Doctor of Philosophy of Science

Studies on efficient prokaryotic production of several cytokines and growth factors

원핵 생물에서 여러 사이토카인과 성장인자의 효율적인 생산에 관한 연구

The Graduate School

of the University of Ulsan

Department of Medical Sciences

Nguyen Thi Kieu Oanh

Studies on efficient prokaryotic production of several cytokines and growth factors

Supervisor Choe Han

Doctor of Philosophy's thesis

Submitted to

the Graduate School of the University of Ulsan

In partial Fulfillment of the Requirements for the Degree of

Doctor of Philosophy of Science

by

Nguyen Thi Kieu Oanh

Department of Medical Sciences

University of Ulsan, Korea

August 2024

Studies on efficient prokaryotic production of several cytokines and growth factors

This certifies that the dissertation
of Nguyen Thi Kieu Oanh is approved

Leem, Chae Hun
Committee Vice-Chair Dr.

Choe, Han
Committee member Dr.

Chang, Suhwan
Committee member Dr.

Kim, Dong-Hou
Committee member Dr.

Hong, Seokmann
Committee member Dr.

Department of Medical Sciences

University of Ulsan, Korea

August 2024

Seoul, August 2024

My deep gratitude for completing this thesis.

Firstly, I would like to express my sincere gratitude to my supervisor, Prof. Choe Han, for his invaluable guidance, support, and encouragement throughout my research. His expertise and insights have been instrumental in shaping this thesis.

I'm also deeply thankful to the faculty of Ulsan University and Asan Medical Center for their financial support and resources, without which this research would not have been possible.

Furthermore, I'm truly grateful for the love, encouragement, and knowledge my family has showered me throughout this journey. Their support has been my guiding light, and I dedicate this thesis to them with profound gratitude and love. Their constant encouragement has been the cornerstone of my success.

Lastly, I wish to acknowledge my colleagues who generously contributed their time and knowledge to this study, their cooperation was essential to completing this research.

Sincerely,

Nguyen Thi Kieu Oanh

Abstract

Studies on efficient prokaryotic production of several cytokines and growth factors

Nguyen Thi Kieu Oanh

The Graduate School of

the University of Ulsan

College of Medicine

Recombinant proteins have comprehensive applications in different fields, including therapeutics, vaccines, industrial productions, agriculture, gene therapies, and research developments. *Escherichia coli (E. coli)* is a common expression system for producing recombinant human proteins. Nevertheless, protein misfolding and aggregation into inclusion bodies often need additional steps in protein purification. This study explores five proteins that tend to misfold and form inclusion bodies in *E. coli*: human granulocyte-macrophage colony-stimulating factor (GM-CSF), interleukin-3 (IL3), interleukin-6 (IL6), keratinocyte growth factor (KGF), and insulin-like growth factor 1 (IGF1).

To conquer these challenges, fusion tags such as maltose binding protein (MBP) or the b' a' domain of protein disulfide isomerase (PDIb'a') were set up at the N-terminus of the fusion proteins. This strategy remarkably improved expression levels and solubility when expressed in *E. coli* strains. The protein constructs exhibited high soluble expression levels, qualifying recombinant protein isolation. Purified proteins were further identified and verified through biological activity assays.

Comprising PDIb'a' tags and using genetically modified ClearColi E. coli strains, coupled with moderate temperature culture conditions (30°C), proved instrumental in successful protein production, and facilitated straightforward purification procedures. The purification process involved three main steps:

primary purification using affinity chromatography, cleavage of the fusion protein with TEV protease, and subsequent ion exchange chromatography to obtain purified proteins. Intriguingly, the fusion proteins exhibited higher bioactivity in *in vitro* assays than their non-fused counterparts.

This study represents the first successful attempt to express and purify soluble human recombinant proteins GM-CSF, IL3, IL6, KGF, and IGF1 in the ClearColi *E. coli* strain using innovative PDIb'a' tags. These findings offer the potential to enhance biopharmaceutical manufacturing by implementing efficient processes, potentially improving scalability and cost-effectiveness. The approach presented here opens new avenues to produce challenging recombinant proteins and may have significant implications for developing novel therapeutics and research applications.

Keywords: *Escherichia coli* (*E. coli*), prokaryotic soluble expression, keratinocyte growth factor (KGF), insulin-like growth factor 1 (IGF1), granulocyte-macrophage colony-stimulating factor (GM-CSF), interleukin-3 (IL3), and interleukin-6 (IL6), fusion tags, the b'a' domain of protein disulfide isomerase (PDIb'a').

Table of Contents

List of Tables

List of Figures

List of Abbreviations

CV	Column volume
DMEM	Dulbecco’s modified Eagle’s medium
DTT	Dithiothreitol
E. coli	*Escherichia coli*
GM-CSF	Granulocyte-macrophage colony-stimulating factor
His	Histidine
IEC	Ion exchange Chromatography
IGF1	Insulin-like growth factor 1
IL3	Interleukin 3
IL6	Interleukin 6
IMAC	Immobilized Metal Affinity Chromatography
IPTG	Isopropyl 1-thio-β-d-galactopyranoside
KGF	Keratinocyte growth factor
LC	Liquid Chromatography
MBP	maltose–binding protein
MS	Mass spectrometry
OD	Optical density
PDIb'a'	b'a' domain of Protein disulfide isomerase
pI	Isoelectric point
POI	Protein of Interest
RPMI	Roswell Park Memorial Institute
SDS-PAGE	Sodium dodecyl sulfate-polyacrylamide gel electrophoresis
TEV	Tobacco etch virus
TEVrs	Tobacco etch virus protease recognition site

I. Introduction

Recombinant protein purification is a fundamental process that underpins a wide range of scientific, medical, and industrial activities, enabling the utilization of specific proteins for various applications. These signaling molecules play crucial roles in regulating cell growth, differentiation, and various physiological processes. Growth factors and cytokines are essential for expanding and maintaining cells in cell culture systems. Growth factors are employed in regenerative medicine to promote tissue repair and regeneration. They play a key role in enhancing the healing process in applications such as wound healing, tissue engineering, and organ transplantation. Cytokines are studied extensively in cancer research for their role in immune responses and inflammation. Several cytokines are used in cancer therapy to stimulate the immune system or inhibit tumor growth. In immunological research, cytokines are constantly demanded to study immune cell behavior and develop immunotherapies. The demand for growth factors and cytokines is likely to continue growing as advancements in these fields lead to the development of new therapies and treatments. Biotechnology, pharmaceutical companies, and research institutions contribute to the ongoing demand for these crucial signaling molecules. The revenue of the global recombinant proteins market size was estimated to be worth USD 2.2 billion in 2023 and is expected to reach USD 3.2 billion by 2028 [1], illustrating the enormous market value and growth potential of recombinant proteins. Similarly, various protein drugs have been successfully marketed, including monoclonal antibodies (mAbs), recombinant vaccines, and hormones, demonstrating that recombinant proteins already play a significant role in the biopharmaceutical field.

Producing recombinant growth factors or cytokines is challenging due to their complex structures, susceptibility to aggregation, and the potential for low yields. Controlling protein folding, preventing aggregation, and addressing expression level issues are critical challenges. Purification from complex cell culture media poses difficulties, necessitating robust processes for high purity and bioactivity. Eukaryotic expression systems, although more effective, were costly and time-consuming [2]. Therefore, prokaryotic expression systems like *Escherichia coli* (*E. coli*) present as an alternative for recombinant protein production due to its fast growth rate, easy

manipulation, and inexpensive culture reagents [3]. However, high-level cytoplasmic expression in *E. coli* frequently often leads to aggregation, resulting in inclusion bodies, which necessitates additional protein refolding procedures [4]. Other approaches, such as cell-free translation systems [5] and prokaryotic secretion mechanisms [6, 7], have also been explored to varying degrees of success.

Addressing the challenges encountered in recombinant protein production holds significant potential for advancing research, biotechnology, and therapeutic development. Various strategies have been developed to overcome these obstacles, such as expression optimization, co-expression of chaperone, *in vitro* refolding, advanced purification techniques, and automated technological advancements. Depending on the prevailing conditions, these strategies can be applied individually or in combination to further enhance the production of soluble recombinant proteins. Therefore, this study focuses on the production of several human cytokines and growth factors, including:

A. Human Granulocyte-Macrophage Colony-Stimulating Factor (GM-CSF).

B. Human Interleukin 3 (IL3).

C. Human Interleukin 6 (IL6).

D. Human Keratinocyte Growth Factor (KGF).

E. Human Insulin-like Growth Factor 1 (IGF1).

Firstly, human granulocyte-macrophage colony-stimulating factor (GM-CSF), also known as colony-stimulating factor 2 (CSF-2), plays a key role in the growth and differentiation of hematopoietic progenitor cells into sub-lineages, including granulocytes, macrophages, eosinophils, megakaryocyte, and red blood cells [8]. Additionally, GM-CSF assumes a paracrine role in tissue inflammation, such as rheumatoid arthritis or multiple sclerosis [9]. Due to its ability to stimulate blood cells from the bone marrow into the blood, recombinant GM-CSF has been employed in treating neutropenia in cancer patients after radiotherapy or chemotherapy. Intriguingly, GM-CSF had induced long-lasting, specific anti-tumor effects in mice [10]. However, subsequent clinical trials failed to demonstrate efficacy in human subjects [11]. The emergence of immune-checkpoint cancer therapy reignited interest in GM-

CSF cancer therapy in combination with the immune-checkpoint antibody and other anti-cancer antibodies [12]. Beyond clinical applications, GM-CSF is useful for inducing hematopoietic cells from pluripotent stem cells. Moreover, GM-CSF also promotes the neuronal differentiation of adult neural stem cells [13]. Owing to both clinical and non-clinical demands for GM-CSF, there is a growing pursuit of efficient production of bioactive GM-CSF.

GM-CSF comprised 127-amino acid monomer with six glycosylation sites [14] and is available in two clinical forms: sargramostim, a glycosylated GM-CSF from *Saccharomyces cerevisiae* [15], and molgramostim, a non-glycosylated GM-CSF from *E. coli [*16]. The glycosylation of GM-CSF interfered with its function [17], reducing its affinity to the receptor, GM-CSFR [18]. GM-CSF possesses four cysteine residues forming two intramolecular disulfide bonds. However, the cytoplasm of *E. coli* provides a reducing environment, leading to improperly folded recombinant GM-CSF, and its aggregation into structures known as inclusion bodies [19-22]. These aggregated proteins require intricate solubilization and refolding processes. Consequently, various methodologies have been explored to achieve the soluble expression of GM-CSF in *E. coli,* including periplasmic expression [23], the use of different *E. coli* strains [24], various additives, and chaperone co-expression [25], thioredoxin-fusion [26], and intein-fusion [27].

Secondly, human interleukin-3 (IL3) is a multifunctional cytokine, ranging from multi-colony stimulating factor and persisting cell stimulating factor to hematopoietic cell growth factor and stem cell activating factor. Categorically, IL3 belongs to the β common chain cytokines, a group that also includes human GM-CSF and human interleukin-5 (IL5). These cytokines orchestrate diverse inflammatory responses, facilitating swift pathogen clearance. They also play a role in the pathogenesis of chronic inflammation [28]. Functionally, IL3 drives the proliferation and differentiation of various hematopoietic cells, including macrophages, dendritic cells, mast cells, basophils, and eosinophils, all of which originated from hematopoietic stem cells [28]. From a clinical perspective, IL3 holds promise for various therapeutic interventions, such as augmenting hematopoietic precursor expansion post bone marrow transplantation and enhancing the recovery of stem cells, $CD34^+$ cells, platelets,

eosinophils, and T cells. IL3 served as an adjunct in aplastic anemia and chemotherapy protocols, offering potential treatments for leukemias and allergic disorders [28]. Moreover, IL3 has been established as an indispensable factor for the *in vitro* maturation of diverse hematopoietic cells from their stem cell precursors [29-31]. Given its pivotal role in hematopoiesis and its potential therapeutic applications, the efficient production of IL3 is crucial for both biomedical research and clinical purposes.

IL3 is a secreted protein comprised of 133 amino acids. Its structure is characterized by a single intramolecular disulfide bond and two putative N-linked glycosylation [32]. Notably, IL3 is recognized as a "difficult-to-express protein" due to its poor yield across various expression hosts. Over the years, it has been successfully expressed in diverse eukaryotic systems including mammalian cells [33], insect cells [34], and *Pichia pastoris* [35-37]. *E. coli* is a frequently chosen system for recombinant protein production. However, expressing IL3 and its derivatives in the cytoplasm of *E. coli* often leads to the aggregation of the resultant protein into inclusion bodies, necessitating complex solubilization and refolding processes [38, 39]. On the other hand, attempts to direct IL3 to the periplasmic space using various signal sequences have resulted in the protein predominantly remaining unprocessed in the cytoplasm [40]. Other prokaryotic hosts, such as *Bacillus subtilis* and *Streptomyces lividans*, have also been explored previously for secretory IL3 expression [41, 42]. In efforts to enhance IL3 solubility within the *E. coli* cytoplasm, several solubility-enhancing fusion tags, such as NusA, GrpE, bacterioferritin, and thioredoxin, have been employed, resulting in varied success [43]. Additionally, several attempts have been made to develop a soluble IL3 variant [44].

Thirdly, human interleukin-6 (IL6) is a multifunctional cytokine involved in immune responses, inflammation, and various other biological processes. Additionally, IL6 serves as an anti-inflammatory myokine produced by muscles, with its secretion increasing in response to muscle contraction [45]. The clinical applications of IL6 are diverse, playing a significant role in both physiological and pathological processes. IL6 is currently under scrutiny as a potential target for various conditions, such as cardiovascular diseases, certain cancers, and neurodegenerative

disorders. Ongoing clinical trials aim to investigate the efficacy and safety of IL6-targeted therapies in these contexts.

Diverse platforms have been employed for the expression and purification of IL6, either *in vitro* or *ex vivo*. From several literature, the production of active recombinant human IL6 from inclusion bodies has yielded up to 3 mg from 1 L of cell culture [46, 47]. Dunstan et al. (1996) demonstrated that intracellularly expressed IL6 showed no *in vivo* effects on immune responses [48]. Therefore, extracellular protein secretion offers distinct advantages. Codon-optimized human IL6 fused with hemolysin (hIL6-HlyAs) was purified through single-step immunoaffinity chromatography, yielding 18 μg/L of culture supernatant with approximately 60% purity [49]. Optimization of cultivation conditions through pH-stat fed-batch in a laboratory-scale bioreactor using the NusA/IL6 expression vector in *E. coli* BL21 enhanced the solubility and yield of IL6 [50]. Notably, Lee et al. (2009) demonstrated the effectiveness of producing soluble and bioactive human IL6 in the *E. coli* BL21 strain by fusing MBP tags and factor Xa cleavage for subsequent purification steps [51]. Alternatively, the co-expression of chaperones and lower cultivation temperatures have been found to enhance IL6 folding capacity [52].

Next, keratinocyte growth factor (KGF), also known as fibroblast growth factor 7 (FGF-7), is a multifaceted growth factor with diverse biological functions. KGF is secreted exclusively by mesenchymal cells and plays a pivotal role in embryonic development, cell proliferation, and differentiation [53]. It is essential for branching morphogenesis and keratinocyte development [54, 55], exerting its major paracrine effect on normal epithelial cell proliferation [56, 57]. KGF is a member of the expansive FGF family, which comprises at least 22 known members [58-60]. It specifically interacts with the b isoforms of FGFR1 and FGFR2 (FGFR1b/FGFR2b) found on epithelial cells [61]. Structurally, KGF consists of 163 amino acid residues, two intramolecular disulfide bonds, and two glycosylation sites [62]. Post-translational truncation at Arg23 can produce a 140-residue, non-glycosylated form with a single disulfide bond, yet both truncated and non-truncated forms are biologically active [62]. Furthermore, deletions at the N-terminus have been shown to significantly impact

receptor binding and biological activity [63-65]. However, despite its importance, the three-dimensional structure of human KGF has not been determined to date.

Recombinant truncated KGF, named Palifermin, has received FDA approval for its efficacy in reducing both the severity and duration of oral mucositis following intensive chemotherapy and radiotherapy treatments [66]. Furthermore, recombinant KGF demonstrates promising applications in the differentiation of human pluripotent stem cells into pancreatic progenitors and islet-like organoids [67], as well as in wound healing and hepatocyte protection [68, 69]. However, the efficient production of recombinant KGF remains a challenge. Previous attempts to produce KGF in *E. coli* resulted in low expression levels and protein aggregation [63, 70]. Various strategies, including the use of different fusion tags and culture media modifications, have yielded only modest improvements [71-74]. Nonetheless, eukaryotic expression systems, although more effective, are costly and time-consuming [68, 69, 75, 76].

Finally, insulin-like growth factor 1 (IGF1), also termed somatomedin-C, serves as a cornerstone in regulating childhood growth and sustains its influential physiological impact into adulthood [77]. The mature form of IGF1 is characterized as a 70-amino acid polypeptide, stabilized by the presence of three intricate disulfide bonds [78]. IGF1 structural maturation and proper folding are orchestrated by the molecular chaperone glucose-regulated protein 94 (GRP94) [79]. Interestingly, IGF1 is thermodynamically unstable once secreted, prone to denaturation under physiological conditions [80]. A family of six homologous, high-affinity IGFBPs, designated as IGFBP-1 through IGFBP-6, exist to fulfill this role. Particularly noteworthy is the formation of a ternary complex comprising IGF1, IGFBP-3, and the acid-labile subunit (ALS); significantly extending the circulatory half-life of IGF1 *in vivo* [81-84]. When bound to IGFBPs, IGF1 is rendered biologically inert; thus, the dissociation of IGFBPs is a prerequisite for IGF1 to engage effectively with its native receptor [85]. The signaling cascade mediated IGF1 was quintessential for the preservation of self-renewal capabilities and the maintenance of pluripotency in human pluripotent stem cells (hPSCs), which encompasses both embryonic stem cells and induced pluripotent stem cells [86, 87]. Moreover, IGF1 held promise for enhancing pain management by contributing to the alleviation of hyperalgesia,

shielding against chemotherapy-induced neuropathy, mitigating neuronal hyperactivity, and enhancing the nociceptive threshold [88].

Multiple expression platforms have been scrutinized for the recombinant production of IGF1, including eukaryotic systems such as rabbit cells [89], plant cells [90], and yeast [91, 92]. While these eukaryotic systems offer specific advantages, they also considerable challenges, notably scalability, and elevated production costs. For enhancing soluble IGF1 expression in the *E. coli* cytoplasm, a myriad of protein fusion partners have been investigated. These include synthetic IgG-binding peptides [93], staphylococcal protein A [94], growth hormone [95], beta-galactosidase [96], DsbA [97], thioredoxin [98], and Fh8 [99]. Remarkably, this body of research lacks exploration into the bioactivity of IGF1 when fused with these tags.

In this study, our focus is on two key strategies: expression optimization and advanced purification techniques. As in expression protein optimization, various strategies such as codon optimization, fusion tags, and inducible promoters are employed to enhance protein expression levels. In advanced purification techniques, chromatography, affinity tags, and other innovative methods are utilized to improve both yield and purity.

Specifically, this study delves into the use of two protein tags, PDIb'a' - the b′a′ domain of protein disulfide isomerase and MBP - maltose binding protein, both of which have been demonstrated the ability to enhance the expression and solubility of various recombinant proteins in *E. coli* [100-106]. We genetically fused PDIb'a', MBP, and His tags to the N-terminus of target proteins to create fusion protein constructs. These fusion proteins are expressed in multiple *E. coli* strains to assess their solubility and expression levels. Afterwards, target proteins are purified by proper purification steps, and investigated their biological activities in relevant systems. From there, the most suitable system of recombinant protein production could be found and safety for further therapeutic applications.

By concentrating our efforts on these specific aspects, we aim to contribute to the advancement of recombinant protein production, thereby facilitating progress in various scientific and industrial applications.

II. Materials and Methods

1. Materials

All reagents were analytical or cell-culture grade standards. Detailed information regarding the sources of these materials can be found in the Tables list.

Table 1. Chemical information.

Chemical	**Company**	**Address**
1-Methoxy-5-methylphenazinium Methyl Sulfate (PMS)	TCI Chemicals	Tokyo, Japan
1-thio-β-d-galactopyranoside (IPTG)	Anaspec	Fremont, CA, USA
2-(N-morpholino) ethane sulfonic acid (MES monohydrate)	Duchefa Biochemie	Haarlem, Netherlands
2-(2-Methoxy-4-nitrophenyl)-3-(4-nitrophenyl)-5-(2,4-disulfophenyl)-2H-tetrazolium sodium salt (XTT)	Biosynth Carbosynth	Staad, Switzerland
Agarose	LPS Solution	Daejeon, South Korea
Bacto Peptone	GIBCO	Carlsbad, CA, USA
Carbenicillin	Biosynth	Berkshire, UK
Cell Counting Kit-8 (CCK-8)	Dojindo	Kumamoto, Japan
Coosmaasie brilliant blue R-250	Amresco	Solon, OH, USA
Dithiothreitol (DTT)	Anaspec	Fremont, CA, USA
DMEM/F12, powder	GIBCO	Carlsbad, CA, USA

Ethylenediaminetetraacetic acid (EDTA)	Samchun Chemical	Pyeongtaek, South Korea
Fetal Bovine Serum (FBS)	Avantor	Radnor, PA, USA
Glycerol	Samchun Chemical	Pyeongtaek, South Korea
Imidazole	Daejung Chemicals	Siheung, South Korea
Lambda integrase (BP reaction)	Elpis Biotec	Daejeon, South Korea
Lambda integrase/exisionase (LR reaction)	Elpis Biotech	Daejeon, South Korea
Maltose monohydrate, 99%	Samchun Chemical	Pyeongtaek, South Korea
Mid-range Prestained Protein Marker	Elpis Biotech	Daejeon, South Korea
Penicillin/Streptomycin solution 100X	Biological Industries	Beit HaEmek, Israel
RPMI1640, powder	GIBCO	Carlsbad, CA, USA
Recombinant human granulocyte-macrophage colony-stimulating factor (GM-CSF)	Peprotech	Rocky Hill, NJ, USA
Recombinant human interleukin-3 (IL3)	R&D Systems	Minneapolis, MN, USA
Recombinant human insulin-like growth factor 1 (IGF1)	R&D Systems	Minneapolis, MN, USA
Sodium chloride	Samchun PureChemical	Pyeongtaek, South Korea
Sodium hydroxide	Samchun PureChemical	Pyeongtaek, South Korea

Silver staining kit Plus	Bio-Rad Laboratories	Hercules, CA, USA
Tris(hydroxymethyl)aminomethane	Daejung Chemicals	Siheung, South Korea
Yeast extract	Samchun PureChemical	Pyeongtaek, South Korea

Table 2. E. coli strains information.

***E. coli* strains**	**Company**	**Address**
BL21	Invitrogen	Carlsbad, CA, USA
BL21 pLysS	Thermo Fisher Scientific	Waltham, MA, USA
BL21 AI	Thermo Fisher Scientific	Waltham, MA, USA
ClearColi	Lucigen	Middleton, WI, USA
ClearColi pLysS	Lucigen	Middleton, WI, USA
LOBSTR	Kerafast	Boston, MA, USA
Origami 2	Novagen	Madison, MI, USA
Rosetta-gami B	Novagen	Madison, MI, USA
SHuffle T7	New England Biolabs	Ipswich, MA, USA
Shuffle T7 pLysS	New England Biolabs	Ipswich, MA, USA

Table 3. Cell lines information.

Cell lines	**Origin**	**Address**
Human leukemia cell line (TF-1)	ATCC	Manassas, Virginia, USA
Human breast cancer cell (MCF-7)	South Korea Cell Line Bank	Seoul, South Korea

Mouse fibroblast cells (NIH/3T3)	South Korea Cell Line Bank	Seoul, South Korea

Table 4. Used consumables information.

Consumable	**Company**	**Address**
Acrodisc syringe filters	Pall	Port Washington, NY, USA
AMICON Ultra (3 kDa, 10 kDa)	Merck	Darmstadt, Germany
Dialysis membrane (10-12 kDa)	Viskase	Darien, IL, USA
HiTrap Q HP, 5 mL column	GE Healthcare	Chicago, IL, USA
HisTrap HP, 5 mL column	GE Healthcare	Chicago, IL, USA
Polyethersulfone (PES) membrane (0.45 µm)	Hyundai Micron	Seoul, South Korea
Syringe filter pore size 0.22 µm	Hyundai Micron	Seoul, South Korea

Table 5. Used machine information.

Machine	**Company**	**Address**
ÄKTA Start system	Cytiva	Marlborough, MA, USA
Ultrasonic cell disruptor JY99-IIDN	Ningbo Scientz Biotechnology	Guangdong, China

2. Construction of plasmids

The expression vectors were assembled using BP and LR reactions within the Gateway cloning system. Specifically, the primary sequences of human gene segments (spanning amino acids 18 to 144 for human GM-CSF, 20 to 152 for human IL3, 30 to 212 for human IL6, and 32 to 194 for human KGF) were optimized for expression in *E. coli* through codon optimization. These sequences were synthesized by Cosmogenetech (Seoul, South Korea). Additionally, a tobacco etch virus recognition site (TEVrs: ENLYFQ/G) sequence was incorporated at the N-terminus of these genes.

The BP reaction facilitated seamless recombination between the attB and attP sites of the synthesized gene and the pDONR207 vector, resulting in the formation of the pENTR-POI entry clone. Subsequently, these entry clones were recombined with Gateway destination vectors (pDEST-H, Addgene_11518; pDEST-H-PDIb'a', Addgene_60556; and pDEST-H-MBP, Addgene_11085) to introduce tags such as His-tag (H-tag), H-PDIb'a' tag, or H-MBP tag at the N-terminus of the target protein. In cases where H-PDIb'a' tag and H-MBP tag were utilized, additional His-segments were incorporated to aid in purification. To ensure the fidelity of the constructed vectors, sequencing was performed by Cosmogenetech, Seoul, South Korea, to verify their sequences.

In the case of IGF1, the human IGF1 gene from amino acids 49 – 118 with three disulfide bonds was amplified by polymerase chain reaction (PCR) from South Korea Human Gene Bank (ID: KU029721). The tobacco etch virus recognition site (TEVrs: ENLYFQ/G) sequence also was inserted at the N-terminus of these genes. The target genes were homogenous recombination into expression vectors by using in vivo cloning, with a specific overlap 20 bp homogenous sequence. The homogenous reaction was performed into iVEC3 *E. coli* (SN1187) by using chemical transformation (TSS method).

3. Expression and solubility of fusion proteins

To evaluate the expression levels of recombinant proteins, the vectors H-POI, H-PDIb′a′-POI, and H-MBP-POI (with POI, protein of interest; as GM-CSF, IL3, IL6,

KGF, and IGF1) were introduced into various *E. coli* strains, namely BL21, ClearColi, LOBSTR, Origami 2, Rosetta-gami B, and SHuffle T7 using the heat shock method. High-expression colonies were identified by growing transformed *E. coli* on LB agar plates supplemented with 100 µg/mL Carbenicillin. The selected colonies were then inoculated overnight in LB medium with the same concentration of Carbenicillin at 37°C, and a shaking speed of 200 rpm. Subsequently, the overnight cultures were diluted (1:100) into 10 mL of fresh LB medium containing 100 µg/mL Carbenicillin. The diluted cultures were incubated under the same conditions until the OD_{600} value 0.6–0.8 was obtained. The *E. coli* cultures were then divided equally into three different round-bottom tubes and induced using 1 mM IPTG. These cultures were then subjected to three unique incubation conditions: 37°C for 4 h, 30°C for 6 h, and 18°C overnight all at a shaking speed of 200 rpm. Following incubation, bacterial pellets were obtained by centrifuging 1 mL samples from each condition at 14,500 x g for 5 min at 4°C. These pellets were then lysed on ice using an ultrasonic cell disruptor at 35% amplitude, operating on an ON/OFF cycle of 5 sec each for 10 repetitions. The disrupted cell suspensions were again centrifuged under the same conditions to segregate the soluble and insoluble protein fractions. To quantify the expression and solubility level of the fusion protein under these experiment conditions, samples were run on a 10% tricine SDS-PAGE gel and analyzed using the Gel Analyzer Software (available at http://www.gelanalyzer.com). The following equations were employed to determine the levels of protein expression and solubility:

Expression level (%) = Fusion protein / Total protein post induction × 100 (Eq. 1)

Solubility level (%) = Soluble fusion protein / Total fusion protein × 100 (Eq. 2)

Soluble expression level (%) = Expression level × Solubility level / 100 (Eq. 3)

4. Purification of recombinant protein with H-MBP-tag

To get the fusion protein, a unique colony was inoculated overnight in LB medium supplemented with 100 µg/mL Carbenicillin at 37°C with 200 rpm shaking speed. In next day, the overnight culture was continued inoculated in 1L of fresh LB

medium supplemented with 100 μg/mL Carbenicillin at 37°C with 200 rpm shaking speed. When *E. coli* culture was in the exponential phase (with OD_{600} reaching 0.6–0.8), the culture was induced to a final concentration of 1mM IPTG, and continued culture at 37°C for an additional 4 h. The cells were then collected through centrifugation at 3,500 × g for 45 min at 4°C. After harvesting, the supernatant was discarded, and the bacterial pellet was stored at −20°C for further steps.

The bacterial pellets were harvested from 0.5 L *E. coli* culture were resuspended in 50 mL of Tris buffer (50 mM Tris-Cl, pH 8) containing a 1X protease inhibitor cocktail. The cell suspension was lysed using an ultrasonic cell disruptor in an ice bath (to prevent overheating). Following sonication, the cell lysate was centrifuged at 11,000 × g for 25 min at 4°C to separate cellular debris. The supernatant was then filtered through a membrane with a pore size of 0.45 μm. Subsequently, the lysate was adjusted to 300 mM NaCl and 0.5 mM EDTA. The completed lysate was purified using amylose affinity chromatography (MBPTrap HP column 5 mL), with an equilibration buffer (50 mM Tris-Cl, pH8, 300 mM NaCl, 0.5 mM EDTA, 5% glycerol (v/v)). After sample loading, the MBPTrap HP column was washed with the equilibration buffer to eliminate loosely bound proteins. The fusion protein H-MBP-POI was eluted with 20 mM Maltose dissolved in the equilibration buffer. Optionally, for further cleaved with TEV protease, the eluted fractions of the fusion protein were subjected to dialysis against a Tris buffer (50 mM Tris-Cl, pH 8) at a dilution ratio of 1:10,000 to remove Maltose. The dialyzed fusion protein was then treated with a custom TEV protease (Addgene: 8827) at a weight ratio of 10:1 (fusion protein: TEV protease) for 16–20 h at 18°C to remove the H-MBP tags.

5. Purification and tag removal of recombinant protein with H-PDIb′a′- tag

To purify the recombinant protein, a single bacterial colony was cultured overnight in 10 mL of LB medium supplemented with 100 μg/mL Carbenicillin, under incubation conditions of 37°C temperature and 200 rpm shaking speed. Subsequently, the overnight cultured was diluted 1:100 into 1 L of fresh LB medium containing 100 μg/mL Carbenicillin and was grown at 37°C with continuous shaking at 200 rpm. Once

the OD_{600} reached 0.6–0.8, the culture was induced using 1 mM IPTG, and the temperature was adjusted to 30°C and maintained for an additional 6 h. The cells were then collected through centrifugation at 3,500 × g for 45 min at 4°C. After centrifugation, the supernatant was discarded, and the resultant bacterial pellets were stored at −20°C for subsequent use.

The bacterial pellets derived from 0.5 L *E. coli* culture were resuspended in 50 mL of Tris buffer (50 mM Tris-Cl, pH 8) containing a 1X protease inhibitor cocktail. The cells were lysed using an ultrasonic cell disruptor in an ice bath (to prevent overheating). After sonication, the lysate was then centrifuged at 11,000 × g for 25 min at 4°C to pellet cellular debris. The clarified supernatant was then passed through a 0.45 μm filter. Following filtration, the salt concentration, and additives of the supernatant H-PDIb′a′-POI were adjusted to 500 mM NaCl and 25 mM imidazole. Subsequently, this treated lysate was loaded onto a pre-equilibrated IMAC column (HisTrap HP 5 mL), employing an equilibration buffer composed of 50 mM Tris-Cl, pH8, 500 mM NaCl, 20 mM imidazole, 5% glycerol (v/v). Following sample loading, the IMAC column was washed with the equilibration buffer to eliminate loosely bound proteins. To further enhance protein purity, an additional washing was conducted using 10 CVs of a 7% elution buffer (50 mM Tris-Cl, pH 8, 500 mM NaCl, 1 M imidazole, 5% glycerol (v/v)). The fusion protein was eluted with 400 mM imidazole, equivalent to 40% of the elution buffer. The eluted fractions of the fusion protein were subjected to dialysis against a Tris buffer (50 mM Tris-Cl, pH 8) at a dilution ratio of 1:10,000 to remove residual NaCl and imidazole. The dialyzed fusion protein was then treated with a custom TEV protease (Addgene: 8827) at a weight ratio of 10:1 (fusion protein: TEV protease) for 16–20 h at 18°C to remove the H-PDIb′a′ tag.

6. Purification of cleaved proteins

6.1. Purification of GM-CSF

The cleaved proteins were buffer-exchanged to MES buffer (pH 6.5) and applied to a HiTrap Q HP 5 mL column equilibrated with 50 mM MES (pH 6.5). After washing with 10 CVs of a 10% elution buffer (50 mM MES, pH 6.5, 1 M NaCl). GM-

CSF was then eluted with 200 mM NaCl in 5 CVs, equivalent to 20% of elution buffer, and then exchanged to PBS (pH 7.4) using a HiTrap Desalting column. To obtain the desired concentration, the purified GM-CSF was centrifuged using a 3 kDa AMICON Ultra filter unit at 3,000 × g and 4°C. Protein fractions at each purification step were harvested and analyzed using 10% tricine SDS-PAGE. Protein concentrations were confirmed by a BCA Protein Assay kit, following the manufacturer's instructions.

6.2. Purification of IL3

Following cleavage, the cleaved proteins were buffer-exchanged to MES buffer (pH 6.5) and were loaded onto a pre-equilibrated HiTrap Q HP 5 mL column using 50 mM MES buffer (pH 6.5). Post-loading, IL3 was predominantly present in the flow-through, which was subsequently exchanged to PBS (pH 7.4) using a HiTrap Desalting column. To attain the desired concentration, the purified IL3 was centrifuged using a 3 kDa AMICON Ultra Centrifuge filter at 3,000 × g and 4°C. The fractions obtained at each purification step were evaluated using 10% tricine SDS-PAGE. Additionally, protein concentrations were calculated using a BCA Protein Assay Kit, following the manufacturer's guidelines.

6.3. Purification of IL6

Following cleavage, the cleaved proteins were buffer-exchanged to MES buffer (pH 6.5) and were loaded onto a pre-equilibrated HiTrap Q HP 5 mL column using 50 mM MES (pH 6.5). After loading, IL6 was predominantly present in the flow-through and subsequently exchanged to PBS (pH 7.4) using a HiTrap Desalting column. To obtain the desired concentration, the purified IL6 was centrifuged using a 3 kDa AMICON Ultra filter at 3,000 × g and 4°C. The fractions obtained at each purification step were evaluated using a 10% tricine SDS-PAGE. Additionally, protein concentrations were established with a BCA Protein Assay Kit, following the manufacturer's instructions.

6.4. Purification of KGF

Following cleavage, the cleaved proteins were adjusted with NaCl to a final concentration of 350 mM before loading into a HiTrap SP HP 5 mL column. The column was pre-equilibrated with 6 CVs of Tris buffer (50 mM Tris-Cl, 350 mM NaCl 5% glycerol, pH 7.8). To remove contaminants, the column was washed with 10 CVs of Tris buffer containing 350 mM NaCl, followed by an additional wash with 10 CVs of Tris buffer containing 400 mM NaCl and 10 CVs of Tris buffer containing 450 mM NaCl. Target proteins were then eluted using a gradient from 500 to 650 mM NaCl in 20 CVs.

The purified KGF was buffer-exchanged into phosphate-buffered saline (PBS) buffer using a HiTrap Desalting column. To obtain the desired concentration, the purified KGF was centrifuged using a 3 kDa AMICON Ultra filter at 3,000 × g and 4°C. Protein concentrations were determined using a BCA Protein Assay kit, following the manufacturer's guidelines.

6.5. Purification of IGF1

The cleaved protein was applied to the HiTrap Q HP column. The column had been pre-equilibrated with 5 CVs of Equilibration buffer (50 mM Tris-Cl, pH 8, and 5% glycerol). After sample application, unbound proteins were washed away with 7 CV of the Equilibration buffer. Based on the distinct isoelectric points (pI), the cleaved IGF1 was purified at a 5% Elution buffer (50 mM Tris-Cl, pH 8, 1 M NaCl, and 5% glycerol) for 5-10 CVs. Elution of PDIb'a' and other impurities was then carried out using isocratic purification at 100% Equilibration buffer, spanning 5 CV.

Subsequently, the purified IGF1 from the IEC was harvested and then exchanged for PBS (pH 7.4) using a HiTrap Desalting column. To obtain the desired concentration, the purified IGF1 was centrifuged using a 3 kDa AMICON Ultra centrifuge filter unit at 3,000 x g and 4°C. The fractions obtained at each purification step were evaluated using 10% tricine SDS-PAGE. Additionally, protein concentrations were determined using a BCA Protein Assay kit, following the manufacturer's instructions.

7. SDS-PAGE and silver staining

Expression analysis of the fusion proteins and all protein samples in the purification procedure was analyzed using 10% tricine SDS-PAGE. For denaturing protein, samples were incubated for 10 min at 100°C in 5X sample buffer (250 mM Tris-Cl, pH 6.8, 30% glycerol, 10% SDS, 0.01% bromophenol blue, 300 mM DTT). For non-reducing conditions, the protein sample was treated in sample buffer without DTT.

Protein samples were subjected to 10% tricine SDS-PAGE. After electrophoresis, proteins were visualized using Coomassie Brilliant Blue R-250 staining. The gel was immersed in a fixative solution composed of 30% methanol, 10% acetic acid, and 60% distilled water for 20–60 min until obtained the clean gel. Subsequently, the gel was subjected to a double rinse in distilled water, each lasting 10 min, to enhance staining sensitivity and contrast.

The gel was processed using the Silver Staining Kit Plus (Bio-Rad Laboratories) for silver staining. Following electrophoresis, the gel was immersed in a fixative solution (50% methanol, 10% acetic acid, 40% distilled water) for 20 min and then washed twice in high-purity water for 10 min. The gel was subjected to a 20-min staining and development process, executed using a freshly prepared silver staining working solution as the manufacturer instructed. The gel was immersed in a 5% acetic acid solution to stop the staining reaction and was subsequently rinsed in high-purity water.

The expression and solubility levels of the fusion proteins and the purity of the target proteins were analyzed using Gel Analyzer software and calculated using Microsoft Excel® following the equations:

$$\text{Expression level (\%)} = \text{Fusion protein} / \text{Total protein post induction} \times 100 \quad \text{(Eq. 1)}$$

$$\text{Solubility level (\%)} = \text{Soluble fusion protein} / \text{Total fusion protein} \times 100 \quad \text{(Eq. 2)}$$

$$\text{Purity (\%)} = \text{Target protein} / \text{Total protein} \times 100 \quad \text{(Eq. 4)}$$

8. Identification of the purified protein using mass analysis

The SDS-PAGE gel bands of purified protein were collected and analyzed by liquid chromatography with tandem mass spectrometry (LC-MS/MS) to accurately identify and quantify the substance in question. LC-MS/MS integrates mass spectrometry (MS), providing a structural characterization of individual components with high specificity and detection sensitivity, along with high-performance liquid chromatography (HPLC) which is capable of separating mixtures with multiple components. In summary, after the gel was washed, it was digested into peptides by trypsin, and then peptides were subsequently collected and desalted if necessary. The protein's identification was confirmed using GelDoc (Protein A280) and analyzed by LC-MS/MS.

During the LC phase, the sample was separated using a PepMap RSLC 75 × 50 columns containing solution A (0.1% formic acid, 2% Acetone in water), and solution B (0.1% formic acid in Acetonitrile) over 200 min with a flow rate 200–280 nL/min. Fractions were eluted in a gradient range from 5% to 90% solution B in 180 min. For MS analysis, two phases were performed: MS1, characterized by a resolution of 70,000 and a maximum fill time of 50 msec, in which the results were analyzed by data-dependent acquisition (DDA) methods; whereas MS2, with a resolution of 17,500 and a maximum fill time of 150 msec, featured an auto gain control for the peaks was 1×10^5.

9. *In vitro* biological activity assay

The purified proteins underwent treatment with suitable cell lines to investigate their biological activity. Cells were seeded in a 96-well plate at an appropriate cell density and exposed to the purified protein for a specified duration. Subsequently, the WST-8 working solution, containing a final concentration of 0.5 mM XTT and 20 μM PMS, was added to each well. The plate was then incubated for 1–4 h (depending on cell types) at 37°C in the absence of light, following which absorbance was estimated at 450 nm with the microplate reader.

The acquired data were analyzed using the following equation and Microsoft Excel Solver:

$$Re = Bl + (Max - Bl) / (1 + (EC_{50}/conc)^{HC})$$

where *Re* represents the cellular response, *Bl* denotes the baseline response at low protein concentrations, *Max* signifies the maximum observed response, *conc* is the protein concentration, and *HC* refers to the Hill coefficient of stimulation.

9.1. Biological activity of purified GM-CSF

The *in vitro* biological activity of the purified GM-CSF was tested to enhance the proliferation of TF-1, the human erythrocytic cell line. TF-1 cells were cultured in RPMI 1640 medium, 10% fetal bovine serum (FBS), 2 ng/mL GM-CSF (Peprotech, 300-03), and 1% Pen/Strep antibiotics at 37°C in a 5% CO_2 atmosphere. TF-1 cells were cleaned up with PBS and plated at a density of 7.5×10^3 cells per well in a 96-well plate, and then maintained without GM-CSF or FBS in the growth medium for 24h before the activity test. TF-1 cells were then treated with different concentrations of GM-CSF (0.01 pM to 10 nM) for 48 h at 37°C in a humidified atmosphere of 5% CO_2. For negative control, TF-1 cells were treated with PBS. The blank control was the culture media containing the highest concentration of GM-CSF without cells. All these experiments were repeated three times. After a 48-h incubation, 10 µL CCK-8 working solution was added to each well and the plate was incubated for 4 h at 37°C in the dark. The absorbance data was measured at 450 nm using a microplate reader.

9.2. Biological activity of purified IL3

TF-1 cells were maintained in RPMI 1640 medium supplemented with 10% FBS, 2 ng/mL GM-CSF (Peprotech, 300-03), and 1% Pen/Strep antibiotics, incubated at 37°C in a 5% CO_2 atmosphere. Before activity testing, the cells were washed twice using RPMI1640 to eliminate residual cytokines. Subsequently, 1×10^4 cells/well were seeded in a 96-well plate and deprived of GM-CSF and FBS for 24 h. For the activity test, cells were exposed to a range of IL3 concentrations (0–2 µM) and incubated for 48 h in a 37°C, 5% CO2 humidified chamber. Control setups included TF-1 cells with PBS (negative control) and culture media containing the maximum IL3 concentration

but devoid of cells (blank control). Each concentration was assayed in triplicate. Post-incubation, 10 µL of WST-8 working solution was introduced to every well, followed by a 4-h incubation in darkness at 37°C. Absorbance was then recorded at 450 nm using a microplate reader.

9.3. Biological activity of purified IL6

The *in vitro* biological activity of IL6 was assessed by its ability to enhance the proliferation of the human erythroleukemia cell line TF-1. TF-1 cells were cultured in RPMI 1640 medium supplemented with 10% fetal bovine serum (FBS), 2 ng/mL GM-CSF (Peprotech, 300-03), and 1% Pen/Strep antibiotics in a 37°C, 5% CO_2 incubator. Before the activity test, TF-1 cells were subjected to a 24-h period of starvation, during which GM-CSF and FBS were withheld. Subsequently, the cells were rinsed twice with RPMI1640 media to eliminate all cytokines and then seeded at a density of 1 x 10^4 cells per well in a 96-well plate.

The TF-1 cells were then exposed to varying concentrations of purified IL6, H-PDIb′a′-IL6, and PDIb'a' (0–1 µM). The experimental plate was incubated for 48 h at 37°C in a humidified atmosphere with 5% CO2. The negative control consisted of TF-1 cells treated with PBS, and the blank control involved culture media without cells. All concentration conditions were repeated three times. Following incubation, 50 µL of a WST-8 working solution was added to each well, and the plate was further incubated for 4 h at 37°C in the absence of light.

9.4. Cellular response of purified KGF

Human breast cancer MCF-7 cells were seeded in a 96-well plate at a density of 5×10^3 cells per well and cultured in DMEM/F12 medium supplemented with 10% fetal bovine serum (FBS) and 1% Pen/Strep antibiotics. The cells were incubated at 37°C in a 5% CO_2 atmosphere for 24 h. The medium was then replaced with fresh DMEM/F12 containing 1% FBS to induce cellular starvation. The following day, cells were treated with varying concentrations (0–1 µM) of purified recombinant KGF or other proteins of interest. After a 72-h incubation period, 10 µL of WST-8 working

solution was added to each well. The plate was further incubated for 1 h at 37°C in the dark and was measured the absorbance at 450 nm using a microplate reader.

9.5. Biological activity of purified IGF1

The human breast cancer cell line MCF-7 was cultured in DMEM/F12 media supplemented with 10% fetal bovine serum (FBS) and 1% Pen/Strep antibiotics. The cells were seeded into 96-well plates at a density of 5×10^3 cells per well. After cell attachments, the culture medium was replaced with serum-free media (DMEM/F12, 15 mM HEPES, 10 µg/mL Transferrin, 200 µg/mL BSA, and 2 mM L-Glutamine) followed by Karey et al. [107, 108] and cells were maintained in this condition for 24h.

Following serum starvation, purified IGF1, PDIb′a′, and other fusion proteins were administered to the MCF-7 cells at various concentrations. After 72 h of incubation, the culture medium was removed, and 10 µL of a WST-8 working solution was added. Following a 2-h incubation, the absorbance at 450 nm was measured using a microplate reader.

10. Wound healing effect of KGF

The wound-healing assay was adapted from the protocol established by Mouritzen et al. [109]. HaCaT keratinocytes were cultured in DMEM supplemented with 2 mM L-glutamine, 10% FBS, and 1% Pen/Strep antibiotics. Cells were then trypsinized and seeded into a 48-well plate at a density of 5×10^4 cells per well to form a monolayer. To inhibit cell proliferation and focus on cell migration, cells were pre-treated with 5 µg/mL Mitomycin C, a DNA synthesis inhibitor [108]. After a 2-h incubation, cells were rinsed with PBS and created a vertical scratch wound by a sterile 200 µL pipette tip. Any dislodged cells were removed by washing with PBS. Post-scratch, cells were incubated in serum-free media (DMEM supplemented with 200 µg/mL BSA, 10 µg/mL Transferrin, and 2 mM L-glutamine) containing varying concentrations of purified KGF (1 nM, 10 nM, 100 nM). The negative control consisted of cells in serum-free media, while the positive control was treated with 80 nM human EGF. Images of designated regions were captured at 0 and 48 h using a 4X objective on the CELENA X high-content imaging system. The scratch area was quantified

using the 'Wound_Healing_size_tools' plugin in the Fiji software [110]. The wound closure percentage was calculated as follows:

$$\text{Wound closure\%} = \frac{\text{Wound area at 0 h} - \text{Wound area at 48 h}}{Wound\ are\ at\ 0\ h} \times 100\%$$

11. Thermal shift assay for protein stability assessment

The CFX96 real-time PCR system was employed in conjunction with SYPRO Orange fluorescent dye to evaluate the thermal stability of the proteins. A 50x working solution of SYPRO Orange was prepared by diluting 5 µL of the 1000x stock solution into 100 µL of deionized water. For the assay, 1.5 µL of the 50x SYPRO Orange working solution was combined with 13.5 µL of the protein sample (each tested in triplicate) in PBS buffer (pH 7.4). The samples were loaded into clear PCR tubes. The protein's melting temperature was indirectly determined by monitoring fluorescence signals across a temperature range of 10°C to 95°C in increments of 0.5°C for 10s.

12. Statistical analysis

Data were presented as means along with standard errors of the mean. All statistical analysis was performed using Microsoft Excel®. For pairwise comparisons, paired t-tests were employed. For comparisons among multiple groups, a one-way Analysis of Variance (ANOVA) was applied, followed by Bonferroni post-hoc tests to adjust for multiple comparisons. The significance level was established at $p < 0.05$ for all tests.

III. Results and Discussion

Part A. Overexpression and purification of Human granulocyte-macrophage colony-stimulating factor (GM-CSF)

A1. Results

Construction of plasmids

The main chain of the GM-CSF sequence was codon-optimized for expression in *E. coli* and chemically synthesized (Figure 1). The TEV protease cleavage site (ENLYFQ/G) was also placed at the N-terminal of the GM-CSF sequence. Using the BP reaction cloned all structures into the pDONR207 vector to create the GM-CSF entry vector (pENTR-GM-CSF). Next, the GM-CSF gene was conducted from the pENTR vector into different destination vectors which included different tag proteins (H-tag, H-MBP tag, or H-PDIb′a′ tag) using the LR reaction to obtain the expression constructs H-GM-CSF, H-MBP-GM-CSF, and H-PDIb′a′-GM-CSF. The expression vectors were certified by sequencing analysis before use.

Expression and solubility analysis

The H-GM-CSF, H-MBP-GM-CSF, and H-PDIb′a′-GM-CSF constructs were transformed into five different *E. coli* strains. IPTG was used for inducing recombinant protein production at three different temperatures. Total lysate, supernatant, and pellet fractions were evaluated with SDS-PAGE (Figure 2) and analyzed as the percentage of expression and solubility using Gel Analyzer software (Table 6). The fusion H-GM-CSF proteins were overexpressed at 37°C and 30°C mainly in an insoluble form, in all the tested *E. coli* strains, but the expression was hardly observed from any culture grown at 18°C (Figure 2; Table 6). When using the H-MBP or H-PDIb′a′ tags in distinct *E. coli* strains, the fusion proteins showed incompatible expression levels in either soluble or insoluble forms (Table 6). In the case of the H-MBP tag, fusion protein was produced at 37°C in the Origami 2 strain (Figure 2, Table 6) with the expression of 25.5% and the highest solubility of 33.3%. In other *E. coli* strains, the expression level was 22–30%, with the mass of overexpressed proteins in insoluble form (Figure 2; Table 6). At 30°C in the ClearColi strain, the H-MBP-GM-CSF protein exhibited an expression level of 33.1% and solubility of 63.5%, compared

to the production in other strains, with expression levels of 14–20%, primarily involving fusion proteins in soluble forms. At 18°C, H-MBP-GM-CSF exhibited a low expression level that was inadequate for large-scale protein purification. Regarding the PDIb′a′ tagged fusion proteins, the ClearColi (Figure 2) and Origami 2 strains had higher expression levels than the other strains at 37°C and 30°C. At 18°C, the ClearColi, Origami 2, and SHuffle T7 strains showed expression levels of around 15% (Table 6), while a small protein expression could be identified in the LOBSTR strain. Overall, our results indicated that at a 30°C incubation temperature and induction with 0.5 mM IPTG, the H-MBP-GM-CSF and H-PDIb′a′-GM-CSF products were mostly overexpressed in soluble forms in the ClearColi and Origami 2 *E. coli* strains.

Growth Rate

The growth rate of bacteria is another crucial factor impacting the overall production of recombinant proteins. Our current analyses evaluated this using standard optical density readings at a 600 nm wavelength (OD_{600}; Figure 3). Without IPTG induction, all the trial *E. coli* strains reached a maximum OD_{600} at 8 h. The BL21 and LOBSTR strains grew fastest, the SHuffle strain was in the middle, and the ClearColi and Origami 2 strains showed the slowest growth rates (Figure 3A). In protein production, protein expression was induced using IPTG when the OD_{600} reached 0.6–0.8. Compared to the control cultures, the growth of all the strains was subsequently reduced (Figure 3B), whereas their relative growth rates remained consistent. Another significant finding was a notable decrease in the maximum OD_{600} for the ClearColi, SHuffle, and Origami 2 strains.

Purification of recombinant GM-CSF

According to our soluble overexpression test results of the recombinant GM-CSF in *E. Coli*, we chose the H-PDIb′a′-GM-CSF from the ClearColi strain, inoculated at 30°C with IPTG induction. The final GM-CSF preparation was purified using a two-step chromatography method involving immobilized metal affinity chromatography (IMAC) followed by ion exchange chromatography (IEC), with TEV protease cleavage after the IMAC step (Figure 4A). Before purification, the large-scale expression of fusion protein H-PDIb′a′-GM-CSF was examined by SDS-PAGE, confirming that the

fusion protein (Figure 4B, Lane 2) exhibited overexpressed compared to the culture without IPTG induction (Figure 4B, lane 3). Because all our tags contain poly Histidine at their N-termini, the H-PDIb′a′-GM-CSF product was able to bind to the IMAC column, effectively isolated using the elution buffer, and subsequently analyzed using SDS-PAGE for protein fractions (Figure 4B, Lane 4).

For recombinant GM-CSF biochemical studies, the fusion tag in our H-PDIb′a′-GM-CSF product was cleaved using tobacco etch virus (TEV) protease. Following dialysis of the IMAC-purified fractions, the fusion protein was cleaved with the TEV protease and subsequently confirmed using SDS-PAGE (Figure 4B, Lane 5).

Subsequently, both the target protein and H-PDIb′a′-tag were purified using anion exchange chromatography, utilizing the isoelectric point of GM-CSF. The eluted GM-CSF (14.6 kDa) displayed a purity of 97.5% (Table 7) when determined by silver-stained (Figure 4C) and was analyzed using Gel Analyzer software. Under reducing SDS-PAGE conditions, the GM-CSF product appeared as a single band with further down of mobility than under non-reducing conditions (Figure 4D). Generally, 0.5 L culture of *E. coli* expressing H-PDIb′a′-GM-CSF could obtain 0.65 mg of GM-CSF with a purity of 97.5%, constituting 13.5% of the total protein in cell lysate (Table 7). To confirm the identity of the purified GM-CSF, LC-MS/MS analysis was conducted (Figure 5). Multiple fragments were detected, with a coverage of 96.09%.

Biological activity of purified GM-CSF

The purified GM-CSF was evaluated for biological effects on the TF-1 human leukemia cell line using the CCK8 assay, which assesses cell viability. Exposure of these cells to our purified GM-CSF stimulated their growth in a dose-dependent manner like that of the commercial GM-CSF (Figure 6). The Hill equation was employed to characterize the sigmoid shape of the dose-response curve. For the purified GM-CSF, the EC_{50} and Hill coefficient were 16.4 ± 2 pM and 0.49 ± 0.08 (n = 3), respectively. Correspondingly, the commercial GM-CSF exhibited the EC_{50} and Hill coefficient of 10.3 ± 3.4 pM and 0.63 ± 0.22 (n = 3), respectively.

A2. Discussion

Many heterologously expressed proteins, particularly those containing disulfide bonds, often misfold in the cytoplasm of *E. coli* and aggregate into inclusion bodies. Recombinant GM-CSF similarly accumulates in this insoluble fraction when expressed in *E. coli*, involving manual solubilization and refolding [19-22]. In our current study, we varied three parameters, such as the *E. coli* strain, expression temperature, and protein tag, to establish the optimal protocol for producing a soluble and substantial GM-CSF preparation in a bacterial expression system. Among these variables, the protein tag emerged as the most influential (Figure 2 and Table 6). In other studies, seven or eight different tags were evaluated in attempts to improve the solubility of passenger proteins [101-104, 111-115]. However, in current studies, only three tags (His, MBP, and PDIb′a′) were assessed. MBP and PDIb′a′ consistently demonstrated favorable outcomes, while His tag was widely used as a control. Notably, both the MBP and PDIb′a′ tags significantly enhanced the solubility of the GM-CSF when expressed in *E. coli* (Figure 2 and Table 6).

The mechanisms underlying the beneficial solubilization effects of the MBP and PDIb′a′ tags have yet to be fully clarified. MBP is a popular enhancer of passenger protein for its ability to enhance the solubility [116, 117]. It has been proposed to function as a holdase, engaging in transient physical interaction where a folded MBP moiety interacts with incompletely folded passenger protein [117]. The presence of hydrophobic residues on its surfaces and its open conformation is also thought to contribute significantly to MBP activity [118]. PDI is a modular polypeptide composed of four homologous domains, including a, b, b′, and a′, plus a C-terminal extension, c [119]. PDIb′a′ consists of two flexibly linked domains, b′ and a′, which represent the minimal region required for enzyme activity [120]. It has been validated that PDIb′a′ also contains hydrophobic patches [121]. Individual domains of PDIb′a′ such as b′ or a′, do not dispose of enhanced solubility effects on passenger proteins (data not shown).

The GM-CSF protein contains two intramolecular disulfide bonds, which are challenging to properly in the naturally reducing environment of the *E. coli* cytoplasm.

The *E. coli* strains SHuffle T7 and Origami 2 were developed to enhance cytoplasmic disulfide bond formation by incorporating genes such as trxB, gor, and ahpC*. In addition to these genes, SHuffle T7 expresses DsbC, further modifying the cytoplasmic redox environment [122]. In previous studies, we observed that using SHuffle T7 or Origami 2 strains improved the solubility of the expressed passenger proteins [104, 105]. Not surprisingly, this effect was more dramatic for albumin than for Oncostatin M, which contains 17 and 2 disulfide bonds, respectively. In our present study, however, neither of these strains improved the solubility nor increased the expression level of GM-CSF (Table 6), suggesting that the lack of disulfide bond formation is not the principal reason for the insolubility of this protein when expressed in bacteria.

Data in Table 6 show that reducing the induction temperature could enhance GM-CSF solubility, consistent with the findings of many previous reports [101, 111, 123]. When the folding of a protein is enzyme-mediated, if the increase in temperatures could improve the enzyme's activity, a higher number of soluble proteins will be produced. However, our results indicated that the folding of GM-CSF in *E. coli* is not mediated enzymatically. Instead, a higher temperature promotes the protein's production but compromises quality control, increasing misfolded protein. Conversely, low-temperature induction could slow the growth rate of *E. coli* [124], therefore reducing protein expression levels (Table 6).

The lesser mobility of GM-CSF in SDS-PAGE under reducing conditions suggests the presence of disulfide bonds, indicating a compact structure of purified protein. A single band in either the reducing or non-reducing conditions confirms that the purified GM-CSF is a monomer. From a 0.5 L flask culture, we successfully purified 0.65 mg of GM-CSF with a 97.5% purity (Table 7), resulting in a yield of 1.3 mg/L. However, this yield is substantially lower than other studies using the flask culture method, which have reported yields ranging from 7 to 20 mg/L for GM-CSF production [21, 25, 27]. Using a fermentation process by bioreactor also increased the yield up to 90 mg/L [22], while using batch-fed culture in *Pichia pastoris* can yield up to 760 mg/L of mouse GM-CSF [125]. The lower yields observed in this study can

be attributed in part to the lower cell density of the ClearColi strain during protein expression (Figure 3B).

Endotoxin contamination has always caused a challenge in the production of human proteins using Gram-negative bacteria such as *E. coli*. Various methods have been employed to alleviate this issue, including ultrafiltration, activated carbon, surfactants, anion-exchange chromatography, histamine- and histidine-immobilized Sepharose, and polymyxin B-immobilized Sepharose [126-132]. In this study, ClearColi has been used, which a the genetically modified *E. coli* strain with non-immunogenic LPS [133]. This eliminates the need for an endotoxin-removal step and importantly, reduces the risk of endotoxin-related toxicity in clinical applications. Based on the initial achievement in improving the production of soluble GM-CSF in *E. Coli*, we anticipate that further optimal protocols can be evolved to give a higher yield and enhanced biological activity.

The purified GM-CSF proteins obtained from *E. coli* exhibited a vigorous stimulatory effect on TF-1 cells, a human leukemia cell line sensitive to various cytokines, such as interleukin-3 and GM-CSF. The EC_{50} of this stimulation was 16.4 ± 2 pM, corresponding to 0.23 ± 0.03 ng/mL (Figure 6). This value is moderately higher than the range of 10–80 pg/mL reported in previous studies [21, 24]. Compared to results from the commercial GM-CSF (Figure 6), inconsistencies suggest potential systematic errors in the biological assays and could also relate to protein stability post-freeze-thaw cycles. Disregarding these issues will be crucial for future improvements.

Part B. Overexpression and purification of human interleukin-3 (IL3)

B1. Results

Construction of plasmids

To enhance the solubility of IL3 in the cytoplasm of *E. coli*, the IL3 sequence, encompassing amino acids 20-152 (excluding the signal peptide), was tagged with two solubility enhancers, H-MBP and H-PDIb'a' (Figure 7A). Additionally, a TEV protease cleavage site was introduced upstream of the IL3 sequence. This codon-optimized sequence was initially cloned into the pDONR207 plasmid using the BP reaction of the Gateway cloning system, generating the pENTR-IL3 construct harboring the IL3 gene (Figure 7B). Subsequently, the IL3 gene was transferred into expression vectors containing the tag proteins (either H-MBP or H-PDIb'a') using the LR reaction. This yielded the expression constructs H-MBP-IL3 and H-PDIb'a'-IL3. The integrity and accuracy of the final expression vectors were validated through sequence analyses.

Expression and solubility analysis

To identify suitable *E. coli* strains and conditions for efficient IL3 production, we introduced the constructs into six different *E. coli* strains (BL21, ClearColi, LOBSTR, Origami 2, Rosetta-gami B, and SHuffle T7). The protein expression was then initiated using IPTG across three distinct temperatures: 37°C, 30°C, and 18°C. Subsequently, the entire bacterial lysate, supernatant, and pellet fractions were analyzed using SDS-PAGE (Figure 8A), quantifying both the expression and solubility levels (Figure 8B and Tables 8). The fusion protein H-MBP-IL3 and H-PDIb'a'-IL3 exhibited varying expression and solubility across different *E. coli* strains. Nevertheless, across all strains, both proteins consistently demonstrated enhanced solubility and expression at the lower temperatures of 18°C and 30°C.

Considering the expression and solubility metrics of the fusion proteins, combined with the purification benefits offered by different *E. coli* strains, we selected H-PDIb′a′-IL3 expressed in ClearColi at 30°C for further study. Monitoring the cell growth and expression profiles in a 1 L flask culture (Figure 9, Table 9), we observed that *E. coli* cells maintained a consistent growth trajectory for approximately 11 h,

beyond which a plateau emerged. The black arrow on the graph indicates the moment of IPTG induction, aligning with the peak cell growth rate. Post-IPTG induction, a noticeable surge in fusion protein expression was observed, peaking merely 2 h after induction.

Purification of recombinant IL3

In initial attempts at purifying H-MBP-IL3, both amylose and immobilized metal affinity columns were applied. However, H-MBP-IL3 exhibited low-affinity binding to these columns, resulting in a substantially low purification yield (Figure 10, Table 10). Given these unfavorable results, the H-MBP-IL3 construct was not to proceed further.

Notably, the highest soluble expression level of H-PDIb'a' was obtained in ClearColi at 30°C (Figure 8B). Therefore, this condition is for further H-PDIb'a'-IL3 production. The overexpressed fusion protein H-PDIb'a'-IL3 was subjected to a two-step chromatographic purification process (Figure 11A). Utilizing an IMAC column (HisTrap HP column), H-PDIb'a'-IL3 was effectively isolated, and subsequent protein fractions were assessed through SDS-PAGE (Figure 11B, Lane 4). Following IMAC, the dialyzed fusion protein fractions were cleaved using the TEV protease (Figure 11B, Lane 5). Exploiting the distinct isoelectric or isoionic points of the target protein and the H-PDIb'a' tag, the two were differentiated using ion exchange chromatography (IEC). Given that IL3 possesses the cysteine residue essential for forming a disulfide bond, the purified protein's integrity was examined under both reducing and non-reducing SDS-PAGE conditions. Notably, protein bands were observed at the expected monomeric size in both conditions (Figure 11C). The purity of the isolated IL3 was assessed through silver staining (Figure 11D), revealing an impressive purity level of 93 ± 4%. A typically 0.5 L *E. coli* culture expressing H-PDIb'a'-IL3 yielded approximately 0.36 ± 0.02 mg of IL3, accounting for 1 ± 0.1% of the total protein from the lysate (Table 11). To further validate the identity of the purified protein, an LC-MS/MS analysis identified masses from various fragments, ensuring a comprehensive 100% coverage (Figure 12).

Biological activity of purified IL3

The functionality of the purified IL3 was confirmed through the TF-1 cell proliferation assay, with commercial IL3 serving as a positive control. The purified IL3 notably stimulated TF-1 cell proliferation in a dose-responsive manner (Figure 13). An interesting observation was the two-step stimulation pattern displayed by both purified IL3 and commercial IL3 in promoting TF-1 cell growth (Figure 13A). In the lower concentration range, purified IL3 exhibited a robust stimulation of 160%, surpassing the stimulation exhibited by commercial IL3 (132%) (Table 12). However, at elevated concentrations, no significant difference in the stimulatory effects between the two IL3 versions was observed.

The biological effect of H-PDIb'a'-IL3 was also analyzed on TF-1 cells. Mirroring the IL3 trend, H-PDIb'a'-IL3 induced a two-step proliferation response in TF-1 cells, although the initial stimulation required higher concentrations than that of pure IL3 (Figure 13B, Table 12).

B2. Discussion

IL3 has been recognized as a "difficult-to-express protein" [134] primarily due to its limited expression when expressed alone in *E. coli* [38, 134]. Nevertheless, in this study, we successfully expressed soluble IL3 within the *E. coli* cytoplasm, attaining a highly pure product. Notably, our purified IL3 demonstrated superior biological activity compared to commercially available IL3 in the TF-1 cell proliferation assessment.

Davis et al. (1999) obtained IL3 from a NusA fusion protein with a yield of 4.3 mg/L of flask culture [43], higher than our yield of 0.36 mg/0.5L (Table 11). However, the yields per gram of cell paste were comparable, with Davis et al. obtaining 0.49 mg IL3 per gram and our result is 0.5 mg/g (calculated from 1.45 g cell paste per 0.5 L culture). The lower volumetric yield in our study is likely due to the slower growth of the ClearColi strain used for expression compared to the JM105 strain used by Davis et al.[106]. Interestingly, Davis et al. predicted that an MBP fusion would result in an insoluble protein using a revised Wilkinson-Harrison solubility model that considered only the numbers of certain amino acids. However, protein solubility in E. coli depends on various other factors not incorporated in the model, such as expression temperature, host strain, and fusion tags. MBP is a well-established versatile solubility-enhancing fusion partner [117]. Indeed, we found the solubility H-MBP-IL3 exceeded 90% when expressed at 18°C in BL21 (Table 8), underscoring the limitations of sequence-based solubility prediction.

Hercus et al. (2013) reported yields as high as 3-7 mg/L for an IL3 variant comprising residues 13-125 (113 amino acids total) [44], whereas our construct included the complete 133-residue IL3 sequence. The difference in construct size complicates a direct comparison of yields. Additionally, Hercus et al. estimated IL3 quantity using analytical size exclusion chromatography (SEC) with theoretical extinction coefficients, while both Davis et al. and our study used the more straightforward BCA protein assay for quantitation.

A key advantage of our H-PDIb′a′ fusion partner (35 kDa) compared to the NusA tag (55 kDa) used by Davis et al. (1999) and Hercus et al. (2013) is its 36% smaller size. A more compact fusion tag may allow more efficient expression by reducing the cellular resources required for protein synthesis. Furthermore, while Davis et al.(1999)

used factor Xa to cleave the fusion tag, both Hercus et al. (2013) and our study employed TEV protease, which offers higher specificity and reduced risk of undesired cleavage compared to factor Xa.

A unique observation in our study was the two-step concentration-dependent increase in TF-1 cell proliferation (Figure 13). Different concentrations of IL3 have been shown to exert varying biological effects. For example, low concentrations of IL3 generally promote the survival and proliferation of hematopoietic progenitor cells, acute myelogenous leukemia cells, basophils, and endothelial cells, while high concentrations induce differentiation, apoptosis, or activation of the cells [135-138]. This phenomenon may be related to the low- and high-affinity IL3 receptor [139-142]. The α subunit (CD123) of the IL3 receptor binds IL3 with low affinity. However, the combination of the α subunit with the β subunit (CD131) results in a high-affinity receptor complex essential for signal transduction [143]. The low-affinity receptor is believed to be primarily involved in the initial capture and concentration of IL3 on the cell surface [144, 145]. Furthermore, the β subunit is shared among the receptors for IL3, IL5, and GM-CSF. Consequently, very high concentrations of these cytokines can bind to the other cytokine receptors, eliciting cellular responses [140, 141, 146-150]. This raises the possibility that a high concentration of IL3 may promote cell proliferation through the GM-CSF and/or IL5 receptors.

The biological activity exhibited by the H-PDIb'a'-IL3 fusion protein was unexpected. To our knowledge, no report of IL3 fusion protein activity has been reported. Intriguingly, H-PDIb'a'-IL3 elicited a two-step stimulation in TF-1 cells, akin to pure IL3 (Figure 13B). However, a notable distinction was the evaluated EC_{50} value associated with the initial stimulation by H-PDIb'a'-IL3, indicating reduced activity when compared to IL3 alone (Table 12). This decreased potency could potentially be attributed to the steric interference caused by the H-PDIb'a' tag, which might impede the binding of IL3 to its designated receptor.

This study, while informative, presents certain limitations that warrant consideration. Firstly, our study focused exclusively on specific fusion tags (MBP and PDIb'a') without conducting a broader comparative analysis with other potential solubility-enhancing tags or approaches. Additionally, while we demonstrated the

production of IL3 with enhanced solubility and biological activity, the scalability of this approach for industrial-level production remains to be explored. Furthermore, the long-term stability and storage conditions of the produced IL3, factors crucial for practical applications, have not been explored in our study. Lastly, it is important to note that the overall yield was relatively low.

Future research endeavors may explore various aspects related to IL3 production. Firstly, the slow growth of the ClearColi strain could potentially be overcome by high cell density bioreactor culture using rich, complex media formulations to boost biomass and productivity [151]. Secondly, the purification processes for scale-up can be optimized and the stability of the recombinant IL3 under various storage conditions can be investigated. Thirdly, the biological activity of IL3 in different cell lines or *in vivo* models could be explored to attain a more comprehensive understanding of its therapeutic potential.

Part C. Overexpression and purification of human interleukin-6 (IL6)

C1. Results

Construction of plasmids

To augment the solubility of IL6 within the cytoplasm of *E. coli*, a modified IL6 sequence spanning amino acids 30 to 212 (excluding the signal peptide) and bearing a single disulfide bond was engineered. This modified IL6 sequence was appended with two soluble enhancer proteins, MBP and PDIb'a' (Figure 14A). Additionally, a recognition site for TEV protease cleavage was introduced immediately preceding the IL6 sequence. The codon-optimized composite sequence was initially inserted into the pDONR207 plasmid through the BP reaction of the Gateway cloning system, resulting in the creation of the pENTR construct housing the modified IL6 gene (Figure 14B). Subsequently, the engineered IL6 gene was homologous recombination into expression vectors containing the tag proteins (either MBP or PDIb'a') using the LR reaction. This process yielded the final expression constructs, denoted as H-MBP-IL6 and H-PDIb'a'-IL6. The fidelity and accuracy of these ultimate expression vectors were verified through comprehensive sequencing analysis.

Expression and solubility analysis

To identify the optimal *E. coli* strains and conditions for IL6 production, the introduction of the constructs into *E. coli* strains (BL21, ClearColi, Origami 2, and SHuffle T7) was carried out. The initiation of protein expression was achieved through the utilization of IPTG at three distinct temperatures: 37°C, 30°C, and 18°C. Subsequently, the analysis of the entire bacterial lysate, supernatant, and pellet fractions was executed using SDS-PAGE (Figure 15A), and the quantification of the expression and solubility levels was performed (Figure 15B and Tables 13).

Varying expression and solubility profiles of the fusion proteins H-MBP-IL6 and H-PDIb′a′-IL6 were observed, contingent upon the specific *E. coli* strain utilized. For the fusion protein H-PDIb'a'-IL6, enhanced solubility and expression were consistently demonstrated at the lower temperatures of 30°C and 18°C. Notably, peak expression and solubility levels were achieved by the H-MBP-IL6 fusion protein when it was expressed in BL21 at 30°C and in the Origami 2 strain at 37°C.

Considering the expression and solubility metrics of the fusion proteins, in conjunction with the purification benefits offered by different *E. coli* strains, H-PDIb′a′-IL6 expressed in ClearColi at 30°C was chosen for further study. In the monitoring of cell growth and expression profiles in a 1 L flask culture (Figure 16), a consistent growth trajectory of *E. coli* cells was maintained for approximately 12 h, after which a plateau emerged. The optimal IPTG induction moment, coinciding with the peak rate of cell growth was marked by the black arrow on the graph. After IPTG induction, a notable surge in fusion protein expression was observed, reaching its zenith a mere 2 h after induction, and becoming saturated after 4 h.

Purification of H-PDIb'a'-IL6 and cleaved IL6

The H-PDIb′a′-IL6 fusion protein underwent a purification process consisting of two sequential steps: immobilized metal affinity chromatography (IMAC) and ion exchange chromatography (IEC). The IMAC step initially yielded a purity level of 56 ± 3% for the H-PDIb′a′-IL6 fusion protein, which was subsequently enhanced to 93 ± 2% following the IEC step (Figure 17 and Table 14).

For the purification of cleaved IL6, the initial IMAC process resulted in a protein purity of 54 ± 2%. Subsequently, using a HisTrap HP column, the fusion protein was effectively isolated, and the ensuing protein fractions were assessed through SDS-PAGE (Figure 18B, Lane 4). Following IMAC, the dialyzed fusion protein fractions underwent cleavage by the TEV protease (Figure 18B, Lane 5). Leveraging the distinct isoelectric or iso-ionic points between the target protein and the PDIb'a' tag, the two components were separated using anion exchange chromatography.

Considering that IL6 possesses a cysteine residue crucial for forming a disulfide bond, the purified protein's integrity was scrutinized under both reducing and non-reducing SDS-PAGE conditions. Significantly, protein bands of the expected monomeric size were observed under both conditions (Figure 18C). The purity of the isolated IL6 was assessed using silver staining (Figure 18D), revealing an impressive purity level of 97.9 ± 0.2%.

Typically, a 0.5 L *E. coli* culture expressing H-PDIb′a′-IL6 yielded approximately 0.39 ± 0.03 mg of IL6, constituting 0.38% of the total protein from the lysate (Table 15).

Thermal stability

The thermal stability of purified IL6 and fusion protein H-PDIb'a'-IL6 was measured using a thermal shift assay (Figure 19). Fusion protein H-PDIb'a'-IL6 exhibited a flatting shape in the melting profile. In contrast to IL6 with a sharp slope, indicating proper folding, with a melting temperature (T_m) of 65.8 ± 0.4°C in PBS buffer. It may be because the hydrophobic residues on fusion proteins are already exposed on the surface and the dye binds strongly to it.

Biological activity of purified proteins

The functionality of the purified IL6 was validated using the TF-1 cell proliferation assay. Our purified IL6 exhibited a pronounced, dose-dependent stimulation of TF-1 cell proliferation (Figure 20). The biological effects of both H-PDIb′a′-IL6 and the isolated PDIb'a' were also assessed on TF-1 cells. IL6 induced TF-1 cell proliferation with an EC_{50} of 1.36 ± 0.13 nM and a Hill coefficient of 0.62 ± 0.22, resulting in a maximum response of 190% in a concentration-dependent manner.

In the case of H-PDIb′a′-IL6, the EC_{50} and Hill coefficient were determined to be 6.7 ± 1.2 nM and 0.6 ± 0.1, respectively, accompanied by a significantly higher maximum response of 380%. Intriguingly, PDIb'a' on its own also demonstrated a stimulatory effect on TF-1 cells, leading to a 150% increase in proliferation. However, the EC_{50} and Hill coefficient for PDIb'a' were measured as 654 ± 90 nM and 0.33 ± 0.03, respectively.

C2. Discussion

IL6 is a glycoprotein of four antiparallel α-helical bundle structures with two intramolecular disulfide bonds. Producing elevated levels of soluble IL6 in *E. coli* poses a challenge due to its pronounced tendency to aggregate [47, 49, 50, 52, 152-157]. While purifying recombinant protein fusions containing tag protein is common, selecting an appropriate tag protein for specific proteins of interest can be a non-trivial task. For example, GST fusion showed negligible impact on IL6 solubility [52]. In contrast, the MBP or PDIb'a' tag increased the solubility to 98% (Table 13). Although exactly how MBP and PDIb'a' promote the solubility of the fusion proteins is unknown, several reasons are suggested as to why they are good chaperones. First, they are highly soluble stable proteins themselves [117, 158]. Second, they are relatively large proteins, providing an extensive surface area. This large surface area is beneficial for interacting with various unfolded polypeptides, preventing their aggregation. Even though PDIb'a' has a smaller size (34 kDa) than MBP (40 kDa), it still showed good solubility-enhancing activity (Table 13). Third, both have somewhat disordered regions, allowing them to be flexible in accommodating various proteins [159, 160]. This flexibility can be crucial in binding and stabilizing different target proteins. Fourth, both have hydrophobic surface patches that can interact with the unfolded proteins which expose some of the hydrophobic interior of the proteins [161-163].

Codon optimization of IL6 for *E. coli* expression had a notable impact on protein secretion [49, 164]. In line with our findings, the codon-optimized IL6 gene significantly enhanced the expression level of the H-MBP-IL6 fusion protein, achieving an impressive 68% improvement compared to the 20% without the codon optimization [50]. Notably, when contrasted with MBP tags, the H-PDIb′a′-IL6 fusion protein tends to have superior expression levels, reaching up to 81% (Table 13).

The final yield of IL6 isolated from the H-PDIb′a′-IL6 fusion protein was as low as 0.38 mg/0.5L (Table 15). This low yield resulted from the low *E. coli* biomass of ClearColi during the flask cultivation procedure and the limited binding of the H-PDIb′a′-IL6 fusion protein to the IMAC column. Despite advancements in free lipopolysaccharide production, the wet biomass obtained from flask cultivation of ClearColi was lower than that of other *E. coli* strains, caused by its low growth rate compared with other strains [106]. The wet biomass yield could be enhanced by

implementing fermentation methods. Another contributing factor to the low final yield of IL6 was the limited binding of the H-PDIb′a′-IL6 fusion protein during the IMAC purification step. Specifically, the H-PDIb′a′-IL6 fusion protein yield obtained from the IMAC step was only 12.1 ± 0.6%. This suggests hindered exposure of the His-tag at the N-terminus of the protein. The low yield can be seen in other reports. For example, hemolysin-tagged IL6 was purified using a single-step immunoaffinity chromatography approach from culture supernatant, resulting in yields of 18 µg/L of culture supernatant with purity in the range of 60% [49]. The folding characteristics and thermal stability of purified IL6 were also evaluated using a thermal shift assay. The melting temperature (T_m) of IL6 was 65.8 ± 0.4°C (Figure 19), slightly higher than the T_m of 60°C reported by Arakawa et al. (2006) [165]. In contrast, the fusion protein did not exhibit a distinct melting temperature, which caused differences in hydrophobic residues exposed on its protein surface. These results strongly reveal the well-folding structure of purified IL6.

The EC_{50} value of IL6 obtained from this study was 1.36 nM which is equivalent to 28.5 ng/mL. This value is relatively high compared to those of 10 - 800 pg/mL [52, 166, 167]. This is probably because our study used TF-1 cells which require IL3 or GM-CSF for efficient proliferation whereas other studies used the B9 mouse hybridoma line, T1165.85.2.1 cell line, and MH60.BSF2 cell line, which requires IL6 for its survival and proliferation. Some studies showed no biological activity test results [49, 50, 154, 168].

The unexpected biological activity observed with the H-PDIb′a′-IL6 fusion protein merits attention. There are no prior accounts of IL6 fusion protein activity. Notably, H-PDIb′a′-IL6 exhibited a more pronounced stimulatory effect on TF-1 cells compared to pure IL6. The precise mechanism underpinning this enhanced activity of H-PDIb′a′-IL6 fusion protein remains elusive and warrants comprehensive exploration in future studies.

Part D.Expression and purification of Human Keratinocyte growth factor (KGF)

D1. Results

Construction of plasmids

To assess the impact of different protein tags on the expression levels and solubility of KGF in an *E. coli* expression system, three distinct tags (PDIb'a', MBP, and His) were genetically fused to the N-terminus of KGF. The signal peptide of KGF was excluded to facilitate cytoplasmic expression in *E. coli* and to ensure the functionality of the purified protein as a growth factor. Additionally, hexahistidine residues were incorporated at the N-terminus of both PDIb'a' and MBP to facilitate purification, resulting in the generation of three constructs, H-KGF, H-PDIb'a'-KGF, and H-MBP-KGF (Figure 21). A TEV protease cleavage site (ENLYFQ/G) was inserted between each tag and the KGF coding sequence to enable the facile removal of these tags during subsequent purification steps.

Expression and solubility analysis

Initially, the expression of H-KGF was evaluated in four *E. coli* strains (BL21, SHuffle T7, ClearColi, and LOBSTR). None of these strains yielded detectable levels of the recombinant protein (Figure 22). Subsequently, the expression and solubility of H-PDIb'a'-KGF and H-MBP-KGF were evaluated across five *E. coli* strains: BL21 pLysS, SHuffle T7 pLysS, ClearColi pLysS, LOBSTR, and BL21 AI (Figure 23A). These strains were selected due to the tight transcriptional and translational control offered by the pLysS plasmid or arabinose induction in BL21 AI, especially for genes that may be toxic. Additionally, LOBSTR was considered for its low expression of endogenous histidine-rich proteins, which is advantageous for His-tag purification.

Protein expression was evaluated at three different temperatures: 37°C, 30°C, and 18°C. Across all strains, KGF fusion proteins displayed varying levels of expression and solubility (Figure 23 and Table 16). The BL21 AI strain was particularly effective, yielding the highest expression levels for both H-MBP-KGF and H-PDIb'a'-KGF. Specifically, the soluble expression levels of H-PDIb'a'-KGF in BL21 AI were 18.0% and 19.2% at 30°C and 18°C, respectively. Furthermore, the soluble expression level of H-MBP-KGF in BL21 AI was as high expression as 21.5% at 30°C.

However, during the initial purification steps, H-MBP-KGF was found to be unstable and susceptible to autocleavage. In contrast, the size of H-PDIb'a'-KGF remained consistent during SDS-PAGE and demonstrated greater stability. Therefore, the H-PDIb'a'-KGF construct was selected for large-scale production and purification.

Purification of H-PDIb'a'-KGF and KGF

The H-PDIb'a'-KGF fusion protein underwent a two-step purification process, involving immobilized metal affinity chromatography (IMAC) and ion exchange chromatography (IEC). The IMAC step using the HisTrap HP column yielded a H-PDIb'a'-KGF fusion protein purity of 56 ± 3%, subsequently elevated to 92.1 ± 0.3% following the IEC step (Figure 24, Table 17). For KGF purification, after the initial IMAC step, the H-PDIb'a'-KGF fusion protein was treated with TEV protease. Subsequently, the PDIb'a' tag was separated from KGF by using IEC (Figure 25A), ultimately achieving g a final KGF purity of 96.9 ± 0.4% (Figure 25B). Silver staining further validated the purity of the final KGF (Figure 25D). To rule out the presence of any intermolecular disulfide bond due to the five cysteines in the KGF, SDS-PAGE was conducted under both non-reducing and reducing conditions; notably, no dimer/multimer size band was observed under non-reducing conditions (Figure 25C). Detailed information regarding the purities and yields of H-PDIb'a'-KGF and KGF throughout the purification stages are delineated in Table 17 and Table 18, respectively. The final purification yields documented H-PDIb'a'-KGF and KGF were 2.4% and 5.0%, respectively. Post-TEV protease cleavage, KGF exhibited efficient separation from other proteins during IEC resulting in higher purity and yield. On the other hand, H-PDIb'a'-KGF was less separable from other proteins during the IEC step, necessitating the collection of a narrow range of eluted fractions, leading to a lower yield. The characteristic of H-PDIb'a'-KGF during the IEC step also necessitated more thorough washing in the earlier IMAC step.

To confirm the identity of the purified KGF protein, LC-MS/MS was employed. The analysis yielded a fragment mass identification coverage of 81% (Figure 26), thereby providing substantial evidence for the successful purification and accurate identification of KGF.

Thermal stability of KGF, H-PDIb'a'-KGF, and H-PDIb'a'

The thermal stability of the purified proteins, KGF, H-PDIb'a'-KGF, and H-PDIb'a' was assessed using a thermal shift assay (Figure 27, Table 19). KGF exhibited low fluorescence at room temperature, indicating proper folding. In contrast, H-PDIb'a'-KGF and H-PDIb'a' displayed higher fluorescence levels than KGF at room temperature. The melting temperatures for these proteins are summarized in Table 6. KGF had a melting temperature of 59.2 ± 0.3°C in PBS buffer. On the other hand, H-PDIb'a'-KGF and H-PDIb'a' had a lower melting temperature compared to KGF, with values of 43.7 ± 0.2°C and 45.7 ± 0.2°C, respectively.

Biological activity of purified KGF

The proliferation activity of purified KGF was evaluated using human breast cancer MCF-7 cells and NIH-3T3 fibroblast cells. Purified KGF increased the proliferation of MCF-7 cells by 50% in a dose-dependent manner with an EC_{50} of 6.4 ± 0.5 pM and a Hill coefficient of 1.0 ± 0.5. However, NIH-3T3 cells did not respond to KGF treatment (Figure 28A). Subsequent experiments assessed the biological activity of H-PDIb'a'-KGF on MCF-7 cells (Figure 28B). H-PDIb'a'-KGF stimulated MCF-7 cell proliferation by 60% in a dose-dependent manner, with an EC_{50} of 79 ± 6 pM and a Hill coefficient of 1.2 ± 0.2. Interestingly, PDIb'a' alone also exhibited a stimulatory effect on MCF-7 cells, increasing proliferation by 40% with an EC_{50} of 1.7 ± 0.2 nM and a Hill coefficient of 1.4 ± 0.1. In addition to its proliferative effects, purified KGF also demonstrated notable wound-healing capabilities. Wound closure assessment 48 h post-scratch revealed marked healing in wounds treated with purified KGF compared to negative controls (Figure 29A). The wound closure induced by purified KGF occurred in a dose-dependent manner, and its efficacy was comparable to that of EGF (Figure 29B, Table 20).

D2. Discussion

In the present study, we successfully tackled the challenge of efficiently producing biologically active recombinant KGF. By employing the PDIb'a' protein tag, we improved the expression and solubility of KGF in E. coli and streamlined the process for its large-scale production and purification.

Our study determined the EC_{50} value for KGF-induced proliferation in MCF-7 cells was 6.4 pM (Figure 28). This represents the most potent KGF preparation to date, outperforming previously reported EC_{50} values ranging from 50 to 100 pM [63, 69, 76], and even exceeding those with substantially higher EC_{50} values [62, 75]. The maximal response elicited was approximately 50% (Figure 28A), which is consistent with findings from similar assays [75, 76]. Importantly, KGF failed to stimulate NIH-3T3 cells, which lack FGFR2-IIIb expression, corroborating earlier observations [76].

Additionally, the EC_{50} value for H-PDIb'a'-KGF was found to be approximately 10-fold higher at 78.6 pM, rendering it 10 times less potent than KGF alone (Figure 28). This reduced potency may be attributed to the potential interference of the PDIb'a' domain with the binding affinity of the KGF moiety to its cognate receptor. Nevertheless, this EC_{50} value remains within the range reported for other KGF preparations [63, 69, 76]. The maximal response for H-PDIb'a'-KGF was measured at approximately 60% (Figure 28B).

In the present study, we employed PDIb'a' and MBP as a fusion tag to enhance both the expression and solubility of KGF in the *E. coli* cytoplasm. While alternative tags such as GST [71], Sumo [72], Trx [73], and Fh8 [74] have been previously investigated, each demonstrated varying degrees of success in improving KGF expression and solubility. Notably, our H-PDIb'a'-KGF construct retained its mitogenic activity, a feature that distinguishes it from some other fusion constructs. For instance, the GST-KGF fusion protein was found to be altogether devoid of mitogenic activity [71]. Similarly, the SUMO-KGF fusion protein exhibited a significantly reduced capacity to stimulate cell proliferation compared to native KGF [72]. For KGF production, the yield in this study, 2.6 mg/L was comparable to the 2.8 mg/L achieved for KGF-1 production [73] but lower than the 4.8 mg/L reported in another study that employed high magnesium chloride [71].

Our thermal shift assay data revealed intriguing characteristics about the folding and stability of KGF and H-PDIb'a'-KGF. Firstly, the melting temperature of KGF was 59.2 ± 0.3°C (Table 19), which aligns closely with the value of 62.0 ± 1.7°C reported previously [169]. The low fluorescence observed for purified KGF (Figure 27) strongly suggests that the protein is well-folded. This inference is supported by the hydrophobic nature of the fluorescent dye used in the assay, which interacts minimally with the generally hydrophilic surface of a well-folded globular protein. In contrast, H-PDIb'a'-KGF exhibited high fluorescence, raising the possibility of either a partially folded state or a well-folded protein with a larger hydrophobic surface area. Structural analyses of PDI have revealed multiple hydrophobic patches on its surface, along with some disordered residues at the C-terminus [162, 170]. The lower melting temperature of H-PDIb'a'-KGF could be attributed to the inherent flexibility of the PDI molecule [171]. This flexibility may induce conformational changes that modulate the hydrophobic patches on the protein's surface [63, 121, 172, 173]. These findings provide valuable insights into the folding and stability of KGF and H-PDIb′a′-KGF, although they also raise new questions that warrant further investigation.

Surprisingly, our data reveal that PDIb'a' exhibits mitogenic activity, as evidenced by Figure 28B. To the best of our knowledge, this is the first report documenting the mitogenic potential of PDIb'a', and indeed of PDI itself. PDI is a multi-domain protein comprising four distinct domains: a, b, b', and a' [174-176]. Of these, only the 'a' domain is known to possess catalytic activity [171, 177]. The mechanism by which PDIb'a' stimulates cellular proliferation remains elusive at this point. One plausible explanation could be related to its chaperone activity [171, 177, 178], which may facilitate the proper folding and maintenance of membrane and extracellular proteins, thereby indirectly promoting cell growth. Alternatively, it is conceivable that PDIb'a' is internalized by cells, where it may exert its mitogenic effect through intracellular mechanisms. Further investigations are warranted to elucidate the precise role of PDIb'a' in cellular proliferation.

Part E. Expression and purification of Human Insulin-Like Growth Factor 1 (IGF1)

E1. Results

Construction of plasmids

Figure 30A illustrates three unique constructs designed for the expression of IGF1, each endowed with one of the following affinity tags: His, PDIb'a', or MBP. To ensure efficient and specific enzymatic cleavage, a Tobacco Etch Virus (TEV) protease recognition site—ENLYFQG—was strategically located at the N-terminus of the IGF1 coding sequences across all constructs. Additionally, homologous overlap sequences of approximately 20 base pairs were integrated at both termini of the IGF1 genes, serving as a foundation for seamless cloning into the destination vectors. For the ease of subsequent purification steps, each construct was further engineered to include either a His6 or His8 peptide tag at the N-terminus. Utilizing an in vivo cloning strategy, we streamlined the cloning process, as detailed in Figures 30B, C.

Expression and solubility analysis

We transformed four *E. coli* strains—namely, BL21, ClearColi, Origami 2, and SHuffle T7—with the designed expression plasmids. Subsequent evaluations of protein expression and solubility were carried out at three different temperatures: 37°C, 30°C, and 18°C, corroborated by SDS-PAGE analysis (Figure 31A). It was observed that the expression efficiencies of IGF1 constructs were influenced by several variables, including the type of affinity tag, the cultivation temperature, and the choice of *E. coli* strain.

Among the constructs, H-IGF1 yielded the least favorable expression rates, falling below the 10% threshold. Most of these fusion proteins were localized within inclusion bodies at 37°C, yet displayed notably enhanced solubility at the reduced temperature of 18°C. Conversely, the H-PDIb′a′-IGF1 fusion construct maintained a relatively consistent expression range between 40-70% across the examined temperatures. Strain-specific differences were evident, with Origami 2 and SHuffle T7 outperforming ClearColi and BL21 in terms of expression levels. Importantly, the solubility of the H-PDIb′a′-IGF1 constructs showed marked improvements at milder temperatures, with values exceeding 85% at both 30°C and 18°C.

In contrast, the H-MBP-IGF1 fusion protein exhibited its peak performance in the Origami 2 strain at 37°C, reaching an expression level of 65 ± 8% and a solubility rate of 75 ± 4% (Figure 31B, Table 21).

Purification of fusion proteins IGF1

Following the evaluation of expression and solubility profiles, we selected the H-MBP-IGF1 construct expressed in the Origami 2 strain for optimization of the purification process. Although attempts were made to purify H-MBP-IGF1 from the ClearColi strain, these efforts did not result in the successful isolation of the desired fusion protein; the details of which are not elaborated upon here. This suggests that the successful folding and solubility of the H-MBP-IGF1 fusion protein are notably influenced by the specific *E. coli* strains used for expression. A one-step purification approach was employed using amylose resin chromatography (MBPTrap HP column)(Figure 32). The resulting purified protein showcased remarkable purity, achieving a yield of 94 ± 1% and a recovery rate of 23 ± 3% (Table 22).

For the purification of the H-PDIb′a′-IGF1 construct, which was expressed in the ClearColi strain (Figure 33), we opted for a two-step purification strategy. Initial purification was facilitated by immobilized metal affinity chromatography (IMAC), capitalizing on the presence of an N-terminal hexahistidine tag. However, given that the IMAC step yielded suboptimal purity levels, a secondary purification stage was incorporated, utilizing ion exchange chromatography (IEC). This additional step significantly enhanced the purity of the protein, elevating it from an initial 66% to 85.5% (Table 22).

In summary, the H-PDIb′a′-IGF1 fusion protein achieved a commendable level of purity at 85.5%; however, it came at the cost of a reduced yield, registering at just 1.3%.

Purification of cleaved IGF1

To isolate the native IGF1 protein, we employed a construct containing the H-PDIb′a′-IGF1 fusion. The purification workflow was comprised of a tripartite strategy (Figure 34A). The process consisted of an initial immobilized metal affinity

chromatography (IMAC) phase, followed by TEV protease-mediated cleavage, and finalized by ion exchange chromatography (IEC). The IMAC step resulted in a partially purified H-PDIb′a′-IGF1 fusion protein with a purity level of 54 ± 1% (Table 23).

After the IMAC stage, the fusion protein was treated with TEV protease to facilitate the excision of the PDIb'a' tag from the core IGF1 protein (Figure 34B, lane 5). The successful cleavage of the PDIb'a' moiety was instrumental in the ensuing IEC-based purification step. The final IEC phase yielded a highly purified IGF1 with an impressive purity level of 91.4 ± 0.9% (Figure 34B, lane 6; Table 23).

Stability of the purified proteins

The thermal stability of IGF1, H-MBP-IGF1, H-PDIb'a'-IGF1, and related proteins was assessed through a thermal shift assay (Figure 35). IGF1 displayed a consistent signal suggesting the heat-stable, with melting temperatures exceeding the range of temperatures under study. In a PBS buffer, the melting temperatures of MBP and H-MBP-IGF1 were 57.3 ± 0.2°C and 55.5 ± 0.3°C, respectively. In contrast, PDIb'a' and H-PDIb′a′-IGF1 exhibited heat sensitivity with a lower melting temperature of 47.3 ± 0.2°C and 31.2 ± 0.1°C, extremely lower than MBP or H-MBP-IGF1.

Biological activity of purified proteins

The in vitro biological efficacy of the various purified IGF1 constructs was evaluated using the MCF-7 human breast carcinoma cell line as a model system [106]. All tested forms of IGF1 were found to stimulate MCF-7 cell proliferation in a dose-responsive manner (Figure 36).

Among the proteins scrutinized, the native, purified IGF1 exhibited the most potent activity, with an EC_{50} value of 0.5 ± 0.1 nM and a Hill coefficient of 0.34 ± 0.01. The maximum response was recorded as 305 ± 16 (Table 24).

Conversely, the fusion proteins H-PDIb'a'-IGF1 and H-MBP-IGF1 showed considerably higher EC_{50} values, indicating reduced potency. Specifically, H-PDIb′a′-IGF1 demonstrated an EC_{50} value of 8.6 ± 2.8 nM and a Hill coefficient of 0.6 ± 0.1 (n = 3), along with a maximum response of 388 ± 16. For H-MBP-IGF1, the EC_{50} and

Hill coefficient were 13.7 ± 3.2 nM and 0.8 ± 0.1 (n = 3), respectively, with the maximum response reaching 338 ± 16.

In summary, while all forms of IGF1 could stimulate cell proliferation, the native, purified IGF1 was the most effective in terms of potency, as evidenced by its lower EC_{50} value and Hill coefficient.

E2. Discussion

The inherently unstable nature of IGF1 under physiological conditions has long been documented [80]. This instability often leads to the formation of inclusion bodies when the protein is expressed in the *E. coli* cytoplasm at 37°C, as corroborated by numerous prior studies [179-182]. Our findings were consistent with these observations, particularly in the case of H-IGF1, which showed substantial insolubility at 37°C. Although lowering the induction temperature improved solubility, the low expression levels (<10%) rendered this construct unsuitable for large-scale purification.

To overcome these limitations, we engineered fusion proteins by attaching two different tags—PDIb'a' and MBP—to the N-terminus of IGF1. Notably, these tagged constructs demonstrated enhanced solubility even at 37°C. This is in line with previous reports that have similarly leveraged tags such as trx or dsbA to improve the solubility of IGF1 fusion proteins [98, 183]. However, when the fusion tags were removed by TEV protease cleavage, the yield of untagged IGF1 remained suboptimal. This prompted us to investigate the biological activities of the fusion proteins with their tags intact.

Intriguingly, both H-PDIb'a'-IGF1 and H-MBP-IGF1 constructs displayed appreciable in vitro biological activities. This suggests that the fusion tags contribute to the stabilization of IGF1 and do not hinder the protein's interaction with its receptor or its ability to elicit cellular responses. These observations underscore the potential utility of fusion tags both as solubility enhancers and as functional modulators, opening new avenues for the large-scale production and application of biologically active IGF1.

Earlier studies reported an EC_{50} value of 21 pM for the action of IGF1 on MCF-7, a breast cancer cell line [107]. Commercial forms of IGF1 have shown EC_{50} values ranging between 39-196 pM in MCF-7 cell proliferation assays. In contrast, our engineered fusion proteins, H-MBP-IGF1 and H-PDIb′a′-IGF1, displayed EC_{50} values of 13.7 ± 3.2 nM and 8.6 ± 2.8 nM, respectively. Considering the sizable tags (~36 kDa) at the N-terminus of these IGF1 constructs, their biological activities are remarkably robust. This resonates with previous work that used long R (3) IGF1—a

fusion protein incorporating 11 residues of growth hormone at the N-terminus of IGF1—to stimulate MCF-7 cell proliferation at a median concentration of 0.5 μg/ml [184].

Structural analyses further support the efficacy of these fusion tags. Cryo-EM studies of the active IGF1R-IGF1 complex have shown that both the N-terminus and C-terminus of IGF1 are oriented away from the IGF1 receptor, potentially allowing for the incorporation of tag proteins without disrupting functionality (Figure 37B). A comparative assessment of the atomic structures of PDIb'a' and MBP suggests that PDIb'a' may be a more suitable fusion partner for IGF1 (Figure 37C). Several reasons underscore this preference: first, PDIb'a' comprises 227 residues, while MBP contains 370 residues, indicating that PDIb'a' would likely exert less steric hindrance compared to MBP. Second, the elongated shape of PDIb'a' contrasts with the globular configuration of MBP [185]. This structural difference could make PDIb'a' a less disruptive partner for IGF1 when binding to its receptor. Third, the PDIb'a' tag consists of b' and a' domains connected by a linker, which could impart added flexibility, further facilitating the interaction between IGF1 and its receptor. Collectively, these factors make PDIb'a' a potentially superior choice as a fusion tag for IGF1, both in terms of structural compatibility and biological activity.

Despite MBP's similar capacity to assist protein folding via hydrophobic patches, PDIb'a' offers an added enzymatic ability for disulfide bond formation, a challenging task in the inherently reducing cytoplasmic milieu of both prokaryotic and eukaryotic cells. While previous studies indicated enhanced solubility using Origami 2 *E. coli* strain [104, 105], this was not consistently observed [106]. Neither Origami 2 nor SHuffle strains significantly improved the expression or solubility of H-MBP-IGF1 and H-PDIb′a′-IGF1 in this study.

Endotoxin contamination is a recurring issue when purifying proteins from Gram-negative bacteria like *E. coli*. To mitigate this, we employed ClearColi, a strain genetically modified to produce non-immunogenic lipopolysaccharides (LPS) [133]. Remarkably, H-PDIb′a′-IGF1, which contains four disulfide bonds, was successfully expressed, and purified from ClearColi while maintaining biological activity.

IV. Conclusion

E. coli stands out as the primary expression system for recombinant protein production, owing to its rapid growth rate, facile genetic manipulation, and ability to attain high cell densities utilizing economical culture media. However, persistent challenges include the formation of inclusion bodies (IBs), dense protein aggregates, and the absence of translational modifications, predisposing to protein misfolding. Hence, this thesis endeavors to mitigate these obstacles by implementing diverse strategies to augment the cytoplasmic production of foreign proteins in *E. coli*.

In this study, three critical parameters were systematically investigated: the selection of fusion tags, the induction temperature, and the specific *E. coli* host strain. These parameters were meticulously optimized to characterize the soluble expression of five distinct human proteins, encompassing three cytokines (GM-CSF, IL3, and IL6) and two growth factors (KGF, IGF1) within the cytoplasm. Results demonstrated that employing PDIb′a′ tags at mild culture temperatures such as 30°C or 18°C in the ClearColi *E. coli* markedly enhances both the expression and solubility of these proteins. Despite significant advancements, challenges endure in attaining optimal yields for industrial-scale production. Specifically, while ClearColi offers advantages in low endotoxin levels, its comparatively low growth rate limits large-scale manufacturing. Furthermore, certain human proteins, such as KGF as identified in this study, are recognized as toxic to standard *E. coli* strains. Addressing this, specific modified E. coli strains like BL21 AI are necessitated for their purification. Therefore, selecting E. coli strains is necessary depending on the protein of interest and purpose.

Moreover, the purified proteins exhibited comparable biological activity to commercially available counterparts, signifying a significant advancement in recombinant protein synthesis. This achievement not only enriches the repertoire of proteins available for basic research but also holds promise for therapeutic applications.

Moreover, our experimental findings unveiled the mitogenic activity exhibited by the PDIb'a' proteins across diverse cell lines, illuminating promising future research directions. Notably, this study represents the inaugural documentation of the

mitogenic potential attributed to PDIb'a' and PDI as a whole. In the future, it is imperative to delineate the precise involvement of PDIb'a' in cellular proliferation.

Tables

Table 6. Expression and solubility levels of fusion proteins GM-CSF.

		Expression level (%) (Mean ± SEM)			Solubility (%) (Mean ± SEM)		
Strains	**Tag**	**37°C**	**30°C**	**18°C**	**37°C**	**30°C**	**18°C**
BL21	His	18 ± 2	17 ± 3	3 ± 1	23 ± 1	16 ± 2	80 ± 3
	MBP	2 ± 4	20 ± 5	7 ± 1	31 ± 5	43 ± 4	88 ± 2
	PDIb'a'	17 ± 1	14 ± 3	4 ± 1	27 ± 5	85 ± 4	78 ± 9
ClearColi	His	66 ± 9	61 ± 6	9 ± 1	9 ± 2	8 ± 2	84 ± 5
	MBP	30 ± 3	33 ± 2	9 ± 2	25 ± 1	64 ± 15	85 ± 8
	PDIb'a'	24 ± 6	27 ± 7	13 ± 4	56 ± 2	95 ± 2	87 ± 6
LOBSTR	His	59 ± 3	39 ± 4	9.3 ± 0.4	11 ± 5	9 ± 1	68 ± 2
	MBP	22 ± 1	15 ± 3	7 ± 1	31 ± 1	68 ± 13	87 ± 7
	PDIb'a'	17 ± 4	14 ± 3	4 ± 1	46 ± 2	90 ± 8	54 ± 11
Origami 2	His	57 ± 6	50 ± 4	26 ± 5	3 ± 1	3 ± 1	19 ± 4
	MBP	25.5 ± 0.3	17 ± 1	17 ± 2	33 ± 3	95 ± 2	93 ± 2
	PDIb'a'	42 ± 11	33 ± 11	22 ± 3	36 ± 7	94 ± 3	95 ± 3
SHuffle T7	His	50 ± 3	39 ± 6	15±2	6 ± 2	13 ± 1	63 ± 1
	MBP	24 ± 7	15 ± 2	6 ± 1	24 ± 5	62 ± 9	94 ± 3
	PDIb'a'	15 ± 4	18.2 ± 0.2	18 ± 2	28 ± 2	97 ± 1	97 ± 1

Table 7. Purity and yield of GM-CSF obtained from a 0.5 L flask culture.

Purification step	Total protein (mg)	Purity (%)	Target protein (mg)	Yield (%)
Supernatant	118.8	23	27.3	100
IMAC	13.3	38	5.05	18.5
IEC	0.7	98	0.68	13.5

Table 8. Expression and solubility level of fusion proteins IL3.

		Expression level (%) (Mean ± SEM)			Solubility (%) (Mean ± SEM)		
Strains	**Tag**	**37°C**	**30°C**	**18°C**	**37°C**	**30°C**	**18°C**
BL21	MBP	32 ± 2	21 ± 3	19 ± 2	46 ± 5	82 ± 11	90 ± 5
	PDIb'a'	29 ± 4	30 ± 2	34 ± 8	53 ± 7	88 ± 5	98.2 ± 0.3
ClearColi	MBP	43 ± 3	33 ± 3	23 ± 2	62 ± 15	70 ± 6	85 ± 9
	PDIb'a'	33 ± 10	43 ± 7	34 ± 5	78 ± 11	97 ± 3	98.1 ± 0.5
LOBSTR	MBP	26 ± 6	24 ± 4	21 ± 4	46 ± 2	65 ± 2	70 ± 8
	PDIb'a'	13 ± 2	12 ± 2	21 ± 3	45 ± 11	67 ± 8	82 ± 6
Origami 2	MBP	38 ± 11	40 ± 12	36 ± 9	61 ± 11	70 ± 8	84 ± 9
	PDIb'a'	16 ± 2	23 ± 2	37 ± 4	71 ± 11	91 ± 7	97 ± 2
Rosetta-gami B	MBP	24 ± 5	24 ± 3	21 ± 4	51 ± 15	80 ± 8	63 ± 9
	PDIb'a'	14 ± 3	22 ± 4	27 ± 8	45 ± 13	93 ± 3	87 ± 7
SHuffle T7	MBP	28 ± 5	37 ± 6	35 ± 7	64 ± 5	95 ± 2	94 ± 3
	PDIb'a'	10 ± 2	25 ± 4	28 ± 8	83 ± 7	97 ± 1	98.6 ± 0.8

Table 9. Production of H-PDIb'a'-IL3 in 1 L flask cultures.

Parameters	
Endpoint OD_{600}	1.31 ± 0.06
Wet cell weight (g)	1.79 ± 0.24
Expression level (%)	22 ± 1
Solubility level (%)	46 ± 2

Table 10. Purity and yield of purified IL3 obtained from a 0.5 L flask culture of H-MBP-IL3

Purification step	Total protein (mg)	Purity (%)	Target protein (mg)	Yield (%)
Supernatant	215	28.9	62.4	100
AC	7.06	24.7	1.7	2.7
IEC	0.02	99	0.02	0.009

Table 11. Purity and yield of purified IL3 obtained from 0.5 L flask cultures of H-PDIb'a'-IL3.

Purification step	Total protein (mg)	Purity (%)	Target protein (mg)	Yield (%)
Supernatant	81.4 ± 14.6	42.4 ± 8.9	33.8 ± 4	100
IMAC	17.8 ± 4.1	30.5 ± 9.1	5.3 ± 1.3	15.9 ± 0.05
IEC	0.38 ± 0.02	95.8 ± 3.6	0.36 ± 0.02	6.6 ± 1.4

Table 12. Parameters governing protein-induced TF-1 cell proliferation.

	Purified IL3	Commercial IL3	H-PDIb′a′-IL3
EC_{50} (1^{st} stimulation)	0.07 ± 0.01 pM	0.18 ± 0.06 pM	195 ± 82 pM
Hill coefficient (1^{st} stimulation)	1.6 ± 0.2	1.17 ± 0.18	1.09 ± 0.1
Maximum response (1^{st} stimulation)	160 ± 18%	132 ± 5%	165 ± 4%
EC_{50} (2^{nd} stimulation)	151 ± 25 nM	262 ± 16 nM	396 ± 47 nM
Hill coefficient (2^{nd} stimulation)	1.6 ± 0.1	1.75 ± 0.35	1.9 ± 0.1
Maximum response (2^{nd} stimulation)	345 ± 14%	419 ± 31%	345 ± 19%

Table 13. Expression and solubility levels of fusion proteins IL6.

		Expression level (%) (Mean ± SEM)			Solubility (%) (Mean ± SEM)		
Strains	**Tag**	**37°C**	**30°C**	**18°C**	**37°C**	**30°C**	**18°C**
BL21	MBP	63.8 ± 0.3	74 ± 4	65 ± 4	86 ± 2	85 ± 5	78 ± 4
	PDIb'a'	82 ± 2	79 ± 4	79 ± 3	72 ± 9	96 ± 1	97 ± 1
ClearColi	MBP	80 ± 3	75 ± 2	66 ± 4	65 ± 3	73 ± 3	72 ± 3
	PDIb'a'	77 ± 6	81 ± 5	83 ± 4	89 ± 7	97 ± 1	96 ± 2
Origami 2	MBP	76 ± 4	77 ± 4	67 ± 1	93 ± 3	91 ± 3	92 ± 1
	PDIb'a'	66 ± 7	76 ± 6	77 ± 1	90 ± 5	97 ± 1	97 ± 1
SHuffle T7	MBP	66 ± 2	70 ± 2	65 ± 2	96 ± 1	93 ± 1	96 ± 1
	PDIb'a'	61 ± 4	67 ± 4	72 ± 4	96 ± 1	95 ± 1	97.9 ± 0.4

Table 14. Purity and yield of H-PDIb'a'-IL6 obtained from 0.5 L flask cultures.

Purification step	Total (mg)	Purity (%)	Target protein (mg)	Yield (%)
Supernatant	137 ± 8	55 ± 3	72 ± 5	100
IMAC	18.3 ± 0.5	56 ± 3	10.3 ± 0.8	14.2 ± 1.1
IEC	0.31 ± 0.04	93 ± 2	0.29 ± 0.01	0.44 ± 0.06

Table 15. Purity and yield of purified IL6 obtained from 0.5 L flask cultures.

Purification step	Total (mg)	Purity (%)	Target protein (mg)	Yield (%)
Supernatant	150 ± 9	57 ± 5	84 ± 3	100
IMAC	20.1 ± 1.3	54 ± 2	11.1 ± 0.2	12.1 ± 0.6
IEC	0.39 ± 0.03	97.9 ± 0.2	0.38 ± 0.03	0.38 ± 0.02

Table 16. Expression and solubility levels of fusion proteins KGF.

		Expression level (%) (Mean ± SEM)			Solubility (%) (Mean ± SEM)		
Strains	**Tag**	**37°C**	**30°C**	**18°C**	**37°C**	**30°C**	**18°C**
BL21 pLysS	MBP	3 ± 1	1 ± 0.4	0	75 ± 12	98.7 ± 0.7	0
	PDIb'a'	4.4 ± 0.1	5.1 ± 0.4	5.0 ± 0.3	87 ± 2	98.7 ± 0.7	99.3 ± 0.6
BL21 AI	MBP	10 ± 2	22 ± 1	12 ± 1	75 ± 6	96 ± 2	99.4 ± 0.1
	PDIb'a'	20 ± 2	24 ± 1	19 ± 2	26 ± 5	76 ± 10	99 ± 0.2
ClearColi pLysS	MBP	24 ± 5	4 ± 1	8 ± 3	51 ± 8	97.8 ± 0.4	98.9 ± 0.4
	PDIb'a'	23 ± 5	9 ± 2	5.8 ± 0.1	17 ± 5	98.4 ± 0.3	99.4 ± 0.3
LOBSTR	MBP	20 ± 2	13 ± 2	10 ± 1	40 ± 3	99 ± 1	99.1 ± 0.8
	PDIb'a'	20 ± 1	14 ± 3	7 ± 1	23 ± 7	50 ± 11	98.6 ± 0.6
SHuffle T7 pLysS	MBP	6 ± 1	2 ± 1	1.1 ± 0.7	59 ± 5	98.5 ± 0.6	99.4 ± 0.3
	PDIb'a'	10 ± 2	9 ± 1	5 ± 1	44 ± 12	98.3 ± 0.9	99 ± 0.4

Table 17. Purity and yield of H-PDIb'a'-KGF obtained from 0.5 L flask cultures.

Purification step	Total (mg)	Purity (%)	Target protein (mg)	Yield (%)
Supernatant	162 ± 9	20 ± 1	32.8 ± 0.4	100 ± 0
IMAC	8.5 ± 0.4	56 ± 3	4.74 ± 0.04	14.6 ± 0.1
IEC	0.85 ± 0.01	92.1 ± 0.3	0.79 ± 0.01	2.43 ± 0.03

Table 18. Purity and yield of purified KGF obtained from 0.5 L flask cultures.

Purification step	Total (mg)	Purity (%)	Target protein (mg)	Yield (%)
Supernatant	124 ± 7	18 ± 1	23 ± 2	100
IMAC	11 ± 1	58 ± 3	6.1 ± 0.5	24 ± 2
IEC	1.3 ± 0.2	96.9 ± 0.4	1.3 ± 0.2	5.0 ± 0.7

Table 19. Melting temperatures of KGF and related proteins in PBS buffer.

Protein	Melting Temperature, °C
KGF	59.2 ± 0.3
H-PDIb′a′-KGF	43.7 ± 0.2
PDIb'a'	45.7 ± 0.2

Table 20. Wound closure capacity of purified KGF.

Protein	Wound closure,%
Control	15 ± 3
EGF 80 nM	57 ± 3
KGF 1 nM	32 ± 1
KGF 10 nM	41 ± 2
KGF 100 nM	61 ± 9

Table 21. Expression and solubility levels of fusion proteins IGF1.

		Expression level (%) (Mean ± SEM)			Solubility (%) (Mean ± SEM)		
Strains	**Tag**	**37°C**	**30°C**	**18°C**	**37°C**	**30°C**	**18°C**
BL21	His	5 ± 1	6.4 ± 0.5	5 ± 1	77 ± 3	65 ± 15	83 ± 1
	MBP	8 ± 2	15 ± 4	26 ± 5	84 ± 2	86 ± 5	87 ± 5
	PDIb'a'	58 ± 5	54 ± 4	45 ± 6	55 ± 4	86 ± 7	91 ± 7
ClearColi	His	5 ± 1	4.2 ± 0.3	3.7 ±0.6	53 ± 16	55 ± 5	75 ± 9
	MBP	42 ± 2	33 ± 3	30 ± 1	51 ± 2	58 ± 8	86 ± 10
	PDIb'a'	57 ± 9	48 ± 4	42 ± 2	46 ± 12	88 ± 7	89 ± 7
Origami 2	His	9 ± 2	9 ± 3	6.7 ± 0.3	67 ± 5	78 ± 5	78 ± 6
	MBP	65 ± 8	51 ± 7	52 ± 3	75 ± 4	66 ± 3	71 ± 8
	PDIb'a'	65 ± 12	64 ± 5	69 ± 7	22 ± 10	88 ± 5	89 ± 3
SHuffle T7	His	8 ± 1	8 ± 1	18 ± 4	11 ± 2	14 ± 1	80 ± 5
	MBP	40 ± 3	47 ± 2	46 ± 2	90 ± 2	74 ± 11	92 ± 2
	PDIb'a'	65 ± 4	67 ± 5	67 ± 7	35 ± 5	85 ± 7	85 ± 7

Table 22. Purity and yield of fusion proteins IGF1 obtained from 0.5 L flask cultures.

Structure	Purification step	Total (mg)	Purity (%)	Target protein (mg)	Yield (%)
H-PDIb′a′-IGF1	Supernatant	101 ± 2	46 ± 5	45.8 ± 4.5	100
	IMAC	9.3 ± 0.4	65 ± 3	6.1 ± 0.3	13.4 ± 1.3
	IEC	0.65 ± 0.05	85.7 ± 1.5	0.55 ± 0.04	1.3 ± 0.2
H-MBP- IGF1	Supernatant	113.5 ± 0.4	31 ± 2	36 ± 3	100
	MBP	9 ± 1	94 ± 1	8 ± 1	23 ± 3

Table 23. Purity and yield of purified IGF1 obtained from 0.5 L flask cultures.

Purification step	Total (mg)	Purity (%)	IGF1 (mg)	Yield (%)
Supernatant	99 ± 2	53 ± 4	53 ± 3	100
IMAC	11 ± 1	54 ± 1	5.9 ± 0.7	11 ± 1
IEC	0.13 ± 0.02	91.4 ± 0.9	0.11 ± 0.01	0.2

Table 24. Parameter governing protein-induced MCF-7 cell proliferation.

Protein	EC_{50} (nM)	Hill coefficient	Max response
Commercial IGF1	1.1 ± 0.4	0.5 ± 0.1	329 ± 11
Purified IGF1	0.5 ± 0.1	0.34 ± 0.01	305 ± 16
H-MBP-IGF1	13.7 ± 3.2	0.8 ± 0.1	338 ± 16
H-PDIb′a′-IGF1	8.6 ± 2.8	0.6 ± 0.1	388 ± 33
H-PDIb′a′	470 ± 41	1.0 ± 0.3	486 ± 14

Figures

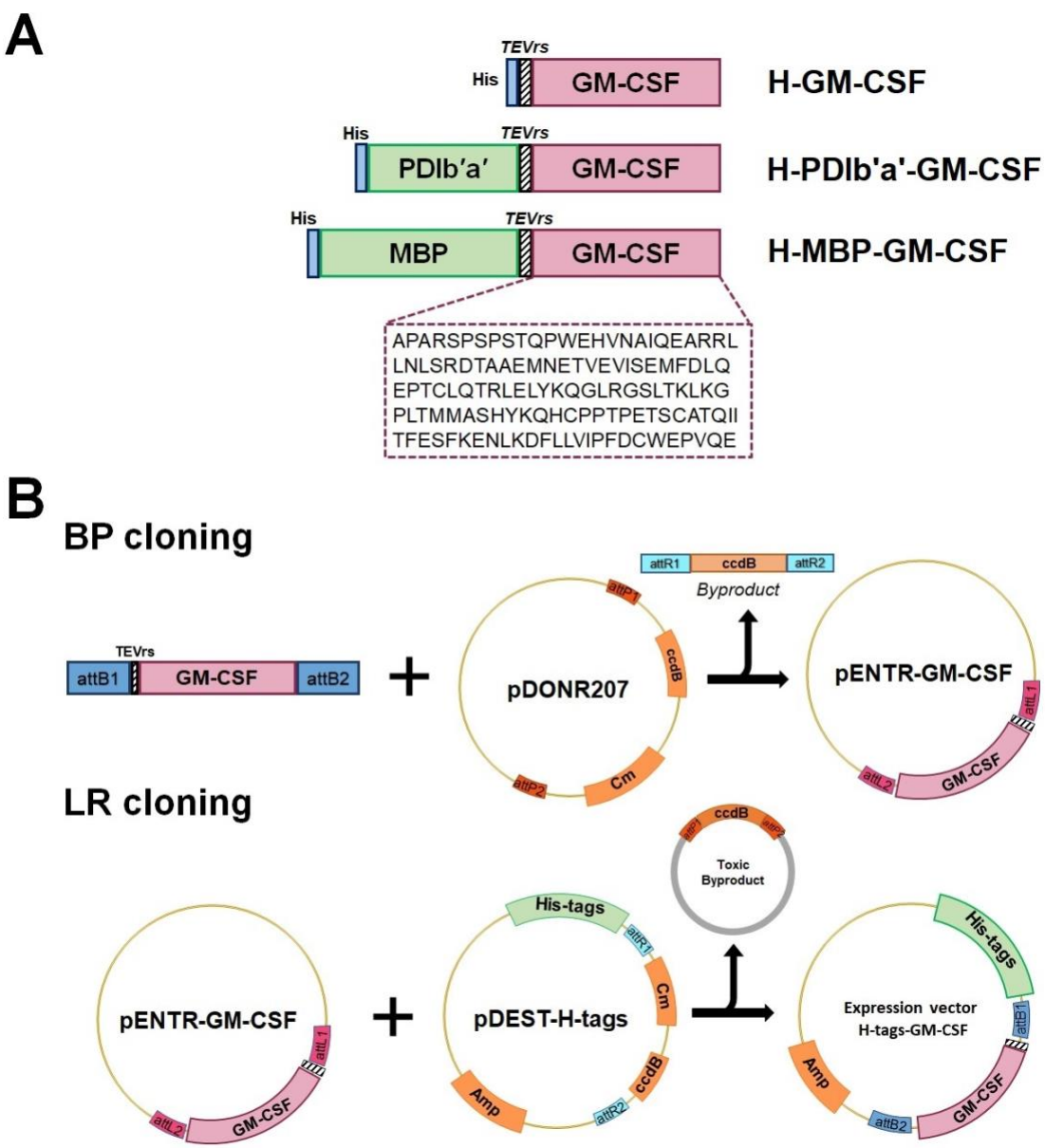
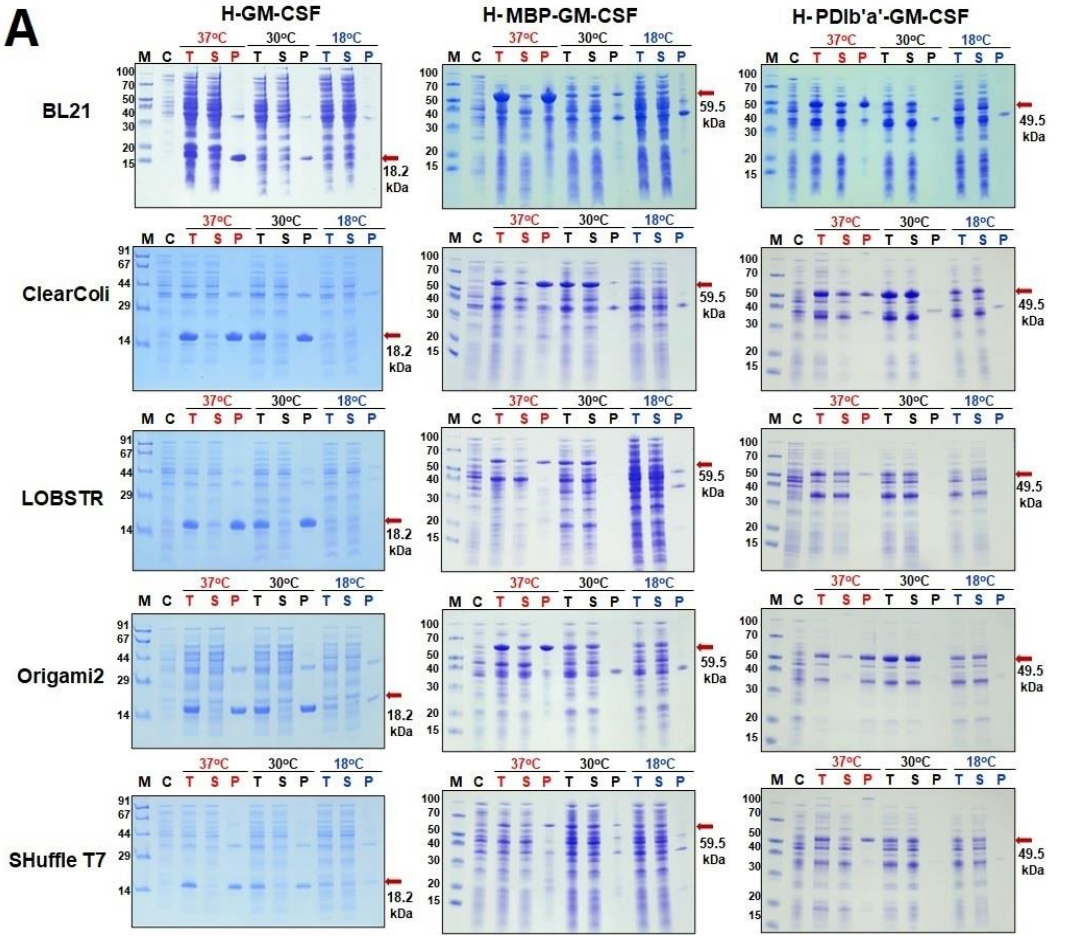


Figure 1. Constructs of the GM-CSF expression vectors.

(A) Structure diagram of H-GM-CSF, H-PDIb′a′-GM-CSF, H-MBP-GM-CSF construct with GM-CSF (spanning amino acid 18-144).

(B) Schematic of Gateway cloning strategy to generate GM-CSF expression vectors.



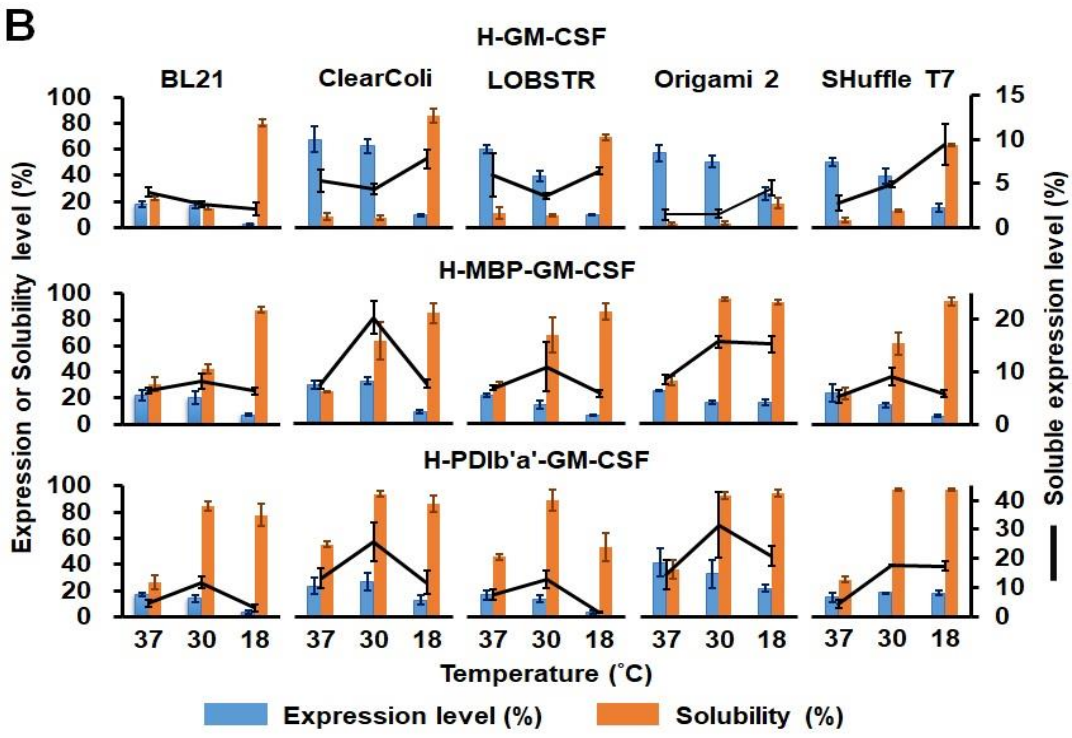


Figure 2. Expression and solubility analysis of fusion proteins GM-CSF across different E. coli strains.

(A) Expression of the proteins was induced by IPTG to the bacterial cultures at 18°C, 30°C, or 37°C using different *E. coli* strains. M, protein molecular weight markers; C, total protein extract of *E. coli* before IPTG induction; T, total protein lysate after IPTG induction; S, supernatant portion of the total lysate after bacterial lysis; P, pellet portion of the total lysate after bacterial lysis. The red arrows indicate the H-GM-CSF (18.2 kDa), H-MBP-GM-CSF (58.5 kDa), and H-PDIb′a′-GM-CSF (49.5 kDa) fusion protein products. The displayed images serve as representative outputs from three separate experimental runs.

(B) Quantitative depiction of fusion protein expression across *E. coli* Strains. The bar plots display expression percentages (in blue) juxtaposed against solubility percentages (in orange) aligned on the primary y-axis. A contrasting black line marks the soluble protein expression, plotted against the secondary y-axis. Expression and solubility metrics were computed utilizing the Gel Analyzer software, rooted in data from three individual experiments.

.

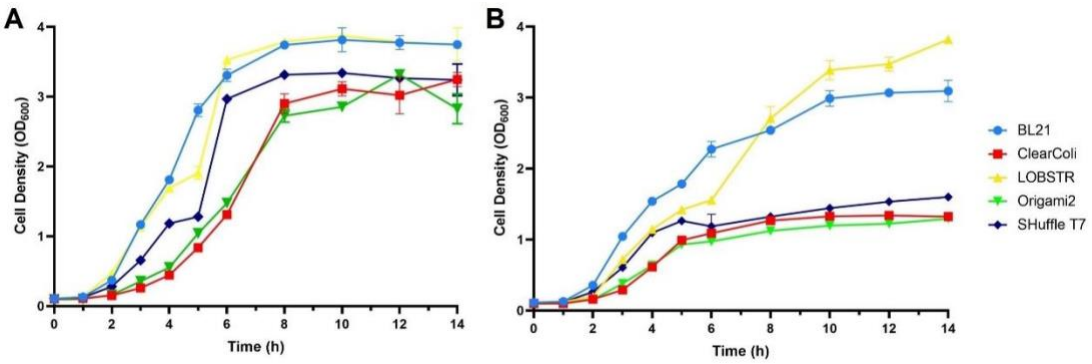


Figure 3. Growth curves of the indicated E. coli strains with and without induction.

(A) *E. coli* strains transformed with the H-PDIb′a′-GM-CSF plasmid were cultured at 30ºC without induction.

(B) IPTG was added to induce the expression of H-PDIb′a′-GM-CSF at 30°C.

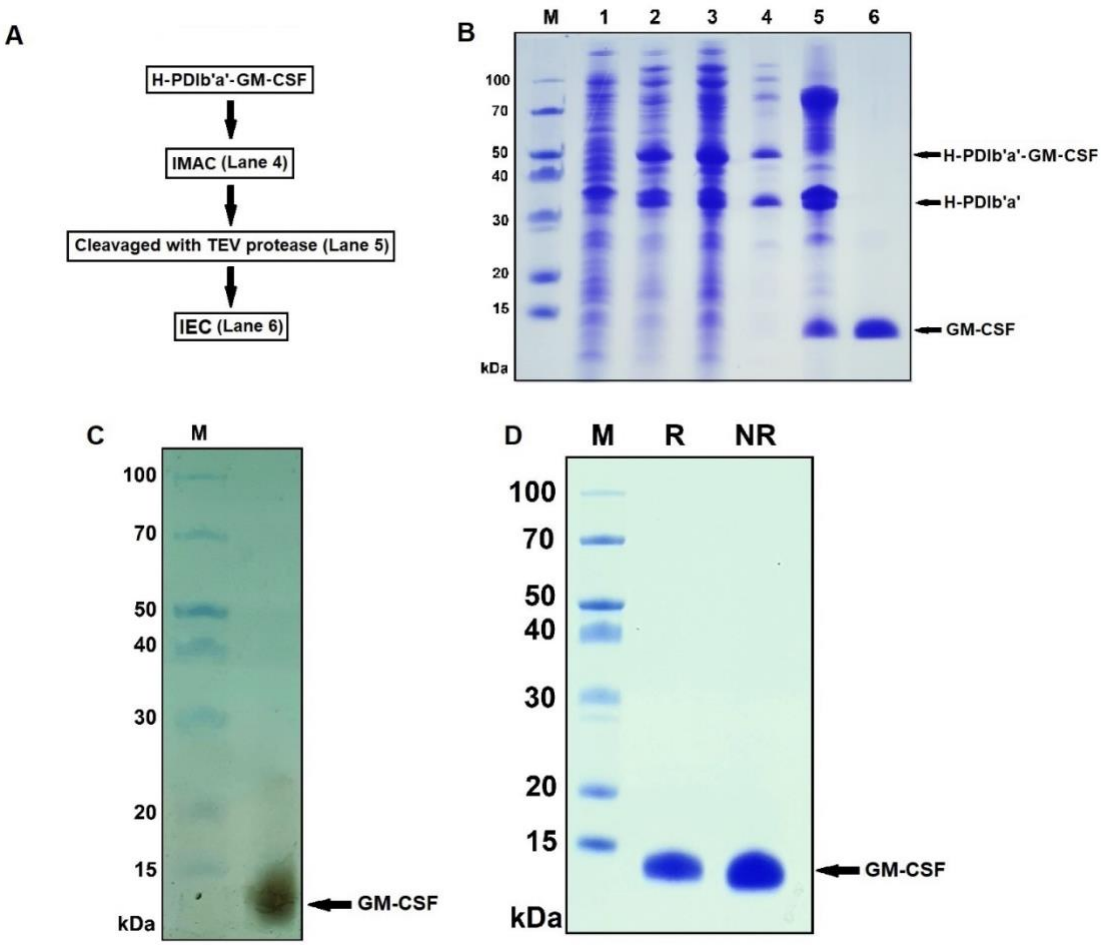


Figure 4. Purification of GM-CSF from ClearColi strain.

(A) Flow chart of the GM-CSF purification steps. (B) GM-CSF was purified by IMAC and anion-exchange chromatography from H-PDIb′a′-GM-CSF produced in ClearColi cultures. M, protein molecular weight marker. Lane 1, total lysate proteins before IPTG induction as a negative control; Lane 2, total proteins induced by IPTG after cell sonication; Lane 3, soluble proteins obtained after cell sonication of the total cell proteins; Lane 4, the H-PDIb′a′-GM-CSF fusion protein (49.5 kDa) purified with an IMAC column; lane 5, the result of TEV protease cleavage (29 kDa) yielding the H-PDIb′a′ (35 kDa) and GM-CSF (14.6 kDa); lane 6, GM-CSF (14.6 kDa) purified using ion-exchange chromatography. (C) Silver staining of GM-CSF under reducing conditions. M, protein molecular weight markers. (D) Purified GM-CSF protein under reducing and non-reducing conditions. M, protein molecular weight marker; R, reducing; NR, non-reducing.

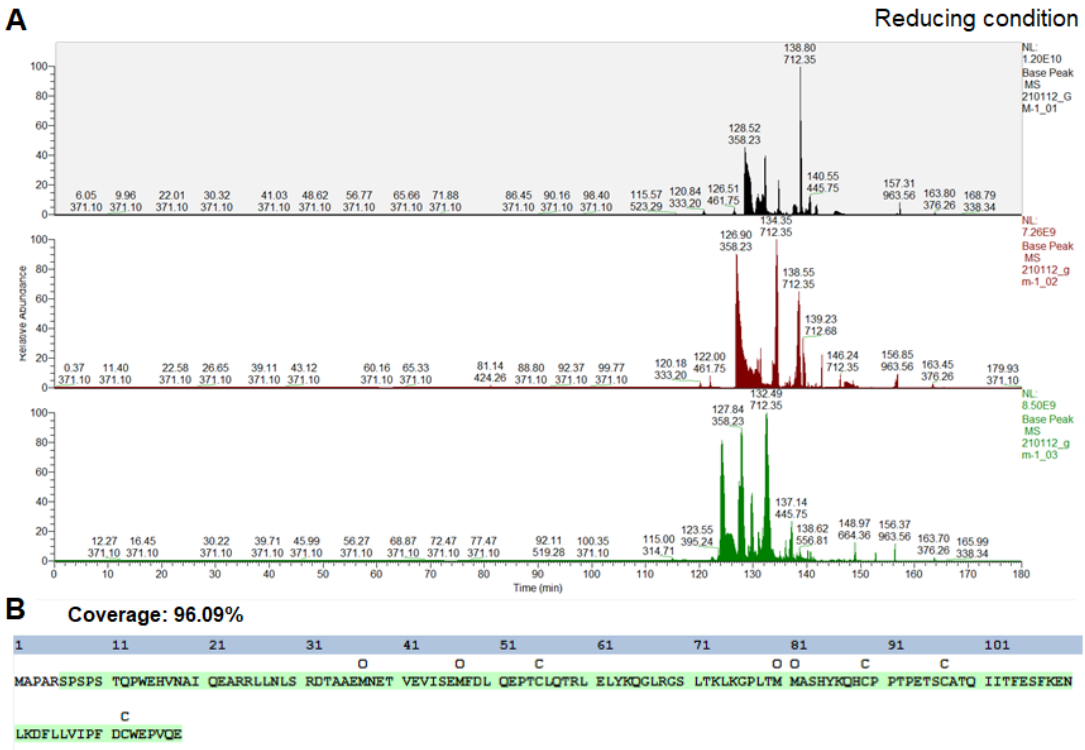


Figure 5. LC-MS/MS analysis of purified GM-CSF.

(A) LC-MS/MS sequencing of peptides for protein identification was performed using trypsin for sample digestion and under reducing conditions. The arrows indicate peaks matching corresponding peptide fragments.

(B) The covered sequence is presented using a color-coded display based on percolator confidence scores (green, high confidence, p-value < 0.01). Coverage was 96.09%

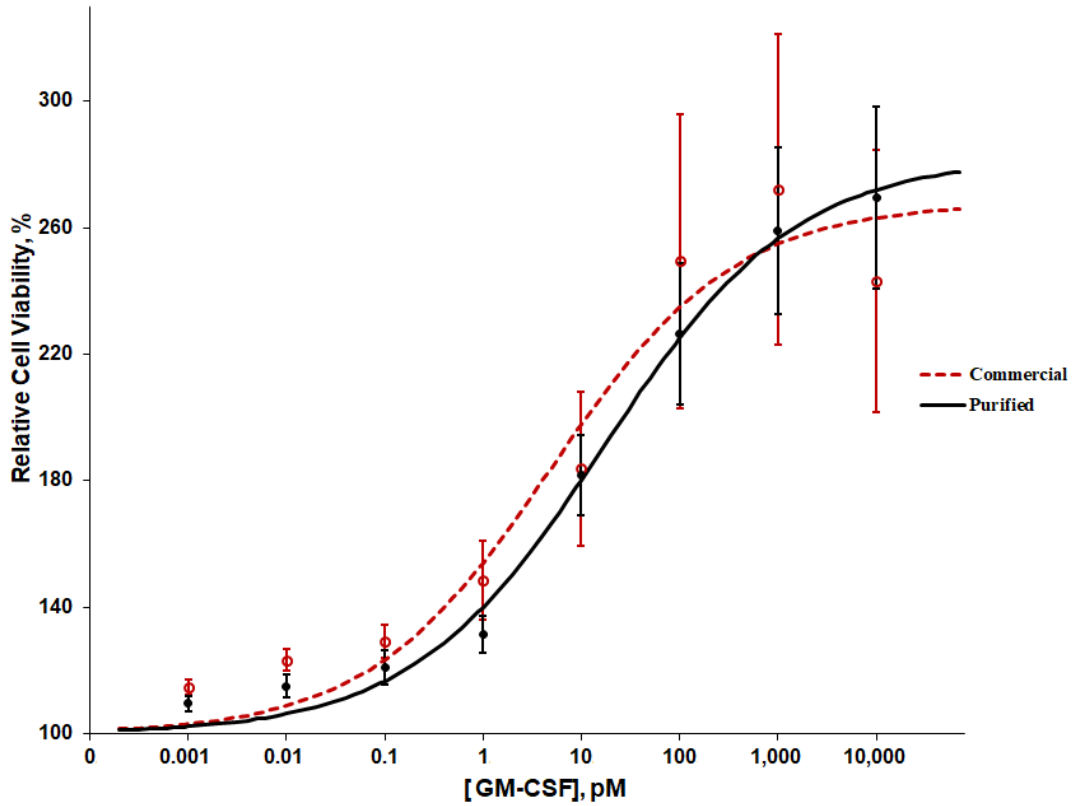


Figure 6. Biological activity of purified GM-CSF.

The biological activities of purified GM-CSF (black solid circle with solid line) and commercial GM-CSF (red open circle with broken line) were tested by their ability to promote the proliferation of TF-1 cells. Cell proliferation was measured after 48 h of treatment using a CCK-8 assay. The EC_{50} is defined as the effective concentration of the growth factor at which cell proliferation reaches 50% maximum growth. The EC_{50} and the Hill coefficient for the purified GM-CSF on TF-1 cells were calculated as 16.4 ± 2 pM and 0.49 ± 0.08 (n = 3), respectively. Some of the standard error bars are hidden by the data symbols.

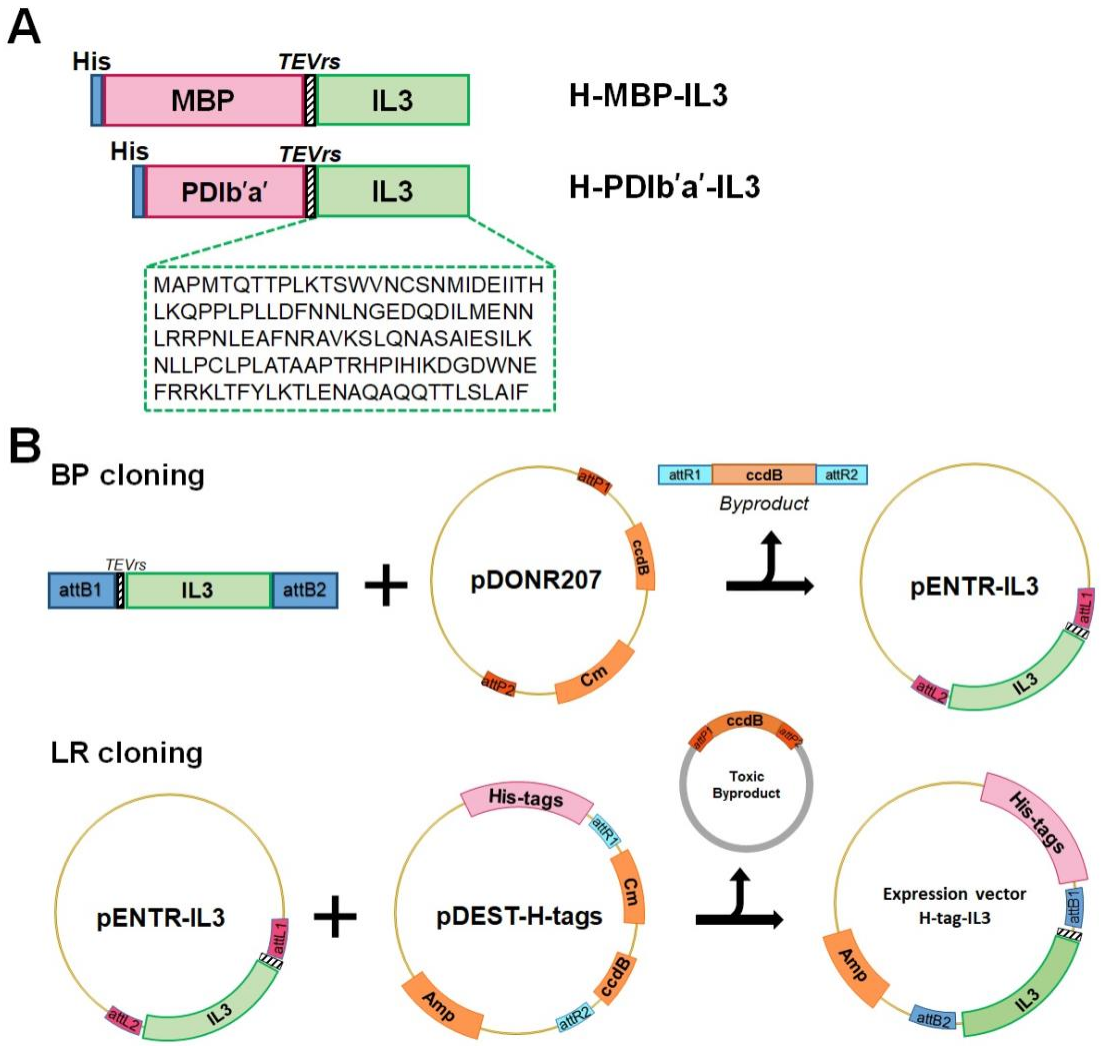
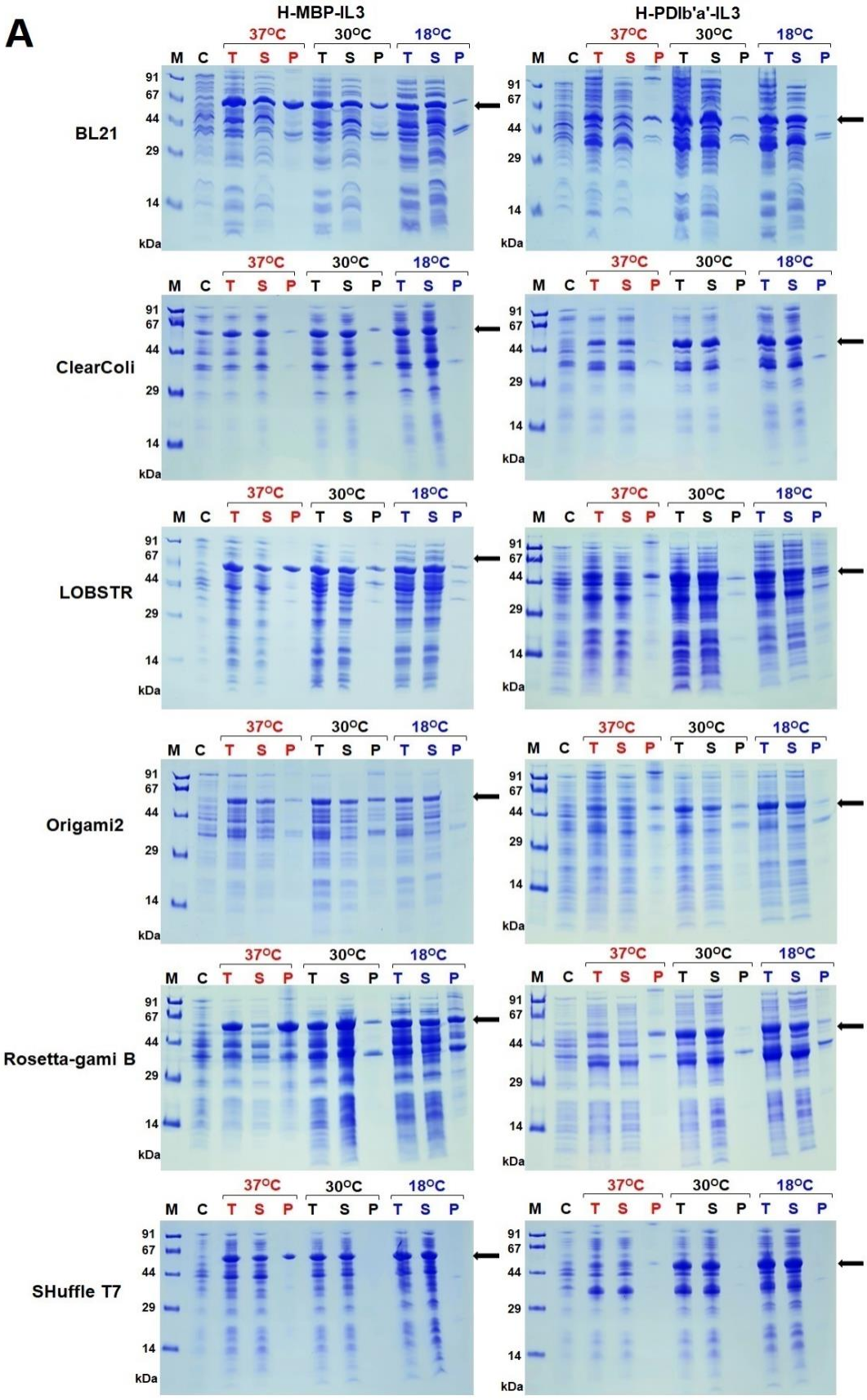


Figure 7. Construct of the IL3 expression vectors.

(A) Structure diagram of H-PDIb′a′-IL3, H-MBP-IL3 construct with IL3 (amino acid 20-152).

(B) Schematic of Gateway cloning strategy to generate IL3 expression vectors



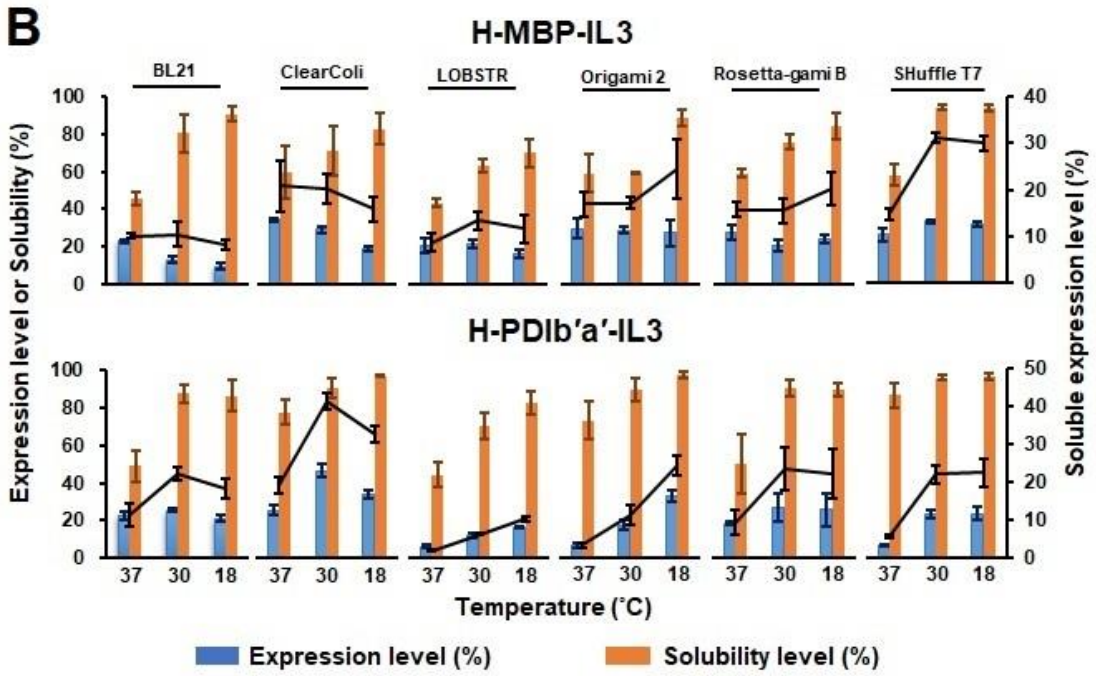


Figure 8. Expression and solubility analysis of fusion proteins IL3 across E. coli strains.

(A) SDS-PAGE analysis of fusion proteins expression. Black arrow, the fusion protein of H-MBP-IL3 (59 kDa), H-PDIb′a′-IL3 (50 kDa); M, protein molecular weight marker; C, control of total protein lysate before induction; T, total protein lysate after IPTG induction; S, soluble fraction; P, pellet fraction of total protein lysate. The images are representative of three independent experiments.

(B) Quantitative analysis of fusion protein expression in *E. coli* strains. The bar graphs show expression percentage (blue bar) and solubility percentage (orange bar) (primary y-axis), and the black lines indicate expression level times solubility (secondary y-axis).

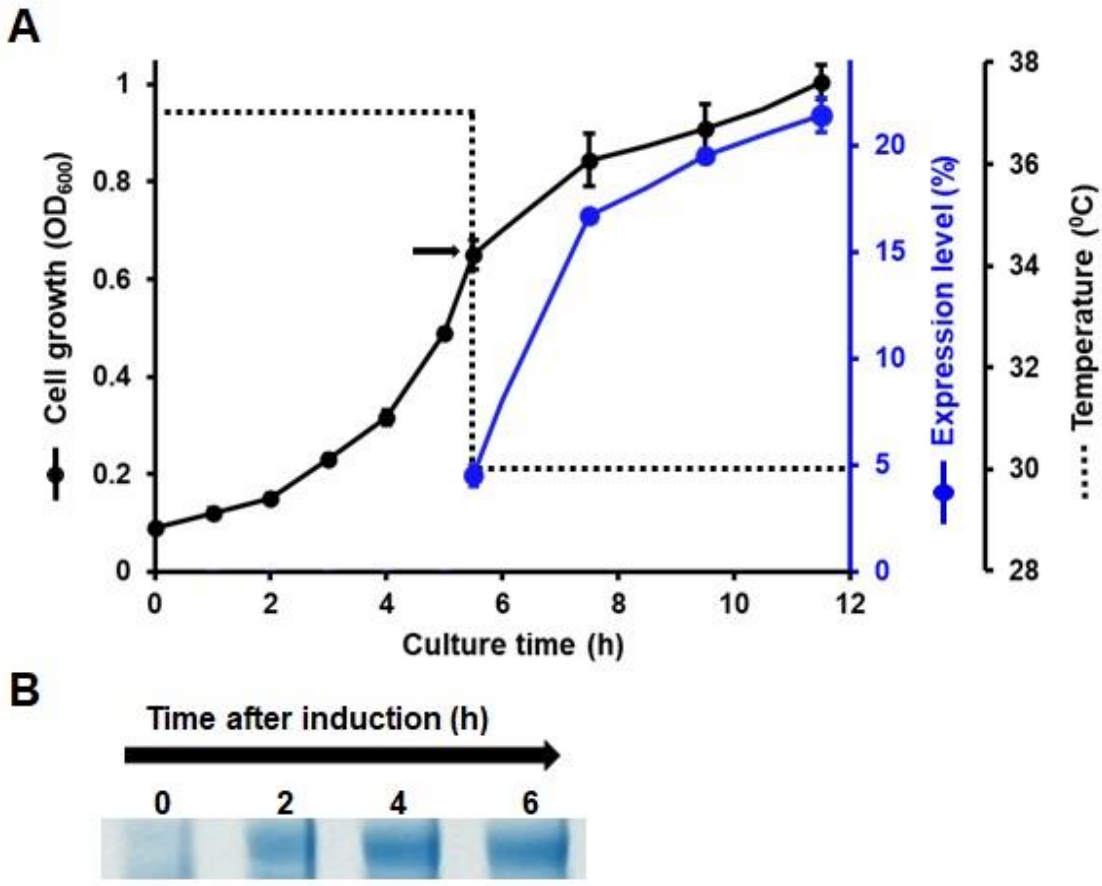


Figure 9. The expression of fusion protein H-PDIb'a'-IL3 in 1 L flask culture.

(A) The expression profile of the H-PDIb′a′-IL3 fusion protein. The x-axis corresponds to the culture time, while the left y-axis portrays the cell growth, as measured by the OD_{600} value, depicted as a black line. Following the induction with 1mM IPTG, the expression levels of the fusion protein are represented by the blue line, plotted against the right y-axis. The dotted line on the secondary right y-axis shows the temperature at which the protein expression occurs. The data points are averages derived from three separate experiments and include the standard error. The induction time is highlighted with a black arrow.

(B) SDS-PAGE images that showcase the expression levels of fusion proteins at different time intervals post-induction.

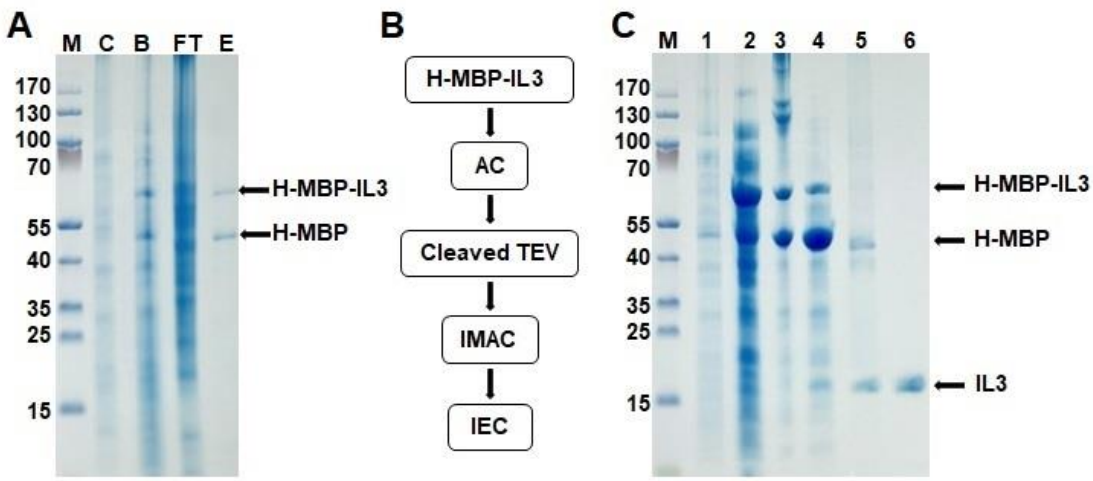


Figure 10. Purification of H-MBP-IL3 using affinity chromatography.

(A) The H-MBP-IL3 fusion protein obtained from the SHuffle strain was purified through immobilized metal affinity chromatography (IMAC).

(B) Flow chart illustrating the sequential purification steps for obtaining IL3 from the H-MBP-IL3 construct.

(C) IL3 was purified from H-MBP-IL3 produced in SHuffle using affinity chromatography and ion exchange chromatography (IEC). M, PageRuler Prestained Protein ladder (Fermentas, #SM0671); Lane 1, total cleavage proteins before IPTG induction (negative control); Lane 2, soluble fraction of the total cell proteins after cell sonication; Lane 3, H-MBP-IL3 fusion protein (59 kDa) purified using amylose column (AC); Lane 4, the result of TEV protease cleavage; Lane 5, IL3 purified using IMAC; Lane 6, IL3 was purified with IEC.

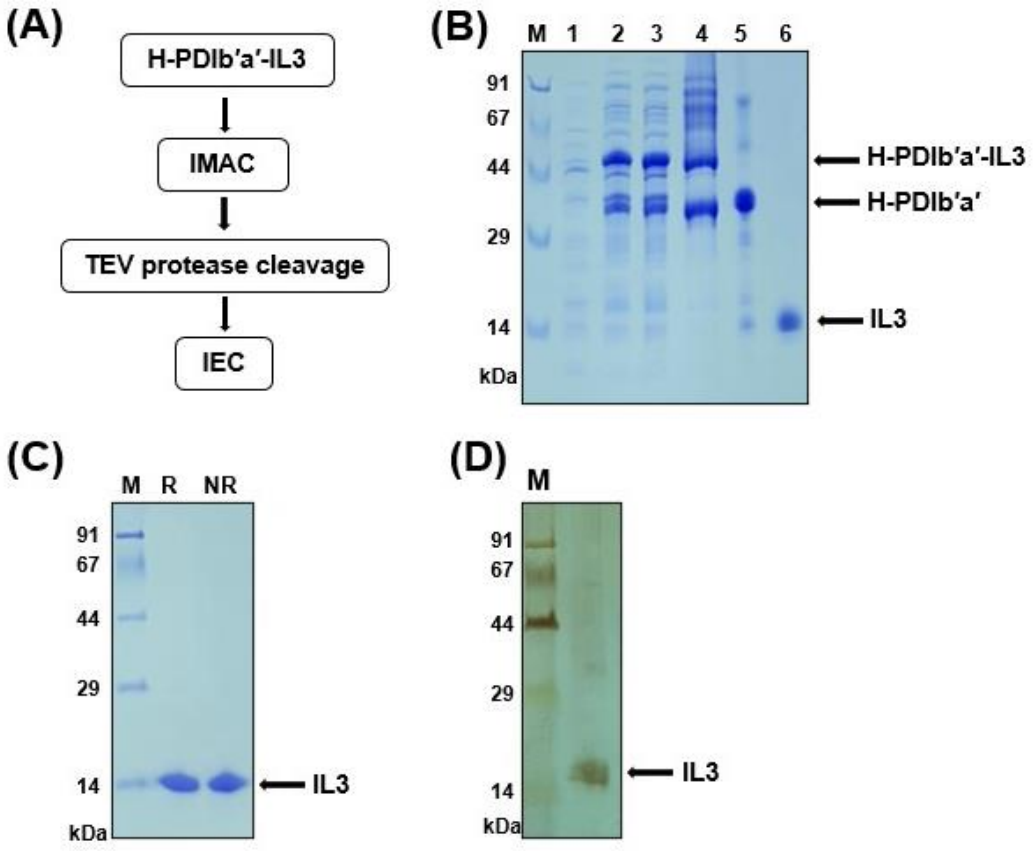


Figure 11. Purification of IL3 in ClearColi strain.

(A) Flow chart of the purification steps.

(B) IL3 was purified from H-PDIb′a′-IL3 produced in ClearColi by IMAC and anion-exchange chromatography. M, protein molecular weight marker; lane 1, total lysate proteins before IPTG induction as negative control; lane 2, Total proteins induced by IPTG after cell sonication; lane 3, soluble fraction after cell sonication from total cell proteins; lane 4, H-PDIb′a′-IL3 fusion protein (50 kDa) purified with IMAC column; lane 5, the result of TEV protease cleavage (29 kDa): H-PDIb'a' (35 kDa) and IL3 (15 kDa); lane 6, IL3 (15 kDa) was purified with ion-exchange chromatography.

(C) Purified IL3 protein under reducing and non-reducing conditions with M, protein molecular weight marker; R, reducing condition (with DTT); NR, non-reducing condition (without DTT).

(D) Silver staining of IL3 under reducing conditions with M, protein molecular weight marker.

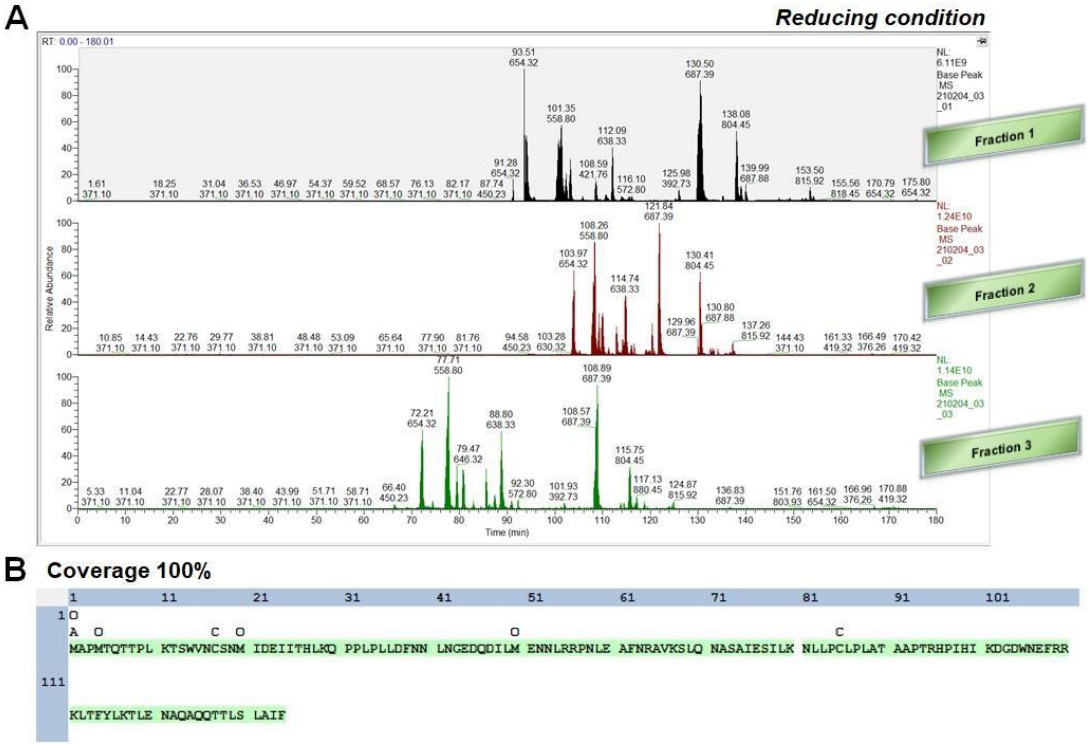


Figure 12. LC-MS/MS analysis of purified IL3.

(A) LC-MS/MS sequencing of peptides for protein identification was performed using trypsin enzyme for sample digestion. The LC-MS/MS sequencing was performed under reducing conditions. The percent coverage for protein identification compared with the whole protein sequence.

(B) IL3 sequence is presented by displaying color-coded based on percolator confidence scores (green, high confidence, p-value < 0.01). Coverage was 100%.

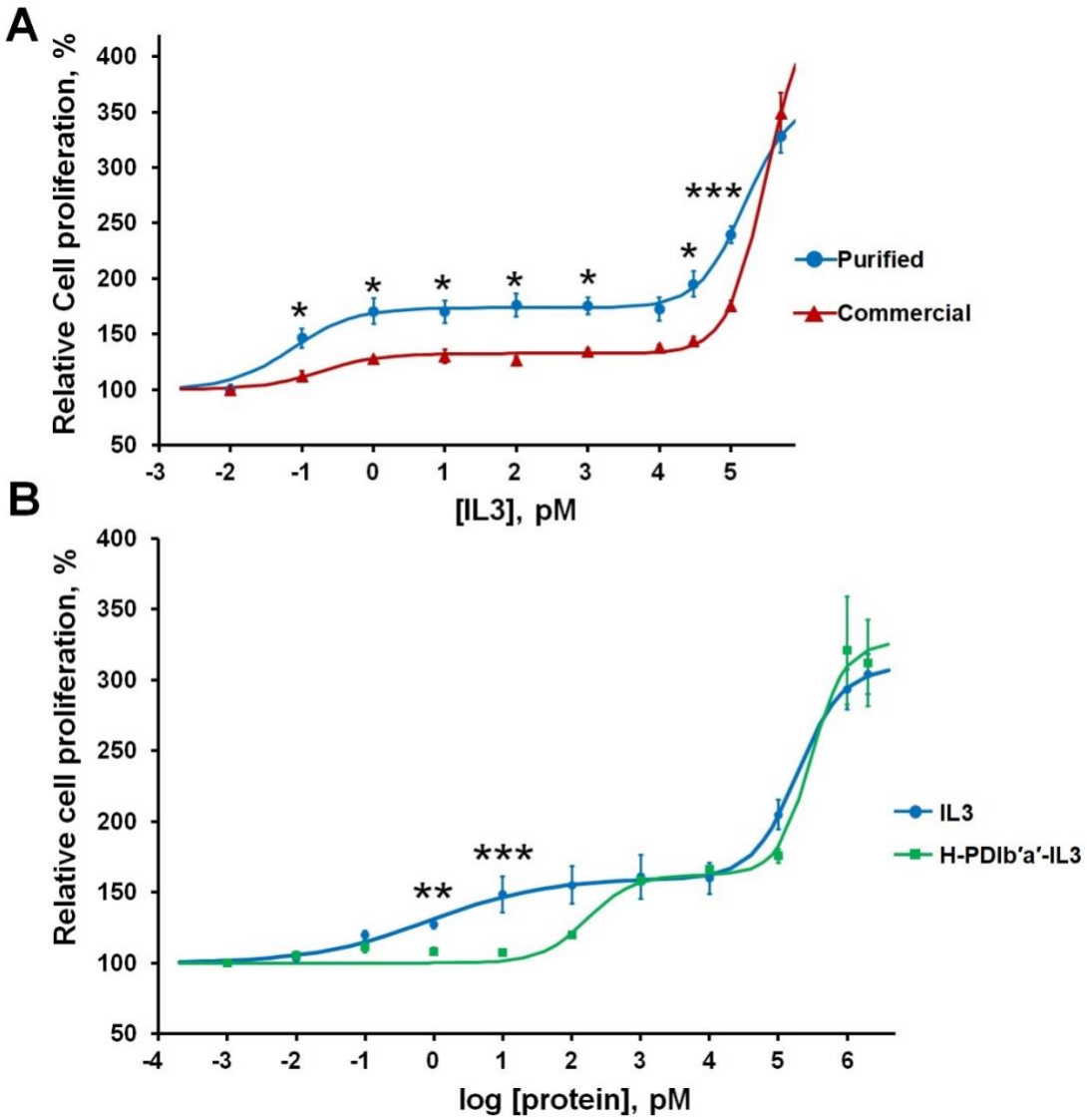


Figure 13. Comparative dose-response effects of IL3 variants on TF-1 cells.

(A) Dose-response curve for TF-1 Cell Proliferation in response to varying concentrations of purified and commercial IL3. A dose-response curve was generated to illustrate the proliferative capacity of TF-1 human erythroleukemic cells in response to varying concentrations (0 to 500 nM) of both purified and commercial IL3. The baseline proliferation level in the absence of IL3 was set as 100%, and all other values were normalized to this reference point (n = 9). Statistical significance between purified IL3 and commercial IL3 was evaluated using a t-test, with significance levels denoted by asterisks (* $p < 0.05$; ** $p < 0.01$; *** $p < 0.001$; **** $p < 0.0001$).

(B) Dose-response curve for TF-1 cell proliferation in response to purified IL3 and H-PDIb′a′-IL3. A dose-response curve demonstrates the differential effects on TF-1 cell proliferation when exposed to purified IL3 and the fusion protein H-PDIb′a′-IL3 at higher concentrations (0 to 2 μM) over a 48-h treatment. The reference point for proliferation, in the absence of any protein, was set at 100%, and all other values were normalized to this reference point (n = 9). Statistical significance was determined using one-way ANOVA, followed by Bonferroni's post-hoc test. Asterisks indicate significant differences between purified IL3 and H-PDIb′a′-IL3 (* $p < 0.05$; ** $p < 0.01$; *** $p < 0.001$). Some error bars are too small to be representative in the graph.

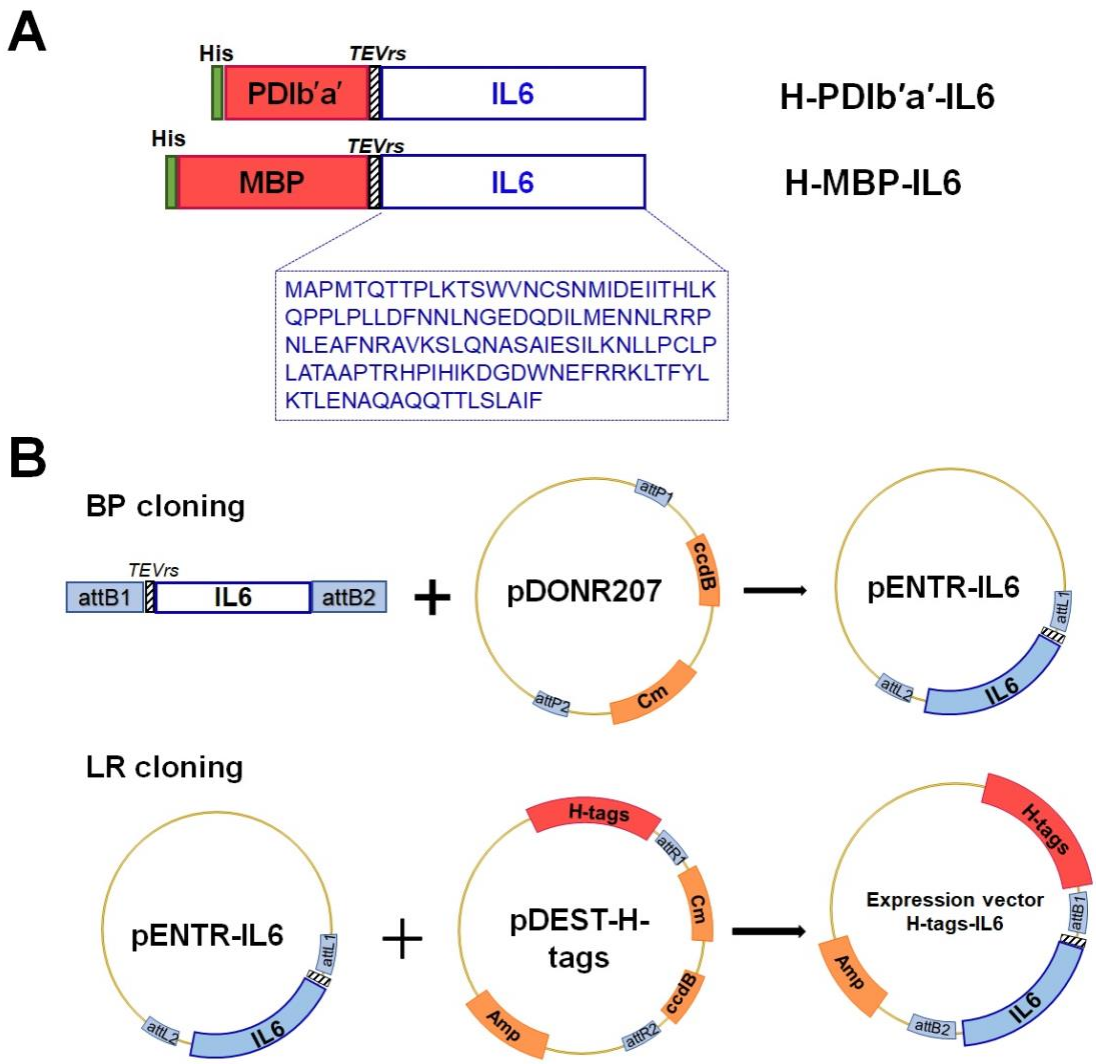
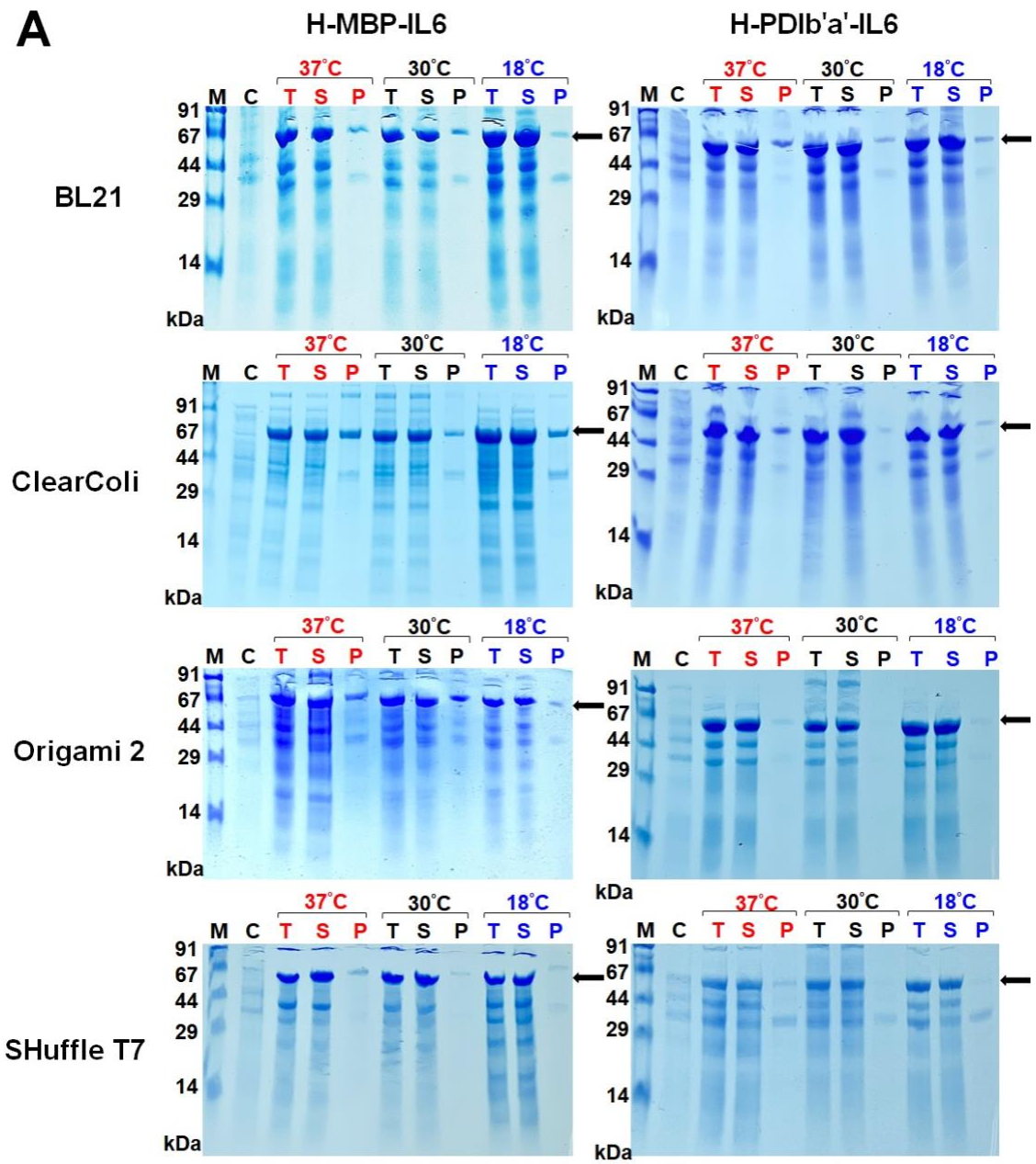


Figure 14. Construct the IL6 expression vectors.

(A) Structure diagram of H-PDIb′a′-IL6, H-MBP-IL6 construct with IL6 (amino acid 30-212).

(B) Schematic of Gateway cloning strategy to generate IL6 expression vectors



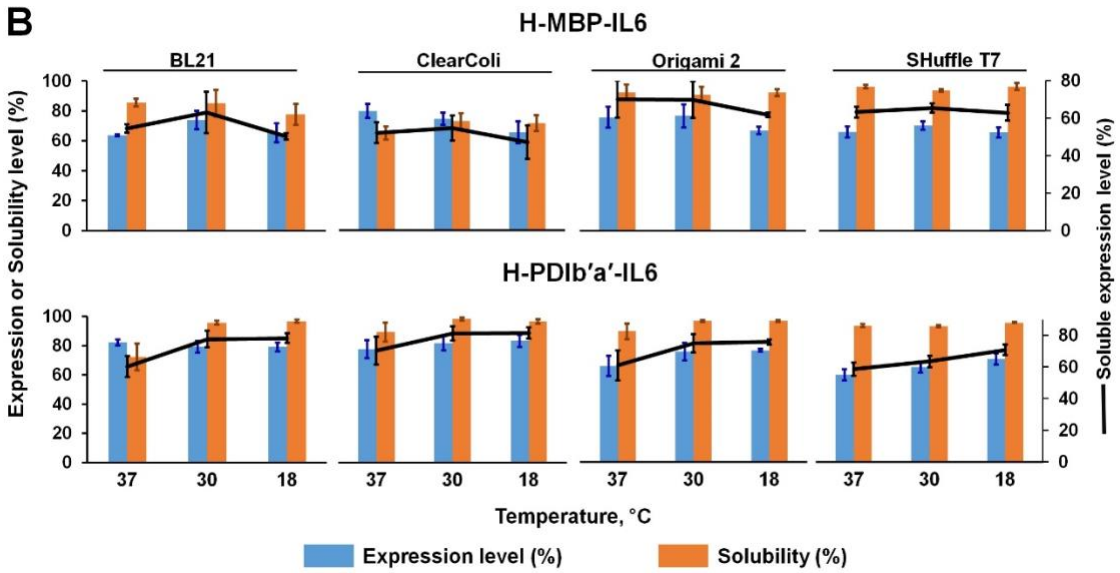


Figure 15. Expression and solubility analysis of fusion proteins IL6 in different E. coli strains.

(A) SDS-PAGE analysis of fusion proteins expression. Black arrow, the fusion protein of H-MBP-IL6 (65 kDa), H-PDIb′a′-IL6 (55 kDa); M, protein molecular weight marker; C, control of total protein lysate before induction; T, total protein lysate after IPTG induction; S, soluble fraction; P, pellet fraction of total protein lysate. The images are representative of three independent experiments.

(B) Quantitative analysis of fusion protein expression in *E. coli* strains. The bar graphs show expression percentage (blue bar) and solubility percentage (orange bar) (primary y-axis), and the black lines indicate expression level x solubility (secondary y-axis).

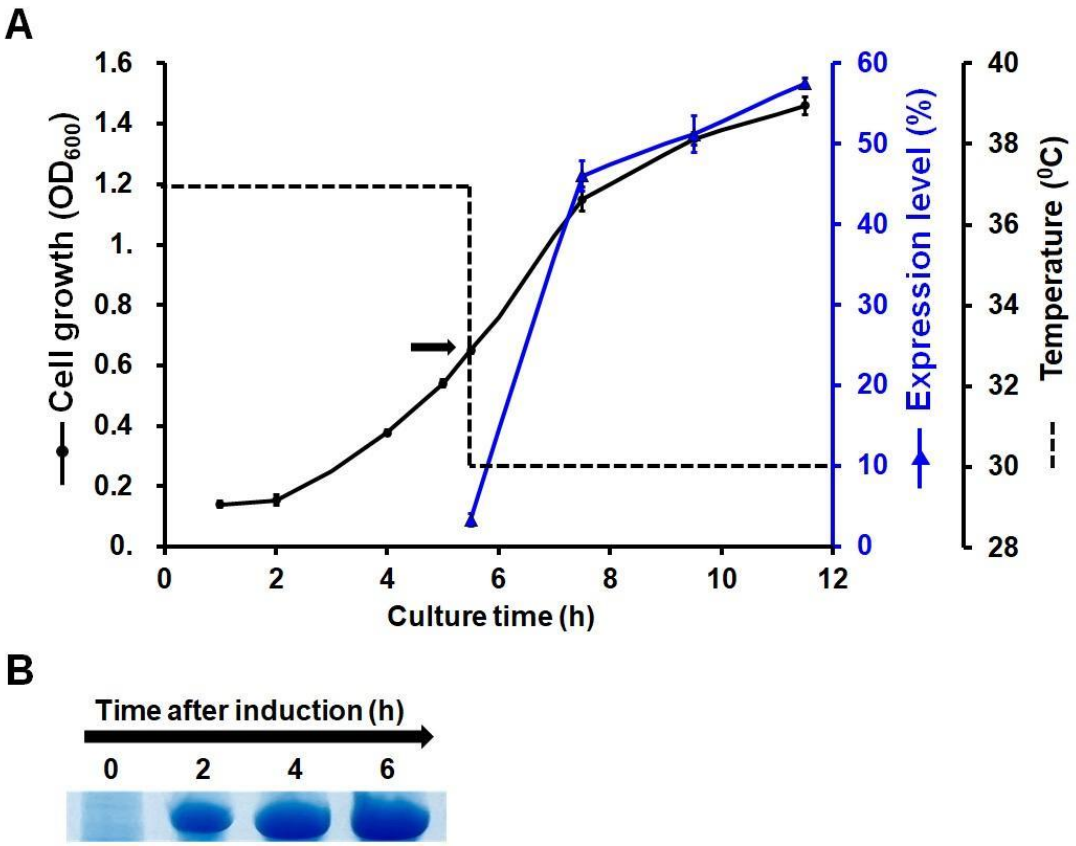


Figure 16. The expression of fusion protein H-PDIb'a'-IL6 in 1 L flask cultures.

(A) The expression profile of the H-PDIb′a′-IL6 fusion protein. The x-axis corresponds to the culture time, while the left y-axis portrays the cell growth, as measured by the OD value, depicted as a black line. Following the induction with 1mM IPTG, the expression levels of the fusion protein are represented by the blue line, plotted against the right y-axis. The dotted line on the secondary right y-axis shows the temperature at which the protein expression occurs. The data points are averages derived from three experiments, including the standard error. The induction time is highlighted with a black arrow.

(B) SDS-PAGE images that showcase the expression levels of fusion proteins at different time intervals post-induction.

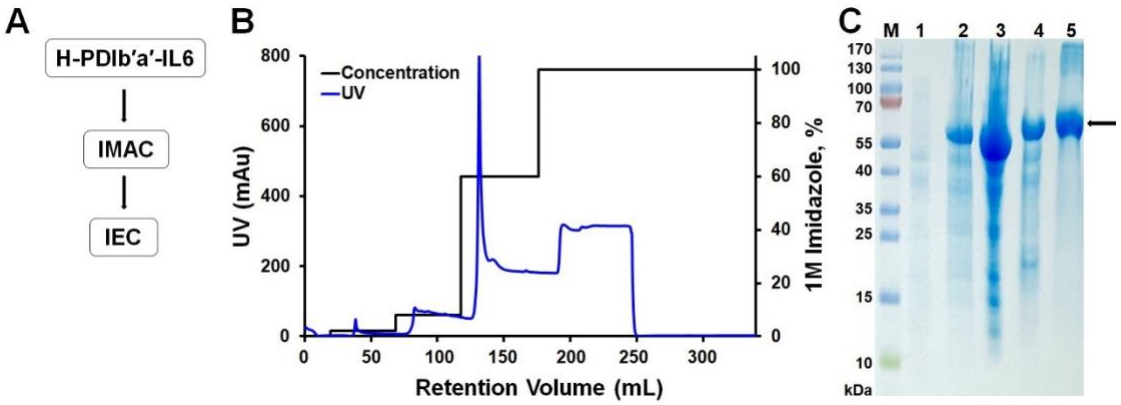


Figure 17. Purification of fusion protein H-PDIb'a'-IL6 in ClearColi strain.

(A) Flow chart of the purification steps.

(B) A typical analytical immobilized metal affinity chromatography (IMAC) profile is shown. The fluorescence trace of the protein is displayed in blue and the elution gradient is displayed in black.

(C) IL6 was purified from H-PDIb′a′-IL6 produced in ClearColi by IMAC and anion-exchange chromatography. M, PageRuler Prestained Protein ladder (Fermentas, #SM0671); lane 1, total lysate proteins before IPTG induction (negative control); lane 2, Total proteins induced by IPTG after cell sonication; lane 3, soluble fraction after cell sonication from total cell proteins; lane 4, H-PDIb′a′-IL6 fusion protein (55 kDa) purified with IMAC column; lane 5, H-PDIb′a′-IL6 (55 kDa) was purified with ion-exchange chromatography to obtain high purity.

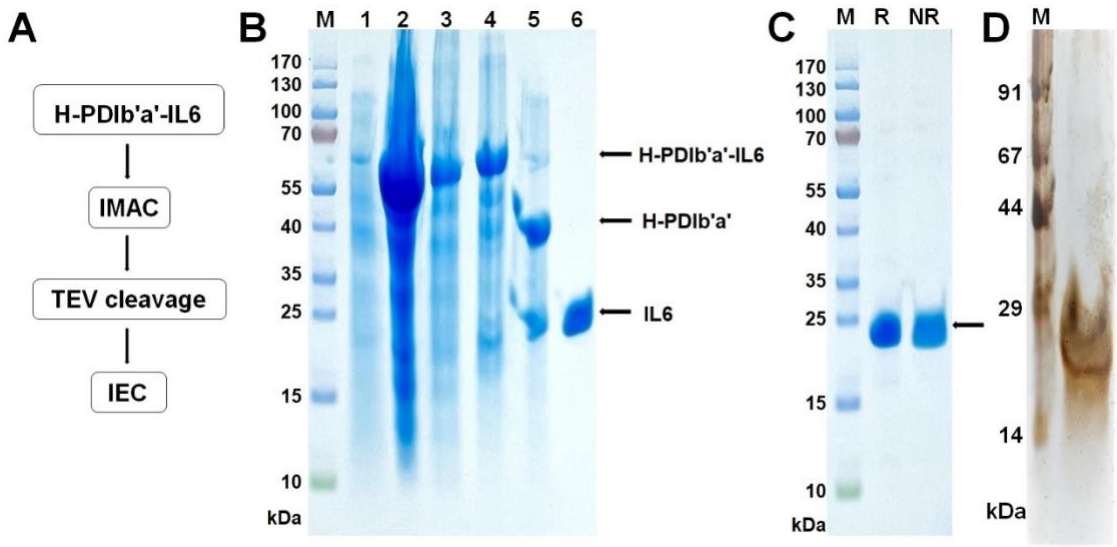


Figure 18. Purification of recombinant IL6 in ClearColi strain.

(A) Flow chart of the purification steps.

(B) IL6 was purified from H-PDIb′a′-IL6 produced in ClearColi by IMAC and anion-exchange chromatography. M, PageRuler Prestained Protein ladder (Fermentas, #SM0671); lane 1, total lysate proteins before IPTG induction as negative control; lane 2, Total proteins induced by IPTG after cell sonication; lane 3, soluble fraction after cell sonication from total cell proteins; lane 4, H-PDIb′a′-IL6 fusion protein (55 kDa) purified with IMAC column; lane 5, the result of TEV protease cleavage (29 kDa): H-PDIb'a' (35 kDa) and IL6 (21 kDa); lane 6, IL6 was purified with ion-exchange chromatography.

(C) Purified IL6 protein under reducing and non-reducing conditions with M, PageRuler Prestained Protein ladder (Fermentas, #SM0671); R, reducing condition (300 mM DTT); NR, non-reducing condition (without DTT).

(D) Silver staining of IL6 under reducing conditions with M, protein molecular weight marker.

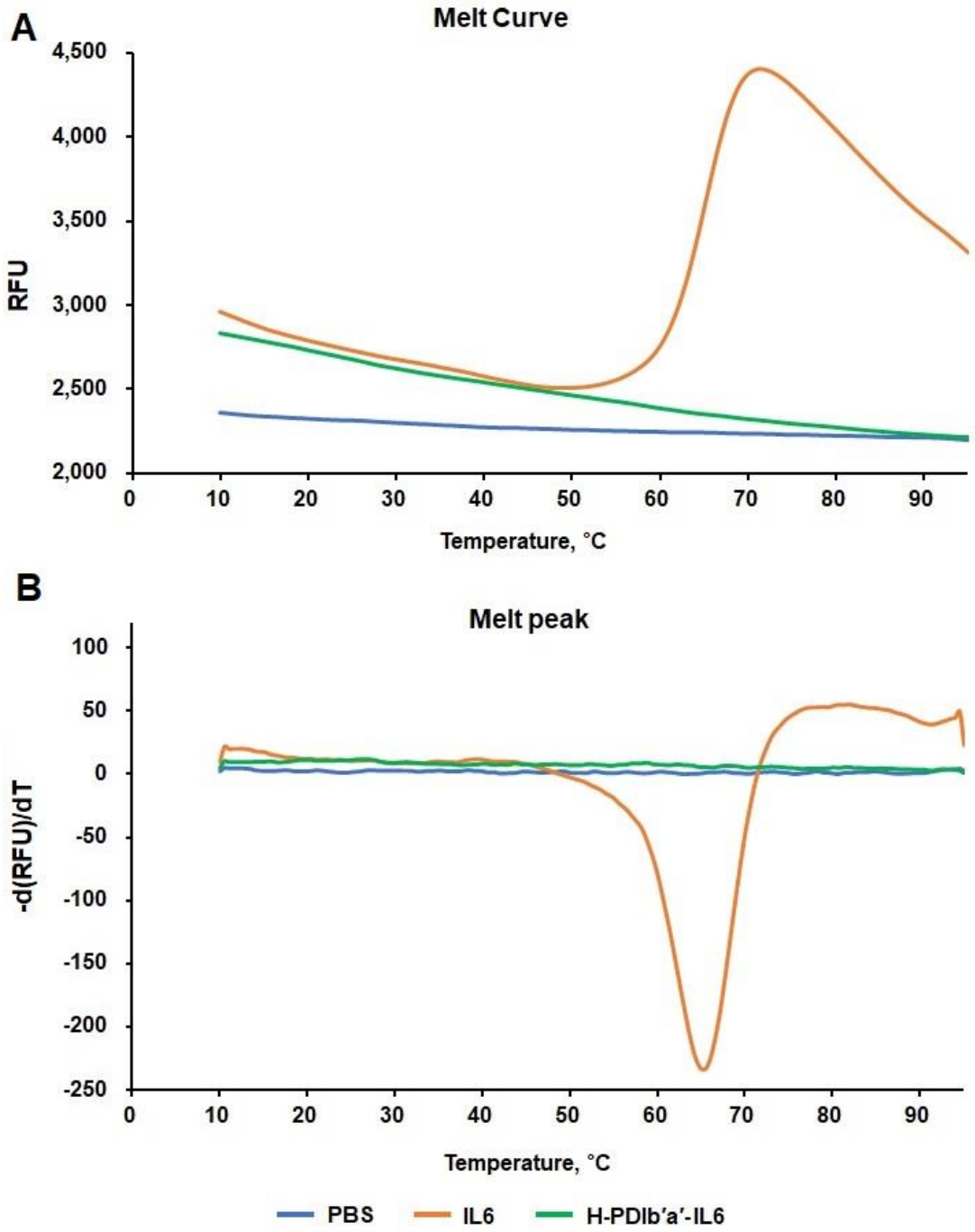


Figure 19. Melting temperature analysis of IL6, H-PDIb'a'-IL6 in PBS buffer.

(A) Melting curve and (B) Melting peak of purified IL6 and H-PDIb'a'-IL6 in PBS buffer, with triplicate repeated experiments. RFU, relative fluorescence units.

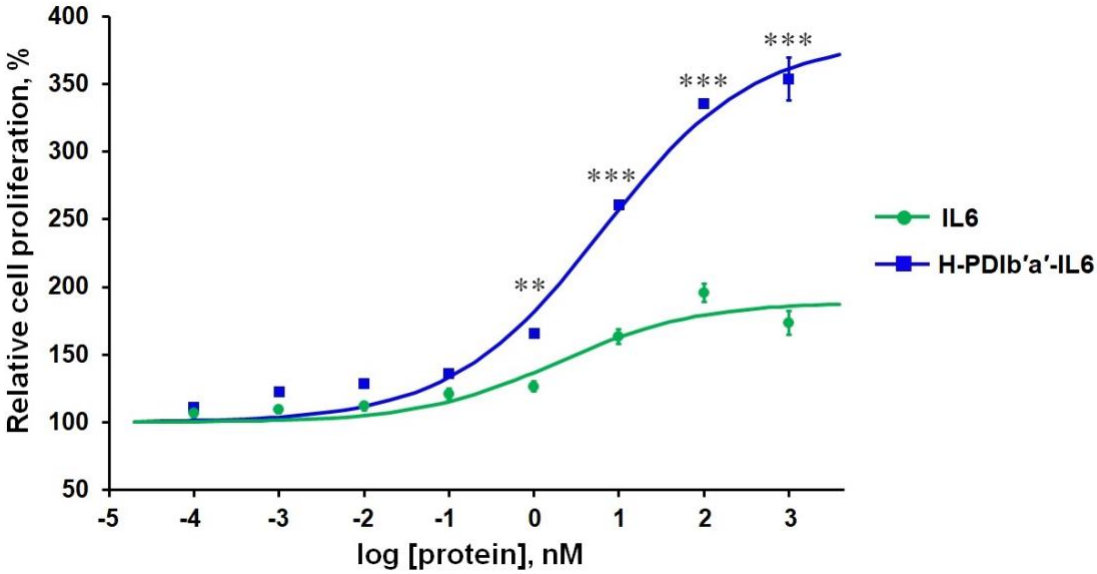
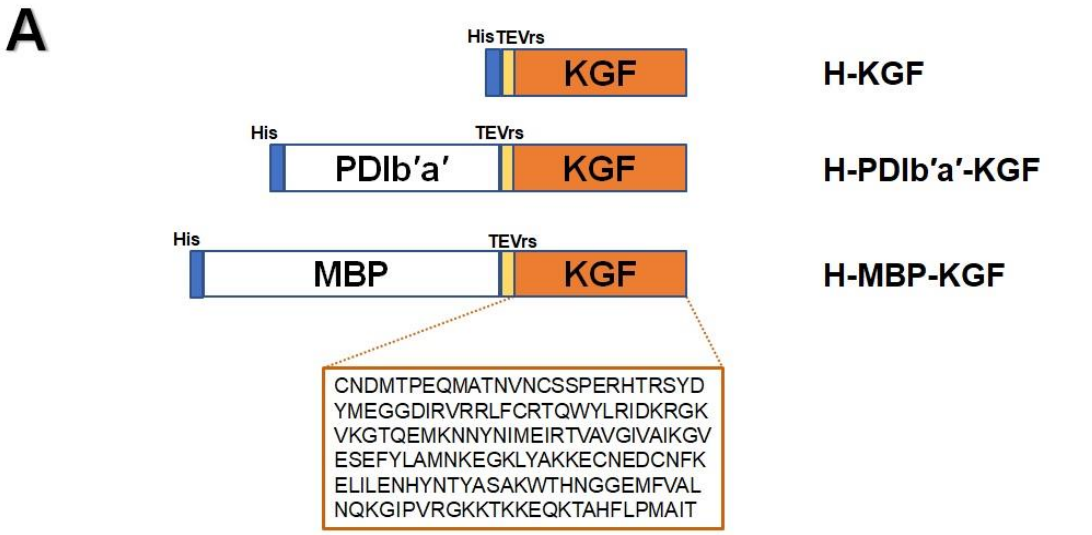


Figure 20. Dose-response effects of IL6 variants on TF-1 cell proliferation.

Dose-response curve for TF-1 cell proliferation in response to purified IL6 and H-PDIb′a′-IL6. A dose-response curve demonstrates the differential effects on TF-1 cell proliferation when exposed to purified IL6 and the fusion protein H-PDIb′a′-IL6 at higher concentrations (0 to 1 μM) over a 48-h treatment. The reference point for proliferation, in the absence of any protein, was set at 100%, and all other values were normalized to this reference point (n = 9). Statistical significance was determined using one-way ANOVA, followed by Bonferroni's post-hoc test. Asterisks indicate significant differences between purified IL6 and H-PDIb′a′-IL6 (* $p < 0.05$; ** $p < 0.01$; *** $p < 0.001$). Some error bars are too small to be representative in the graph.



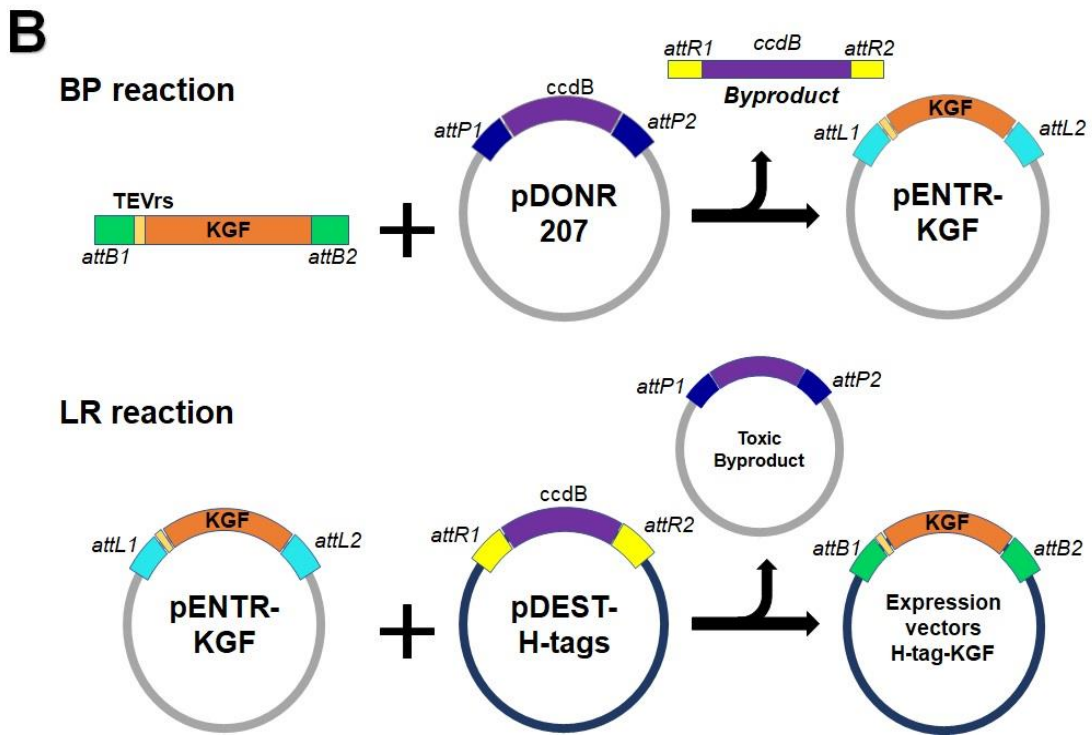


Figure 21. Construct the KGF expression vectors.

(A) Structure diagram of H-PDIb′a′-KGF, H-MBP-KGF construct with KGF (amino acid 32-194).

(B) Schematic of Gateway cloning strategy to generate KGF expression vectors.

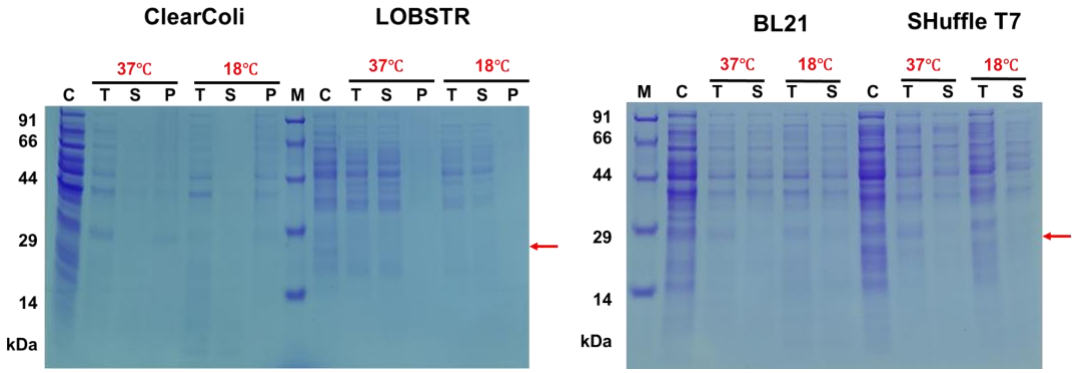
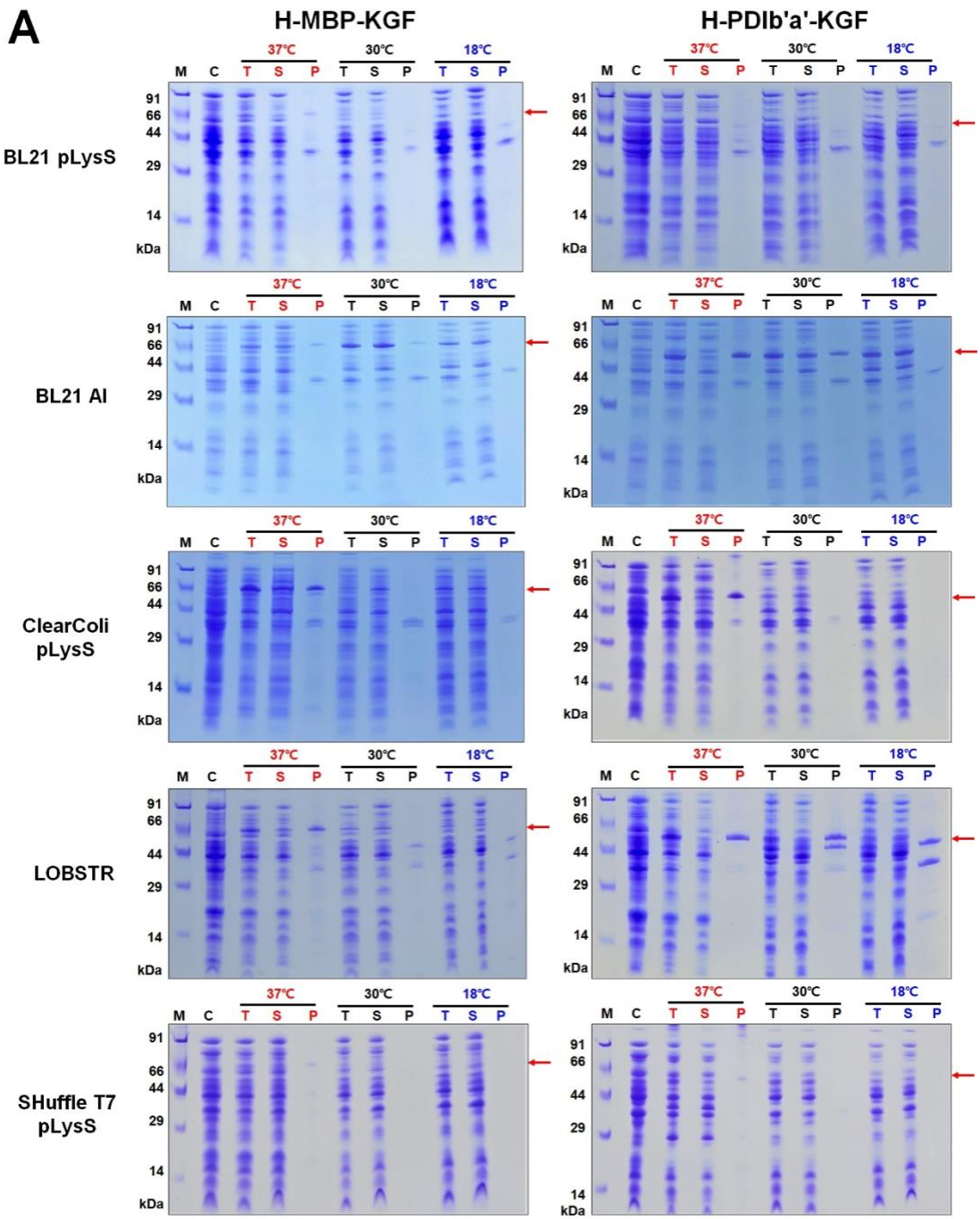


Figure 22. Expression and solubility of H-KGF in different E. coli strains at two distinct temperatures.

M, protein molecular weight marker; C, filtered supernatant of the cell lysate before IPTG induction; T, total cellular protein after IPTG induction; S, soluble protein after sonication; P, insoluble protein in the form of inclusion bodies in the pellet. The red arrows on the right-hand side of each SDS-PAGE gel highlight the expected size of H-KGF (23 kDa).



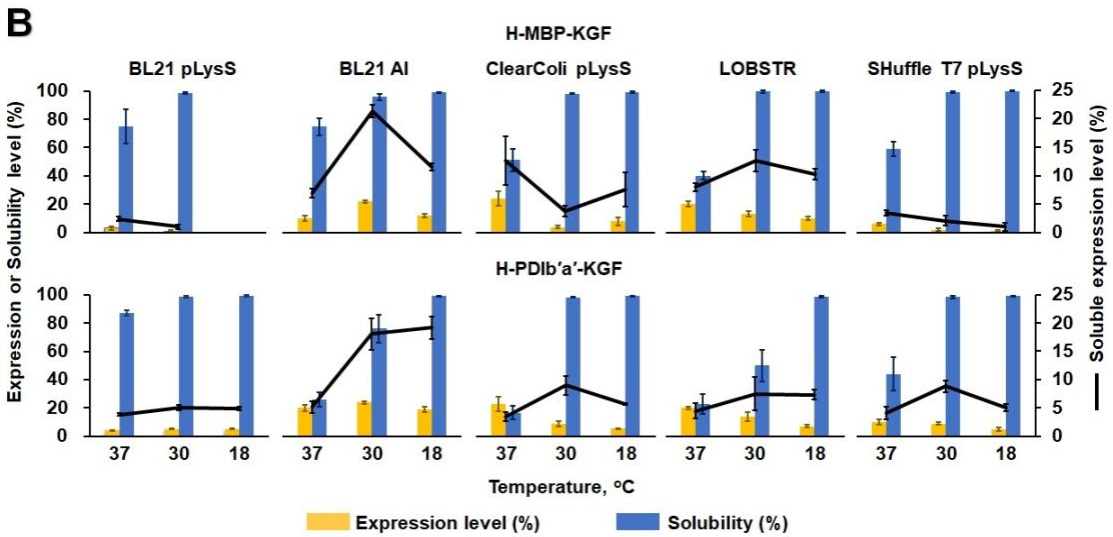


Figure 23. Expression and solubility analysis of fusion proteins KGF in different E. coli strains.

(A) SDS-PAGE analysis illustrating the expression profiles of different fusion proteins in the designated *E. coli* strains. Red arrows distinctly mark the target fusion proteins: H-PDIb'a'-KGF (53.5 kDa), and H-MBP-KGF (63 kDa). M, protein molecular weight marker; C, the control comprising total protein lysate before induction; T, the post-IPTG induction total protein lysate; S, the soluble fraction; and P, the pellet fraction. The displayed images serve as representative outputs from three experiment replicates.

(B) Quantitative analysis of fusion protein expression in *E. coli* strains. The bar graphs show expression percentage (yellow bar) and solubility percentage (blue bar) (primary y-axis), and the black lines indicate expression level times solubility (secondary y-axis).

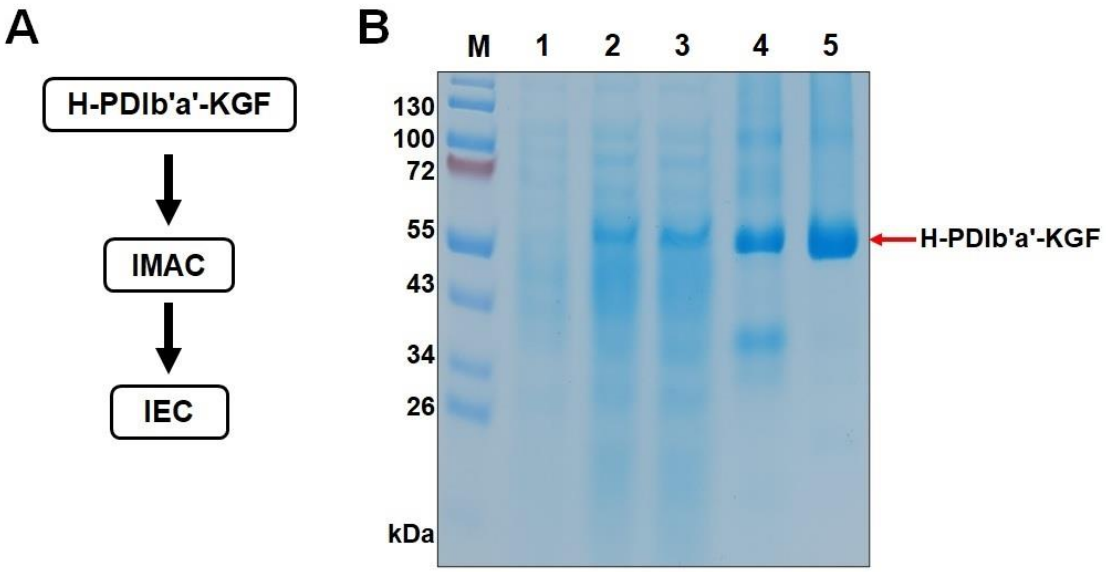


Figure 24. Purification and characterization of H-PDIb'a'-KGF expressed in BL21 AI strain.

(A) Schematic flowchart outlining the multi-step purification process employed for the isolation of H-PDIb′a′-KGF.

(B) H-PDIb′a′-KGF, produced in the BL21 AI strain, was purified through a dual chromatographic approach involving IMAC followed by IEC. The gel lanes are designated as follows: M, PAGERuler Prestained Protein marker (Fermentas, SM0671); Lane 1, total cellular proteins before IPTG induction, serving as a negative control; Lane 2, total protein lysate post-IPTG induction and subsequent cell sonication; Lane 3, soluble protein fraction isolated after cell sonication; Lane 4, H-PDIb′a′-KGF fusion protein (53.5 kDa) purified via IMAC; Lane 5, H-PDIb′a′-KGF fusion protein (53.5 kDa) further purified through IEC to remove impurities.

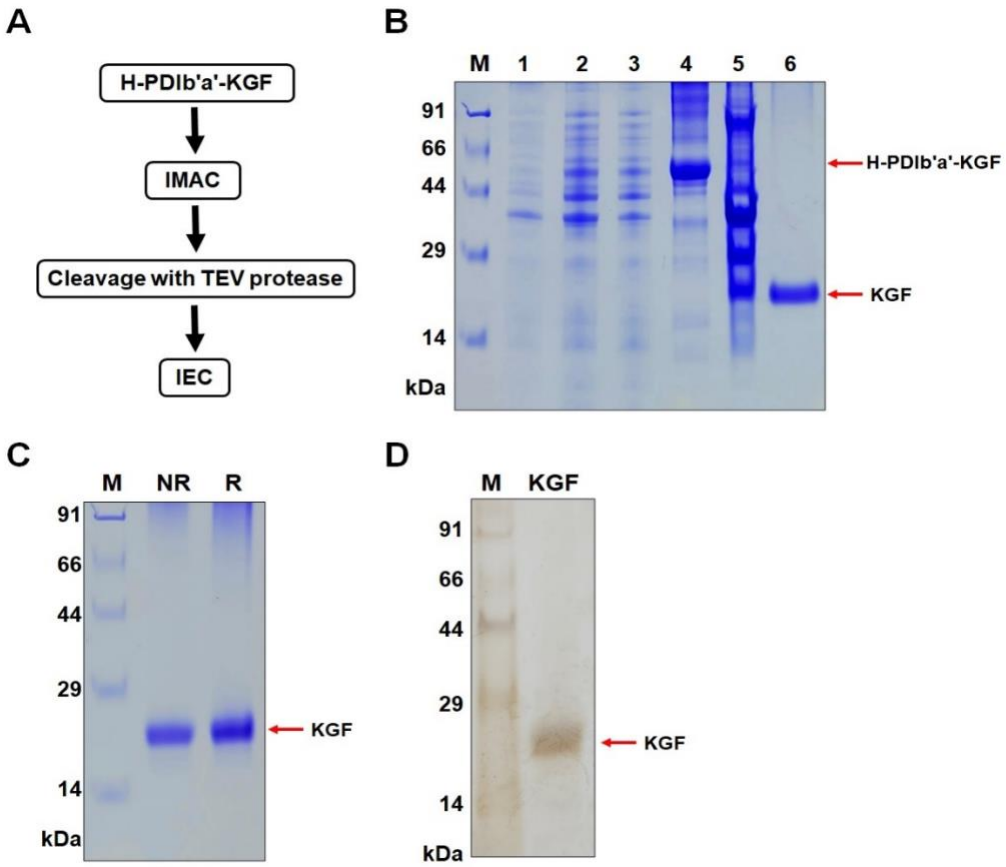


Figure 25. Purification and analysis of KGF expressed in BL21 AI strain.

(A) Schematic flowchart detailing the sequential purification steps employed for isolating KGF. (B) KGF was extracted from H-PDIb′a′-KGF, produced in the BL21 AI strain, utilizing a two-step chromatographic process involving IMAC and IEC. Gel lanes are designated as follows: M, protein molecular weight marker; Lane 1, total cellular proteins before IPTG induction (serving as a negative control); Lane 2, post-IPTG induction total proteins following cell sonication; Lane 3, soluble protein fraction post-sonication; Lane 4, IMAC-purified H-PDIb′a′-KGF fusion protein (53.5 kDa); Lane 5, products of TEV protease cleavage, including H-PDIb'a' (35 kDa) and KGF (19 kDa); Lane 6, KGF isolated via cation ion-exchange chromatography (IEC). (C) The purified KGF protein was subjected to SDS-PAGE analysis under both reducing and non-reducing conditions. Labels are as follows: M, protein molecular weight marker; R, reducing condition (treated with 300 mM DTT); NR, non-reducing condition (absence of DTT). (D) Silver-stained SDS-PAGE gel of KGF under reducing conditions. M, protein molecular weight marker.

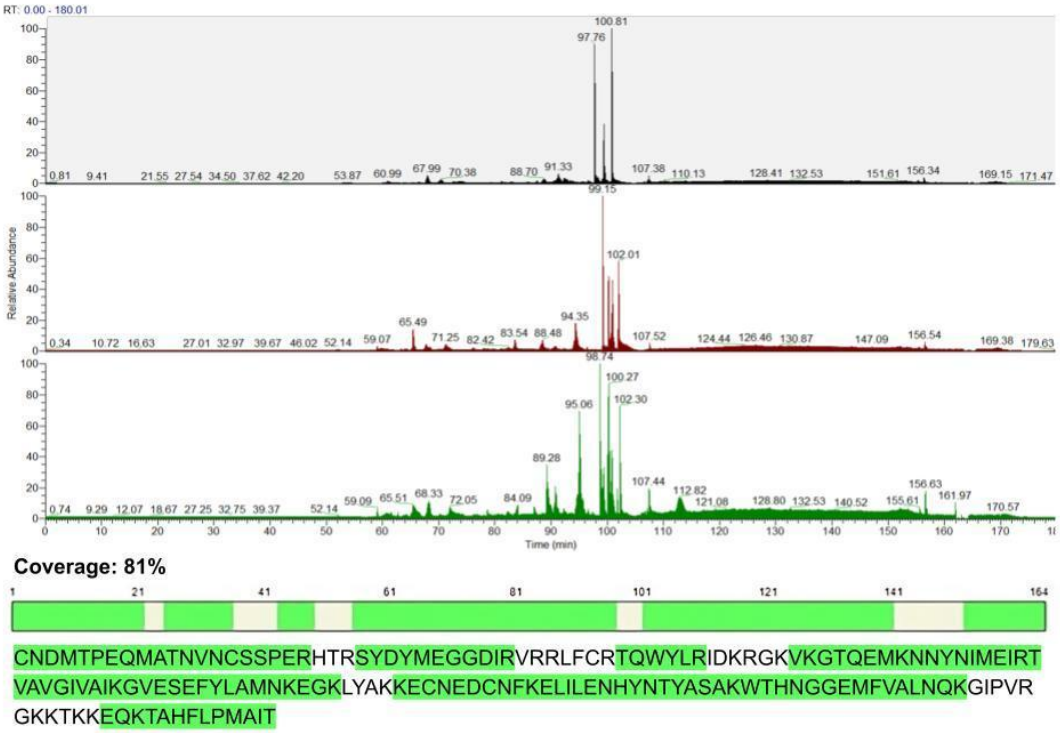


Figure 26. LC-MS/MS analysis of purified KGF.

The figure presents the results of protein identification achieved through LC-MS/MS following trypsin-mediated digestion of the sample. The LC-MS/MS analysis was performed under reducing conditions. Protein sequence coverage is visually represented by green shading, which is based on percolator confidence scores. A green hue indicates high confidence (p-value < 0.01), and the coverage was determined to be 81% relative to the complete protein sequence.

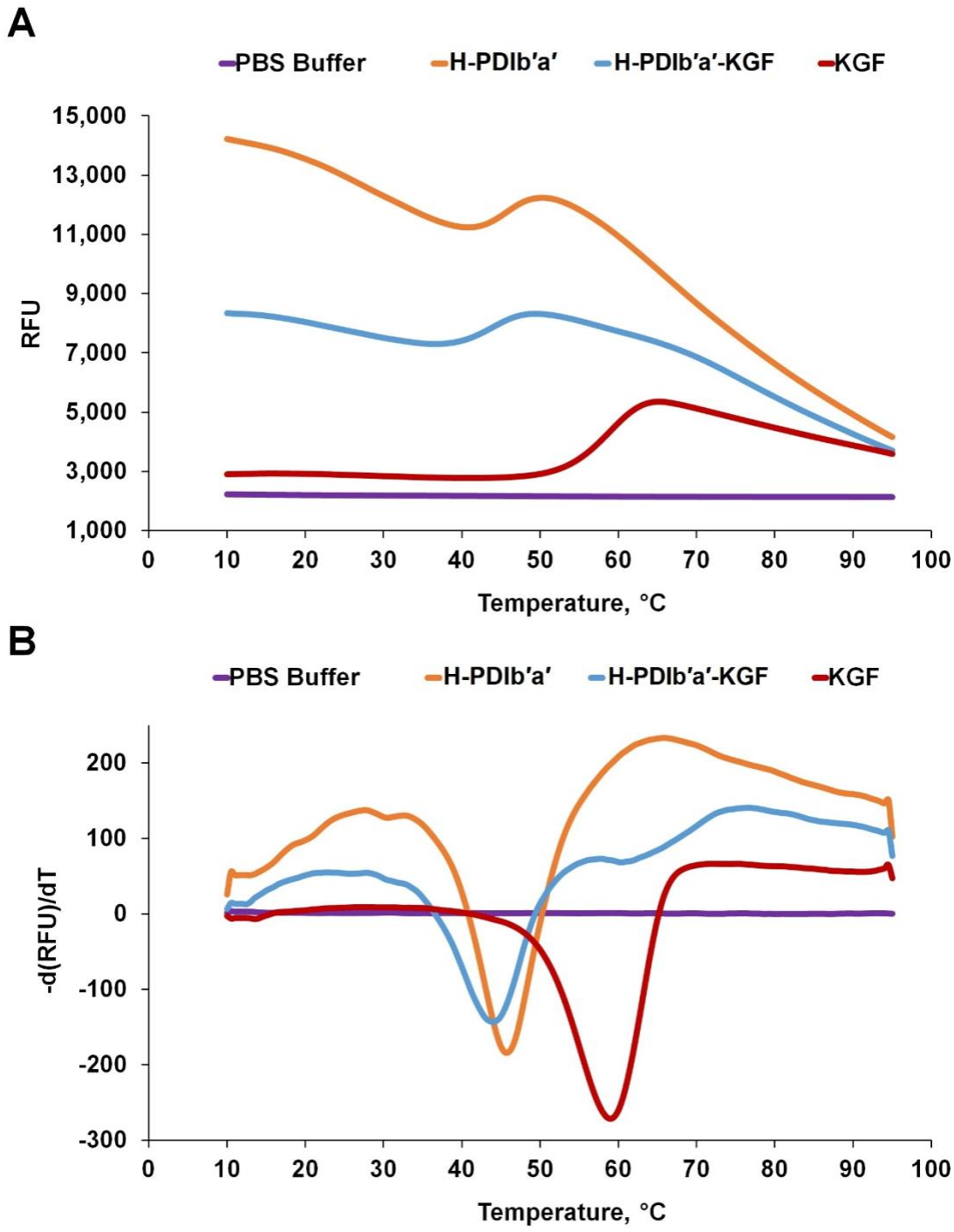


Figure 27. Melting temperature analysis of KGF, H-PDIb'a'-KGF, and H-PDIb'a'.

(A) Melting curve and (B) melting peak of KGF, H-PDIb′a′-KGF, and H-PDIb'a' in PBS buffer, with each data point representing the mean of triplicates. RFU denotes relative fluorescence units.

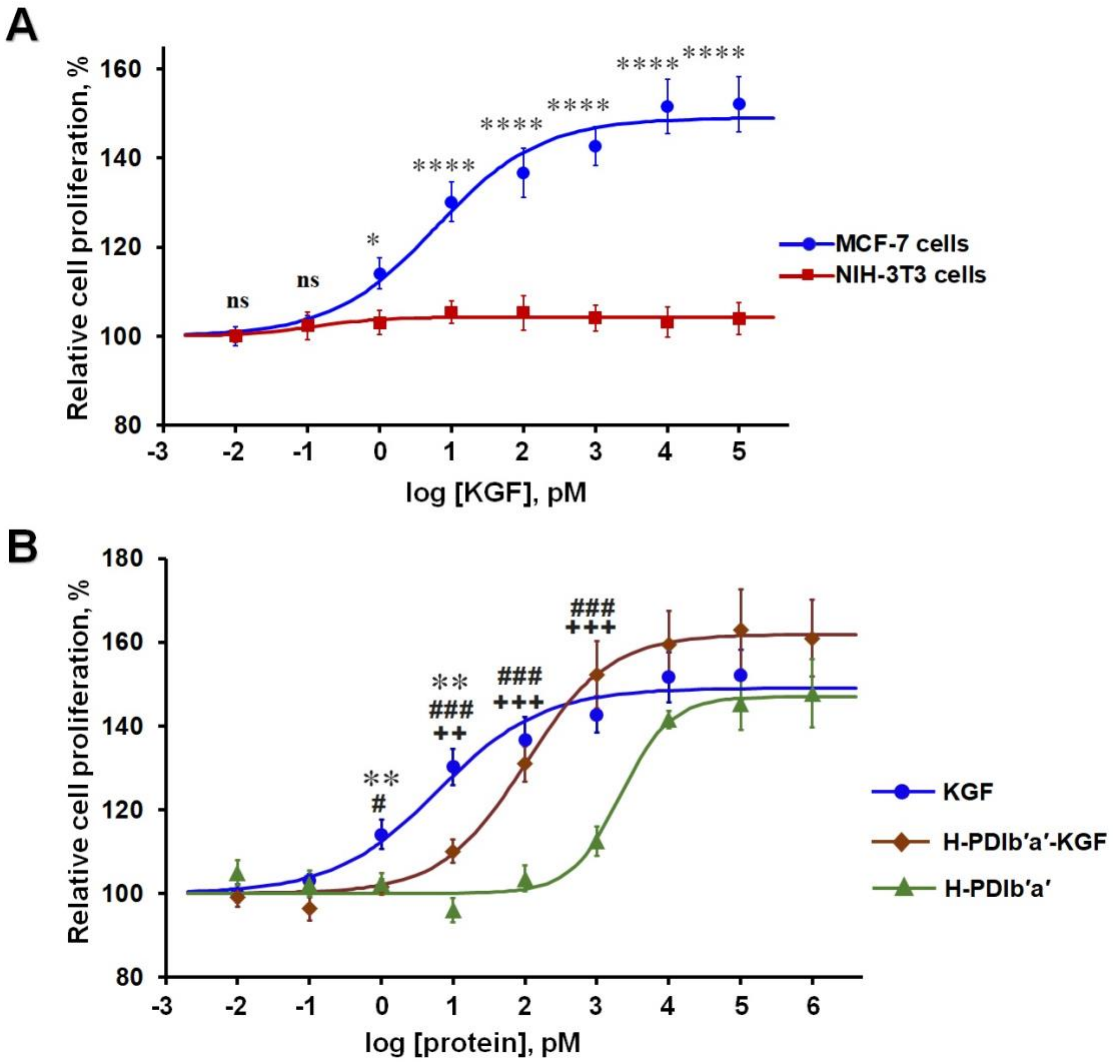


Figure 28. Proliferation activity of purified KGF and related proteins.

(A) MCF-7 and NIH-3T3 cells were treated with varying concentrations of purified KGF for 72 h. Cell proliferation was assessed using a WST-8 assay. The proliferation level in the absence of KGF was set as the baseline (100%), and all other values were normalized to this (n = 9). Statistical significance between the two cell lines was evaluated using a t-test. Significance levels are denoted by asterisks (* $p < 0.05$; ** $p < 0.01$; *** $p < 0.001$; **** $p < 0.0001$); 'ns' indicates non-significance.

(B) MCF-7 cells were treated with varying concentrations of purified KGF, H-PDIb′a′-KGF fusion protein, and H-PDIb'a' for 72 h. The proliferation level without any protein was used as the reference point (100%), and all other values were normalized to this (n = 9). Statistical significance was determined using one-way ANOVA, followed by

Bonferroni's post-hoc test. Asterisks indicate significant differences between purified KGF and H-PDIb′a′-KGF (* $p < 0.05$; ** $p < 0.01$; *** $p < 0.001$). Hashtags denote significance levels when comparing purified KGF to PDIb'a' (# $p < 0.05$; ## $p < 0.01$; ### $p < 0.001$). Plus signs represent significant differences between H-PDIb′a′-KGF and H-PDIb'a' (+ $p < 0.05$; ++ $p < 0.01$; +++ $p < 0.001$).

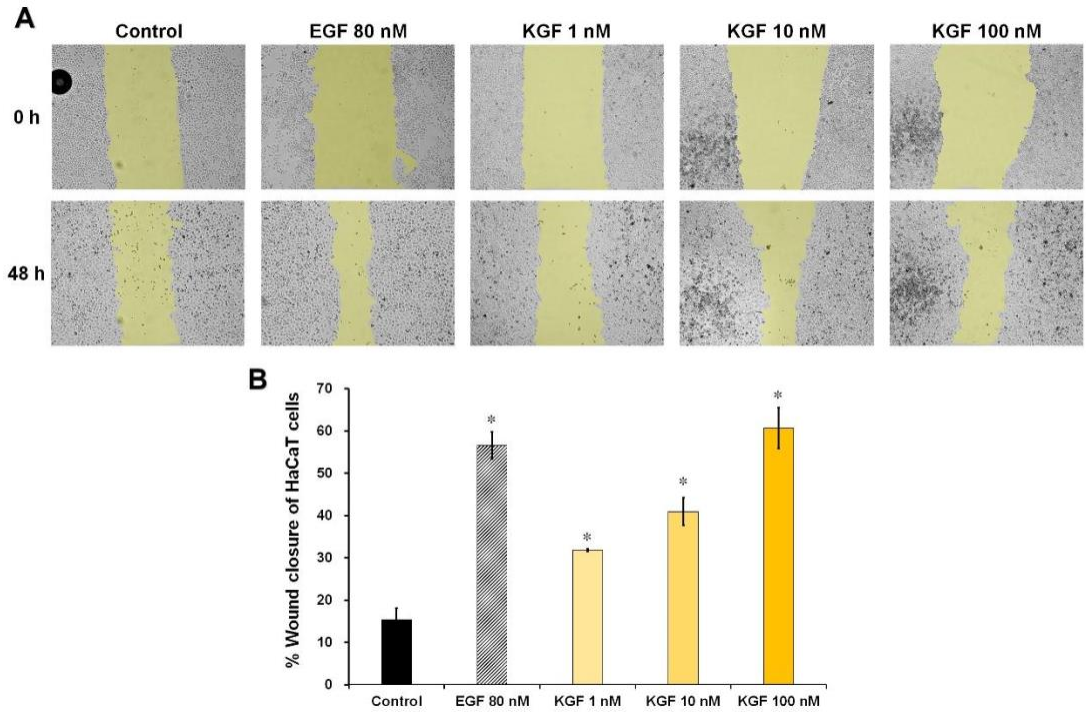


Figure 29. Wound healing efficacy of purified KGF in HaCaT cells.

(A) Scratch wound healing assay comparing HaCaT cells treated with control (untreated), EGF (80 nM), and varying concentrations of purified KGF (1 nM, 10 nM, 100 nM). Images were acquired at both 0- and 48-h post-scratch.

(B) Quantification of wound closure at 48-h timepoint, presented as mean ± SEM (n=3). The asterisk denotes statistical significance (* $p < 0.05$).

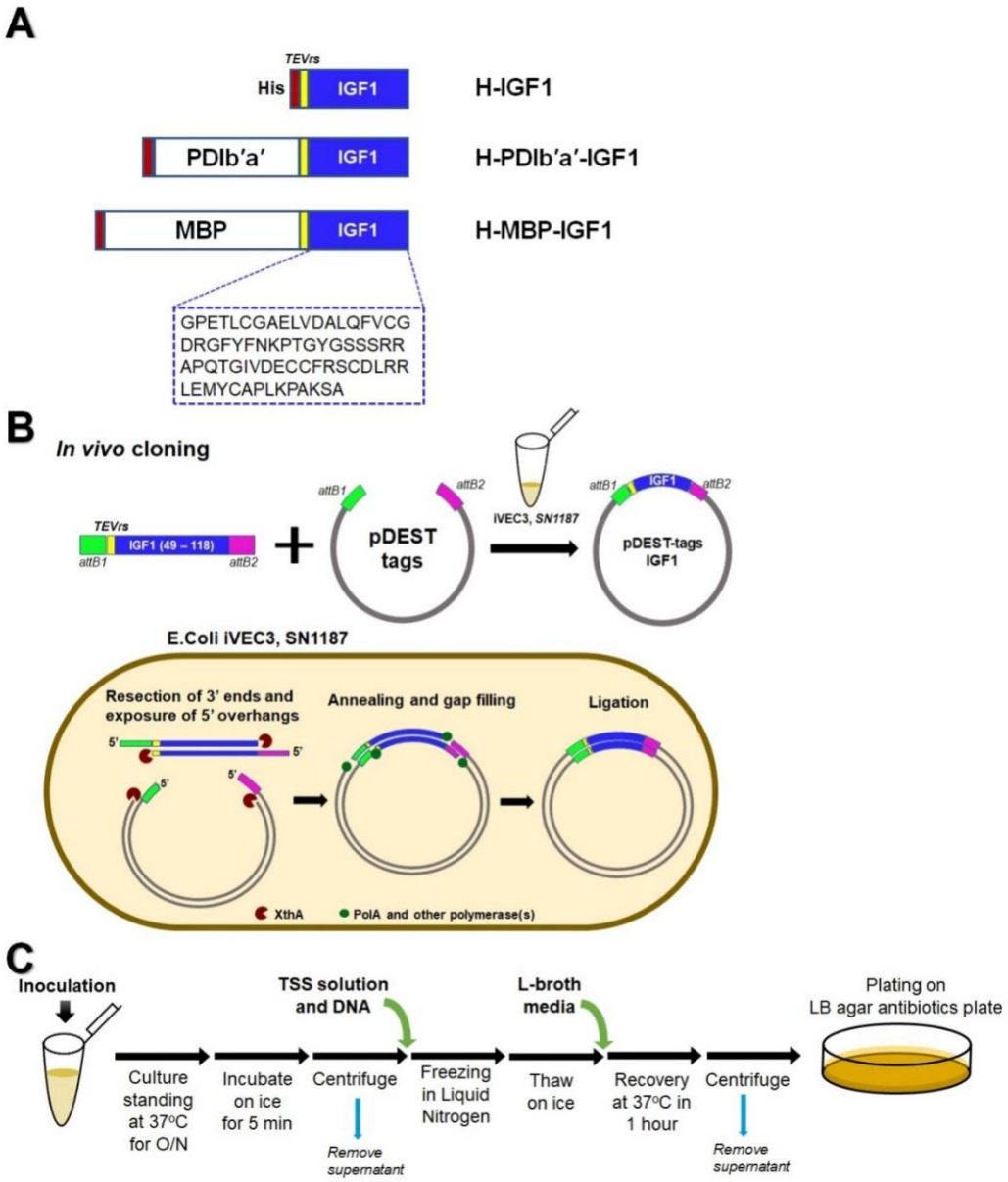


Figure 30. Constructs of the IGF1 expression vectors.

(A) Structure diagram of H-IGF1, H-PDIb′a′-IGF1, and H-MBP-IGF1 construct with IGF1 (amino acid 49-118).

(B) Schematic depiction of the cloning procedure for the IGF1 expression vector and the mechanistic model of iVEC3 *E. coli* (SN1187).

(C) Outline of the TSS (transformation and storage solution) method utilized for *in vivo* cloning.

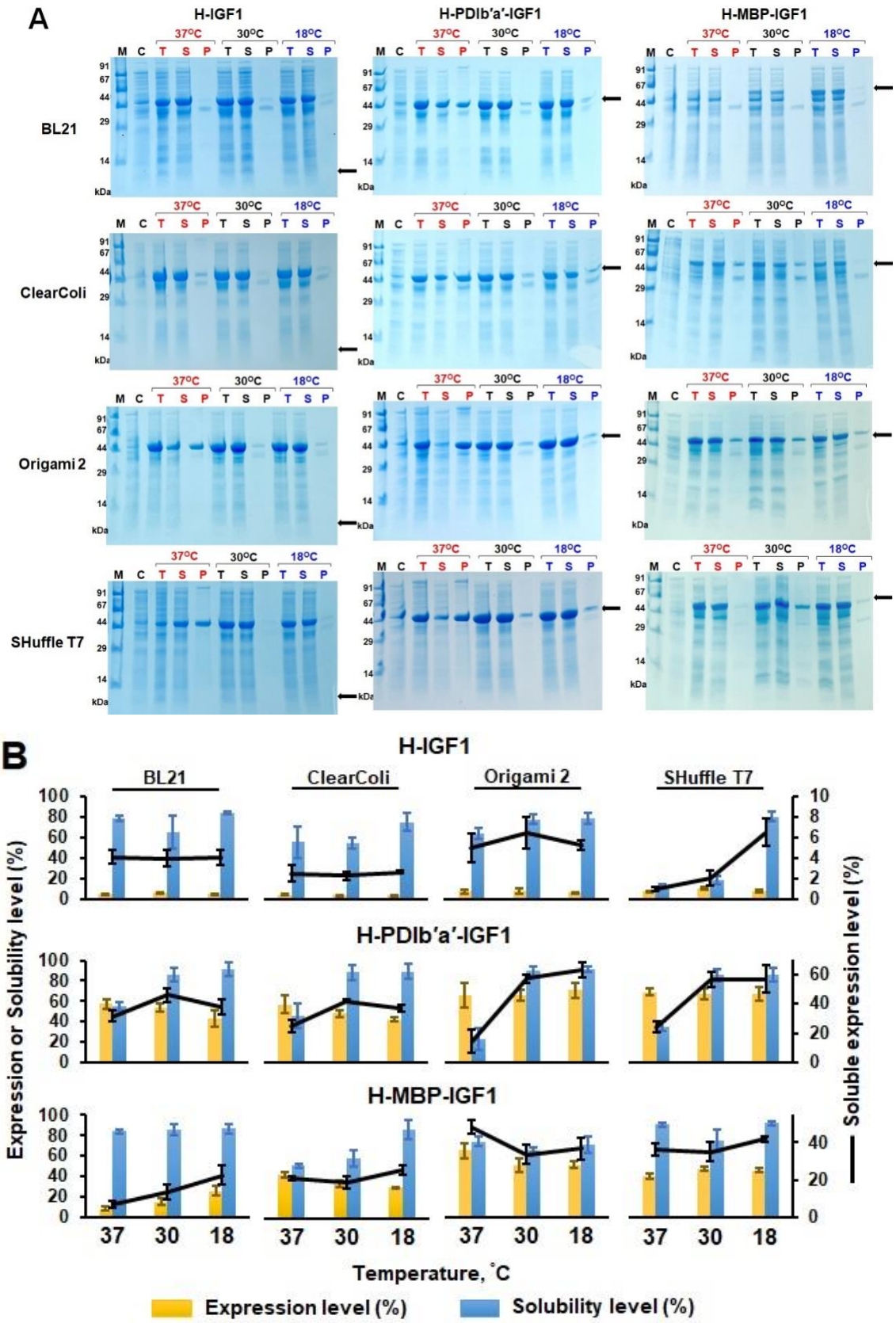


Figure 31. Expression and solubility analysis of fusion proteins IGF1 in different E. coli strains.

(A) SDS-PAGE analysis illustrating the expression profiles of different fusion proteins in the designated *E. coli* strains. Black arrows distinctly mark the target fusion proteins: H-IGF1 (11 kDa), H-PDIb′a′-IGF1 (43 kDa), and H-MBP-IGF1 (52 kDa). Annotations are as follows: M, protein molecular weight marker; C, the control comprising total protein lysate before induction; T, post-IPTG induction total protein lysate; S, the soluble fraction; and P, the pellet fraction. The displayed images serve as representative outputs from three separate experimental runs.

(B) Quantitative Depiction of Fusion Protein Expression across a Range of *E. coli* Strains. The bar plots display expression percentages (in yellow) juxtaposed against solubility percentages (in blue) aligned on the primary y-axis. A contrasting black line marks the product of expression multiplied by solubility, plotted against the secondary y-axis. Expression and solubility metrics were computed utilizing the Gel Analyzer software, rooted in data from three individual experiments.

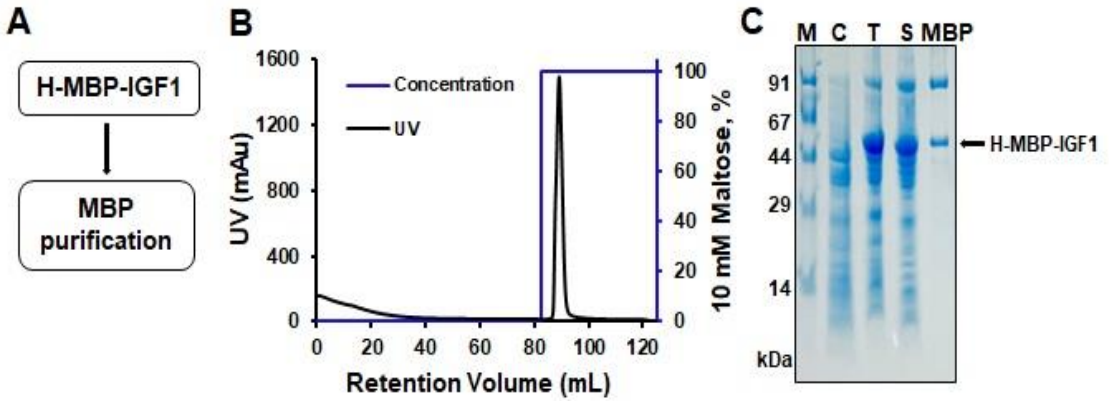


Figure 32. Purification of fusion protein H-MBP-IGF1 in Origami 2 strain.

(A) Diagram detailing the H-MBP-IGF1 purification steps.

(B) Chromatogram showcasing the amylose resin chromatography for H-MBP-IGF1. The fluorescence trace of the protein is displayed in black and the elution concentration is displayed in blue.

(C) SDS-PAGE gel analysis capturing the H-MBP-IGF1 purification stages: M, molecular weight marker; C, total lysate proteins before IPTG induction, serving as a negative control; T, total proteins post-IPTG induction following cell sonication; S, the soluble fraction utilized for purification; and MBP highlights the purified fusion protein H-MBP-IGF1 (52 kDa) isolated via amylose resin chromatography.

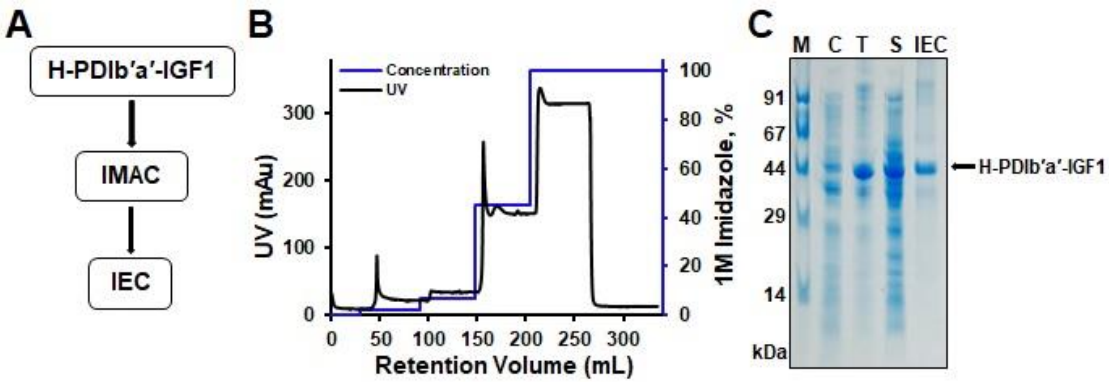


Figure 33. Purification of fusion protein H-PDIb'a'-IGF1 in ClearColi strain.

(A) Diagram detailing the H-PDIb′a′-IGF1 purification steps.

(B) Chromatogram showcasing the affinity column purification of H-PDIb′a′-IGF1. The fluorescence trace of the protein is displayed in black and the elution gradient is displayed in blue.

(C) SDS-PAGE gel analysis detailing the H-PDIb′a′-IGF1 purification process: M, molecular weight marker; C, total lysate proteins before IPTG induction, used as a negative control; T, total proteins after IPTG induction and cell sonication; S, the soluble fraction post-sonication; and IEC highlights the fusion protein H-PDIb′a′-IGF1 (43 kDa) purified using both immobilized metal affinity chromatography and ion-exchange chromatography.

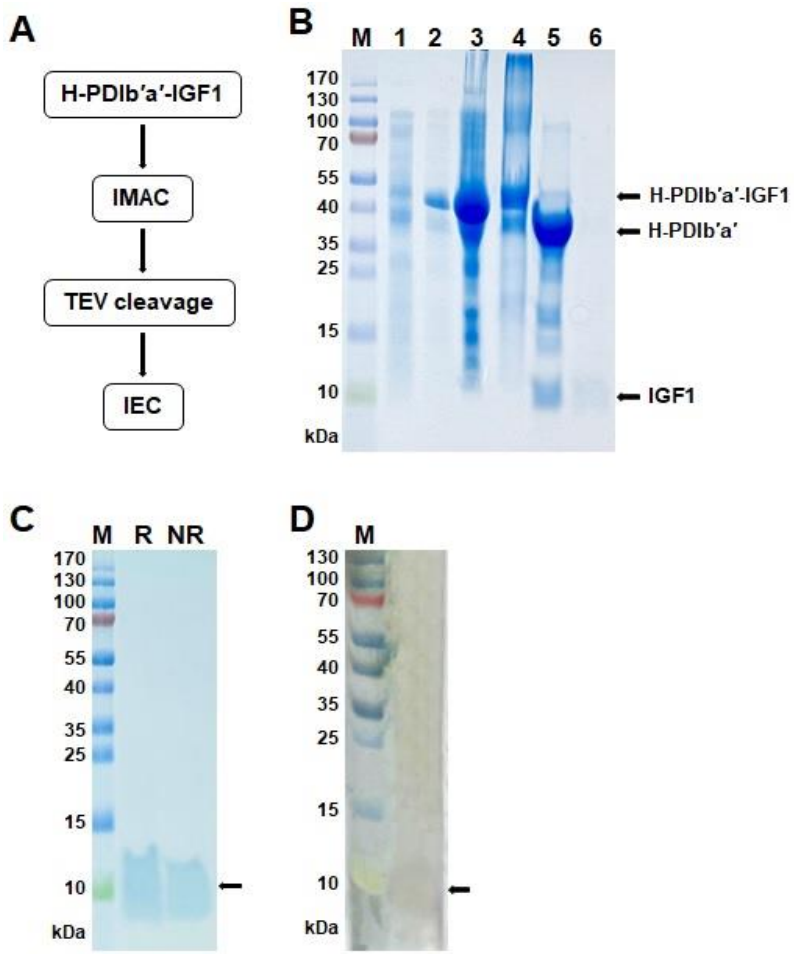


Figure 34. Purification of recombinant IGF1 in ClearColi strain.

(A) Depiction of the IGF1 purification process. (B) SDS-PAGE analysis of the IGF1 purification stages. M, PageRuler Prestained Protein ladder (Fermentas, #SM0671). Lane 1, the total lysate proteins before IPTG induction, serves as a negative control. Lane 2, proteins post-IPTG induction following cell sonication. Lane 3, the soluble fraction obtained after sonication. Lane 4, the elution product from immobilized metal affinity chromatography. Lane 5, TEV protease cleavage products, which include H-PDIb′a′ (35 kDa) and IGF1 (8 kDa). Lane 6, IGF1 isolated via ion exchange chromatography. (C) SDS-PAGE analysis of the purified IGF1 protein under different conditions. M, PageRuler Prestained Protein ladder; R, the reducing condition where the sample was treated with 300 mM DTT, while NR signifies the non-reducing condition where DTT was omitted. (D) Silver-stained SDS-PAGE gel of IGF1 under reducing conditions. M, PageRuler Prestained Protein ladder.

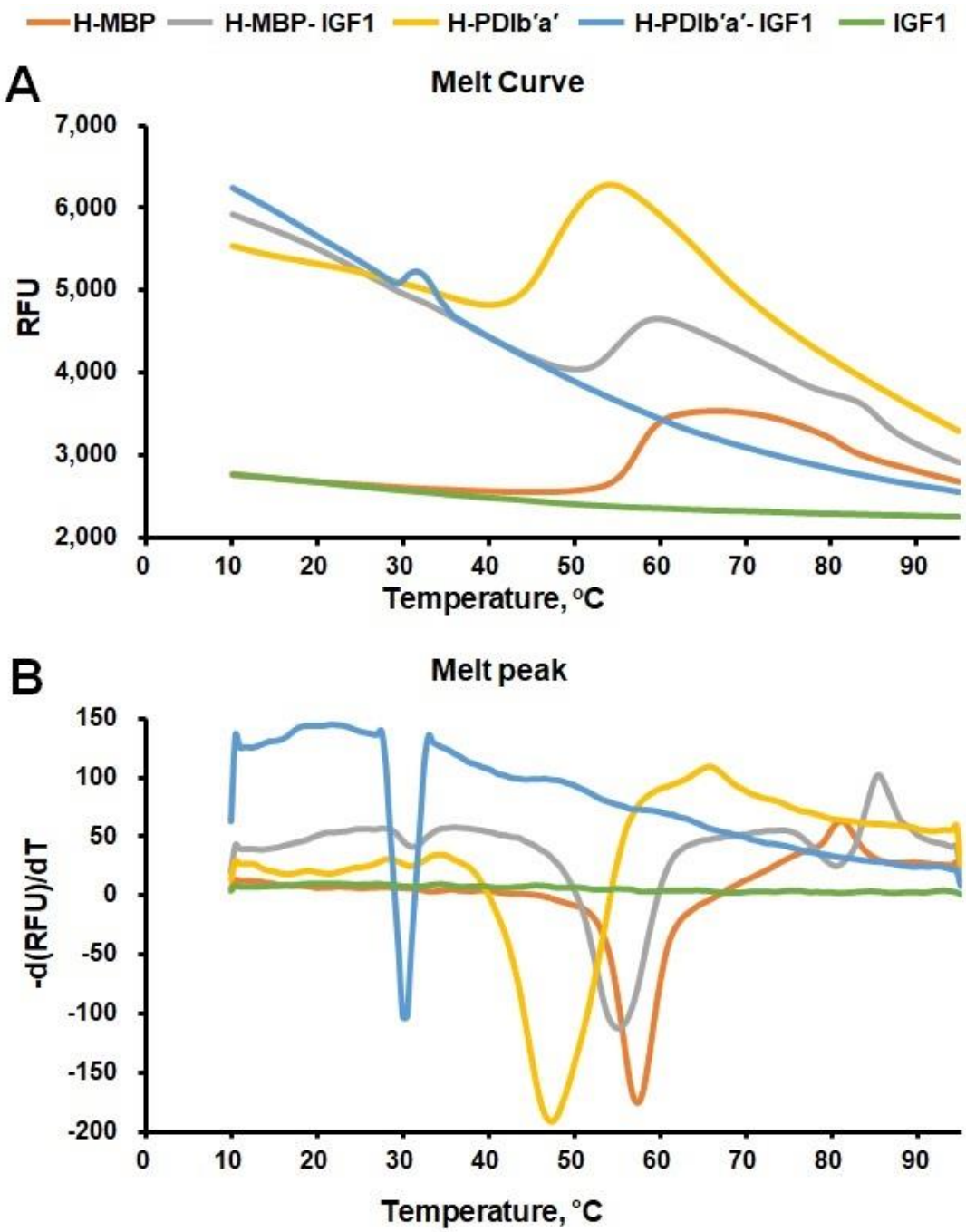


Figure 35. Melting temperature analysis of IGF1 and related proteins.

(A) Melting curve and (B) melting peak of IGF1, H-MBP-IGF1, H-PDIb′a′-IGF1, and protein tags in PBS buffer, with each data point representing the mean of triplicates. RFU, relative fluorescence units.

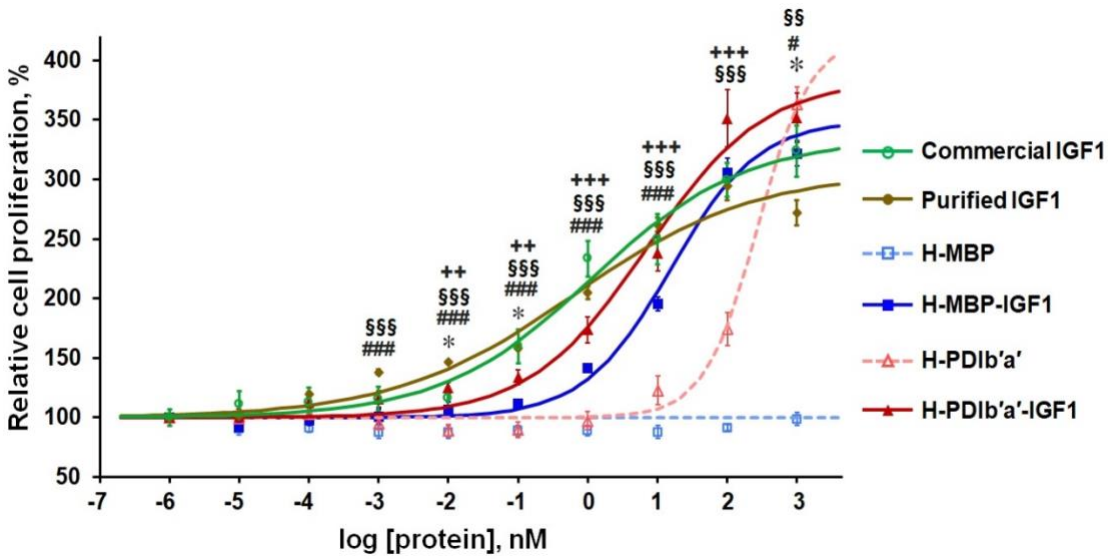


Figure 36. Proliferation activity of IGF1 and related proteins in MCF-7 cells

MCF-7 cell proliferation was evaluated after 72-h exposure to graded concentrations of commercial IGF1 and purified proteins: IGF1, H-PDIb′a′-IGF1, H-MBP-IGF1, H-PDIb'a', and H-MBP. Baseline proliferation, without any protein, was designated 100%, and results were normalized accordingly (n = 9). Significance was ascertained using one-way ANOVA, with Bonferroni's post-hoc test. Significance markers: *****, Purified IGF1 vs. H-PDIb′a′-IGF1 ($p < 0.05$, ** $p < 0.01$, *** $p < 0.001$); **#**, Purified IGF1 vs. H-MBP-IGF1 (# $p < 0.05$, ## $p < 0.01$, ### $p < 0.001$); **+**, H-PDIb'a' vs. H-PDIb′a′-IGF1 (+ $p < 0.05$, ++ $p < 0.01$, +++ $p < 0.001$); **§**, Purified IGF1 vs. H-PDIb'a' (§ $p < 0.05$, §§ $p < 0.01$, §§§ < 0.001).

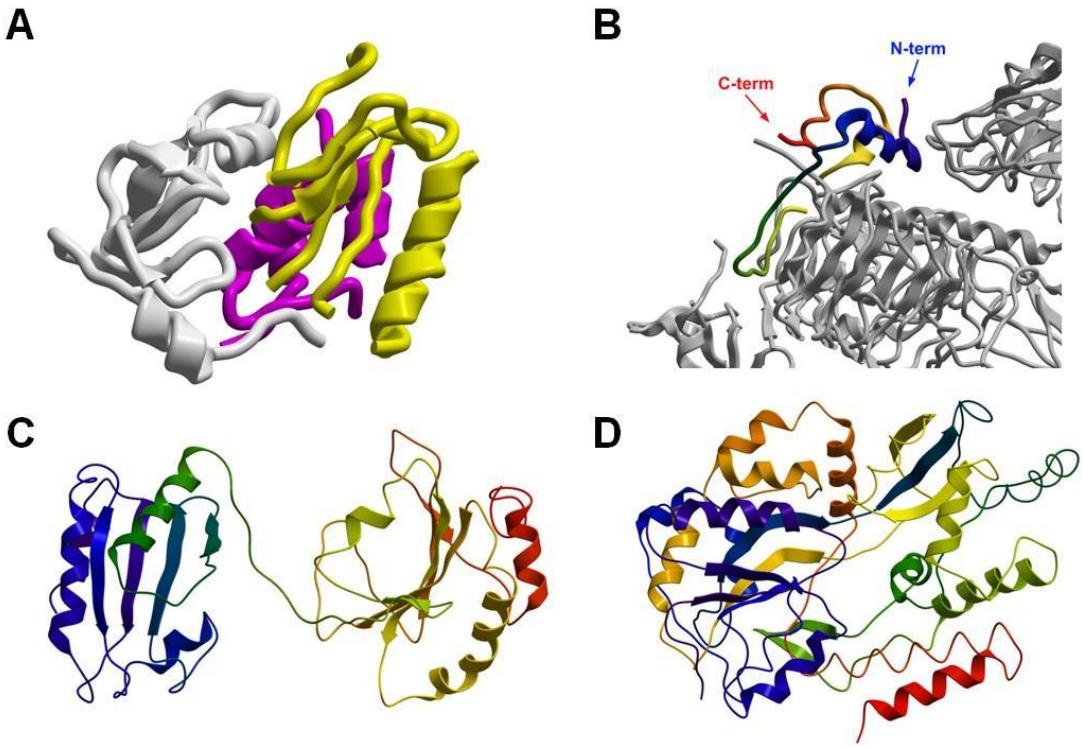


Figure 37. Structural representations of IGF1, IGF1R complex, PDIb'a', and MBP.

(A) Crystal structure showcasing the IGF1 and IGFBP complex with IGF1 in magenta, IGFBP1 in yellow, and IGFBP4 in gray. PDB code: 2DSQ. (B) Structure of the IGF1 and IGF1R complex. Arrows highlight the N-terminus and C-terminus of IGF1, while IGF1R is depicted in gray. PDB code: 6PYH. (C) Structural representation of PDIb'a' where a flexible linker interconnects the b' and a' domains. PDB code: 5CRW. (D) MBP structure with the corresponding PDB code: 1ANF.

References

1. Industrial Enzymes Market. https://www.marketsandmarkets.com/Market-Reports/industrial-enzymes-market-237327836.html. Published on Jun 2023.
2. Metcalf, D., *Hematopoietic cytokines.* Blood, 2008. **111**(2): p. 485-91.
3. Zhang, Z.-X., et al., *Strategies for efficient production of recombinant proteins in Escherichia coli: alleviating the host burden and enhancing protein activity.* Microbial Cell Factories, 2022. **21**(1): p. 191.
4. Meng, H., B.D. Burleigh, and G.M. Kelly, *Reduction studies on bacterial recombinant somatomedin C/insulin-like growth factor-1.* J Chromatogr, 1988. **443**: p. 183-92.
5. Swartz, J., *Developing cell-free biology for industrial applications.* J Ind Microbiol Biotechnol, 2006. **33**(7): p. 476-85.
6. Obukowicz, M.G., et al., *Secretion and export of IGF-1 in Escherichia coli strain JM101.* Mol Gen Genet, 1988. **215**(1): p. 19-25.
7. Wong, E.Y., et al., *Expression of secreted insulin-like growth factor-1 in Escherichia coli.* Gene, 1988. **68**(2): p. 193-203.
8. Becher, B., S. Tugues, and M. Greter, *GM-CSF: From Growth Factor to Central Mediator of Tissue Inflammation.* Immunity, 2016. **45**(5): p. 963-973.
9. Lee, K.M.C., A.A. Achuthan, and J.A. Hamilton, *GM-CSF: A Promising Target in Inflammation and Autoimmunity.* Immunotargets Ther, 2020. **9**: p. 225-240.
10. Dranoff, G., et al., *Vaccination with irradiated tumor cells engineered to secrete murine granulocyte-macrophage colony-stimulating factor stimulates potent, specific, and long-lasting anti-tumor immunity.* Proc Natl Acad Sci U S A, 1993. **90**(8): p. 3539-43.
11. Lawson, D.H., et al., *Randomized, Placebo-Controlled, Phase III Trial of Yeast-Derived Granulocyte-Macrophage Colony-Stimulating Factor (GM-CSF) Versus Peptide Vaccination Versus GM-CSF Plus Peptide Vaccination Versus Placebo in Patients With No Evidence of Disease After Complete Surgical Resection of Locally Advanced and/or Stage IV Melanoma: A Trial of the Eastern Cooperative Oncology Group-American College of Radiology Imaging Network Cancer Research Group (E4697).* J Clin Oncol, 2015. **33**(34): p. 4066-76.

12. Markham, A., *Naxitamab: First Approval.* Drugs, 2021. **81**(2): p. 291-296.
13. Krüger, C., et al., *The hematopoietic factor GM-CSF (granulocyte-macrophage colony-stimulating factor) promotes neuronal differentiation of adult neural stem cells in vitro.* BMC Neurosci, 2007. **8**: p. 88.
14. Schwanke, R.C., et al., *Molecular cloning, expression in Escherichia coli and production of bioactive homogeneous recombinant human granulocyte and macrophage colony stimulating factor.* Int J Biol Macromol, 2009. **45**(2): p. 97-102.
15. Waller, E.K., *The role of sargramostim (rhGM-CSF) as immunotherapy.* Oncologist, 2007. **12 Suppl 2**: p. 22-6.
16. Correale, P., et al., *Recruitment of dendritic cells and enhanced antigen-specific immune reactivity in cancer patients treated with hr-GM-CSF (Molgramostim) and hr-IL-2. results from a phase Ib clinical trial.* Eur J Cancer, 2001. **37**(7): p. 892-902.
17. Moonen, P., et al., *Increased biological activity of deglycosylated recombinant human granulocyte/macrophage colony-stimulating factor produced by yeast or animal cells.* Proc Natl Acad Sci U S A, 1987. **84**(13): p. 4428-31.
18. Cebon, J., et al., *Granulocyte-macrophage colony stimulating factor from human lymphocytes. The effect of glycosylation on receptor binding and biological activity.* J Biol Chem, 1990. **265**(8): p. 4483-91.
19. Burgess, A.W., et al., *Purification and properties of bacterially synthesized human granulocyte-macrophage colony stimulating factor.* Blood, 1987. **69**(1): p. 43-51.
20. Doherty, D.H., et al., *Site-specific PEGylation of engineered cysteine analogues of recombinant human granulocyte-macrophage colony-stimulating factor.* Bioconjug Chem, 2005. **16**(5): p. 1291-8.
21. Thomson, C.A., et al., *A simplified method for the efficient refolding and purification of recombinant human GM-CSF.* PLoS One, 2012. **7**(11): p. e49891.
22. Malekian, R., et al., *High-yield Production of Granulocyte-macrophage Colony-stimulating Factor in E. coli BL21 (DE3) By an Auto-induction Strategy.* Iran J Pharm Res, 2019. **18**(1): p. 469-478.

23. Taherian, E., et al., *Cloning, Optimization of Periplasmic Expression and Purification of Recombinant Granulocyte Macrophage-Stimulating Factor in Escherichia coli BL21 (DE3).* Adv Biomed Res, 2019. **8**: p. 71.
24. Soheili, S., A. Jahanian-Najafabadi, and V. Akbari, *Evaluation of soluble expression of recombinant granulocyte macrophage stimulating factor (rGM-CSF) by three different E. coli strains.* Res Pharm Sci, 2020. **15**(3): p. 218-225.
25. Malekian, R., et al., *Improvement of soluble expression of GM-CSF in the cytoplasm of Escherichia coli using chemical and molecular chaperones.* Protein Expr Purif, 2019. **160**: p. 66-72.
26. Das, K.M., et al., *Cloning, soluble expression and purification of high yield recombinant hGMCSF in Escherichia coli.* Int J Mol Sci, 2011. **12**(3): p. 2064-76.
27. Srinivasa Babu, K., et al., *Single step intein-mediated purification of hGMCSF expressed in salt-inducible E. coli.* Biotechnol Lett, 2009. **31**(5): p. 659-64.
28. Dougan, M., G. Dranoff, and S.K. Dougan, *GM-CSF, IL-3, and IL-5 Family of Cytokines: Regulators of Inflammation.* Immunity, 2019. **50**(4): p. 796-811.
29. Nitsche, A., et al., *Interleukin-3 promotes proliferation and differentiation of human hematopoietic stem cells but reduces their repopulation potential in NOD/SCID mice.* Stem Cells, 2003. **21**(2): p. 236-44.
30. Ren, Z., et al., *Influence of IL-3 functional fragment on cord blood stem cell ex vivo expansion and differentiation.* Stem Cell Investig, 2016. **3**: p. 6.
31. Ackermann, M., et al., *A 3D iPSC-differentiation model identifies interleukin-3 as a regulator of early human hematopoietic specification.* Haematologica, 2021. **106**(5): p. 1354-1367.
32. Broughton, S.E., et al., *A dual role for the N-terminal domain of the IL-3 receptor in cell signalling.* Nat Commun, 2018. **9**(1): p. 386.
33. Yang, Y.C., et al., *Human IL-3 (multi-CSF): identification by expression cloning of a novel hematopoietic growth factor related to murine IL-3.* Cell, 1986. **47**(1): p. 3-10.
34. Ding, H., et al., *Molecular cloning, expression, purification, and characterization of soluble full-length, human interleukin-3 with a baculovirus-insect cell expression system.* Protein Expr Purif, 2003. **31**(1): p. 34-41.

35. Dagar, V.K., et al., *Bioprocess development for extracellular production of recombinant human interleukin-3 (hIL-3) in Pichia pastoris.* J Ind Microbiol Biotechnol, 2016. **43**(10): p. 1373-86.
36. Dagar, V.K. and Y.P. Khasa, *Combined effect of gene dosage and process optimization strategies on high-level production of recombinant human interleukin-3 (hIL-3) in Pichia pastoris fed-batch culture.* Int J Biol Macromol, 2018. **108**: p. 999-1009.
37. Li, H., et al., *High level expression of active recombinant human interleukin-3 in Pichia pastoris.* Protein Expr Purif, 2011. **80**(2): p. 185-93.
38. Lokker, N.A., et al., *Structure-activity relationship study of human interleukin-3. Identification of residues required for biological activity by site-directed mutagenesis.* J Biol Chem, 1991. **266**(16): p. 10624-31.
39. Boyle, D.M., et al., *Evaluation of refolding conditions for a recombinant human interleukin-3 variant (daniplestim).* Biotechnol Appl Biochem, 1999. **30**(2): p. 163-70.
40. Dagar, V.K., Adivitiya, and Y.P. Khasa, *High-level expression and efficient refolding of therapeutically important recombinant human Interleukin-3 (hIL-3) in E. coli.* Protein Expr Purif, 2017. **131**: p. 51-59.
41. Westers, L., et al., *Secretion of functional human interleukin-3 from Bacillus subtilis.* J Biotechnol, 2006. **123**(2): p. 211-24.
42. Nowruzi, K., et al., *Development of a minimal defined medium for recombinant human interleukin-3 production by Streptomyces lividans 66.* Biotechnol Bioeng, 2008. **99**(1): p. 214-22.
43. Davis, G.D., et al., *New fusion protein systems designed to give soluble expression in Escherichia coli.* Biotechnol Bioeng, 1999. **65**(4): p. 382-8.
44. Hercus, T.R., et al., *High yield production of a soluble human interleukin-3 variant from E. coli with wild-type bioactivity and improved radiolabeling properties.* PLoS One, 2013. **8**(8): p. e74376.
45. Febbraio, M.A. and B.K. Pedersen, *Contraction-induced myokine production and release: is skeletal muscle an endocrine organ?* Exerc Sport Sci Rev, 2005. **33**(3): p. 114-9.

46. Grenett, H.E., et al., *Isolation and characterization of biologically active murine interleukin-6 produced in Escherichia coli.* Gene, 1991. **101**(2): p. 267-71.
47. Spiridonova, V.A., et al., *Preparation of functionally active recombinant human interleukin-6.* Biochemistry (Mosc), 2007. **72**(4): p. 424-9.
48. Dunstan, S.J., A.J. Ramsay, and R.A. Strugnell, *Studies of immunity and bacterial invasiveness in mice given a recombinant salmonella vector encoding murine interleukin-6.* Infect Immun, 1996. **64**(7): p. 2730-6.
49. Li, Y., et al., *Cloning and hemolysin-mediated secretory expression of a codon-optimized synthetic human interleukin-6 gene in Escherichia coli.* Protein Expr Purif, 2002. **25**(3): p. 437-47.
50. Kim, T.W., B.H. Chung, and Y.K. Chang, *Production of soluble human interleukin-6 in cytoplasm by fed-batch culture of recombinant E. coli.* Biotechnol Prog, 2005. **21**(2): p. 524-31.
51. Lee, E.G., et al., *Efficient proteolytic cleavage by insertion of oligopeptide linkers and its application to production of recombinant human interleukin-6 in Escherichia coli.* Enzyme and Microbial Technology, 2009. **44**(5): p. 254-262.
52. Nausch, H., et al., *Recombinant production of human interleukin 6 in Escherichia coli.* PLoS One, 2013. **8**(1): p. e54933.
53. Jang, J.H., F. Wang, and M. Kan, *Heparan sulfate is required for interaction and activation of the epithelial cell fibroblast growth factor receptor-2IIIb with stromal-derived fibroblast growth factor-7.* In Vitro Cell Dev Biol Anim, 1997. **33**(10): p. 819-24.
54. Zhang, X., et al., *Receptor specificity of the fibroblast growth factor family. The complete mammalian FGF family.* J Biol Chem, 2006. **281**(23): p. 15694-700.
55. Ornitz, D.M., et al., *Receptor specificity of the fibroblast growth factor family.* J Biol Chem, 1996. **271**(25): p. 15292-7.
56. Alpdogan, O., et al., *Keratinocyte growth factor (KGF) is required for postnatal thymic regeneration.* Blood, 2006. **107**(6): p. 2453-60.
57. Hille, A., et al., *Effect of tumour-cell-derived or recombinant keratinocyte growth factor (KGF) on proliferation and radioresponse of human epithelial tumour cells (HNSCC) and normal keratinocytes in vitro.* Radiat Environ Biophys, 2010. **49**(2): p. 261-70.

58. Ornitz, D.M. and N. Itoh, *Fibroblast growth factors.* Genome Biol, 2001. **2**(3): p. Reviews3005.
59. Beenken, A. and M. Mohammadi, *The FGF family: biology, pathophysiology and therapy.* Nat Rev Drug Discov, 2009. **8**(3): p. 235-53.
60. Yun, Y.R., et al., *Fibroblast growth factors: biology, function, and application for tissue regeneration.* J Tissue Eng, 2010. **2010**: p. 218142.
61. Zinkle, A. and M. Mohammadi, *Structural Biology of the FGF7 Subfamily.* Front Genet, 2019. **10**: p. 102.
62. Hsu, Y.R., et al., *Human keratinocyte growth factor recombinantly expressed in Chinese hamster ovary cells: isolation of isoforms and characterization of post-translational modifications.* Protein Expr Purif, 1998. **12**(2): p. 189-200.
63. Ron, D., et al., *Expression of biologically active recombinant keratinocyte growth factor. Structure/function analysis of amino-terminal truncation mutants.* J Biol Chem, 1993. **268**(4): p. 2984-8.
64. Nybo, R.E., E.N. Everton, and C.F. Morris, *Mitogenic activity of keratinocyte growth factor amino-terminal truncation mutants: deletion of amino acid residues 1-15 through 1-27.* In Vitro Cell Dev Biol Anim, 1997. **33**(8): p. 606-7.
65. Hsu, E., et al., *Enhanced stability of recombinant keratinocyte growth factor by mutagenesis.* Protein Eng Des Sel, 2006. **19**(4): p. 147-53.
66. Spielberger, R., et al., *Palifermin for oral mucositis after intensive therapy for hematologic cancers.* N Engl J Med, 2004. **351**(25): p. 2590-8.
67. Yoshihara, E., et al., *Immune-evasive human islet-like organoids ameliorate diabetes.* Nature, 2020. **586**(7830): p. 606-611.
68. Feng, Z.G., et al., *Recombinant keratinocyte growth factor 1 in tobacco potentially promotes wound healing in diabetic rats.* Biomed Res Int, 2014. **2014**: p. 579632.
69. Xue, P., et al., *Highly efficient expression of functional recombinant human keratinocyte growth factor 1 and its protective effects on hepatocytes.* Appl Microbiol Biotechnol, 2014. **98**(9): p. 3933-45.
70. Bare, L.A., et al., *Effect of cysteine substitutions on the mitogenic activity and stability of recombinant human keratinocyte growth factor.* Biochem Biophys Res Commun, 1994. **205**(1): p. 872-9.

71. Luo, Y., et al., *Improved production of recombinant fibroblast growth factor 7 (FGF7/KGF) from bacteria in high magnesium chloride.* Protein Expr Purif, 2004. **33**(2): p. 326-31.
72. Wu, X., et al., *Expression and purification of human keratinocyte growth factor 2 by fusion with SUMO.* Mol Biotechnol, 2009. **42**(1): p. 68-74.
73. Kim, Y.S., et al., *Production of a 135-residue long N-truncated human keratinocyte growth factor 1 in Escherichia coli.* Microb Cell Fact, 2023. **22**(1): p. 98.
74. Kim, Y.S., et al., *High-level production of keratinocyte growth factor 2 in Escherichia coli.* Protein Expr Purif, 2023. **204**: p. 106229.
75. Bahadori, Z., H.R. Kalhor, and S.J. Mowla, *Producing functional recombinant human keratinocyte growth factor in Pichia pastoris and investigating its protective role against irradiation.* Enzyme Microb Technol, 2018. **111**: p. 12-20.
76. Dudognon, B., et al., *Production of functional active human growth factors in insects used as living biofactories.* J Biotechnol, 2014. **184**: p. 229-39.
77. LeRoith, D., J.M.P. Holly, and B.E. Forbes, *Insulin-like growth factors: Ligands, binding proteins, and receptors.* Mol Metab, 2021. **52**: p. 101245.
78. Raschdorf, F., et al., *Location of disulphide bonds in human insulin-like growth factors (IGFs) synthesized by recombinant DNA technology.* Biomed Environ Mass Spectrom, 1988. **16**(1-12): p. 3-8.
79. Argon, Y., et al., *Glucose-Regulated Protein 94 (GRP94): A Novel Regulator of Insulin-Like Growth Factor Production.* Cells, 2020. **9**(8).
80. Francis, G.L., et al., *Purification and partial sequence analysis of insulin-like growth factor-1 from bovine colostrum.* Biochem J, 1986. **233**(1): p. 207-13.
81. Guler, H.P., et al., *Insulin-like growth factors I and II in healthy man. Estimations of half-lives and production rates.* Acta Endocrinol (Copenh), 1989. **121**(6): p. 753-8.
82. Baxter, R.C., J.L. Martin, and V.A. Beniac, *High molecular weight insulin-like growth factor binding protein complex. Purification and properties of the acid-labile subunit from human serum.* J Biol Chem, 1989. **264**(20): p. 11843-8.
83. Scharf, J., et al., *Synthesis of insulinlike growth factor binding proteins and of the acid-labile subunit in primary cultures of rat hepatocytes, of Kupffer cells,*

and in cocultures: regulation by insulin, insulinlike growth factor, and growth hormone. Hepatology, 1996. **23**(4): p. 818-27.

84. Ueki, I., et al., *Inactivation of the acid labile subunit gene in mice results in mild retardation of postnatal growth despite profound disruptions in the circulating insulin-like growth factor system.* Proc Natl Acad Sci U S A, 2000. **97**(12): p. 6868-73.
85. Sitar, T., et al., *Structural basis for the inhibition of insulin-like growth factors by insulin-like growth factor-binding proteins.* Proc Natl Acad Sci U S A, 2006. **103**(35): p. 13028-33.
86. Bendall, S.C., et al., *IGF and FGF cooperatively establish the regulatory stem cell niche of pluripotent human cells in vitro.* Nature, 2007. **448**(7157): p. 1015-21.
87. Wang, L., et al., *Self-renewal of human embryonic stem cells requires insulin-like growth factor-1 receptor and ERBB2 receptor signaling.* Blood, 2007. **110**(12): p. 4111-9.
88. Jin, S. and J. Cheng, *Insulin-like Growth Factor-1 (IGF-1) Related Drugs in Pain Management.* Pharmaceuticals (Basel), 2023. **16**(5).
89. Wolf, E., et al., *Human insulin-like growth factor I (IGF-I) produced in the mammary glands of transgenic rabbits: yield, receptor binding, mitogenic activity, and effects on IGF-binding proteins.* Endocrinology, 1997. **138**(1): p. 307-13.
90. Panahi, M., et al., *Recombinant protein expression plasmids optimized for industrial E. coli fermentation and plant systems produce biologically active human insulin-like growth factor-1 in transgenic rice and tobacco plants.* Transgenic Res, 2004. **13**(3): p. 245-59.
91. Vai, M., et al., *Improved secretion of native human insulin-like growth factor 1 from gas1 mutant Saccharomyces cerevisiae cells.* Appl Environ Microbiol, 2000. **66**(12): p. 5477-9.
92. Gellerfors, P., et al., *Isolation and characterization of a glycosylated form of human insulin-like growth factor I produced in Saccharomyces cerevisiae.* J Biol Chem, 1989. **264**(19): p. 11444-9.

93. Forsberg, G., et al., *Separation and characterization of modified variants of recombinant human insulin-like growth factor I derived from a fusion protein secreted from Escherichia coli.* Biochem J, 1990. **271**(2): p. 357-63.
94. Nilsson, B., G. Forsberg, and M. Hartmanis, *Expression and purification of recombinant insulin-like growth factors from Escherichia coli.* Methods Enzymol, 1991. **198**: p. 3-16.
95. King, R., et al., *Production and characterization of recombinant insulin-like growth factor-I (IGF-I) and potent analogues of IGF-I, with Gly or Arg substituted for Glu3, following their expression in Escherichia coli as fusion proteins.* J Mol Endocrinol, 1992. **8**(1): p. 29-41.
96. Kim, S.O. and Y.I. Lee, *High-level expression and simple purification of recombinant human insulin-like growth factor I.* J Biotechnol, 1996. **48**(1-2): p. 97-105.
97. Zhang, Y., et al., *Expression of eukaryotic proteins in soluble form in Escherichia coli.* Protein Expr Purif, 1998. **12**(2): p. 159-65.
98. Aboutalebi, F., et al., *An efficient method for bacterial production and activity assessment of recombinant human insulin like growth factor 1.* Mol Biol Rep, 2018. **45**(6): p. 1957-1966.
99. Kim, Y.S., et al., *Effective production of human growth factors in Escherichia coli by fusing with small protein 6HFh8.* Microb Cell Fact, 2021. **20**(1): p. 9.
100. Lee, E., et al., *Novel Bacterial Production of Two Different Bioactive Forms of Human Stem-Cell Factor.* Int J Mol Sci, 2021. **22**(12).
101. Nguyen, A.N., et al., *Prokaryotic soluble expression and purification of bioactive human fibroblast growth factor 21 using maltose-binding protein.* Sci Rep, 2017. **7**(1): p. 16139.
102. Vu, T.T., et al., *Soluble Prokaryotic Expression and Purification of Human Interferon Alpha-2b Using a Maltose-Binding Protein Tag.* J Mol Microbiol Biotechnol, 2016. **26**(6): p. 359-368.
103. Nguyen, M.T., et al., *Prokaryotic Soluble Overexpression and Purification of Human VEGF165 by Fusion to a Maltose Binding Protein Tag.* PLoS One, 2016. **11**(5): p. e0156296.

104. Nguyen, M.T., et al., *Prokaryotic soluble overexpression and purification of oncostatin M using a fusion approach and genetically engineered E. coli strains.* Sci Rep, 2019. **9**(1): p. 13706.
105. Nguyen, M.T., et al., *Bacterial overexpression and purification of soluble recombinant human serum albumin using maltose-binding protein and protein disulphide isomerase.* Protein Expr Purif, 2020. **167**: p. 105530.
106. Nguyen, T.K.O., et al., *Soluble Prokaryotic Overexpression and Purification of Human GM-CSF Using the Protein Disulfide Isomerase b'a' Domain.* Int J Mol Sci, 2021. **22**(10).
107. Karey, K.P. and D.A. Sirbasku, *Differential responsiveness of human breast cancer cell lines MCF-7 and T47D to growth factors and 17 beta-estradiol.* Cancer Res, 1988. **48**(14): p. 4083-92.
108. Ogasawara, M. and D.A. Sirbasku, *A new serum-free method of measuring growth factor activities for human breast cancer cells in culture.* In Vitro Cell Dev Biol, 1988. **24**(9): p. 911-20.
109. Vang Mouritzen, M. and H. Jenssen, *Optimized Scratch Assay for In Vitro Testing of Cell Migration with an Automated Optical Camera.* J Vis Exp, 2018(138).
110. Suarez-Arnedo, A., et al., *An image J plugin for the high throughput image analysis of in vitro scratch wound healing assays.* PLoS One, 2020. **15**(7): p. e0232565.
111. Do, B.H., et al., *Soluble Prokaryotic Expression and Purification of Bioactive Tumor Necrosis Factor-Related Apoptosis-Inducing Ligand.* J Microbiol Biotechnol, 2017. **27**(12): p. 2156-2164.
112. Do, B.H., et al., *Soluble prokaryotic overexpression and purification of bioactive human granulocyte colony-stimulating factor by maltose binding protein and protein disulfide isomerase.* PLoS One, 2014. **9**(3): p. e89906.
113. Vu, T.T., et al., *Soluble prokaryotic expression and purification of crotamine using an N-terminal maltose-binding protein tag.* Toxicon, 2014. **92**: p. 157-65.
114. Vu, T.T., et al., *Soluble overexpression and purification of bioactive human CCL2 in E. coli by maltose-binding protein.* Mol Biol Rep, 2015. **42**(3): p. 651-63.

115. Nguyen, M.T., et al., *Prokaryotic soluble overexpression and purification of bioactive human growth hormone by fusion to thioredoxin, maltose binding protein, and protein disulfide isomerase.* PLoS One, 2014. **9**(3): p. e89038.
116. Kapust, R.B. and D.S. Waugh, *Escherichia coli maltose-binding protein is uncommonly effective at promoting the solubility of polypeptides to which it is fused.* Protein Sci, 1999. **8**(8): p. 1668-74.
117. Waugh, D.S., *The remarkable solubility-enhancing power of Escherichia coli maltose-binding protein.* Postepy Biochem, 2016. **62**(3): p. 377-382.
118. Nallamsetty, S. and D.S. Waugh, *Mutations that alter the equilibrium between open and closed conformations of Escherichia coli maltose-binding protein impede its ability to enhance the solubility of passenger proteins.* Biochem Biophys Res Commun, 2007. **364**(3): p. 639-44.
119. Kozlov, G., et al., *A structural overview of the PDI family of proteins.* Febs j, 2010. **277**(19): p. 3924-36.
120. Pirneskoski, A., et al., *Domains b' and a' of protein disulfide isomerase fulfill the minimum requirement for function as a subunit of prolyl 4-hydroxylase. The N-terminal domains a and b enhances this function and can be substituted in part by those of ERp57.* J Biol Chem, 2001. **276**(14): p. 11287-93.
121. Nguyen, V.D., et al., *Alternative conformations of the x region of human protein disulphide-isomerase modulate exposure of the substrate binding b' domain.* J Mol Biol, 2008. **383**(5): p. 1144-55.
122. Ren, G., N. Ke, and M. Berkmen, *Use of the SHuffle Strains in Production of Proteins.* Curr Protoc Protein Sci, 2016. **85**: p. 5.26.1-5.26.21.
123. Sørensen, H.P. and K.K. Mortensen, *Soluble expression of recombinant proteins in the cytoplasm of Escherichia coli.* Microb Cell Fact, 2005. **4**(1): p. 1.
124. Shaw, M.K. and J.L. Ingraham, *Synthesis of macromolecules by Escherichia coli near the minimal temperature for growth.* J Bacteriol, 1967. **94**(1): p. 157-64.
125. Jacobs, P.P., et al., *Fed-batch fermentation of GM-CSF-producing glycoengineered Pichia pastoris under controlled specific growth rate.* Microb Cell Fact, 2010. **9**: p. 93.

126. Bordier, C., *Phase separation of integral membrane proteins in Triton X-114 solution.* J Biol Chem, 1981. **256**(4): p. 1604-7.
127. Aida, Y. and M.J. Pabst, *Removal of endotoxin from protein solutions by phase separation using Triton X-114.* J Immunol Methods, 1990. **132**(2): p. 191-5.
128. Liu, S., et al., *Removal of endotoxin from recombinant protein preparations.* Clin Biochem, 1997. **30**(6): p. 455-63.
129. Magalhães, P.O., et al., *Methods of endotoxin removal from biological preparations: a review.* J Pharm Pharm Sci, 2007. **10**(3): p. 388-404.
130. Mack, L., et al., *Endotoxin depletion of recombinant protein preparations through their preferential binding to histidine tags.* Anal Biochem, 2014. **466**: p. 83-8.
131. London, A.S., et al., *A novel method to determine residual detergent in biological samples post endotoxin reduction treatment and evaluation of strategies for subsequent detergent removal.* Int Immunopharmacol, 2016. **37**: p. 16-22.
132. Boratyński, J. and B. Szermer-Olearnik, *Endotoxin Removal from Escherichia coli Bacterial Lysate Using a Biphasic Liquid System.* Methods Mol Biol, 2017. **1600**: p. 107-112.
133. Mamat, U., et al., *Detoxifying Escherichia coli for endotoxin-free production of recombinant proteins.* Microb Cell Fact, 2015. **14**: p. 57.
134. Jong, W.S., et al., *Application of an E. coli signal sequence as a versatile inclusion body tag.* Microb Cell Fact, 2017. **16**(1): p. 50.
135. Sui, X., et al., *gp130 and c-Kit signalings synergize for ex vivo expansion of human primitive hemopoietic progenitor cells.* Proc Natl Acad Sci U S A, 1995. **92**(7): p. 2859-63.
136. Testa, U., et al., *Elevated expression of IL-3Ralpha in acute myelogenous leukemia is associated with enhanced blast proliferation, increased cellularity, and poor prognosis.* Blood, 2002. **100**(8): p. 2980-8.
137. Didichenko, S.A., et al., *IL-3 induces a Pim1-dependent antiapoptotic pathway in primary human basophils.* Blood, 2008. **112**(10): p. 3949-58.

138. Brizzi, M.F., et al., *Interleukin 3 stimulates proliferation and triggers endothelial-leukocyte adhesion molecule 1 gene activation of human endothelial cells.* J Clin Invest, 1993. **91**(6): p. 2887-92.
139. Urdal, D.L., et al., *Molecular characterization of colony-stimulating factors and their receptors: human interleukin-3.* Ann N Y Acad Sci, 1989. **554**: p. 167-76.
140. Budel, L.M., et al., *Common binding structure for granulocyte macrophage colony-stimulating factor and interleukin-3 on human acute myeloid leukemia cells and monocytes.* Blood, 1990. **75**(7): p. 1439-45.
141. Park, L.S., et al., *Heterogeneity in human interleukin-3 receptors. A subclass that binds human granulocyte/macrophage colony stimulating factor.* J Biol Chem, 1989. **264**(10): p. 5420-7.
142. Lopez, A.F., et al., *Reciprocal inhibition of binding between interleukin 3 and granulocyte-macrophage colony-stimulating factor to human eosinophils.* Proc Natl Acad Sci U S A, 1989. **86**(18): p. 7022-6.
143. Broughton, S.E., et al., *The βc receptor family - Structural insights and their functional implications.* Cytokine, 2015. **74**(2): p. 247-58.
144. Robb, L., et al., *Hematopoietic and lung abnormalities in mice with a null mutation of the common beta subunit of the receptors for granulocyte-macrophage colony-stimulating factor and interleukins 3 and 5.* Proc Natl Acad Sci U S A, 1995. **92**(21): p. 9565-9.
145. Muto, A., et al., *High affinity chimeric human granulocyte-macrophage colony-stimulating factor receptor carrying the cytoplasmic domain of the beta subunit but not the alpha subunit transduces growth promoting signals in Ba/F3 cells.* Biochem Biophys Res Commun, 1995. **208**(1): p. 368-75.
146. Park, L.S., et al., *Interleukin-3, GM-CSF, and G-CSF receptor expression on cell lines and primary leukemia cells: receptor heterogeneity and relationship to growth factor responsiveness.* Blood, 1989. **74**(1): p. 56-65.
147. Gesner, T.G., et al., *Specific binding, internalization, and degradation of human recombinant interleukin-3 by cells of the acute myelogenous, leukemia line, KG-1.* J Cell Physiol, 1988. **136**(3): p. 493-9.

148. Elliott, M.J., et al., *Recombinant human interleukin-3 and granulocyte-macrophage colony-stimulating factor show common biological effects and binding characteristics on human monocytes.* Blood, 1989. **74**(7): p. 2349-59.
149. Lopez, A.F., et al., *Human interleukin-3 inhibits the binding of granulocyte-macrophage colony-stimulating factor and interleukin-5 to basophils and strongly enhances their functional activity.* J Cell Physiol, 1990. **145**(1): p. 69-77.
150. Taketazu, F., et al., *IL-3 specifically inhibits GM-CSF binding to the higher affinity receptor.* J Cell Physiol, 1991. **146**(2): p. 251-7.
151. Cardoso, V.M., et al., *ClearColi as a platform for untagged pneumococcal surface protein A production: cultivation strategy, bioreactor culture, and purification.* Appl Microbiol Biotechnol, 2022. **106**(3): p. 1011-1029.
152. Ahmed, N., et al., *Matrix-assisted refolding and purification of placenta-derived recombinant human interleukin-6 produced in Escherichia coli.* Biotechnol Appl Biochem, 2014. **61**(5): p. 541-8.
153. Ejima, D., et al., *High yield refolding and purification process for recombinant human interleukin-6 expressed in Escherichia coli.* Biotechnol Bioeng, 1999. **62**(3): p. 301-10.
154. Barthelemy, I., et al., *Production and secretion of human interleukin 6 into the periplasm of Escherichia coli: efficient processing of N-terminal variants of hIL6 by the E. coli signal peptidase.* J Biotechnol, 1993. **27**(3): p. 307-16.
155. Zhang, J.G., et al., *Purification and characterization of a recombinant murine interleukin-6. Isolation of N- and C-terminally truncated forms.* Eur J Biochem, 1992. **207**(3): p. 903-13.
156. Arcone, R., et al., *Single-step purification and structural characterization of human interleukin-6 produced in Escherichia coli from a T7 RNA polymerase expression vector.* Eur J Biochem, 1991. **198**(3): p. 541-7.
157. Yasueda, H., et al., *High-level direct expression of semi-synthetic human interleukin-6 in Escherichia coli and production of N-terminus met-free product.* Biotechnology (N Y), 1990. **8**(11): p. 1036-40.
158. Ki, M.R. and S.P. Pack, *Fusion tags to enhance heterologous protein expression.* Appl Microbiol Biotechnol, 2020. **104**(6): p. 2411-2425.

159. Telmer, P.G. and B.H. Shilton, *Insights into the conformational equilibria of maltose-binding protein by analysis of high affinity mutants.* J Biol Chem, 2003. **278**(36): p. 34555-67.
160. Römer, R.A., et al., *The flexibility and dynamics of protein disulfide isomerase.* Proteins, 2016. **84**(12): p. 1776-1785.
161. Fox, J.D., R.B. Kapust, and D.S. Waugh, *Single amino acid substitutions on the surface of Escherichia coli maltose-binding protein can have a profound impact on the solubility of fusion proteins.* Protein Sci, 2001. **10**(3): p. 622-30.
162. Tian, G., et al., *The crystal structure of yeast protein disulfide isomerase suggests cooperativity between its active sites.* Cell, 2006. **124**(1): p. 61-73.
163. Yagi-Utsumi, M., T. Satoh, and K. Kato, *Structural basis of redox-dependent substrate binding of protein disulfide isomerase.* Sci Rep, 2015. **5**: p. 13909.
164. Terpe, K., *Overview of bacterial expression systems for heterologous protein production: from molecular and biochemical fundamentals to commercial systems.* Appl Microbiol Biotechnol, 2006. **72**(2): p. 211-22.
165. Arakawa, T., et al., *Aggregation Suppression of Proteins by Arginine During Thermal Unfolding.* Protein & Peptide Letters, 2006. **13**(9): p. 921-927.
166. Shiga, Y., et al., *Efficient production of N-terminally truncated biologically active human interleukin-6 by Bacillus brevis.* Biosci Biotechnol Biochem, 2000. **64**(3): p. 665-9.
167. Nordan, R.P., J.G. Pumphrey, and S. Rudikoff, *Purification and NH2-terminal sequence of a plasmacytoma growth factor derived from the murine macrophage cell line P388D1.* J Immunol, 1987. **139**(3): p. 813-7.
168. Gąciarz, A., et al., *Efficient soluble expression of disulfide bonded proteins in the cytoplasm of Escherichia coli in fed-batch fermentations on chemically defined minimal media.* Microb Cell Fact, 2017. **16**(1): p. 108.
169. Buchtova, M., et al., *Instability restricts signaling of multiple fibroblast growth factors.* Cell Mol Life Sci, 2015. **72**(12): p. 2445-59.
170. Kemmink, J., et al., *Structure determination of the N-terminal thioredoxin-like domain of protein disulfide isomerase using multidimensional heteronuclear 13C/15N NMR spectroscopy.* Biochemistry, 1996. **35**(24): p. 7684-91.

171. Tian, G., et al., *The catalytic activity of protein-disulfide isomerase requires a conformationally flexible molecule.* J Biol Chem, 2008. **283**(48): p. 33630-40.
172. Freedman, R.B., et al., *'Something in the way she moves': The functional significance of flexibility in the multiple roles of protein disulfide isomerase (PDI).* Biochim Biophys Acta Proteins Proteom, 2017. **1865**(11 Pt A): p. 1383-1394.
173. Peng, L., et al., *Probing the structure of human protein disulfide isomerase by chemical cross-linking combined with mass spectrometry.* J Proteomics, 2014. **108**: p. 1-16.
174. Flaumenhaft, R., *Advances in vascular thiol isomerase function.* Curr Opin Hematol, 2017. **24**(5): p. 439-445.
175. Wang, C., et al., *Structural insights into the redox-regulated dynamic conformations of human protein disulfide isomerase.* Antioxid Redox Signal, 2013. **19**(1): p. 36-45.
176. Yang, S., et al., *Compact conformations of human protein disulfide isomerase.* PLoS One, 2014. **9**(8): p. e103472.
177. Rahman, N.S.A., et al., *Functions and mechanisms of protein disulfide isomerase family in cancer emergence.* Cell Biosci, 2022. **12**(1): p. 129.
178. Freedman, R.B., P. Klappa, and L.W. Ruddock, *Protein disulfide isomerases exploit synergy between catalytic and specific binding domains.* EMBO Rep, 2002. **3**(2): p. 136-40.
179. Choi, S.P., et al., *Effects of L-arginine on refolding of lysine-tagged human insulin-like growth factor 1 expressed in Escherichia coli.* Bioprocess Biosyst Eng, 2012. **35**(1-2): p. 255-63.
180. Jafari, S., et al., *Recombinant production of mecasermin in E. coli expression system.* Res Pharm Sci, 2014. **9**(6): p. 453-61.
181. Iranpoor, H., et al., *Expression of Recombinant Human Insulin-like Growth Factor Type 1 (rhIGF-1) in Escherichia coli.* Avicenna J Med Biotechnol, 2015. **7**(3): p. 101-5.
182. Ranjbari, J., et al., *Enhanced Production of Insulin-like Growth Factor I Protein in Escherichia coli by Optimization of Five Key Factors.* Iran J Pharm Res, 2015. **14**(3): p. 907-17.

183. Emamipour, N., et al., *Soluble expression of IGF1 fused to DsbA in SHuffle™ T7 strain: optimization of expression and purification by Box-Behnken design.* Appl Microbiol Biotechnol, 2019. **103**(8): p. 3393-3406.
184. Vink-van Wijngaarden, T., et al., *Inhibition of insulin- and insulin-like growth factor-I-stimulated growth of human breast cancer cells by 1,25-dihydroxyvitamin D3 and the vitamin D3 analogue EB1089.* Eur J Cancer, 1996. **32a**(5): p. 842-8.
185. Biterova, E.I., et al., *The crystal structure of human microsomal triglyceride transfer protein.* Proc Natl Acad Sci U S A, 2019. **116**(35): p. 17251-17260.

국문요약

원핵 생물에서 여러 사이토카인과 성장인자의 효율적인 생산에 관한 연구

웬 티 키에우 오안

울산대학교

의과대학 대학원

재조합 단백질은 치료제, 백신, 산업 생산, 농업, 유전자 치료, 연구 및 개발 등 다양한 분야에서 광범위하게 적용되어 왔다. 대장균(Escherichia coli)은 재조합 인간 단백질을 생산하는 데 일반적으로 사용되는 발현 시스템이다. 그러나 단백질 잘못 접힘과 봉입체 형성은 종종 어려움을 야기하며, 후속 정제 단계를 복잡하게 만들었다. 이 연구는 대장균에서 잘못 접히고 봉입체를 형성하는 경향이 있는 것으로 알려진 다섯 가지 단백질인 인간 과립구-대식세포 집락 자극 인자(GM-CSF), 인터루킨-3(IL3), 인터루킨-6(IL6), 각질세포 성장 인자(KGF), 인슐린 유사 성장 인자 1(IGF1)을 조사했다.

이러한 어려움을 해결하기 위해 맥아당 결합 단백질(MBP)이나 단백질 이황화물 이성질체(PDIb'a')의 b' a' 도메인과 같은 가용화 태그를 융합 단백질의 N-말단에 도입했다. 이 접근법은 다양한 대장균 균주에서 발현될 때 발현 수준을 크게 높이고 용해도를 개선했다. 단백질 구성체는 높은 수준의 가용성 발현을 산출하여 재조합 단백질의 분리를 가능하게 했다. 정제된 단백질은 이후 생물학적 활성 분석을 통해 확인되고 검증되었다.

PDIb'a' 태그의 도입과 유전적으로 변형된 ClearColi 대장균 균주의 사용은 온화한 온도(30°C) 배양 조건과 결합되어 성공적인 단백질 생산에 중요한 역할을 하였으며 간단한 정제 절차를 가능하게 했다. 정제 과정은 친화성 크로마토그래피를 사용한 1 차 정제, TEV 프로테아제를 사용한 융합 단백질 절단, 이온 교환 크로마토그래피를 통한 정제된 단백질 획득의 세 가지 주요 단계를 포함했다. 흥미롭게도, 융합 단백질은 비융합 대응물에 비해 시험관 내 분석에서 더 높은 생물 활성을 나타냈다.

이 연구는 혁신적인 PDIb'a' 태그를 사용하여 ClearColi 대장균 균주에서 가용성 인간 재조합 단백질 GM-CSF, IL3, IL6, KGF 및 IGF1 을 발현하고 정제하는 최초의 성공적인 시도였다. 이러한 발견은 효율적인 공정을 구현함으로써 생물 의약품 제조를 향상시킬 가능성을 제공하며, 잠재적으로 확장성과 비용 효율성을 개선할 수 있다. 여기에 제시된 접근법은 난제 재조합 단백질 생산을 위한 새로운 길을 열어주며, 새로운 치료제 개발과 연구 응용에 상당한 영향을 미칠 수 있다.

중심단어: 대장균(E. coli), 원핵생물 가용성 발현, 과립구-대식세포 집락 자극 인자 (GM-CSF), 인터류킨-3 (IL3), 인터류킨-6 (IL6), 각질세포 성장 인자 (KGF), 인슐린 유사 성장 인자 1 (IGF1), 융합 태그, 단백질 이황화 이성질화효소 (PDIb'a')의 b'a' 도메인.