

High-level Expression and Integrated Purification of Recombinant Human Type I+III Collagen in *Pichia pastoris*

Lu Li

Erasync Biochemical Products Co., Ltd.

Yue Huang

South China University of Technology

Xin Fu

Erasync Biochemical Products Co., Ltd.

He Xiao

Erasync Biochemical Products Co., Ltd.

Yun Zhang

Guangdong Marubi Biotechnology Co., Ltd.

Huaiqing Sun

Guangdong Marubi Biotechnology Co., Ltd.

Chaowan Guo

Guangdong Marubi Biotechnology Co., Ltd.

Ying Lin

feylin@scut.edu.cn

South China University of Technology

Research Article

Keywords: Recombinant collagen, Type I + III, *Pichia pastoris*, Multi-copy integration, High-density fermentation, Integrated purification, Hydrophobic interaction chromatography

Posted Date: July 1st, 2026

DOI: <https://doi.org/10.21203/rs.3.rs-10138169/v1>

License: © ⓘ This work is licensed under a Creative Commons Attribution 4.0 International License.

[Read Full License](#)

Additional Declarations: No competing interests reported.

Abstract

Recombinant human type I and type III collagens are widely applied in skin repair, wound healing, anti-aging cosmetics and biomedical materials. However, traditional production methods face challenges of low yield, difficult scale-up and complicated purification. In this study, we developed a stable and scalable process for high-level secretory expression of recombinant human type I + type III collagen (I + III collagen) in *Pichia pastoris* GS115 using multi-copy integration strategy, combined with high-density fermentation and integrated purification. The codon-optimized fusion gene named WM was cloned into the pHKA vector under the control of AOXm promoter with α CC signal peptide, and multi-copy strains (1-copy to 5-copy) were constructed via Cre/loxP-mediated iterative integration at Int1 locus. Fermentation was systematically optimized in 5 L bench-top and 500 L pilot-scale bioreactors with four-stage control: batch growth, glycerol feeding, methanol adaptation and induction expression. The induction temperature was adjusted to 22°C, which significantly reduced protein degradation. After fermentation optimization, the expression level of I + III collagen reached 10.42 mg/mL in 500 L culture. An integrated purification procedure including centrifugation, 750 kD hollow fiber clarification, 10 kD ultrafiltration buffer exchange, Phenyl Sepharose 6FF hydrophobic interaction chromatography (HIC) and 3 kD desalting was established. The final product achieved purity $\geq 94.7\%$ and activity recovery $\geq 90\%$ with intact structure and no obvious degradation. This study provides a robust, scalable and cost-effective manufacturing platform for recombinant I + III collagen, which is suitable for industrial production and high-value applications.

Introduction

Collagen is the most abundant structural protein in human extracellular matrix, which plays critical roles in tissue strength, elasticity, cell adhesion and tissue regeneration (Guo, et al. 2024; He, et al. 2025; Meganathan, et al. 2022). Type I collagen provides mechanical support and tensile strength (Amirrah, et al. 2022; Naomi, et al. 2021; Selvaraj, et al. 2024), while type III collagen dominates tissue flexibility and repair capacity (Nikolov and Popovski 2022; Shan, et al. 2025). The natural coexistence of type I and III collagens in skin, blood vessels and other tissues determines their synergistic effects in clinical and cosmetic applications (Bernardo, et al. 2024; Botor, et al. 2021). Compared with animal-derived collagen, recombinant human collagen avoids pathogen contamination and immune rejection, showing higher safety and biocompatibility (Cao, et al. 2024; Jie, et al. 2025; Sun, et al. 2025).

Pichia pastoris expression system has become a preferred host for industrial production of recombinant proteins due to its high expression level, easy scale-up, strict regulation of AOX promoter and efficient secretory expression (Eskandari, et al. 2023; Liu, et al. 2022; Pan, et al. 2022; Unver and Dagci 2024). However, the current production of recombinant collagen still has several bottlenecks (Long, et al. 2025):

1. Most studies focus on single-type collagen expression, which cannot simulate the synergistic function of type I + III collagens.
2. Low expression yield and unstable fermentation restrict large-scale production.

3. Traditional multi-step chromatography leads to low recovery, high cost and high endotoxin level.
4. Lack of stable and consistent process from lab scale to industrial pilot scale.

To overcome these limitations, we constructed a 5-copy multi-copy recombinant *Pichia pastoris* strain for expression of human type I + III collagen. A four-stage high-density fermentation process was established and successfully scaled up from 5 L to 500 L. Meanwhile, an integrated purification workflow was designed to achieve high purity and high recovery. The final product exhibited excellent stability and biological activity, meeting the requirements of high-end cosmetics and medical materials.

Materials and methods

Materials

Restriction enzymes, DNA ligase and PCR reagents were purchased from Thermo Fisher Scientific (USA). *Pichia pastoris* GS115, pHKA vector and Cre/loxP integration system were preserved in our laboratory. Hydrophobic chromatography medium Phenyl Sepharose 6FF was obtained from GE Healthcare (USA). Hollow fiber membranes (750 kD) and ultrafiltration membrane packages (10 kD, 3 kD) were from Millipore (USA). All other chemicals were of analytical grade.

Construction of multi-copy recombinant *Pichia pastoris*

The fusion gene WM of human type I and III collagen was codon-optimized for *Pichia pastoris*. The target gene was inserted into pHKA vector with AOXm promoter and αCC signal peptide to generate single-copy plasmid. Then 2-copy, 4-copy and 5-copy plasmids were constructed through iterative integration using *Bgl*I/*Bam*H digestion and seamless cloning. The 5-copy expression cassette was integrated into the Int1 locus of GS115 genome via Cre/loxP-mediated recombination. Hygromycin was used for screening, and the resistance gene was knocked out after positive clone identification. Positive strains were verified by PCR and SDS-PAGE.

High-density fermentation (5 L and 500 L scale)

Medium preparation

BSM basal medium: CaSO₄ 0.65 g/L, MgSO₄ 8.0 g/L, K₂SO₄ 12.74 g/L, KOH 2.9 g/L, H₃PO₄ 26.7 mL/L, glycerol 40 g/L, supplemented with PTM1 4.35 mL/L. Fed-batch medium: 50% glycerol (with PTM1 12 mL/L), methanol (with PTM1 12 mL/L).

Four-stage fermentation control

1. **Batch growth phase:** 30°C, pH 5.5, DO 25–35%, cultured until glycerol depletion and DO jump.
2. **Glycerol feeding phase:** fed 50% glycerol at 30 L/h, maintained DO 25–35%, until OD₆₀₀ ≈ 250, then starved for 60 min.

3. **Methanol adaptation phase:** initial methanol feeding rate 3 g/h/L, increased by 1 g/h/L every 2 h, 25°C, pH 6.0.
4. **Induction expression phase:** 22–25°C, DO 10–20%, gradient methanol feeding, induction for 96–166 h.

For 500 L pilot scale, induction temperature was optimized to **22°C** to reduce protease degradation and improve protein stability.

Integrated purification process

1. **Clarification:** fermentation broth was centrifuged at 12,000 rpm, 4°C for 15 min, then filtered by 750 kD hollow fiber and 0.45 µm membrane.
2. **Ultrafiltration buffer exchange:** 10 kD membrane package was used to exchange buffer to 20 mM PB + 0.2 M (NH₄)₂SO₄, pH 7.0.
3. **Hydrophobic interaction chromatography (HIC):** sample was adjusted to 1 M (NH₄)₂SO₄, loaded onto Phenyl Sepharose 6FF column. Equilibrated with 20 mM PB + 1 M (NH₄)₂SO₄, washed with 0.8 M (NH₄)₂SO₄, eluted with 0.4 M (NH₄)₂SO₄.
4. **Desalting and filtration:** eluted protein was desalted and concentrated by 3 kD ultrafiltration, then filtered through 0.22 µm sterile membrane.

Protein analysis

Protein concentration was determined by BCA assay. Purity was analyzed by SDS-PAGE and HPLC. Protein integrity and degradation were evaluated by SDS-PAGE.

Results

Construction and identification of multi-copy strains

To investigate the correlation between target gene copy number and collagen secretory expression level, strains carrying 1 to 5 copies of the fusion gene were cultured in shake flasks, with blank GS115 strain (9K) set as negative control. SDS-PAGE analysis of culture supernatants (Fig. 1) showed no target protein band in the blank host strain. Obvious bands corresponding to recombinant human type I + III collagen were detected in all multi-copy strains, and the band intensity exhibited a clear positive correlation with gene copy number. The 1-copy strain (1C) yielded faint target bands, while the band brightness gradually increased in 2C, 3C and 4C strains, reaching the maximum expression level in the 5-copy strain (5C). This result verified that elevating integrated gene copy number effectively enhances the secretory production of dual-type collagen in *Pichia pastoris*, and the 5-copy engineered strain was selected as the optimal strain for subsequent tank fermentation optimization.

Optimization of high-density fermentation

To optimize the induction temperature for the multi-copy *Pichia pastoris* strain expressing human type I + III collagen, three induction temperatures (20°C, 22°C and 25°C) were evaluated in 5 L bench-top bioreactors. Cell growth was tracked via continuous OD₆₀₀ detection throughout the 168 h induction period (Fig. 2). The maximum OD₆₀₀ values reached 475 at 25°C, 529 at 22°C and 501 at 20°C, indicating that 22°C supported the highest biomass accumulation. SDS-PAGE analysis of fermentation supernatants collected at 46 h, 58 h and 70 h induction (Fig. 3) revealed distinct expression differences among groups: no visible target collagen band was detected at 20°C, demonstrating this temperature failed to trigger effective secretory expression. Induction at 25°C led to unstable target protein expression; several repeated batches failed to generate detectable recombinant collagen. In contrast, the 22°C group exhibited the thickest, brightest intact target collagen band without prominent degradation fragments across all sampling time points. Collectively, these results confirm that 22°C is the optimal induction temperature, which balances robust cell growth, high-level collagen secretion and minimal protein degradation.

Cell density was determined by monitoring OD₆₀₀ at different induction time points. Black square, 25°C induction; red circle, 22°C induction; blue triangle, 20°C induction. The maximum OD₆₀₀ values at 168 h induction for 25°C, 22°C and 20°C groups were 475, 529 and 501, respectively.

We further scaled up the optimized fermentation process from lab scale to 500 L pilot bioreactor with two independent batches. The cell growth profiles monitored by OD₆₀₀ were presented in Fig. 4A and 4C for two fermentation runs, both showing steady cell proliferation throughout the whole induction period and the maximum OD₆₀₀ reached approximately 580 and 480 at the endpoint, respectively. SDS-PAGE analysis of fermentation supernatants sampled at serial induction time points (0–166 h) was displayed in Fig. 4B and 4D. The target band of recombinant human type I + III collagen gradually intensified with extended induction time in both batches, and consistent expression trends were observed between two independent 500 L pilot fermentations, which verified the excellent reproducibility and scalability of the established high-density fermentation platform for co-production of dual-type collagen in *Pichia pastoris*.

Integrated purification efficiency

An integrated multi-step purification workflow combining hollow fiber clarification, ultrafiltration buffer exchange, hydrophobic interaction chromatography (HIC) and desalting was established to purify recombinant human type I + III collagen from 500 L fermentation supernatant. Figure 5A and 5B display the appearance of crude fermentation supernatant and final desalted protein solution, respectively. The full HIC chromatogram is shown in Fig. 5C, with distinct peaks corresponding to flow-through fraction (FT), wash fraction (WF), elution fraction (EF) and desalted product (CIP). SDS-PAGE (Fig. 5D) confirmed that most heterologous impurities were removed in FT and WF fractions, while the target dual-type collagen was concentrated in EF. Subsequent HPLC quantification (Fig. 5E) demonstrated that the crude collagen supernatant possessed a purity of 86%. After HIC purification, the elution fraction EF achieved a purity of 94.7% with an overall activity recovery of 90%; by contrast, the wash fraction WF only reached 71% purity at merely 10% recovery. These data validated that the integrated purification strategy

efficiently removes miscellaneous proteins and retains high yields of high-purity target collagen, meeting the quality requirements for medical and cosmetic high-end applications.

Discussion

In this study, we developed a complete manufacturing process for recombinant human type I + III collagen in *Pichia pastoris*, from strain engineering to industrial-scale fermentation and integrated purification. Multi-copy integration is an effective strategy to enhance gene dosage and expression level (Ahn et al., 2016). The 5-copy strain constructed by Cre/loxP system avoided residual resistance markers and maintained high genetic stability.

High-density fermentation is key to high-yield production. The four-stage control strategy ensured sufficient cell growth and targeted protein expression. Lowering induction temperature to 22°C in 500 L scale reduced protease activity and protein degradation, which is critical for large-scale production of collagen (Jungo et al., 2007).

Traditional collagen purification often involves multiple chromatography steps with low recovery. Our integrated process combined clarification, ultrafiltration, HIC and desalting, simplifying operation and improving recovery to over 90%. HIC effectively removed impurities without harsh denaturation, preserving protein structure and activity (Lee et al., 2019).

The expression level of 10.42 mg/mL in 500 L fermentation and purity of 94.7% are significantly higher than most reported processes. The stable scale-up from 5 L to 500 L confirms the industrial potential of this technology. The recombinant I + III collagen mimics natural composition and shows great promise in skin repair, wound dressing and tissue engineering.

Conclusion

We established a scalable and efficient platform for production of recombinant human type I + III collagen in *Pichia pastoris*. The 5-copy multi-copy strain achieved high-level secretory expression. The optimized 500 L high-density fermentation yielded 10.42 mg/mL. The integrated purification process obtained high-purity ($\geq 94.7\%$) and high-recovery ($\geq 90\%$) product with intact structure. This process is stable, cost-effective and suitable for industrial manufacturing, supporting wide applications in cosmetics and biomedicine.

Declarations

Conflict of interest

The authors declare no conflict of interest.

Ethical statement

This study does not involve human participants or animals.

Author Contribution

Author contributions are as follows: Lu Li drafted and submitted the manuscript; Yue Huang constructed multi-copy strains and performed shake-flask cultivation; Xin Fu completed 5 L to 500 L lab-scale fermentation trials; He Xiao carried out the integrated purification of recombinant collagen; Yun Zhang designed all collagen sequences; Huaqing Sun coordinated experimental resources; Chaowan Guo managed the overall project progress; Ying Lin revised and proofread the entire manuscript.

Acknowledgements

This work was supported by [Funding Information the National Natural Science Foundation of China (31470159)].

Data Availability

All raw experimental data including SDS-PAGE gel images, fermentation OD₆₀₀ growth curves, chromatograms, quantitative purity and recovery values generated and analyzed during this study are included within the published article. Additional raw datasets, original electrophoresis pictures and bioreactor recording files are available from the corresponding author upon reasonable request.

References

1. Amirrah, I. N., Y. Lokanathan, I. Zulkiflee, Mfmr Wee, A. Motta, and M. B. Fauzi (2022) A Comprehensive Review on Collagen Type I Development of Biomaterials for Tissue Engineering: From Biosynthesis to Bioscaffold. *Biomedicines* 10(9). 10.3390/biomedicines10092307
2. Bernardo, R. T. R., R. C. G. Oliveira, K. M. S. Freitas, J. R. Albergaria-Barbosa, and C. M. Rizzatti-Barbosa (2024) Effect of poly-L-lactic acid and polydioxanone biostimulators on type I and III collagen biosynthesis. *Skin Res Technol* 30(4):e13681. 10.1111/srt.13681
3. Botor, M., A. Fus-Kujawa, M. Uroczynska, K. L. Stepień, A. Galicka, K. Gawron, and A. L. Sieron (2021) Osteogenesis Imperfecta: Current and Prospective Therapies. *Biomolecules* 11(10). 10.3390/biom11101493
4. Cao, L., Z. Zhang, D. Yuan, M. Yu, and J. Min (2024) Tissue engineering applications of recombinant human collagen: a review of recent progress. *Front Bioeng Biotechnol* 12:1358246. 10.3389/fbioe.2024.1358246

5. Eskandari, A., N. G. Nezhad, T. C. Leow, M. B. A. Rahman, and S. N. Oslan (2023) Current achievements, strategies, obstacles, and overcoming the challenges of the protein engineering in *Pichia pastoris* expression system. *World J Microbiol Biotechnol* 40(1):39. 10.1007/s11274-023-03851-6
6. Guo, X., Y. Ma, H. Wang, H. Yin, X. Shi, Y. Chen, G. Gao, L. Sun, J. Wang, Y. Wang, and D. Fan (2024) Status and developmental trends in recombinant collagen preparation technology. *Regen Biomater* 11:rbad106. 10.1093/rb/rbad106
7. He, H., M. Ye, G. Cui, and J. Xiao (2025) Recombinant collagen in regenerative medicine: Expression strategies, structural design, and translational applications. *Mater Today Bio* 35:102452. 10.1016/j.mtbio.2025.102452
8. Jie, I. W. K., K. W. A. Lee, S. E. Yoon, J. K. Song, L. K. W. Chan, C. H. Lee, E. Jeong, J. H. Kim, and K. H. Yi (2025) Advancements in Clinical Utilization of Recombinant Human Collagen: An Extensive Review. *Life (Basel)* 15(4). 10.3390/life15040582
9. Liu, C., J. S. Gong, C. Su, H. Li, H. Li, Z. M. Rao, Z. H. Xu, and J. S. Shi (2022) Pathway engineering facilitates efficient protein expression in *Pichia pastoris*. *Appl Microbiol Biotechnol* 106(18):5893–5912. 10.1007/s00253-022-12139-y
10. Long, C., M. Zhang, J. Ren, H. Liu, Y. Chen, Z. Zhao, H. Zhang, and D. Lei (2025) [Progress in Application of Recombinant Collagen in Medical Devices]. *Zhongguo Yi Liao Qi Xie Za Zhi* 49(6):640–646. 10.12455/j.issn.1671-7104.250152
11. Meganathan, I., M. Pachaiyappan, M. Aarthy, J. Radhakrishnan, S. Mukherjee, G. Shanmugam, J. You, and N. Ayyadurai (2022) Recombinant and genetic code expanded collagen-like protein as a tailorable biomaterial. *Mater Horiz* 9(11):2698–2721. 10.1039/d2mh00652a
12. Naomi, R., P. M. Ridzuan, and H. Bahari (2021) Current Insights into Collagen Type I. *Polymers (Basel)* 13(16). 10.3390/polym13162642
13. Nikolov, A., and N. Popovski (2022) Extracellular Matrix in Heart Disease: Focus on Circulating Collagen Type I and III Derived Peptides as Biomarkers of Myocardial Fibrosis and Their Potential in the Prognosis of Heart Failure: A Concise Review. *Metabolites* 12(4). 10.3390/metabo12040297
14. Pan, Y., J. Yang, J. Wu, L. Yang, and H. Fang (2022) Current advances of *Pichia pastoris* as cell factories for production of recombinant proteins. *Front Microbiol* 13:1059777. 10.3389/fmicb.2022.1059777
15. Selvaraj, V., S. Sekaran, A. Dhanasekaran, and S. Warriar (2024) Type 1 collagen: Synthesis, structure and key functions in bone mineralization. *Differentiation* 136:100757. 10.1016/j.diff.2024.100757
16. Shan, Y., T. Wang, and H. Lin (2025) Applications of recombinant type III collagen in tissue engineering. *Biomed Eng Online* 24(1):114. 10.1186/s12938-025-01447-9
17. Sun, W., M. H. Shahrajabian, K. Ma, and S. Wang (2025) Advances in Molecular Function and Recombinant Expression of Human Collagen. *Pharmaceuticals (Basel)* 18(3). 10.3390/ph18030430
18. Unver, Y., and I. Dagci (2024) *Komagataella phaffii* (*Pichia pastoris*) as a Powerful Yeast Expression System for Biologics Production. *Front Biosci (Elite Ed)* 16(2):19. 10.31083/j.fbe1602019

Figures

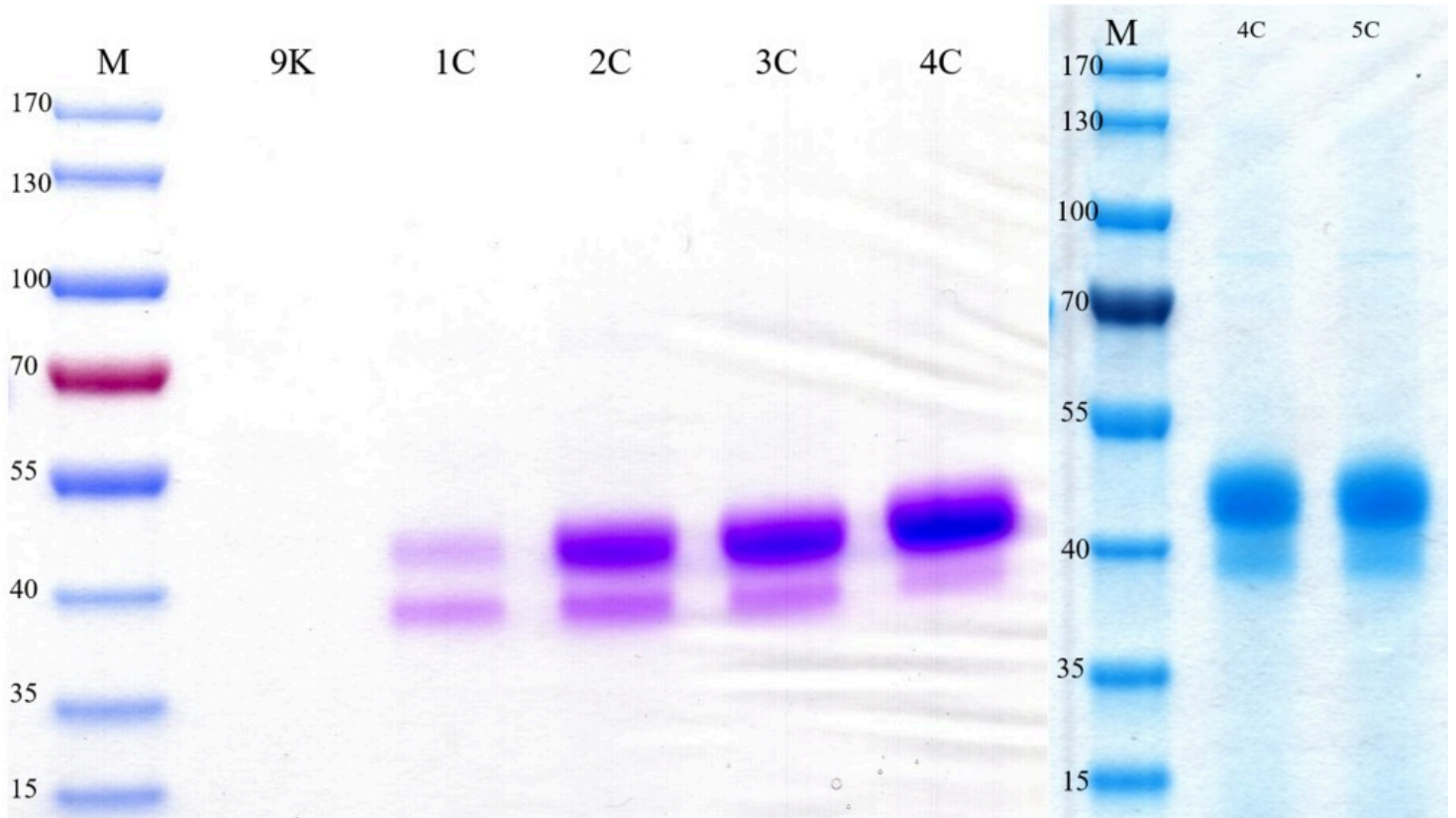


Figure 1

SDS-PAGE analysis of secreted recombinant human type I + III collagen from *Pichia pastoris* strains with different gene copy numbers in shake flasks.

M: Protein molecular weight marker (kDa); 9K: Blank host strain GS115 negative control; 1C, 2C, 3C, 4C, 5C: Fermentation supernatant of 1-, 2-, 3-, 4-, 5-copy engineered strains expressing type I+III collagen. Target collagen bands (~40–55 kDa) show gradually increased expression intensity with elevated gene copy number.

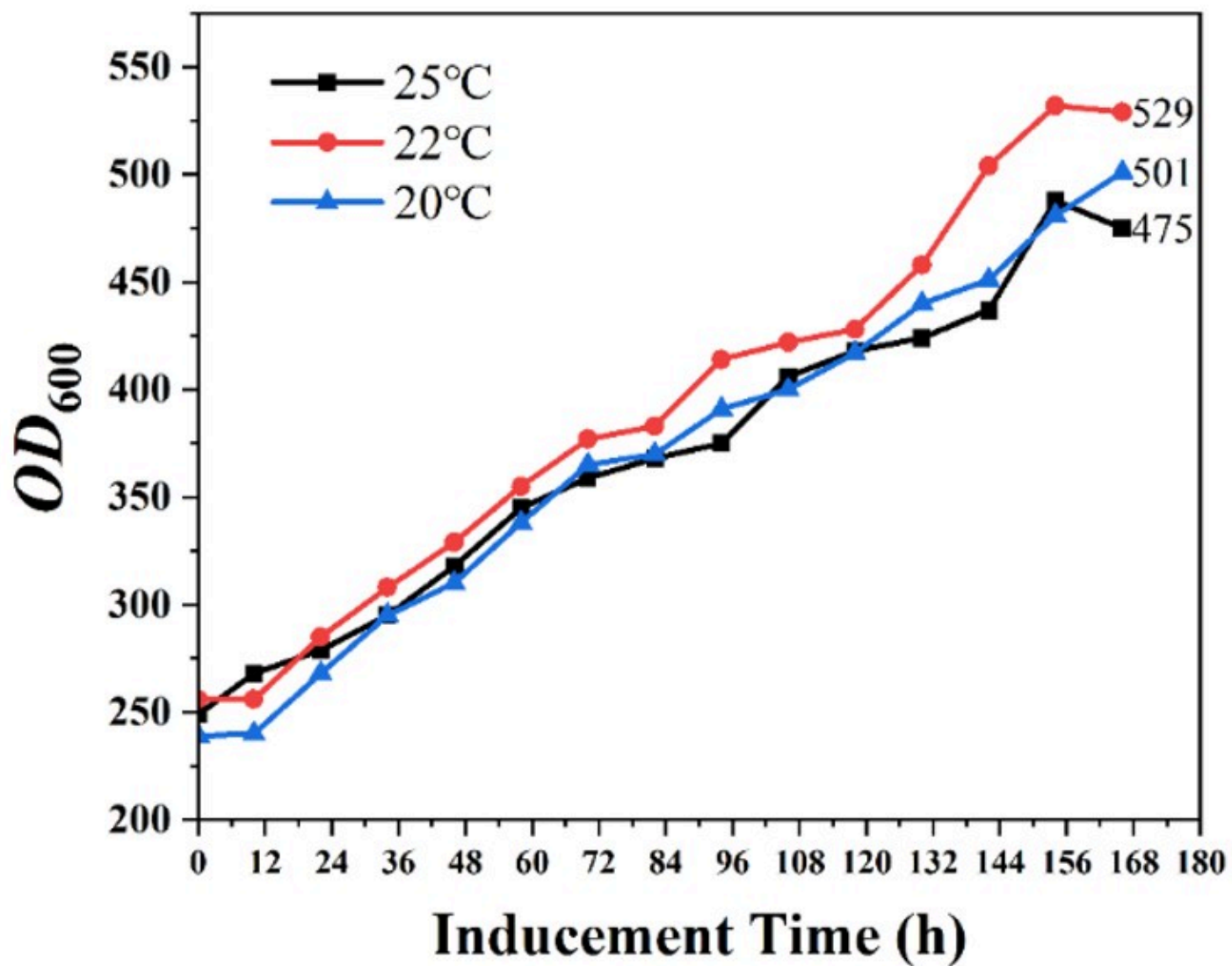


Figure 2

Growth curve of multi-copy *Pichia pastoris* strain under different induction temperatures in 5L bioreactor.

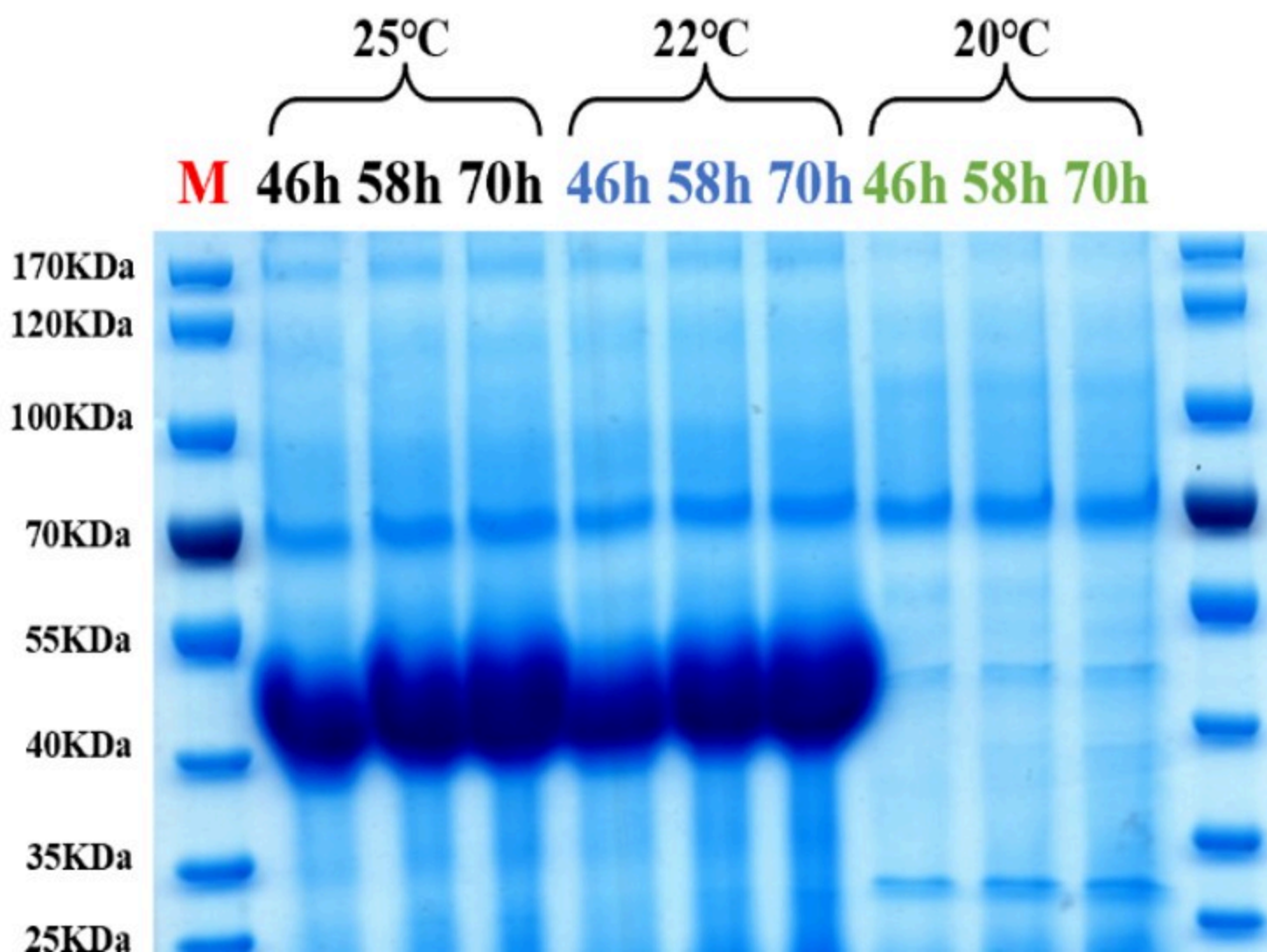


Figure 3

SDS-PAGE analysis of recombinant human type I + III collagen secreted by *Pichia pastoris* induced at different temperatures for various durations in 5L bioreactor.

M: Protein molecular weight marker (kDa). Lanes under 25 °C, 22 °C and 20 °C correspond to induction times of 46 h, 58 h and 70 h, respectively. The target collagen bands show the highest expression level at 22 °C, accompanied by minimal protein degradation.

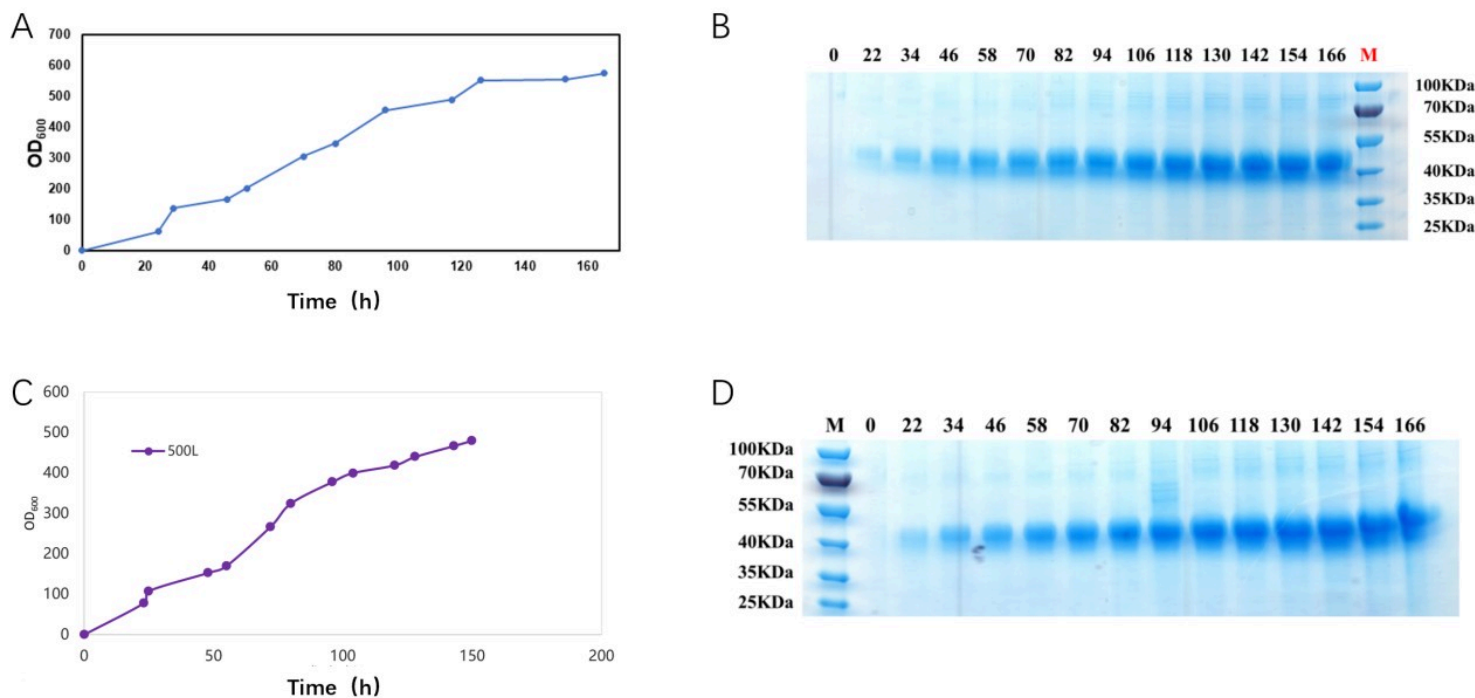


Figure 4

Cell growth and collagen expression profiles of two independent 500 L pilot-scale fermentations of multi-copy *Pichia pastoris* strain induced at 22 °C.

A and C. Time-course OD₆₀₀ curves representing cell density of Batch 1 and Batch 2, respectively. B and D. SDS-PAGE analysis of fermentation supernatants collected at serial induction time points (0–166 h) from Batch 1 and Batch 2. M: protein molecular weight marker (kDa).

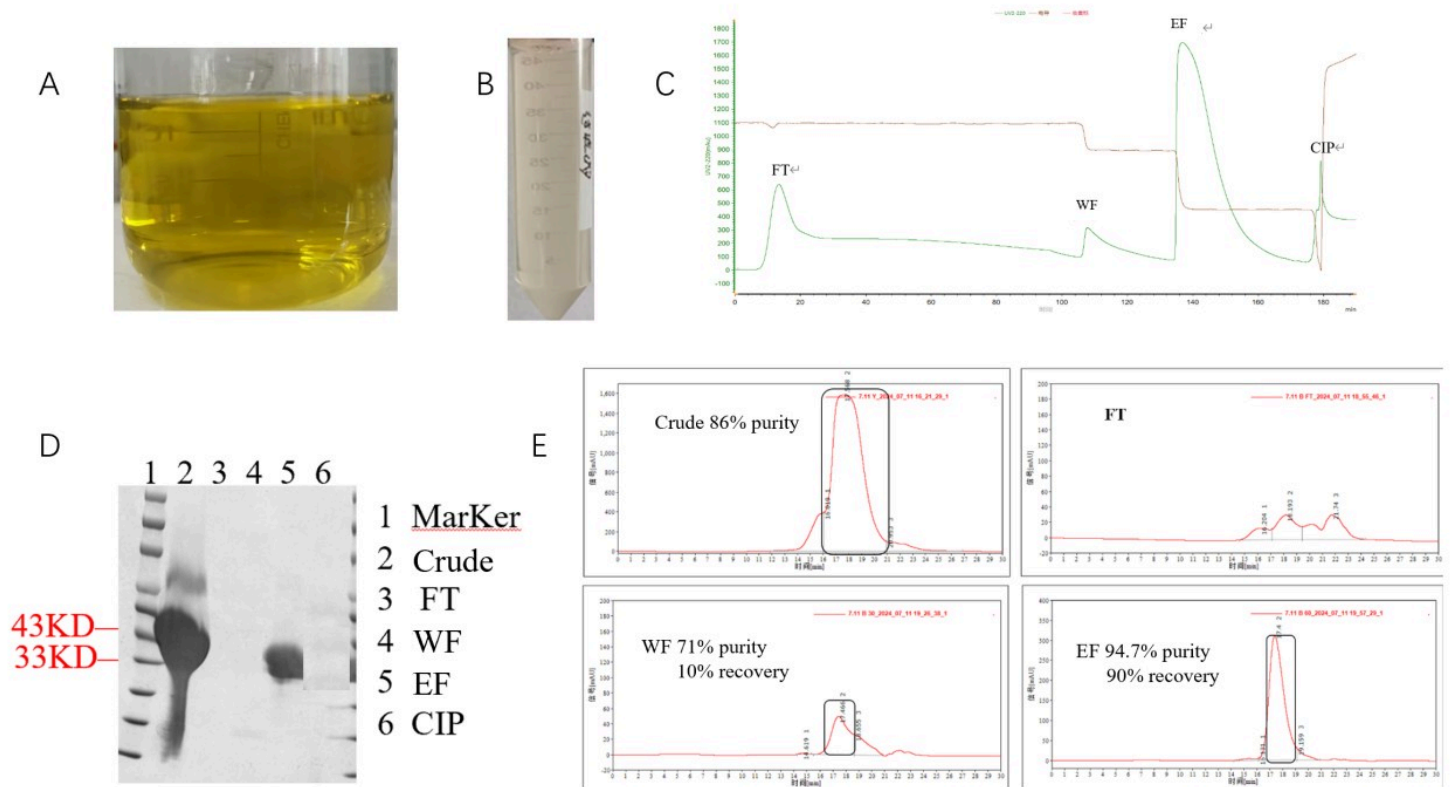


Figure 5

Characterization of the integrated purification process for recombinant human type I + III collagen.

(A) Visual image of crude fermentation supernatant; (B) Visual image of final desalted collagen solution after full purification. (C) Full hydrophobic interaction chromatography (HIC) elution profile monitored at UV 280 nm. FT: flow-through fraction; WF: wash fraction; EF: target elution fraction; CIP: cleaning-in-place peak. (D) SDS-PAGE analysis of samples collected at different purification stages. Lane 1: protein molecular weight marker; Lane 2: crude supernatant; Lane 3: FT fraction; Lane 4: WF fraction; Lane 5: EF fraction; Lane 6: CIP sample. (E) HPLC purity detection results of crude liquid, FT, WF and EF fractions, with corresponding purity and recovery values labeled in each subgraph.