Journal of Pharmacy

The challenges of using *Escherichia coli* as a host in recombinant insulin production

Muhammad Taher

Department of Pharmaceutical Technology, Kulliyyah of Pharmacy, International Islamic University Malaysia, Jalan Sultan Ahmad Shah, 25200 Kuantan, Pahang, Malaysia.

Editorial

Article history:

Received: 14 January 2025
Accepted: 28 January 2025
Published: 31 January 2025

doi: 10.31436/jop.v5i1.380

Diabetes mellitus is a metabolic disease characterised by elevation of blood glucose level which leads to serious damage to the blood vessels, eyes, heart, kidney, and nerves affecting about 830 million people worldwide. The most common diabetes is type 2 which usually happened in adults when the body becomes resistant to insulin, or the body does not produce enough insulin. Type 1 diabetes mellitus is dependent on insulin which required accessible and affordable insulin (Diabetes, 2025).

In Malaysia, the prevalence of diabetes mellitus (DM), depends on factors such as gender, age, and ethnicity, with women, the elderly, and the Indian community having the highest prevalence of DM. In the 103,063 participants that made up the study's sample, the combined prevalence of diabetes by gender in the population-based studies was 13.80% for men and 14.54% for women, while the combined prevalence of prediabetes was 11.40% for women and 10.98% for men (Akhtar et al., 2022). For age, from this study, it can be observed that the prevalence of diabetes showed a notable upward trend as people aged, rising from 3.16% in the 20–29 age group to 13.71% in the 30-45 age group, 25.66% in the 46–59 age group, and 33.45% in the 60 and older age group (Akhtar et al., 2022). Ethnicity and races can also affect the prevalence of DM. The subpopulation of Indian had the greatest prevalence of diabetes which is 25.10%, among all ethnic groups, followed by Malay with 15.25%, Chinese with 12.87%, Bumiputera with 8.62%, and others with 6.91%. The prevalence demands oral hypoglycaemic agents (OHAs) market size in Malaysia at USD282.22 million in 2025 with a CAGR of greater than 3% during forecast period (2025-2030). The drugs are mainly fall under the following segment: biguanides, alpha-glucosidase inhibitors, dopamine-d2 receptor agonists, sodium-glucose cotransport-2 (SGLT-2) inhibitor, dipeptidyl peptidase-4 (DPP-4) inhibitors, sulfonylureas, and meglitinides (*Malaysia Oral Anti-Diabetic Drug Market Size | Mordor Intelligence*, 2025).

*Corresponding author's email: mtaher@iium.edu.my

Beside of OHAs, insulin is a very important agent in diabetic management. Insulin as a biopharmaceuticals has become a high demand biotherapeutic which is valued USD 310 billion and projected to reach USD 550 billion by 2029, growing at a compound annual growth rate (CAGR) of 10% (Research and Markets 4 min read, 2025). The therapeutic insulin was first derived from animal sources, however, due to a shortage of time to produce enough insulin and immunogenicity, scientists developed a new technique, which is using recombinant deoxyribonucleic acid (rDNA) technology. The steps involved in the synthesis of human insulin using rDNA include the identification and insertion of the insulin gene into suitable vectors followed by transfection into *Eschericia coli (E. coli)* as the expression system.

The discovery of the DNA cloning technique by Stanley Cohen and Herbert Boyer marked the beginning of genetic engineering, which led to the development of recombinant proteins with therapeutic applications from DNA recombinant technology. Several downstream processes are also involved until the final product of active human insulin is produced. Additionally, the advancement of molecular biology allows for the modification of human insulin to become more pharmacologically favourable, thus resulting in insulin analogues such as insulin isophane (intermediate-acting), short-acting insulin, ultra-rapid-acting insulin, long-acting insulin, inhaled insulin, and U-500 insulin.

Throughout the years, insulin production has gone through a massive evolution using *E. coli* as a host. In both research and the biotechnological field, *E. coli* remains a widely used host for the production of recombinant proteins. Despite the advantages offered by using *E. coli*, such as high yield, low cost, ease of management, minimal media requirement, and fast growth rate, researchers faced a lot of obstacles to improve the production of insulin using the *E. coli* expression system (Baeshen et al., 2014) compared to other expression systems, like mammalians cells and yeasts.

According to Kim et al., this system exhibited a number of drawbacks, including mispairing disulfide bonds, which leads to the production of inactive proteins, incorrect folding, and protein expression in inclusion bodies. These, in turn, will result in low production yields. In this case, *E. coli* host strains have been genetically modified to regulate the cellular redox environment in order to prevent the mispairing of disulfide bonds (Kim et al., 2021).

For effective secretion, drug efficacy, reducing immunogenicity and increased stability, the majority of proteins employed in drug therapy need to undergo post-translational modifications such as glycosylation, oxidation of methionine, misfolding and aggregation, proteolysis, and glutamine deamidation of asparagine and glutamine (Jenkins, 2007). However, *E. coli* expression system lacks post-translational modifications (e.g., is incapable of forming disulfide bonds). Arya et al. stated that *E. coli* is unable to perform N- or O-linked glycosylation as it lacks the machinery that involves hundreds of genes and proteins located in the Golgi apparatus and endoplasmic reticulum. Furthermore, it also has no capacity to perform other modifications like phosphorylation, acetylation, and disulfide bridges for proper

protein folding (Arya et al., 2008). Glycosylation is crucial in the pharmaceutical industry because it is frequently required for proper protein glycosylation, and it is desirable to pursue cost-effective bioproduction by using the simplest purification techniques (Ferrer-Miralles & Villaverde, 2013). Moreover, glycosylation will also increases the bioactivity of the foreign protein and its stability, with the closer glycosylation pattern to the native one resulting in the production of a more active and stable protein (Waegeman & Soetaert, 2011). If it is not possible to use *E. coli* as the expression host, other less convenient expression systems need to be applied (Table 1). Simple eukaryotic cells like yeast have the ability to be glycosylated, but they frequently use different glycosidic residues than what is needed for that particular circumstance (Flaschel & Friehs, 1993). Waegeman and Soetaert have also mentioned that yeast strains are particularly ideal for the synthesis of more complex mammalian proteins because, unlike bacteria, they are capable of performing eukaryotic glycosylation patterns (Waegeman & Soetaert, 2011).

Table 1. The advantages and challenges of insulin production in different hosts*

Host	**Advantages**	**Challenges**
Escherichia coli	High yield, rapid growth, well-understood genetics	Inclusion body formation, complex refolding required
Saccharomyces cerevisiae	Post-translational modifications, well-established system	Requires optimization of fermentation conditions
Pichia pastoris	High cell density growth, efficient secretion, fewer impurities	Stability of expression cassettes, optimization needed
CHO cells	Extensive post-translational modifications, high yield	High production costs, longer development times
Transgenic plants	Cost-effective, scalable, long-term stability	Regulatory hurdles, extensive purification required

*The table was generated by Scopus AI (https://www.scopus.com/pages/ai?query=insulin%20production&isExample=false)

Another issue that can be highlighted resulting from the lack of posttranslational modification in *E. coli* is proteolytic maturation. Based on Kamionka, preproinsulin required proteolytic modification to cleave the peptide in the endoplasmic reticulum, leaving the B-chain attached to the A-chain by two disulfide bridges to become fully active (Kamionka, 2011). This issue presented a significant challenge for recombinant human insulin expression, which led to various engineering strategies being implemented. Using a single expression construct, complete proinsulin is produced. For instance, Eli Lilly and Aventis designed an innovative technique that involves generating a single chemically synthesised cDNA encoding for human proinsulin in *E. coli*, purifying, and then removing the resultant C-peptide using proteolytic digestion (Baeshen et al., 2014). It has been supported by Kamionka, who discovered that following the purification of

proinsulin, the development of a disulfide bridge, the elimination of proteolytic C-peptide, and N-terminal methionine cleavage results in the formation of a physiologically active recombinant human insulin. Since this method was thought to be more practical and effective for producing insulin on a wide scale, humulin, the first recombinant insulin was invented to treat diabetes mellitus (Kamionka, 2011).

Furthermore, the strain's failure to secrete the recombinant protein into the culture medium presents another challenge when employing *E. coli* as an expression host (Kleiner-Grote et al., 2018). Inclusion bodies (IBs), which are related to insoluble misfolded protein aggregates, may result from the accumulation of protein in *E. coli* (Baeshen et al., 2014). Furthermore, according to Bathwa et al. (2021), the target protein starts to misfold and finally aggregates into IBs when the rate of recombinant protein expression exceeds the host cell's capacity to control post-translational modifications and protein folding (Bhatwa et al., 2021). To overcome this problem, a few approaches have been tested, as suggested by Baeshen et al. (Baeshen et al., 2014) to use molecular chaperones to help in proper protein folding and increase protein solubility. In this cooperative context, some chaperones inhibit protein aggregation, while others aid in the refolding and solubilisation of misfolded proteins. Not only that, chaperones such as GroEL, GroES, DnaK, DnaJ, and Grp can also reduce aggregation by facilitating the breakdown of proteins that are unable to fold correctly (Carrió & Villaverde, 2003).

In conclusion, the emergence of recombinant DNA technology has become one of the greatest turning points in the development of insulin production. Multiple hosts have been used to express insulin with their significant characteristic, both advantages and drawbacks. However, the drawbacks of using *E.coli* have opened many opportunities to create innovation in molecular biology to improve the product.

References

Akhtar, S., Nasir, J. A., Ali, A., Asghar, M., Majeed, R., & Sarwar, A. (2022). Prevalence of type-2 diabetes and prediabetes in Malaysia: A systematic review and meta-analysis. *PLOS ONE*, *17*(1),e0263139.https://doi.org/10.1371/journal.pone.0263139

Arya, R., Bhattacharya, A., & Saini, K. S. (2008). *Dictyostelium discoideum*—A promising expression system for the production of eukaryotic proteins. *The FASEB Journal*, *22*(12), 4055–4066. https://doi.org/10.1096/fj.08-110544

Baeshen, N. A., Baeshen, M. N., Sheikh, A., Bora, R. S., Ahmed, M. M. M., Ramadan, H. A. I., Saini, K. S., & Redwan, E. M. (2014). Cell factories for insulin production. *Microbial Cell Factories*, *13*(1),141.https://doi.org/10.1186/s12934-014-0141-0

Bhatwa, A., Wang, W., Hassan, Y. I., Abraham, N., Li, X.-Z., & Zhou, T. (2021). Challenges Associated With the Formation of Recombinant Protein Inclusion Bodies in Escherichia coli and Strategies to Address Them for Industrial Applications. *Frontiers in Bioengineering and Biotechnology*, *9*, 630551.https://doi.org/10.3389/fbioe.2021.630551

Carrió, M. M., & Villaverde, A. (2003). Role of molecular chaperones in inclusion body formation. *FEBS Letters*, *537*(1–3), 215–221. https://doi.org/10.1016/S0014-5793(03)00126-1

Diabetes. (2025). https://www.who.int/health-topics/diabetes

Ferrer-Miralles, N., & Villaverde, A. (2013). Bacterial cell factories for recombinant protein production; expanding the catalogue. *Microbial Cell Factories*, *12*(1), 113. https://doi.org/10.1186/1475-2859-12-113

Flaschel, E., & Friehs, K. (1993). Improvement of downstream processing of recombinant proteins by means of genetic engineering methods. *Biotechnology Advances*, *11*(1), 31–77. https://doi.org/10.1016/0734-9750(93)90409-G

Jenkins, N. (2007). Modifications of therapeutic proteins: Challenges and prospects. *Cytotechnology*, *53*(1–3), 121–125. https://doi.org/10.1007/s10616-007-9075-2

Kamionka, M. (2011). Engineering of therapeutic proteins production in Escherichia coli. *Current Pharmaceutical Biotechnology*, *12*(2), 268–274. https://doi.org/10.2174/138920111794295693

Kim, Y. S., Lee, H.-J., Han, M., Yoon, N., Kim, Y., & Ahn, J. (2021). Effective production of human growth factors in Escherichia coli by fusing with small protein 6HFh8. *Microbial Cell Factories*, *20*(1), 9. https://doi.org/10.1186/s12934-020-01502-1

Kleiner-Grote, G. R. M., Risse, J. M., & Friehs, K. (2018). Secretion of recombinant proteins from *E. coli*. *Engineering in Life Sciences*, *18*(8), 532–550. https://doi.org/10.1002/elsc.201700200

Malaysia Oral Anti-Diabetic Drug Market Size | Mordor Intelligence. (2025). https://www.mordorintelligence.com/industry-reports/malaysia-oral-anti-diabetic-drug-market

Research and Markets 4 min read. (2025, January 13). *Biotherapeutics Market Insights 2025, Analysis and Forecast to 2030: Insulin Demand Soars Amid Diabetes Epidemic, Sees $40 Billion Opportunity*. Yahoo Finance. https://finance.yahoo.com/news/biotherapeutics-market-insights-2025-analysis-114000447.html

Waegeman, H., & Soetaert, W. (2011). Increasing recombinant protein production in Escherichia coli through metabolic and genetic engineering. *Journal of Industrial Microbiology & Biotechnology*, *38*(12), 1891–1910. https://doi.org/10.1007/s10295-011-1034-4