Medical Research
Volume 7, Issue 2, 2025, Page 46-53
https://doi.org/10.6913/mrhk.070207

Medical Research

Recombinant Expression of Hypericin Synthase Hyp1 in Escherichia coli and Elucidation of Its Solubility Mechanisms

Fengxia ZHOU, Hanyang DU, Jiaxin LÜ,Jun WANG*

*School of Lifescience, Northwestern Polytechnical University Dongxiang Road, Chang'an District, Xi'an Shaanxi,710129, P.R.China.*Correspondence: Jun WANG Email:wangj@nwpu.edu.cn.*

Abstract

Hypericin, the principal bioactive naphthodianthrone from *Hypericum perforatum*, exhibits broad-spectrum antibacterial, antitumor, antidepressant, and antiviral properties. However, its low natural abundance and suboptimal purity hinder large-scale applications. In this study, two hypericin synthase isoforms, **Hyp1A** and **Hyp1B**, were recombinantly expressed in *Escherichia coli*. Following IPTG induction, Hyp1A accumulated to 94.2 $\mu g \cdot g^{-1}$ wet cell mass, which increased to 1.41 $mg \cdot g^{-1}$ upon fusion with a solubility-enhancing tag. In contrast, Hyp1B achieved 7.53 $mg \cdot g^{-1}$ wet cell mass without modification, highlighting distinct expression phenotypes. Molecular dynamics simulations revealed that Hyp1A harbors a higher proportion of random coil and fewer α-helices relative to Hyp1B, correlating with reduced structural stability. *In vitro* assays confirmed that Hyp1B catalyzes emodin dimerization to hypericin, albeit with limited activity under current conditions. Collectively, these findings demonstrate the feasibility of emodin-to-hypericin bioconversion and establish a foundation for engineering more efficient biosynthetic processes and advancing the clinical application of hypericin.

Keywords Hypericin; Hypericin Synthase; Recombinant Expression; Enzymatic Catalysis; Molecular Dynamics Simulation; *Hypericum perforatum*

To Cite This Article Fengxia ZHOU,et al. (2025). Recombinant Expression of Hypericin Synthase Hyp1 in Escherichia coli and Elucidation of Its Solubility Mechanisms. *Medical Research*, 7(2), 46-53. https://doi.org/10.6913/mrhk.070207

Medical Research, ISSN 2664-0333 (print), ISSN 2664-0341 (online), a bimonthly, founded on 2019, published by Creative Publishing Co., Limited. Email: wtocom@gmail.com, https://mrhk.cc, https://cpcl.hk.

1 Introduction

Hypericin, the principal active constituent of St. John's Wort (Hypericum perforatum L.), is a naphthodianthrone-type compound[1,2]. It exhibits a broad spectrum of biological activities, including antidepressant, antimicrobial, and antineoplastic effects[3,4]. However, endogenous levels of hypericin in H. perforatum are exceedingly low, yielding only 0.03%-0.09% of dry weight[5]. Such scant accumulation precludes bulk applications, as both yield and purity remain insufficient for industrial demands. Consequently, plant extraction persists as the dominant source of hypericin, despite its limitations.

Chemical approaches have sought to address the low natural abundance of hypericin, yet prevailing synthetic routes typically involve multiple steps and harsh conditions that undermine large-scale feasibility.

For instance, Falk et al. demonstrated a three-step semisynthesis from emodin, achieving an overall yield of approximately 51.6%, but this strategy relies on photochemical coupling in organic media and remains time-intensive[6,7]. Hence, devising streamlined, high-efficiency, and environmentally benign synthetic paradigms continues to represent a formidable challenge.

Biosynthetic methods, by contrast, offer potential advantages in terms of yield, cost-efficiency, and environmental sustainability. However, the complete *de novo* biosynthetic pathway of hypericin in *Hypericum perforatum* has not yet been fully elucidated [8]. In 2003, Bais et al. identified emodin in suspension-cultured *H. perforatum* cells and proposed that it functions as a direct biosynthetic precursor to hypericin [9,10]. Subsequent research confirmed that emodin undergoes sequential oxidative coupling reactions to form hypericin and led to the identification of a putative hypericin synthase, Hyp-1, whose gene product catalyzes the dimerization of emodin **ref11**. Nevertheless, Hyp-1 expression *in planta* does not consistently correlate with the spatial pattern of hypericin accumulation, and its catalytic activity under *in vitro* conditions remains a subject of ongoing debate.

Given its relative abundance—and its market price (less than 1% that of hypericin, depending on supplier)—*emodin* is an attractive substrate for the biocatalytic production of *hypericin*. In this study, we generated recombinant constructs for two distinct *Hyp-1* isoforms and expressed them in *Escherichia coli*. Both isoforms yielded soluble protein, albeit with markedly different expression levels and solubilities. Through sequence alignment and secondary-structure modeling, we identified molecular features that likely account for their divergent expression phenotypes. *In vitro* enzymatic assays then demonstrated that each isoform can indeed convert *emodin* to *hypericin*. Although the observed catalytic efficiencies warrant optimization, our findings establish the feasibility of *emodin*-to-*hypericin* conversion via recombinant *Hyp-1* and lay the groundwork for future engineering of more robust biocatalysts to enhance *hypericin* biosynthetic throughput.

2 Materials and Methods

2.1 Construction of Recombinant Expression Plasmids

Hyp1A and *Hyp1B* coding sequences were cloned into either *pET-28a* or *pET-28a-SUMO*. Each vector and insert were digested with *BamH I* and *Xho I* (TaKaRa, Beijing) and subsequently ligated using the Ligation Mix kit (TaKaRa, Beijing, China). Ligation products were transformed into *E. coli* DH5α (Tsingke, Beijing, China), and colonies were selected on LB agar containing 50 μg/mL kanamycin (Sangon Biotech, Shanghai, China). Plasmids from kanamycin-resistant clones were sequenced at Tsingke Biotech (Xi' an, China). Only clones confirmed to carry correct inserts were propagated, and plasmid DNA was purified for downstream expression.

2.2 Induction of Recombinant Protein Expression

Recombinant plasmids were introduced into *E. coli* BL21(DE3) (Tsingke, Beijing, China) via the calcium chloride method. Transformants were selected on LB agar containing 50 μg/mL kanamycin. A single colony was inoculated into 10 mL LB medium (kanamycin, 50 μg/mL) and grown overnight at 37 °C to produce seed culture. This culture was diluted into 1 L of fresh LB medium (kanamycin, 50 μg/mL) and incubated at 37 °C with agitation until OD_{600} reached 0.6 0.8. Protein expression was induced by adding IPTG (final concentration as indicated in the figure legend), and cultures were incubated overnight under the specified temperature conditions.

2.3 Purification of Recombinant Histidine-Tagged Protein

Cells were harvested by centrifugation (6,000 rpm, 4 °C, 10 min) and resuspended in TieChui *E. coli* Lysis Buffer (ACE Biotech, Changzhou, China). Cells were lysed, and lysates were clarified by centrifugation (10,000 rpm, 4 °C, 30 min). The supernatant was loaded onto a Ni-NTA pre-packed column (Smart Lifescience, Changzhou, China). The column was washed sequentially: (1) drain residual 20% ethanol; (2) rinse with five column volumes of deionized water; (3) equilibrate with five column volumes of PBS. Clarified lysate containing His-tagged protein was applied to the column, washed with 20 column volumes of PBS to remove non-specific proteins, and eluted with PBS containing 300 mM imidazole. Eluates were desalted using a PD-10 column (Cytiva, Michigan, USA) to remove imidazole. Glycerol was added to 10% (v/v), and purified protein was stored at –20 °C.

2.4 Enzyme Activity Assay for Hypericin Synthase

Hypericin synthase activity was measured in 15 mL reaction tubes. Each reaction contained 3 mL of enzyme buffer (Tris, 1.211 g/100 mL; KCl, 1.245 g/100 mL; $MgCl_2 \cdot 6H_2O$, 2.033 g/100 mL), 1 mL of purified enzyme solution, and 60 μL of substrate stock (1 mg/mL *emodin*). Reactions were incubated at 4 °C and 200 rpm in the dark for 4 h. Subsequently, 4 mL of *n*-hexane was added for extraction; mixtures were shaken at 30 °C and 200 rpm for 6 h. After centrifugation (11,000 rpm, room temperature, 20 min), the upper *n*-hexane layer was transferred to a new tube. The solvent was evaporated using a rotary evaporator, and the residue was resuspended in 1 mL methanol. Samples were sonicated for 10 min, and absorbance was recorded at 560 nm. Each experimental group (enzyme + substrate) was run in octuplicate. Controls included substrate without enzyme and enzyme without substrate (each also in octuplicate)[1].

2.5 Molecular Dynamics Simulations of Hypericin Synthase

All-atom molecular dynamics (MD) simulations were performed using `GROMACS` 2019.6 under periodic boundary conditions. The AMBER99SB-ILDN force field and TIP3P water model were employed. Each Hyp-1 crystal structure was placed at the center of a cubic simulation box, maintaining a minimum distance of 1 nm between the protein surface and the box boundaries. The systems were solvated with TIP3P water molecules, and Na^+ ions were added to neutralize the net charge.

Energy minimization was conducted using the steepest descent algorithm until convergence was achieved (i.e., the maximum force was less than 1000 $kJ \cdot mol^{-1} \cdot nm^{-1}$). The minimized systems were equilibrated in two phases: a 100 ps NVT ensemble at 300 K, followed by a 100 ps NPT ensemble at 1 bar, both with positional restraints applied to protein heavy atoms.

Production MD simulations were then performed for 30 ns at 300 K and 1 bar with a 2 fs integration time step. All bonds involving hydrogen atoms were constrained using the LINCS algorithm, and long-range electrostatic interactions were computed using the Particle-Mesh Ewald (PME) method.

Secondary structure evolution during the simulations was analyzed using the Define Secondary Structure of Proteins (DSSP) algorithm. Changes in folding free energy ($\Delta\Delta G$) resulting from point mutations were calculated using the `FoldX` program.

3 Results

3.1 Recombinant Expression of *Hyp1A* in *E. coli*

To evaluate *Hyp1A* expression, we employed the *E. coli* BL21(DE3) pET-28a expression system. The *hyp1A* gene was inserted into the pET-28a plasmid via *Bam*HI and *Xho*I restriction sites (Fig. 1A), and the resulting construct was transformed into *E. coli* BL21(DE3) for IPTG induction. SDS PAGE analysis

revealed negligible levels of soluble Hyp1A protein (Fig. 1C). Since SUMO is known to enhance protein solubility, we recloned *hyp1A* into the pET-28a SUMO vector (Fig. 1B). Fusion with SUMO significantly improved the soluble expression yield, increasing it from 94.2 μg/g to 1.41 mg/g of wet cell mass. However, when purified SUMO Hyp1A protein was incubated with emodin, no hypericin was detected, indicating that enzymatic activity was lost despite the improved solubility.

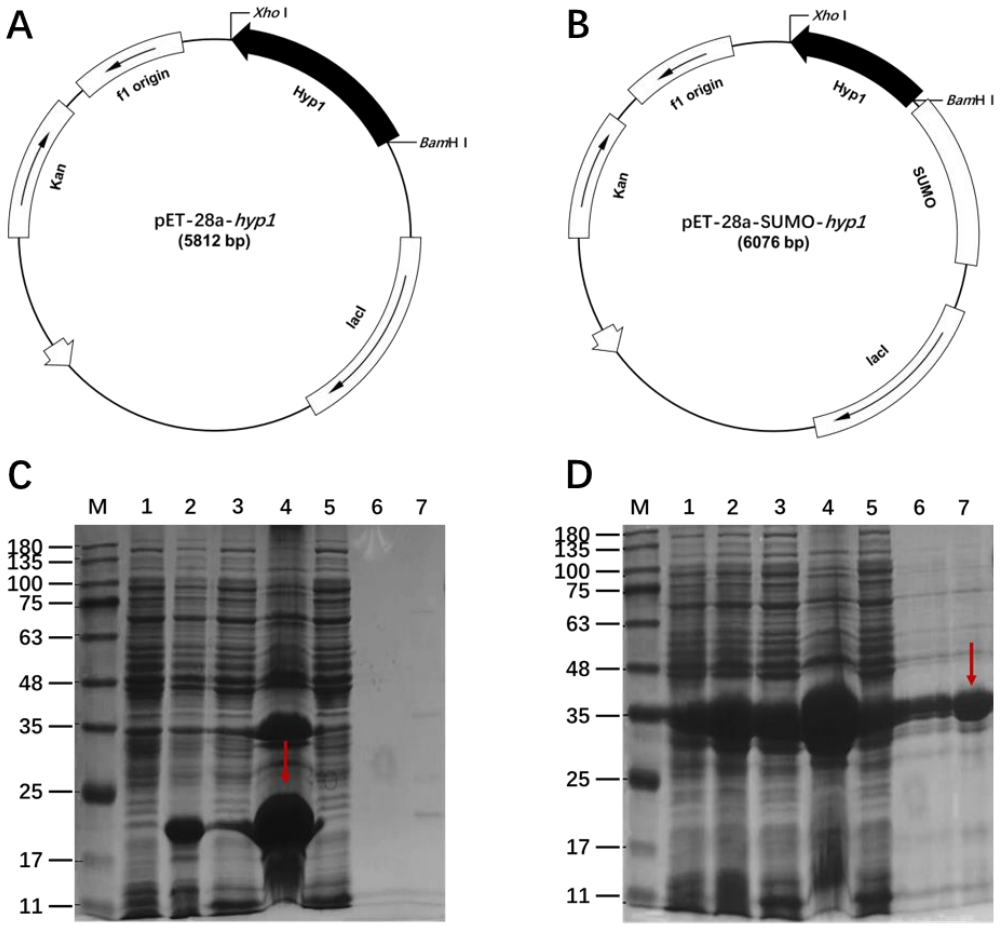

Figure 1. Recombinant expression of hypericin synthase Hyp1A in *Escherichia coli*. (A) Schematic of the Hyp1A expression vector; (B) Schematic of the SUMO-Hyp1A expression vector carrying a solubility tag; (C) Expression of Hyp1A in *E. coli*; (D) Expression of SUMO-Hyp1A in *E. coli*. M: Marker; Lane 1: uninduced cells;Lane 2: induced cells; Lane 3: lysate supernatant (soluble fraction); Lane 4: lysate pellet (insoluble fraction); Lane 5: flow-through 1; Lane 6: flow-through 2; Lane 7: affinity-purified eluate.

3.2 Structural Basis for Differential Solubility of *Hyp1* Isoforms

A second isoform, *Hyp1B*, was identified in the literature as exhibiting robust soluble expression in *E. coli*. Sequence alignment (Fig. 2A) revealed divergence between *Hyp1A* and *Hyp1B* at residues 14, 86, and 119 (Thr → Ile, Ile → Leu, and Ser → Thr, respectively). To investigate how these substitutions affect folding, we performed 30 ns all-atom MD simulations and analyzed secondary-structure dynamics via DSSP. Relative to *Hyp1B*, *Hyp1A* displayed a marked reduction in 3-helix content near residue 86 (replaced by Bend) and an absence of an α-helix near residue 128 (forming Turn instead). Because 3-helix and α-helix elements stabilize protein cores, these alterations suggest that *Hyp1A* is inherently less stable. Indeed, throughout the MD trajectory, the 3-helix fraction of *Hyp1A* declines sharply in later frames, accompanied by an increase in destabilizing Bend segments (Fig. 2D, 2E), whereas *Hyp1B* maintains a consistent

secondary-structure profile.

Free-energy calculations using FoldX yielded a $\Delta\Delta G$ difference of 4.36 kcal/mol (Hyp1A vs. Hyp1B), corroborating Hyp1A's reduced structural stability. Structural inspection of the final simulation frames highlighted two key regions: residues 26–36, which adopt an α-helix in Hyp1B but remain a random coil in Hyp1A; and residues 90–95, which form a 3_{10}-helix in Hyp1B but are also random coil in Hyp1A. Michalska *et al.*[11] noted that Hyp1B's hydrophobic cavity is delineated by a β-sheet concave face and three α-helices. The three residue substitutions ablate one of these helices in Hyp1A, replacing it with an unstable coil (Fig. 2F–H), and constrict one entrance to the hydrophobic cavity (Fig. 2I, 2J). Together, these observations explain Hyp1A's propensity for misfolding and precipitation during expression, and suggest that alterations in the hydrophobic cavity disrupt substrate binding and catalytic function.

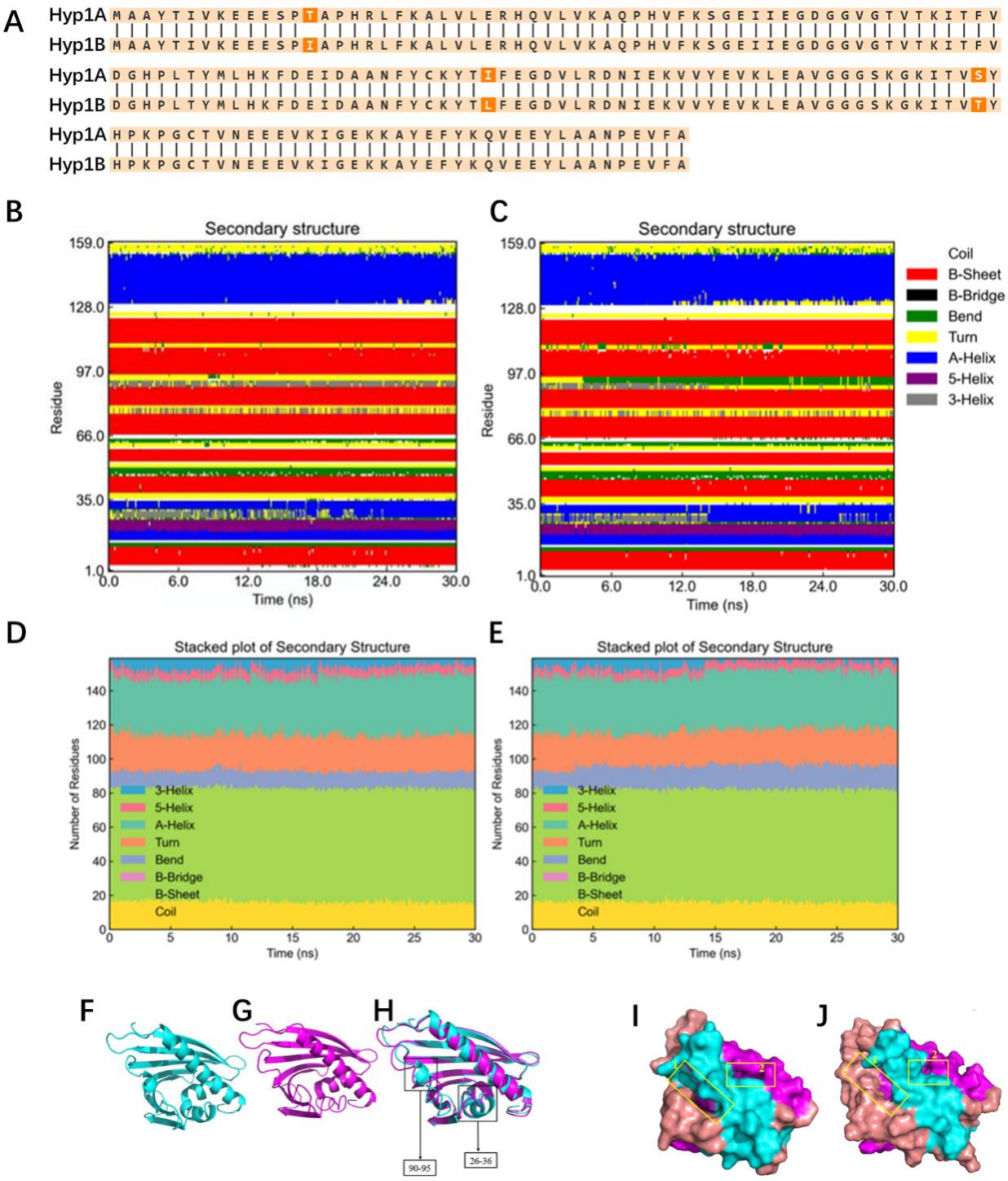

Figure 2. Structure analysis of Hyp1A and Hyp1B. (A) Sequence alignment of Hyp1A and Hyp1B; (B) and (D) Secondary structure analysis of Hyp1B; (C) and (E) Secondary structure analysis of Hyp1A; (H) Structural alignment of the Hyp1A(G) and Hyp1B (F); (I) and (J) Surface models of Hyp1B and Hyp1A, respectively.

3.3 Optimization of *Hyp1B* Expression and Storage

Guided by the above insights, we expressed Hyp1B in *E. coli* and systematically optimized induction conditions. Maximum soluble yield (7.53 mg/g wet cell mass, without any solubility tag) was achieved at 25 °C with 0.1 mM IPTG (Fig. 3C). To ensure long-term stability, we evaluated glycerol-mediated cryoprotection. Hyp1B remained fully soluble through multiple freeze thaw cycles when formulated in 10% glycerol; lower concentrations led to precipitation (Table 1).

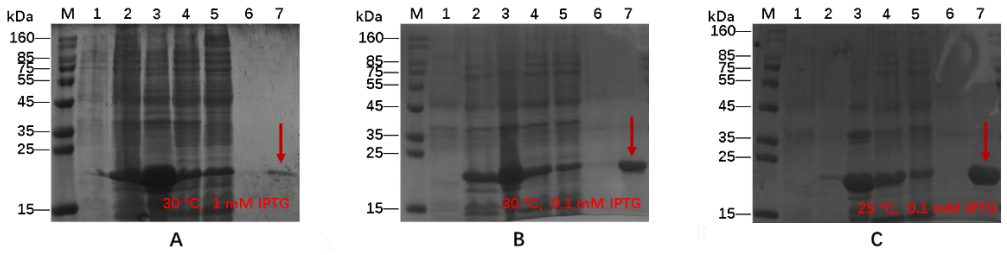

Figure 3. SDS-PAGE of Hyp1B expression under different conditions. M: Marker; Lane 1: uninduced cells;Lane 2: induced cells; Lane 3: lysate supernatant (soluble fraction); Lane 4: lysate pellet (insoluble fraction); Lane 5: flow-through 1; Lane 6: flow-through 2; Lane 7: affinity-purified eluate.

Table 1. Glycerol-Dependent Stability of Hyp1B Solutions

Percent of glycerinum	Precipitability
1%	+++
2%	++
5%	+
10%	-

*"+": precipitation observed (greater "+" indicates greater precipitation); "-": no precipitation.

3.4 Enzymatic Conversion of *Emodin* to *Hypericin* by *Hyp1B*

Emodin and *hypericin* exhibit distinctive absorbance spectra. A UV vis scan (300 700 nm, 10 nm increments) of 0.1 mg/mL standards revealed that *emodin* peaks at 430 nm, whereas *hypericin*'s λ_{max} shifts to 500 600 nm (Fig. 4A). When purified *Hyp1B* was incubated with *emodin*, we detected *hypericin* formation by measuring absorbance at 560 nm. Notably, reactions conducted in the dark yielded approximately 1.9-fold higher *hypericin* levels than those under ambient light, suggesting photolability during enzymatic turnover.

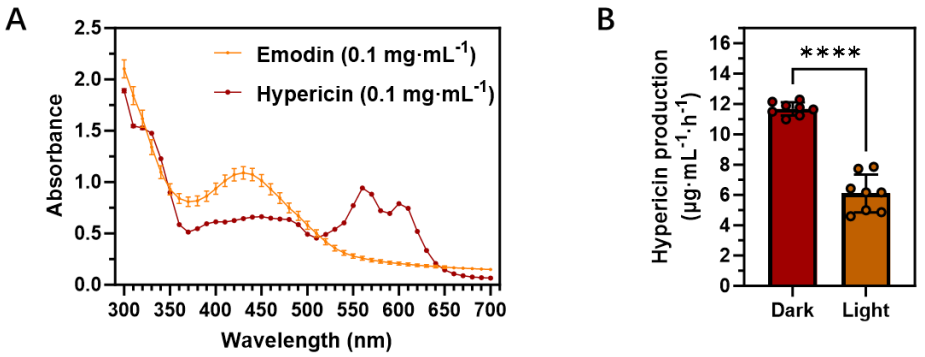

Figure 4. Activity of Hyp1B in catalyzing the conversion of emodin to hypericin.
(A) Full-wavelength scan of emodin and hypericin. (B) Hypericin production by Hyp1B under dark and light conditions.

4 Discussion

In this work, we delineated the distinct behaviors of two hypericin synthase isoforms, *Hyp1A* and *Hyp1B*, in *Escherichia coli*, benchmarking our findings against analogous enzyme-engineering studies.

Strikingly, Hyp1A yielded negligible soluble protein when expressed from an unmodified pET-28a vector; fusing a SUMO tag elevated the soluble yield to 1.41 $mg \cdot g^{-1}$ wet cell mass, yet SUMO Hyp1A remained catalytically inert toward emodin. This underscores that solubility enhancement *per se* does not guarantee functional restoration. Indeed, Smith *et al.*[12] reported that NusA, Trx, or GST fusions improved the solubility of human lysyl oxidase (LOX), but failed to boost—and in some cases even abrogated—its enzymatic activity *in vitro*. Likewise, Xu *et al.*[13] employed machine-learning designed peptide tags to enhance the folding of model enzymes, but did not observe uniform gains in catalytic performance. These precedents mirror our observations: Hyp1A's improved solubility does not translate into enzymatic competence.

By contrast, Hyp1B expresses robustly in the absence of any solubility tag, achieving 7.53 $mg \cdot g^{-1}$ wet cell mass, and remains stable in 10% glycerol. Sequence analysis revealed that three non-conserved residues in Hyp1A—positions 14, 86, and 119—induce local conformational perturbations. All-atom molecular dynamics simulations and DSSP analysis demonstrate that Hyp1A lacks stabilizing α-helical and 3_{10}-helix motifs in critical regions (residues 26 36, ~86, and ~128), which are replaced instead by disordered coils and Turn/Bend elements. `FoldX` calculations further quantify Hyp1A's compromised structural stability ($\Delta\Delta G \approx 4.36$ $kcal \cdot mol^{-1}$), rationalizing its propensity for aggregation. In contrast, Hyp1B's hydrophobic cavity is sculpted by a β-sheet concave face and three intact α-helices, facilitating substrate accommodation —features absent in Hyp1A.

Importantly, purified *Hyp1B* catalyzes *emodin* dimerization to *hypericin*, as evidenced by the characteristic red-shift ($\lambda_{max} \approx 560$ nm). Notably, dark-incubated assays yielded approximately 1.9-fold more *hypericin* than light-exposed controls, echoing Zobayed et al.'s findings that light enhances *hypericin* formation but also accelerates photodegradation[14]. This highlights the necessity of stringent light management during process development.

Although *Hyp1B* exhibits bona fide catalytic activity, turnover remains modest. There are already numerous available methods to enhance the catalytic activity of natural enzymes[15]. We propose that structure-guided mutagenesis of *Hyp1B*'s hydrophobic pocket may improve polarity or geometric complementarity to strengthen substrate binding. Additionally, high-throughput screening or directed evolution could be employed to isolate variants with enhanced substrate affinity and catalytic efficiency. Finally, strategic incorporation of antioxidants (e.g., glutathione), cofactors (e.g., NADPH), or metal ions (e.g., Fe^{2+}, Cu^{2+}), coupled with systematic titration of pH (5.5 7.5) and temperature (20 35 °C), may further maximize *hypericin* yields.

Collectively, our comparative analysis elucidates how subtle sequence variations dictate the solubility, stability, and function of hypericin synthase isoforms. By integrating insights from related enzyme-engineering studies, we lay a roadmap for the rational design of next-generation biocatalysts tailored for scalable *hypericin* production.

Article History

Received: April 25, 2025 **Accepted:** May 10, 2025 **Published:** June 30, 2025

References

[1] Bais, H. P., Vepachedu, R., Lawrence, C. B., Stermitz, F. R., & Vivanco, J. M. (2003). Molecular and biochemical characterization of an enzyme responsible for the formation of hypericin in St. John's wort (*Hypericum perforatum* L.). *The Journal of Biological Chemistry*, **278**(34), 32413 32422.

[2] Fernandes, H., Konieczna, M., Kolodziejczyk, R., Bujacz, G., Sikorski, M., & Jaskolski, M. (2008). Crystallization and preliminary crystallographic studies of Hyp-1, a St John ' s wort protein implicated in the biosynthesis of hypericin. *Acta Crystallographica Section F: Structural Biology and Crystallization Communications*, **64**(5), 405-408.

[3] Karppinen, K., Hokkanen, J., Mattila, S., Neubauer, P., & Hohtola, A. (2008). Octaketide-producing type III polyketide synthase from *Hypericum perforatum* is expressed in dark glands accumulating hypericins. *The FEBS Journal*, **275**(17), 4329-4342.

[4] Butterweck, V., & Schmidt, M. (2007). St. John's wort: role of active compounds for its mechanism of action and efficacy. *Wiener Medizinische Wochenschrift*, **157**(13 14), 356-361.

[5] Murthy, H. N., Kim, Y. S., Park, S. Y., & Paek, K. Y. (2014). Hypericins: biotechnological production from cell and organ cultures. *Applied Microbiology and Biotechnology*, **98**(22), 9187-9198.

[6] Falk, H. (1999). From the photosensitizer hypericin to the photoreceptor stentorin—The chemistry of phenanthroperylene quinones. *Angewandte Chemie International Edition*, **38**(21), 3116-3136.

[7] Kusari, S., Lamshöft, M., Zühlke, S., & Spiteller, M. (2008). An endophytic fungus from *Hypericum perforatum* that produces hypericin. *Journal of Natural Products*, **71**(2), 159 162.

[8] Rizzo, P., Altschmied, L., Ravindran, B. M., Rutten, T., & D'Auria, J. C. (2020). The biochemical and genetic basis for the biosynthesis of bioactive compounds in *Hypericum perforatum* L., one of the largest medicinal crops in Europe. *Genes*, **11**(10), 1210.

[9] Wu, J. J., Zhang, J., Xia, C. Y., Ding, K., Li, X. X., Pan, X. G., Xu, J. K., He, J., & Zhang, W. K. (2023). Hypericin: A natural anthraquinone as promising therapeutic agent. *Phytomedicine*, **111**, 154654.

[10] Kucharíková, A., Kusari, S., Sezgin, S., Spiteller, M., & Čellárová, E. (2016). Occurrence and distribution of phytochemicals in the leaves of 17 in vitro cultured *Hypericum* spp. adapted to outdoor conditions. *Frontiers in Plant Science*, **7**, 1616.

[11] Michalska, K., Fernandes, H., Sikorski, M., & Jaskolski, M. (2010). Crystal structure of Hyp-1, a St. John's wort protein implicated in the biosynthesis of hypericin. *Journal of Structural Biology*, **169**(2), 161-171.

[12] Smith, M. A., Gonzalez, J., Hussain, A., Oldfield, R. N., Johnston, K. A., & Lopez, K. M. (2016). Overexpression of soluble recombinant human lysyl oxidase by using solubility tags: Effects on activity and solubility. *Enzyme Research*, **2016**, 5098985.

[13] Han, X., Ning, W., Ma, X., Wang, X., & Zhou, K. (2020). Improving protein solubility and activity by introducing small peptide tags designed with machine learning models. *Metabolic Engineering Communications*, **11**, e00138.

[14] Jendželovská, Z., Jendželovský, R., Kuchárová, B., & Fedoročko, P. (2016). Hypericin in the light and in the dark: Two sides of the same coin. *Frontiers in Plant Science*, **7**, 560.

[15] Zhou, J., & Huang, M. (2024). Navigating the landscape of enzyme design: From molecular simulations to machine learning. *Chemical Society Reviews*, **53**(16), 8202-8239.