BIOLOGIA PLANTARUM 54 (4): 725-729, 2010

BRIEF COMMUNICATION

Isolation and functional analysis of cDNAs similar to *Hyp-1* involved in hypericin biosynthesis from *Hypericum erectum*

M.L. JIN[1], J.C. AHN[2], B. HWANG[1], H.-S. PARK[3], H.S. LEE[4] and D.-W. CHOI[5]*

Department of Biology, Chonnam National University, Kwangju, 500-757, Korea[1]
Department of Bioscience, Seonam University, Namwon, 590-711, Korea[2]
Korea Research Institute of Bioscience and Biotechnology, Daejeon, 305-333, Korea[3]
Biotechnology division, Korea Forest Research Institute, Suwon, 441-350, Korea[4]
Department of Biology Education, Chonnam National University, Kwangju, 500-757, Korea[5]

Abstract

Hypericin, a naphthodianthrone, has been identified as the principal active compound found in St. John's wort (*Hypericum perforatum* L.). To generate a gene resource for hypericin and other valuable metabolites, we generated expressed sequence tags (ESTs) from *H. erectum*. Analyses of the ESTs enabled us to select three cDNAs, *HeHyp1, HeHyp2,* and *HeHyp3*, evidencing significant sequence homology to *Hyp-1* that were involved in hypericin biosynthesis from *H. erectum*. The deduced amino acid sequence of *HeHyp1* cDNA exhibits 95 % identity with *Hyp-1*. The HeHyp2 and HeHyp3 polypeptides also exhibit 81.1 % identity with Hyp-1. The transcripts of *HeHyp1, HeHyp2,* and *HeHyp3* were detected in the root, stem, leaf, flower, and callus cells. Study using recombinant protein suggests that Hyp-1, HeHyp2, and HeHyp3 may be involved in the biosynthetic of hypericin or other emodin derivatives.

Additional key words: amino acid sequence, ESTs, HPLC, RT-PCR, TLC.

Plants generate a variety of secondary metabolites which perform crucial biological functions in their environmental adaptation, operating as defensive or attractive compounds for other living organisms, or as UV-protective compounds. The plant secondary metabolites are also used in the creation of dyes, flavours, drugs and many other useful products for humans. St. John's wort (*Hypericum perforatum* L.) is a medicinal plant generally used for the treatment of neurological disorders and depression (Deltito and Beyer 1998). Hypericin, a naphthodianthrone, a major active compound found in St John's wort has been determined to function as a monoamine oxidase inhibitor (Linde *et al*. 1996) and has been detected also in other plant species in genus *Hypericum,* including *H. erectum* (Kim *et al*. 2005). Little is currently known regarding the biosynthesis of hypericin, other than that it involves the polyketide pathway, presumably through emodin, with proto-hypericin as the penultimate precursor (Dewick 2002). Synthetically, it can be easily prepared by treating emodin dianthrone with ferrous sulfate and pyridine-N-oxide in pyridine, followed by irradiation of the formed protohypericin (Dewick 2002). Hypericin production by plant tissue cultures has been previously assessed (Kirakosyan *et al*. 2000, Zobayed and Saxena 2003). Recently, Bais *et al.* (2003) cloned a gene (*Hyp-1*) encoding for the enzyme governing the conversion of emodin to hypericin from *H. perforatum*.

In order to develop a gene resource database for plant secondary metabolites, we generated ESTs from *H. erectum*, and identified three cDNAs evidencing significant sequence homology with *Hyp-1*. We assessed

Received 18 January 2009, *accepted* 18 August 2009.
Abbreviations: BA - 6-benzylaminopurine; cDNA - complementary deoxyribonucleic acid; ESTs - expressed sequence tags; HPLC - high performance liquid chromatography; IAA - indole-3-acetic acid; IPTG - isopropyl-thio-β-D-galactoside; MS - Murashige and Skoog medium; LB - Luria-Bertani medium; PCR - polymerase chain reaction; RT-PCR - reverse transcription polymerase chain reaction; SDS-PAGE - sodium dodesyl sulfate polyacrylamide gel electrophoresis; TLC - thin layer chromatography.
Acknowledgement: This work was supported by a grant (No. 2005030-1034486 for DWC) from BioGreen 21 Program, Rural Development Administration (RDA), a grant from ARPC (Agricultural R&D promotion center) for DWC, and a grant (No. M10752000001 for HSP) from the Ministry of Education, Science and Technology of the Korean Government.
* Corresponding author; fax: (+82) 62 530 2509, e-mail: dwchoi63@chonnam.ac.kr

the expression pattern and enzymatic function of the three cDNAs, *HeHyp1, HeHyp2,* and *HeHyp3*.

Hypericum erectum Thunb. plants were cultured in a greenhouse. For *in vitro* propagation, the seeds were surface-sterilized and germinated on Murashige and Skoog's (MS) agar medium. Seedling shoots were cut and transferred to MS medium. Callus tissues were induced from stem explants, which were transferred onto MS medium supplemented with 0.57 μM indole-3-acetic acid (IAA) and 2.22 μM 6-benzylaminopurine (BA). Hypericin was extracted from approximately 0.1 g of fresh tissues, as described by Southwell and Campbell (1991). In brief, ground plant tissues were extracted with 10 cm^3 of diethylether and re-extracted with 10 cm^3 of EtOH after all diethylether was evaporated. The same volume of the hypericin extract from each tissue was spotted onto thin layer chromatography (TLC) plates (*Silica gen 60 F254*, *Merck*, Whitehouse Station, NJ, USA) and analyzed in toluene-ethylacetate-formic acid (5:4:1 v/v). Hypericin content was assessed *via* the TLC-densitometer method as described in previous studies (Mulinacci *et al.* 1999, Kim *et al*. 2005).

In order to identify the cDNA involved in hypericine biosynthesis, expressed sequence tags (ESTs) from *H. erectum* were searched by the amino acid sequence of *Hyp-1* (Bais *et al*. 2003). In addition, the individual ESTs from *H. erectum* were searched against the *GenBank nr* database, using a *BLASTX* algorithm. The putative *Hyp-1* homologous cDNAs were identified *via* key-word searches with the *BLASTX* as described in a previous paper (Choi *et al*. 2005). Multiple sequence alignment, sequence editing and amino acid sequence prediction were conducted from the selected ESTs by the *Sequencher* program (*Gene Code Corporation*, Ann Arbor, MI, USA).

In an effort to assess the organ-specific expression of the selected *HeHyp1, HeHyp2* and *HeHyp3* cDNAs, RNA samples were prepared from the roots, stems, leaves, flowers from soil grown plant, and callus cells, as described above. Leaves were harvested at three different positions on the same branch: young but fully expanded leaves from the top, the mature leaves from the middle and senescing leaves from the bottom. Total RNA (1 μg) was reverse-transcribed in a 100 mm^3 reaction volume using *oligo(dT)*$_{17}$ as a primer with a *QuantiTech* reverse transcription kit, in accordance with the manufacturer's recommendations (*Qiagen*, Hilden, Germany). For PCR-amplification reaction 2 mm^3 cDNA was utilized in a 20 mm^3 solution, which contained 5 mm^3 of 10× PCR buffer (200 mM Tris-HCl, pH 8.4, 500 mM KCl), 1.5 mm^3 of 50 mM $MgCl_2$, 1 mm^3 of 10 mM dNTPs, 1 mm^3 of each gene-specific primer (10 pmol cm^{-3}), and 2.5 units of ExTaq polymerase. The PCR reaction (94 °C, 30 s; 58 °C, 30 s; 72 °C, 30 s) was run for 35 cycles. The gene-specific primers for the *HeHyp1, HeHyp2* and *HeHyp3* cDNAs were as follows: 5'-TTCTGAATA TGGCGGCGTACACTA-3' and 5'-TTACTTGAGAAT CAGTACACGACC-3' for *HeHyp1*, 5'-ACACAGTTA GCAAGGAAGACGAA-3' and 5'-CGTGTGCTGACG ACGACTATT-3' for *HeHyp2*, 5'-ACACAGTTAG CAAGGAAGACGAATC-3' and 5'-ACACCAAGC ACCACAAAATAGGC-3' for *HeHyp3*. As positive controls, a primer set was designed on the basis of the *Arabidopsis* 18S rRNA gene sequence and then employed as an internal positive control. The 3 mm^3 of the RT-PCR reaction mix was loaded onto 1 % agarose gel and separated.

Open reading frame of the *HeHyp1* and *HeHyp3* were amplified *via* PCR using the primer, 5'-GCG CATATGTCGGTGTACACTATTGTTAAG-3' (sense) harboring the *Nde* I restriction enzyme site and 5'-GAGCTCGAGAGCGAAAACTTCAGGATTAGC-3' (antisense) harboring the *Xho* I restriction enzyme site for *HeHyp1*, and 5'-GCGCATATGGCAGTGTACACAGTT AGCAAG-3' (sense) harboring the *Nde* I restriction enzyme site and 5'-ATACTCGAGAGCGAAAGCTTC AGGGTGGGC-3' (antisense) harboring the *Xho*I restriction enzyme site for *HeHyp3*. The PCR products were then cloned into *Nde* I/*Xho* of the pET21a expression vector (*Novagen*, Madison, WI, USA) and then transformed into *Escherichia coli* BL21 (DE3) CodonPlus RIL cells. Isopropyl-thio-β-D-galactoside (IPTG) was added to a final concentration of 0.4 mM in order to induce the expression of the fusion protein. The recombinant HeHyp1 and HeHyp3 proteins were purified *via* Ni-chelate chromatography. The recombinant protein was analyzed on SDS-PAGE.

In order to evaluate the conversion of emodin to hypericin by recombinant HeHyp1 and HeHyp3, IPTG-induced *E. coli* cells were harvested *via* centrifugation and resuspended in fresh Luria-Bertani (LB) medium containing emodin (20 μg cm^{-3}) as well as antibiotics and IPTG and incubated for an additional 4 h. The recombinant HeHyp1 and HeHyp3 proteins were incubated for 1 h in 0.05 M phosphate buffer (pH 6.5) containing 20 μM emodin at 25 °C. The conversion of emodin to hypericin was assessed initially by colour change according to the method of Bais *et al.* (2003). The hypericin formed was identified by HPLC.

Previously, we reported that hypericin is detected in other plant species of genus *Hypericum,* including *H. erectum* (Kim *et al*. 2005). Hypericin was detected principally in the leaves and flowers, however, traces were detected in the roots and stems of *H. erectum* (Table 1). Walker *et al.* (2002) observed high hypericin contents in the dark grown cells and leaf tissue of *H. perforatum,* however, trace amounts in the stems and

Table 1. Hypericin content [μg g^{-1}(d.m.)] in different tissues of *H. erectum.*

Tissue	Hypericin	Tissue	Hypericin
Root	4.4 ± 0.71	leaf 1	206.0 ± 6.66
Stem	3.0 ± 1.00	leaf 2	200.7 ± 13.6
Flower	352.1 ± 6.52	leaf 3	103.4 ± 13.3
Callus	16.4 ± 3.68		

roots. These results show that the accumulation of hypericin in *H. erectum* is similar to that in *H. perforatum* (Zobayed *et al*. 2006).

The biosynthetic pathway of hypericin remains unclear, although it involves the polyketide pathway. It has been suggested that hypericin might be converted from protohypericin, which is synthesized by subsequent oxidative reactions of emodin dianthrone (Fig. 1). Emodin dianthrone may be formed *via* the oxidative coupling of two emodin anthrones or the condensation of an emodin anthrone and emodin, which is converted from emodin anthrone by oxidation (Dewick 2002). Recently, Bais *et al.* (2003) suggested that hypericin is converted directly from emodin *via* an enzymatic dimerization reaction (Fig. 1). They reported the cloning of the gene, *Hyp-1,* which mediates the direct conversion of emodin to hypericin. The *Hyp-1* encodes for a polypeptide of 159 amino acids and is abundantly expressed in cell cultures in darkness. Bais *et al.* (2003) suggested that Hyp-1 catalyzes an initial condensation reaction to form emodin dianthrone, followed by a sequence of oxidative cyclizations to yield hypericin. Despite the considerable amount of commercial interest in hypericin, the *Hyp-1* is the first isolated gene, to the best of our knowledge, and the biosynthetic pathway of hypericin remains to be characterized at the molecular level.

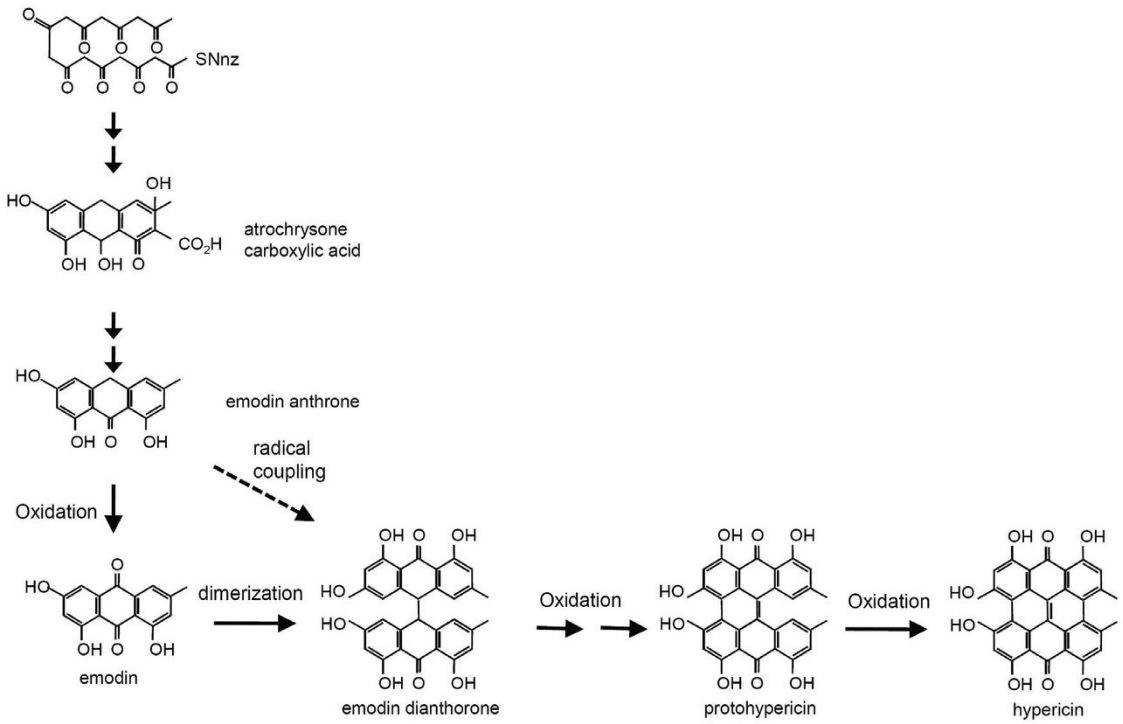

Fig. 1. Putative biosynthetic pathway of hypericin.

```
HYP-1       MAAYTIVKEEESPTAPHRLFKALVLERHQVLVKAQPHVFKSGEIIEGDGGVGTVTKITFV
HeHYP1      MSVYTIVKEEESPIAPHRLFKALVLERHQVLVKAQPHVFKSGEIIEGDGGVGTVTKITFV
HeHYP2      MAVYTVSKEDESPVAAHRLFKALVLERHNVLVKAQPHVFSTGEIIEGDGGVGTVTKITFA
HeHYP3      MAVYTVSKEDESPVAAPRLFKALVLERHNVVVKAQPHVFSTGEIIEGDGGVGTVTKITFA
            *:.**: **:*** *. **********:*:********.:******************.
HYP-1       DGHPLTYMLHKFDEIDAANFYCKYTIFEGDVLRDNIEKVVYEVKLEAVGGGSKGKITVSY
HeHYP1      DGHPFTYMLHKFDEIDAANFYCKYTLFEGDVLRDNIEKVVYEVKLESVGGGSKGKITVTY
HeHYP2      PGHPFKYMLHRFDELDVANCYCKYTLFEGDVLRDNLEKVVYEVKVESVGTGSKGKITVHY
HeHYP3      PGHPFKYMLHRFDELDVANCYCKYTLFEGDVLRDNIEKIVYEVKVESVGTGSKGKITVHY
             ***:.****:***:*.** *****:*********:**:*****:*:** ******** *
HYP-1       HPKPGCTVNEEEVKIGEKKAYEFYKQVEEYLAANPEVFA
HeHYP1      HPKAGCTVNEEEVKIGEKKAYEFYKQVEEYLAANPEVFA
HeHYP2      HPKPGCSVNEEEIKLGQEKAHALYKQVEEYLVANPDVFA
HeHYP3      HPKPGCTVNEEEIKLGQEKAHAFYKQVEEYLVAHPEAFA
            ***.**:*****:*:*::**: :********.*:*:.**
```

Fig. 2. Comparison of the deduced amino acid sequences of HeHyp1, HeHyp2 and HeHyp3 from *H. erecta* and Hyp-1 from *H. perforatum*. Amino acid residues are designated *via* single-letter codes. ***** and **.** indicate identical and homologous amino acids, respectively; **:** indicates places in which a sequence has been expanded to allow for optimal sequence alignment. Alignments of the deduced amino acid sequences were conducted with the *CLUSTAL V* program.

In order to construct a database for the gene resources involved in secondary metabolites, we generated approximately 4 000 ESTs from *H. erectum* (unpublished data). *H. erectum* ESTs data analysis allowed us to identify 3 cDNAs that were similar to *Hyp-1*. All of these cDNAs harbor an open reading frame of 159 amino acid residues. We designated these cDNAs as *HeHyp1* (*H. erectum Hyp-1* like protein-1), *HeHyp2,* and *HeHyp3* (Fig. 2). The deduced amino acid sequence of HeHyp1 exhibits 95 % identity with Hyp-1. These findings show that *HeHyp1* may be an orthologue of *Hyp-1* and may be involved in the hypericin biosynthetic pathway in *H. erectum*. The HeHyp2 and HeHyp3 polypeptides also exhibit 81.1 % identity with Hyp-1. Both polypeptides have approximately 84 % amino acid sequence identity with HeHyp1. HeHyp2 and HeHyp3 evidence 94.4 % amino acid sequence identity with each other, although both cDNAs differ significantly in terms of their cDNA sequences. These results also indicate that HeHyp2 and HeHyp3 may have similar biochemical functions, including the biosynthesis of hypericin or other emodin derivatives. All three of the genes, *HeHyp1, HeHyp2,* and *HeHyp3* are basically expressed in the roots, stems, leaves, flowers, and callus cells at a similar level, although the contents of hypericin were different in the different tissues (Fig. 3). The transcript of *Hyp-1* from *H. perforatum* was detected in the roots, stems and suspension-cultured cells. *Hyp-1* gene expression was more abundant in the cells cultured under dark conditions than those grown under light (Bais *et al*. 2003). Our results indicate that *HeHyp1*, a putative orthologue of *Hyp-1*, was expressed in the majority of the tissues and the whole developmental program at a similar level.

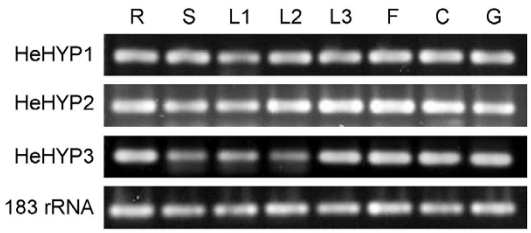

Fig. 3. Expression of the *HeHyp1, HeHyp2* and *HeHyp3* gene in different tissues. Total RNA was isolated from root (R), stem (S), top leaf (L1), middle leaf (L2), bottom leaf (L3), flower (F) and callus cell (C). The RNAs were reverse-transcribed using the 3'-primer, and then amplified using gene-specific primer sets. Genomic DNA (G) was utilized as a PCR control. The RT-PCR products were separated on 1.0 % agarose gel.

In order to assess the biochemical function of the selected gene, *HeHyp1* and *HeHyp3* cDNAs were expressed in *E. coli*. SDS-PAGE yielded a single protein band of approximately 20 kDa, which corresponded to the molecular masses of HeHyp1 and HeHyp3 deduced from the cDNA sequence (Fig. 4*A*). During the first step, we assessed the HeHyp1 and HeHyp3 activity by a change in colour to red, as described by Bais *et al.* (2003). This test is based on the fact that the emodin substrate is yellow but the enzymatic reaction product, hypericin, is red. When the transformed bacteria cells to which *HeHyp1* or *HeHyp3* cDNA was introduced were cultured in medium containing emodin, the transformed bacteria became red (Fig. 4*B*). However, the untransformed bacteria and transformed bacteria with the empty vector, pET21a, exhibit no red colour. These results suggest that HeHyp1 and HeHyp3 converted emodin to hypericin. The activity of HeHyp1 was substantially higher than that of the HeHyp3 (Fig. 4*B*). The purified HeHyp1 and HeHyp3 recombinant proteins were incubated with 20 μM emodin, and the reaction mixtures were extracted with methanol. The recombinant HeHyp1 and HeHyp3 protein reactions showed similar red colour as hypericin in methanol (Fig. 4*C*). Thus HeHyp1 and HeHyp3 can be involved in the conversion of emodin to hypericin. However, we failed to detect the hypericin in HeHyp3 as well as the HeHyp1 recombinant protein reaction

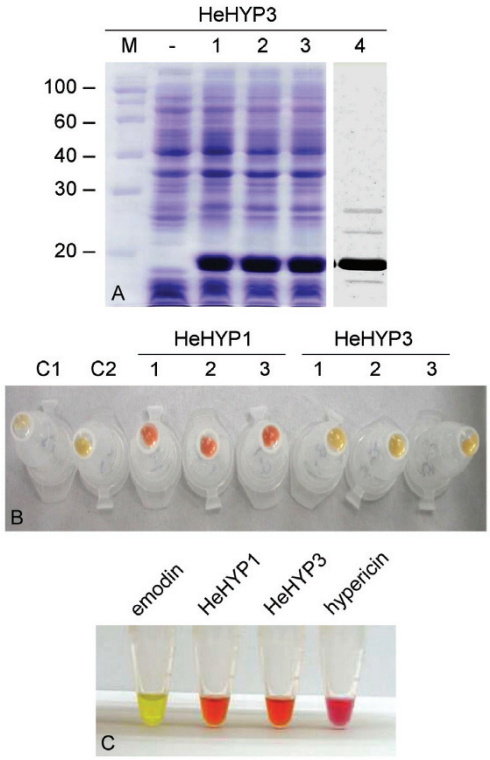

Fig. 4. Heterologous expression of *HeHyp1* and *HeHyp3* cDNA in *E. coli* and the color assay of the recombinant protein for hypericin biosynthesis. *A* - The expression of the fusion protein was induced *via* the addition of IPTG. M: molecular size marker, -: no IPTG, *lanes* 1 - 3: IPTG induced different clone of HeHyp3, *lane* 4: recombinant HeHyp3 proteins purified using Ni^{2+} chelated chromatography. Proteins were separated on 12 % SDS-PAGE. *B* - Transformed *E. coli* with *HeHyp1* and *HeHyp3* cDNA was cultured in LB medium containing 20 μM emodin and then harvested. C1: no IPTG and emodin; C2: no IPTG and add emodin; 1 - 3: different clones with IPTG and emodin. *C* - Purified recombinant HeHyp1 and HeHyp3 proteins were incubated in phosphate buffer including emodin, after which the reactions were extracted with methanol. Emodin is yellow, while hypericin is red.

mixtures by HPLC. These results indicate that HeHyp1 and HeHyp3 may not be involved in hypericin biosynthesis, or that the enzymatic activity of these proteins may be too low for hypericin to be detected under our HPLC conditions. Another possibility is that the red colour in the recombinant protein reaction mixture may not be attributable to hypericin, rather to the presence of other emodin derivatives.

The results of this study indicate that there are several genes with high degrees of sequence homology with *Hyp-1* that are involved in the hypericin biosynthesis in the genome of the *Hypericum* genus. These genes, which include *HeHyp2* and *HeHyp3,* may be involved in the biosynthetic pathway of the other secondary metabolites from emodin or emodin derivatives.

References

Bais, H.P., Vepachedu, R., Lawrence, C.B., Stermitz, F.R., Vivanco, J.M.: Molecular and biochemical characterization of an enzyme responsible for the formation of hypericin in St. John's wort (*Hypericum perforatum* L.). - J. biol. Chem. **278**: 32413-32422, 2003.

Choi, D.W., Jung, J.D., Ha, Y.I., Park, H.W., In, D.S., Chung, H.J., Liu, J.R.: Analysis of transcripts in methyl jasmonate-treated ginseng hairy roots to identify genes involved in the biosynthesis of ginsenosides and other secondary metabolites. - Plant Cell Rep. **23**: 557-566, 2005.

Deltito, J., Beyer, D.: The scientific, quasi-scientific and popular literature on the use of St. John's wort in the treatment of depression. - J. Affect. Disord. **51**: 345-351, 1998.

Dewick, P.M.: Medicinal Natural Products, a Biosynthetic Approach. 2nd Ed. - John Wiley and Sons, New York 2002.

Kim, S.H., Jung, Y.J., Ahn, J.C., Hwang, B.: Hypericin contents of *Hypricum erectum* Thunberg. - Kor. J. med. Crop Sci. **13**: 101-104, 2005.

Kirakosyan, A., Hayashi, H., Inoue, K., Charchloglyan, A., Vardapetyan, H.: Stimulation of the production of hypericins by mannan in *Hypericum perforatum* shoot cultures. - Phytochemistry **53**: 345-348, 2000.

Linde, K., Ramirez, G., Mulrow, C.D., Pauls, A., Weidenhammer, W., Melchart, D.: St John's wort for depression – an overview and meta-analysis of randomized clinical trials. - Brit. med. J. **313**: 253-258, 1996.

Mulinacci, N., Bardazzi, C., Romani, A., Pinelli, P., Vincieri, F.F., Costantini, A.: HPLC-DAD and TLC-densitometry for quantification of hypericin in *Hypericum perforatum* L. extracts. - Chromatography **49**: 197-201, 1999.

Southwell, I.A., Campbell, M.A.: Hypercin content variation in *Hypericum perforatum* in Australia. - Phytochemistry **30**: 475-478, 1991.

Walker, T.S., Bais, H.P., Vivanco, J.M.: Jasmonic acid-induced hypericin production in cell suspension cultures of *Hypericum perforatum* L. (St. John's wort). - Phytochemistry **60**: 289-293, 2002.

Zobayed, S.M.A., Saxena, P.K.: *In vitro*-grown roots: a superior explants for prolific shoot regeneration of St. John's wort (*Hypericum perforatum* L. cv 'New Stem') in a temporary immersion bioreactor. - Plant Sci. **165**: 463-470, 2003.

Zobayed, S.M.A., Afreen, F., Goto, E., Kozai, T.: Plant-environment interaction: Accumulation of hypericin in dark glands of *Hypericum perforatum*. - Ann. Bot. **98**: 793-804, 2006.