag), is usually the first choice to get recombinant protein because His-tag is a smaller affinity tag with high levels of expression due to the presence of a strong promoter and the presence of robust purification system like Ni-NTA. When a histidine tagged gene does not express the recombinant protein, the gene of interest should be cloned in pGEX system of GST tag. It has been found that maltose binding protein (MBP) tag and GST tag

increases solubility of fused proteins and expression of genes. Tags may affiliate with the structure and function of the target protein; therefore, provision tags must be taken out after the expression and purification of the protein. Multiple cleavage sites can be engineered into the expression construct to remove tags. Enterokinase (cleavage site DDDDK:), Factor Xa (cleavage site IEGR:), SUMO protease (recognizes SUMO tertiary structure and cleaves at the C-terminal end of the conserved Gly–Gly sequence in SUMO), Thrombin (cleavage site LVPR:GS) are the common proteases used to remove tags often from recombinant protein (Table 2). SUMO tag recombinant protein has a tendency to be soluble in the *E. coli* host. pSUMO vector has an ability to get recombinant protein without any extra amino acid after cleavage of the tag by SUMO protease. N-utilization substance A (NusA) tag also increases solubility of the tagged protein. Solubility modeling of *E. coli* protein shows that NusA tag has the highest potential for solubility. pET-43 and pET-44 vectors can be used for NusA tagged recombinant protein. Fusion of 109-aa long Trx.tag thioredoxin protein also increases the solubility of conjugated protein. The pET-Trx plasmid vector can be used for this tag. Fusion tags like NusA, Trx and SUMO needs extra affinity tag for purification. A vector should be chosen in such a way with consideration that the tag increase solubility but does not disrupt the activity of the recombinant protein or the cleavage of the tag.

Tag	Protein	Size
GST	Glutathione S- transferase	211 aa
MBP	Maltose- binding protein	396 aa
NusA	N- utilization substance	495 aa
Sumo	Small ubiquitin modifier	100 aa
Trx	Thioredoxin	109 aa

Table 1 : Some common tags used in enhancing protein solubility

Changing The Host Strain

In the genome of *E. coli* strain BL21(DE3) a T7 polymerase encoding gene is introduced and *Lon* and *OmpT* protease deleted. *Lon* and *OmpT* protease deficient strain of *E. coli* are unable to degrade foreign protein. Leaky expression of the desired gene in BL21(DE3) cell is possible. To minimize leaky expression of toxic gene, the BL21 host strain was improved. The improved strain is BL21(DE3)pLys S and BL21(DE3)pLys E. Both the strain has lysozyme coding plasmid. Lysozyme is inhibitor of T7 polymerases which inhibits residual T7 polymerase and thus prevents leaky expression. Many times, expression does not occur in *E. coli* due to a difference in codons present in gene of interest and preferential codon use by the host *E. coli* strain. To solve this codon bias in host cell machinery, a few companies provide modified *E. coli* hosts that have extra tRNA coding genes (AUA, AGG, AGA, CUA, CCC and GGA) compensating for the scarcity of rare tRNA (Fig. 9). Rosetta host strains are BL21(DE3) derivatives designed to enhance the expression of heterologous proteins by containing codons rarely used in *E. coli*. Lemo21 (DE3) is derivative of BL21(DE3) strain and are genetically engineered in such manner that it can act as C43(DE3) self-mutant strain. T7 polymerase is cloned with the well-titrable rhamnose promoter and its activity can be precisely controlled by its natural inhibitor T7 lysozyme found in this strain. This strain is also used for the expression of difficult proteins such as membrane protein, toxic protein or the protein that aggregates in *E. coli* without compromising their stability and function. The flexibility of this strain makes it potent, time-saving and cost-effective to identify the optimal conditions for the overexpression of protein of interest. Strains of *E. coli* frequently used in recombinant protein production are shown and summarized in Fig. 9 (25).

9103

Through relatively small set of mutations a dramatic metabolic shift from heterotrophy to carbon fixation and sugar synthesis can occur.

In this review, the improved understanding of the molecular mechanism of *E.coli* pathogenesis has also disclosed new aspects of host response to infecting pathogens. Once, dismissed as a harmless inhabitant of the intestinal tract, *E.coli* is now seen as a pathogenic species with remarkable versatility to cause disease in humans and animals. Studies of several pathotypes, EPEC, EHEC, EIEC, EAEC and DAEC have been particularly defined and the molecular and cellular basis of pathogenesis of both of these pathotypes is emerging. Phylogenetic group B2 and D were prevented in ExPEC, within the *E.coli* species, the B2 phylogenetic group which is different from the other strains and represent the main strains that are involved in various infection in Extraintestinal *E.coli*. Knowledge of the pathogenic mechanism of *E.coli* pathotypes has shown to the development of treatment and prevention of *E.coli* induced disease. The apparent evolutionary success of derived a virulence has some interesting outcome on the evolution of virulence and commensalism.

E. coli have proven to be versatile vectors and able to deliver a wide range of therapeutic payloads including DNA for vaccination and for gene therapy. Microbes have been individually engineered to diagnose and treat a variety of diseases and disorders, including inflammatory disorders, and metabolic disorders. *E.coli* is a commensal bacterium of the human gastrointestinal tract. It is found in the mucus layer lining of the epithelium of the intestine and is generally symbiotic with its host. However, the high plasticity of the genome of this bacterial species gives it a tremendous capacity to evolve, leading to the emergence of pathogenic strains from the commensal strains. The production of proteins in sufficient amounts is key for their study or use as biotherapeutic agents. Production of recombinant protein using *E.coli* is less costly than using other hosts and the handling is also easier. Therefore, it is beneficial to utilize *E.coli* as a host for production of desired protein. The information described here regarding to protein expression can subsequently be combined and utilized to express protein in *E.coli*, minimizing the current problems researches face and may also eliminate barriers and allow for successful implementation and management of multi parallel expression strategies. The discussed strategies have had a significant impact on recombinant protein expression in *E.coli*. This is of utmost importance in the future for enhancing production of recombinant proteins.

Unquestionably, *E.coli* has long been the workhouse of molecular biology and a preferred industrial host organism for genetic and metabolic engineering applications. Availability of genetic and molecular tools to engineer existing native pathways or to create a synthetic new pathway has made *E.coli* as the microorganism of best choice in order to produce biofuels from renewable energy source. As *E.coli* has been object of research since decades, its versatility still challenges new possibilities for investigation also in the future.

REFERENCES

- Forster, Andreas H., and Johannes Gescher. "Metabolic engineering of *Escherichia coli* for production of mixed-acid fermentation end products." *Frontiers in bioengineering and biotechnology*, 2, 16.
- <https://about-ecoli.com/>
- <https://paramedicsworld.com/escherichia-coli/morphology-culture-characteristics-of-escherichia-coli/medical-paramedical-studynotes>
- Bennett, P. (2021). Genome Plasticity. Retrieved 5 April 2021.
- Watanabe, H., Mori, H., Itoh, T., & Gojobori, T. (1997). Genome plasticity as a paradigm of eubacteria evolution. *Journal of Molecular Evolution*, 44(S1), S57-S64. doi: 10.1007/pl00000052
- LiselotDewachter, Natalie Verstraeten, Maarten Fauvart, Jan Michiels An integrative view of cell cycle control in *Escherichia coli*. *FEMS Microbiology Reviews* 2006; Volume 42, Issue 2: Pages 116–136.
- Nielsen HJ, LiY, Youngren B et al. Progressive segregation of the *Escherichia coli* chromosome. *MolMicrobiol* 2006; Volume 61, Issue 2: Pages 383–393.
- Inoue I, InoR, NishimuraA. New model for assembly dynamics of bacterial tubulin in relation to the stages of DNA replication. *Gene Cell* 2009; Volume 14, Issue 3: Pages 435–44.
- Bram Van den Bergh, Joran E. Michiels, TomWenseleers, EthelM. Windels, Pieterjan Vanden Boer, DonaatKestemont. Frequency of antibiotic application drives rapid evolutionary adaptation of *Escherichia coli* persistence. *Nature Microbiology* 2016 march 7; Article No. 16020.
- Olivier Clermont, Stephane Bonacorsi, and EdourdBingene, Rapid & Simple determinationof the *E. coli* phylogenetic group. *Appl Environ Microbial*. 2000 Oct;
- Olivier Clermont, ManiWennOlier, Olier, Claire Hoede, Laure Diancourt, Sylvain Braise, Monique Keroudean, Jeremy Glodt, Bertrand Picard, Eric Oswald, Eric Denamur, Animal and human pathogenic *Escherichia coli* strains share common gene backgrounds, O, Clermont et.al/ *Infection, Genetic and Evolution* 11(2011) 654-662.



12. Guide to *E. coli* genotype and Genetic marker nomenclature by Fernanda Ruiz, PhD
13. Campbell NA, Reece JB. Biology, Pearson Education 2. Inc, San Francisco, 2002.
14. Vargas- Maya Naura Idalia and Franco Bernardo (July 2017), "*E.coli* as a model organism and its application in biotechnology." DOI: 10.5772/67306.
15. Tof and Ilanit 1994, "Recombinant DNA Technology In the synthesis of human insulin." Little Tree Pty.Ltd.
16. Koppolu and Vasigala, " Role of *E.coli* in biofuel production." Microbiology Insights 2016;9 29-35 doi: 10.4137/MBI.S10878.
17. Kay S, Hahn S, Marois E, Hause G, Bonas U. 2007. A bacterial effector acts as a plant transcription factor and induces a cell size regulator. Science 318:648–651.
18. Atsumi S, Cann AF, Connor MR, et al, "Metabolic engineering of *E.coli* for 1-butanol production." *Metab Eng.* 2008;10 (6):305-311.
19. Mukesh Saini, Ze Win Wang, Chung- Jen Chiang and Yun-Peng Chao, "Metabolic engineering of *E.coli* for production of n-butanol from crude glycerol." Saini et al. *Biotechnol Biofuels* (2017) 10: 173.
20. Sauer M, Porro D, Mattanovich D, Branduardi P, "Microbial production of organic acids." *Expanding the markets, Trends Biotechnol* 2008; 26: 100-8.
21. Xianzhong Chen, Li Zhou, Kangming Tian, Ashwani Kumar, Suren Singh, Bernard A. Prior, Zhengxiang Wang, "Metabolic engineering of *E.coli*: A sustainable industrial platform for bio-based chemical production." 2013 Elsevier Inc.
22. Chao Yu, Yujin Cao, Huibin Zou, Mo Xian, "Metabolic engineering of *E.coli* for biotechnological production of high-value organic acids and alcohols." *Appl Microbial Biotechnol* (2011) 89: 573-583.
23. Dandan Niu, Kangming Tian, Bernard A Prior, Min Wang, Zhengxiang Wang, Fuping Lu and Suren Singh, "Highly efficient L- lactate production using engineered *E.coli* with dissimilar temperature optima for L-lactate formation and cell growth." Niu et.al. *Microbial Cell Factories* 2014, 13:78.
24. Zhou L, Niu DD, Tian KM, Shen W, Shi GY, Singh S, et al, "Improvement of D-lactate productivity in recombinant *E.coli* by coupling production with growth." *Biotechnol Lett* 2012a; 34: 1123-30.
25. Gopal Jee Gopal. Awanish Kumar, "Strategies for the production of recombinant protein in *E.coli*." *Protein J.*
26. Graslund S, Nordlund P, Weigelt J, Gunsalus KC (2008) *Nat Methods* 5 (2): 135-146.
27. deMarco A, Vigh L, Diamant S, Goloubinoff P(2005) *Cell stress Chaperones* 10:329-339.



**Impact Factor:
7.569**



INTERNATIONAL JOURNAL OF INNOVATIVE RESEARCH

IN SCIENCE | ENGINEERING | TECHNOLOGY

 **9940 572 462**  **6381 907 438**  **ijirset@gmail.com**



www.ijirset.com

Scan to save the contact details