Volume 330, number 2, 169–173 FEBS 12957 September 1993

cDNAs sequences encoding cytochrome P450 (CYP71 family) from eggplant seedlings

Naoyuki Umemoto, Osamu Kobayashi, Osamu Ishizaki-Nishizawa, Toshihiro Toguri*

Central Laboratory for Key Technology, Kirin Brewery Co., 1-13-5 Fukuura Kanazawa-ku, Yokohama 236, Japan

Received 24 July 1993

Three cytochrome P450 (P450) cDNAs were isolated from an eggplant hypocotyl cDNA library using eggplant *CYP75* cDNA as a probe. These cDNAs have greater than 65% identity in their amino acid sequences, indicating that they belong to the same family. Comparison of these with P450 proteins from other sources showed that the protein with the greatest degree of homology is CYP71, isolated from avocado fruits (~ 48%). We concluded that they are novel members of the *CYP71* gene family (*CYP71A2, 3* and *4*). We have examined the level of mRNA transcripts from the *CYP71* family in eggplant hypocotyl tissues and petunia flower buds, and found that the level of transcripts is developmentally regulated in the flower bud.

Cytochrome P450; cDNA sequence; Hypocotyl; *Solanum melongena*; *Petunia hybrida*

1. INTRODUCTION

Cytochrome P450 (P450)-dependent monooxygenases have been studied extensively in animals, bacteria, yeasts, and fungi. Many members of the P450 superfamily, mostly from animals, have been characterized with respect to their primary structure and substrate specificity [1]. In contrast, with the exception of *CYP71* from ripening avocado fruits [2], *CYP72* from cultured periwinkle cells [3] and *CYP73* from mungbean and Jerusalem artichoke [4,5], there is scant information about P450s in higher plants. In plants, P450 enzymes are known to catalyze major oxidative reactions leading to the synthesis of secondary metabolites, such as phenolics, membrane sterols, phytoallexins and terpenoids [6,7]. A number of pathways in which the P450 enzymes participate produce only a small amount of end product, for example the gibberellins, and as the result of the low level of enzyme activity analysis is difficult. Moreover, since P450 enzymes are tightly bound to the membranes, purification and analysis of their structure and function is also difficult. In fact, only CYP73 has been assigned with a precise function [4,5], while the physiological functions of CYP71 and CYP72 remain unclear [2,3,8]. We have cloned a P450 cDNA (*CYP75*) from eggplant hypocotyl tissues, and it was observed that the accumulation of anthocyanins and the RNAs involved in their synthesis, including CYP75, was controlled by UV light [9].

Here, we describe the structure of novel P450 cDNAs which are members of the *CYP71* family, and describe their expression in the eggplant and petunia.

*Corresponding author. Fax: (81) (45) 788 4042.

Abbreviations: cDNA, DNA complementary to RNA; SDS, sodium dodecyl sulfate; SSC, 0.15 M NaCl, 0.015 M Na_3 citrate pH 7.0; SSPE, 0.15 M NaCl, 10 mM Na phosphate, 1 mM EDTA pH 7.4.

2. MATERIALS AND METHODS

2.1. *Materials*

Eggplant seeds (*Solanum melongena* cv. Sinsadoharanasu) and petunia seeds (*Petunia hybrida* cv. Blue star) were purchased from Tohoku seeds Co., Japan and Sakata seeds Co., Japan, respectively. [α-^{32}P]dCTP (specific activity 1.1×10^{14} Bq/mmol) was obtained from New England Nuclear. Nylon membranes were obtained from Amersham. A cDNA synthesis kit and an in vitro packaging kit were purchased from Stratagene. Restriction endonucleases were purchased from Toyobo Biochemicals, Tokyo and New England Biolabs.

2.2. *Cloning and sequencing*

Poly(A)$^+$ RNA was isolated from the hypocotyl tissues of eggplant seedlings as described previously [9]. A cDNA library was constructed from the poly(A)$^+$ RNA using the vector λ ZAPII according to the Stratagene cDNA synthesis kit. The 0.6 kb fragment covering the 3′ region of the *CYP75* cDNA (from the *Apa*I site in Fig. 3e to the 3′ terminus: EMBL accession number X70824) [9] was labelled with [α-^{32}P]dCTP, and was used as a probe. Approximately 2×10^5 plaques were screened under hybridization conditions of low stringency (5 × SSPE/5 × Denhardt's solution/1% SDS/100 μg per ml of denatured salmon sperm DNA at 42°C for 24 h; washing in 2 × SSC/0.1% SDS at 42°C). In the case of high stringency screening, hybridization and washing were performed at 65°C in the solution mentioned above. The DNAs isolated from positive plaques were subcloned in vivo into pBluescript, using the helper phage, VSCM13 (Stratagene). The DNA was deleted for sequencing by digestion with appropriate restriction endonucleases, or exonuclease III (Takara Shuzo Co., Kyoto) [10]. DNA sequencing was performed for both DNA strands by dideoxy chain-termination methods, using a Sequenase II kit (United States Biochemical) with [α-^{32}P]dCTP, or a fluorescent-dye-based cycle sequence system (Applied Biosystems). Sequence comparison was car-

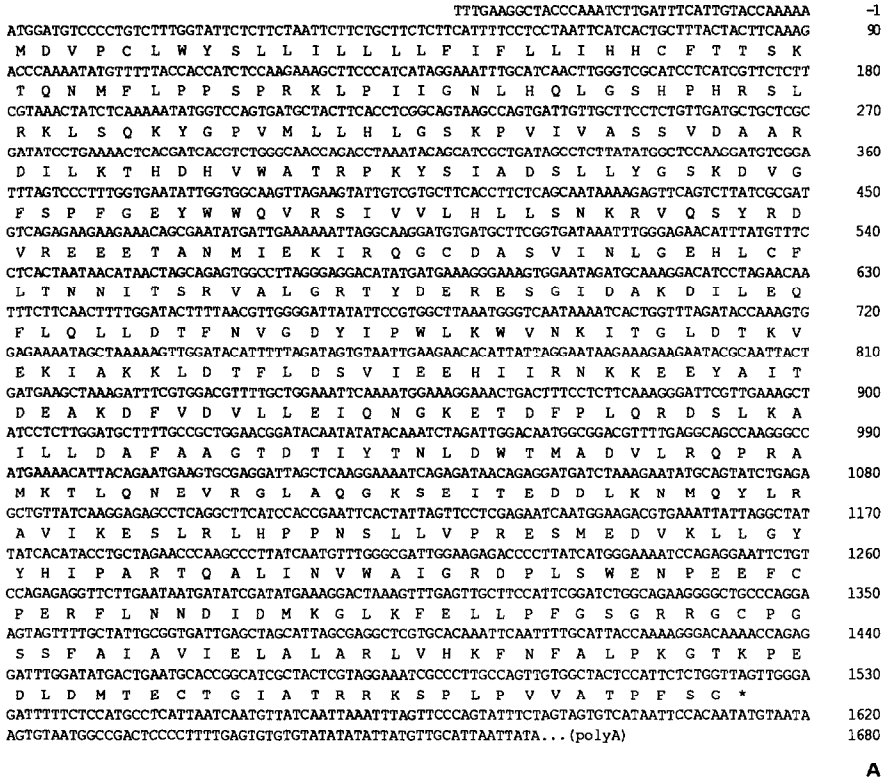

Fig. 1. The nucleotide sequence and the deduced amino acid sequence of *CYPEG2/CYP71A4* (A) and *CYPEG4/CYP71A2* (B). The amino acid sequence deduced from an open reading frame is shown below the nucleotide sequences. Both cDNAs are polyadenylated. The nucleotide sequences have been submitted to the DDBJ, EMBL and GenBank Nucleotide Sequence Databases with the accession numbers X70981 (*CYPEG2/CYP71A4*), X70982 (*CYPEG3/CYP71A3*; only the deduced amino acid sequence is shown in Fig. 2) and D14990 (*CYPEG4/CYP71A2*), respectively.

ried out using GENETYX software (Software Development Co., Ltd., Tokyo).

2.3. *Southern and Northern blot analysis*

Genomic DNA from eggplant (10 μg) was digested with restriction enzymes, separated by electrophoresis, and transferred to a Hybond N^+ membrane (Amersham) with 0.4 N NaOH. The membrane was hybridized with a ^{32}P-labelled DNA probe made from each clone at 65°C, washed at 65°C in 2 × SSPE/0.1% SDS (low stringency) or in 0.1 × SSPE/0.1% SDS (high stringency), and autoradiographed.

Northern blot analysis was carried out as described previously [9]. Briefly, mRNAs from non-induced and white light-induced hypocotyl tissues, in addition to mRNAs isolated from petunia leaves, and flower buds at three developmental stages, were prepared. The developmental stage of the flower buds was determined by the length of the flower buds (early, less than 20% of maximal growth; intermediate, 40% of maximal growth; late, more than 80% of maximal length, which was about 5 cm). Next, the mRNAs were separated on a 1% agarose gel containing 2.2 M formaldehyde, and transferred with 10 × SSC to a nylon membrane. The transferred RNA was hybridized with a ^{32}P-labelled *CYPEG2* DNA probe.

3. RESULTS AND DISCUSSION

A cDNA library constructed from poly(A)$^+$ RNA from eggplant hypocotyl tissues grown under white light followed by red light irradiation [9], was screened with a ^{32}P-labelled *CYP75* DNA probe by hybridization under conditions of low stringency. One clone that was isolated and named pE138 had a significant degree of homology with CYP75. Using this clone as a probe, the cDNA library was rescreened under conditions of high stringency. Several positive clones were isolated and were classified into three groups, based on their restriction maps and the intensity of their hybridization signals. The longest sequence isolated in each group was sequenced. The 1720-bp insert of one clone CYPEG2 contains an open reading frame coding for a protein consisting of 508 amino acid residues (Fig. 1A). The insert of the clone CYPEG4 is 1743 bp in length, and contains an open reading frame encoding a protein of 506 amino acid residues (Fig. 1B). The primary structure of the CYPEG4 protein is 83% homologous to the CYPEG2 protein. The insert of the third clone CYPEG3 is truncated at both the N- and C-termini, as determined by comparison with the CYPEG2 sequence, but this truncated sequence has about 70% homology with CYPEG2 and CYPEG4. This indicates that these P450s belong to the same gene family. A computer search of SWISSPROT and GenBank databases showed that avocado fruit CYP71 has the highest degree of homology with these three eggplant P450s, sequence homology is approximately 48%. However, they show less than 35% homology with CYP75 isolated from the same plant. The amino acid sequences with the greatest homology were found to be mainly located in

```
                                 : .   ..     .    .  :.::::.:::::::::::::  ::::: .:. ..::..:::.
CYPEG2       1 : MDVPCLWYSLLILLLLFIFLLIHHC-F----TTSKTQNMFLPPSPRKLPIIGNLHQLGSHPHRSLRKLSQKYGPVMLLHL
CYPEG4       1 : MDVPCPWYSLLIPLFVFIFLLIHHCFF----TTSKKQNMLLLPSPRKLPIIGNLHQLGSLPHRSLHKLSQKYGPVMLLHF
CYP71        1 :           MAILVSLLFLAIALTFFLLKLNEKREKKPNLPPSPPNLPIIGNLHQLGNLPHRSLRSLANELGPLILLHL
CYPEG3       1 :                 LVPLFVFILFLHKCFFTTSNNNKKLPPSPRKLPIIGNLHQLGLHPHRSLHKLSKKYGPVMLLHL

                  :..:....:.. ....: ::.: ... ::.   .... :...:: ::: :::::.. ..: ::.:::..::.::: :.::
CYPEG2      76 : GSKPVIVASSVDAARDILKTHDHVWATRPKYSIADSLLYGSKDVGFSPFGEYWWQVRSIVVLHLLSNKRVQSYRDVREEE
CYPEG4      77 : GSKPVIVASSVDAARDIMKTHDVVWASRPKSSIVDRLSYGSKDVGFSPFGEYWRRAKSITVLHLLSNTRVQSYRNVRAEE
CYP71       71 : GHIPTLIVSTAEIAEEILKTHDLIFASRPSTTAARRIFYDCTDVAFSPYGEYWRQVRKICVLELLSIKRVNSYRSIREEE
CYPEG3      65 : GSKPVIVASSVEAVRDIMKTNDLVWSNRPKSKMADRLIYGSKDVSFSPHGEYWRQIRSITVLHLLSNKRVQSYRAAREEE

                  . .:.... : .   .  .: . .. :  ... :::.:  :  : .           .  :.. :.:::::.: . ::..
CYPEG2     156 : TANMIEKIRQGC-DAS-VINLGEHLCFLTNNITSRVALGRTYD-ERESGIDAKDILEQFLQLLDTFNVGDYIPWLKWVNK
CYPEG4     157 : TANMIGKIRQGC-DSS-VINLGEHLCSLTNNIISRVALGRTYD-EKESGIE--HIIEQFVELLGIFNVGDYIPRLEWVNK
CYP71      151 : VGLMMERISQSC-STGEAVNLSELLLLLSSGTITRVAFGKKYEGEEERKNKFADLATELTTLMGAFFVGDYFPSFAWVDV
CYPEG3     145 : TSNMIEKLKQMSNSSSSATDLRDLFCWLAYNIINRVALGKKYNDEID----AKATLDKFVELLGTFNVGDYIPCLEWVNK

                   :: : .. . .. :: ..  ::. : .  ....     : :: ::::::.:..... ....:..::: .:: : .::::
CYPEG2     233 : ITGLDTKVEKIAKKLDTFLDSVIEEHIIRNKKEEYAITDEAKDFVDVLLEIQNGKETDFFLQRDSLKAILLDAFAAGTDT
CYPEG4     232 : FTGLDAKVKKVAKELDMFLEIVIEEHIIRKKKEEYTSTGEAKDFVDVLLEIQNGNETDFPLQRDSLKAILLDSFAAGTDT
CYP71      230 : LTGMDARLKRNHGELDAFVDHVIDDHLLSRKANG-SDGVEQKDLVDVLLHLQKDSSLGVHLNRNNLKAVILDMFSGGTDT
CYPEG3     221 : ITGFDSKVDKVAKDLDTFLEFVIEAHMIRNEKEE-NRAGESKDLVDVLLEIQNGKETGFPIQRDSLKALLLDPFSAGTDT

                     : :.:..... :...  .: ::....: : ...:.:: .. :: ..::: ::::::    ::.:::: ::   : :
CYPEG2     313 : IYTNLDWTMADVLRQPRAMKTLQNEVRGLAQGKSEITEDDLKNMQYLRAVIKESLRLHP-PNSLLVPRESMEDVKLLGYY
CYPEG4     312 : TFATLDWTMAELLRQPRALKTLQDEVRGLAQGKSEITEDDLKNMQYLRAVIKESLRLHPTQESLLVPRESMEDVNLLGYY
CYP71      309 : TAVTLEWAMAELIKHPDVMEKAQQEVRRVVGKKAKVEEEDLHQLHYLKLIIKETLRLHP-VAPLLVPRESTRDVVIRG-Y
CYPEG3     300 : IYTVLEWTMTELLRHPRAMEKLQNEVQGLAQEKAEITEDDLPNMHYLKAVIKETLRLHPPIPLLLP

                  :::: :   :: ::::::: :::: ::  :::: :   : ::  : : ::: ::::*::  : :   :  :: :   :
CYPEG2     392 : HIPARTQALINVWAIGRDPLSWENPEEFCPERFLNNDIDMKGLKFELLPFGSGRRGCPGSSFAIAVIELALARLVHKFNF
CYPEG4     392 : HIPARTQAIINAWAIGRDPLSWENPEEYQPERFLNSDADVKGLNFKLLPFGAGRRGCPGSSFAIAVIELALARLVHKFDF
CYP71      387 : HIPAKTRVFINAWAIGRDPKSWENAEEFLPERFVNNSVDFKGQDFQLIPFGAGRRGCPGIAFGISSVEISLANLLYWFNW

                  ::
CYPEG2     472 : ALPKGTKPEDLDMTECTGIATRRKSPLPVVATPFSG
CYPEG4     472 : ALPEGIKPEDLDMTETIGITTRRKLPLLVVATPC
CYP71      467 : ELPGI
```

Fig. 2. Multiple alignment of amino acid sequences for CYPEG2, CYPEG3, CYPEG4 and CYP71. Matching sequences were determined using the software mentioned in section 2. The period above the sequence indicates that three out of the four sequences are the same, the colon indicates the position of residues conserved in all genes, and an asterisk shows the position of cysteine, conserved in P450 monooxygenases in the heme-binding site.

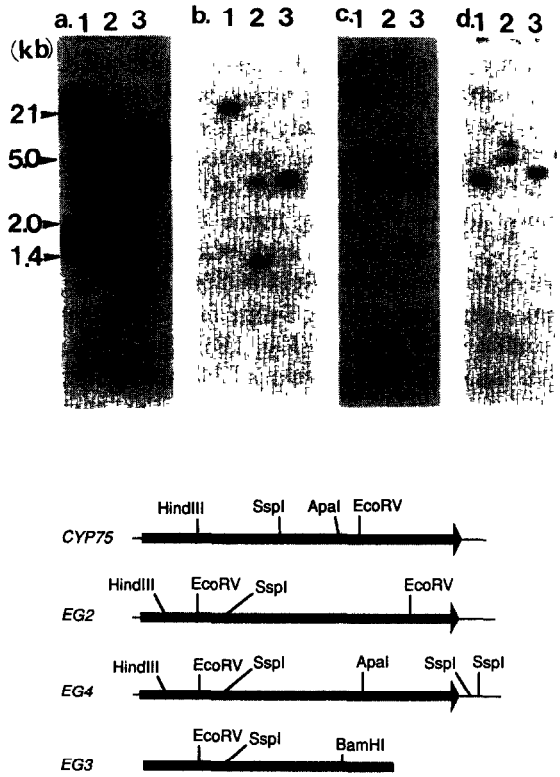

Fig. 3. (a–d) Genomic Southern analysis under hybridization conditions of high and low stringency. Genomic DNA from the eggplant (10 μg) was digested with *Bam*HI (lane 1), *Eco*RV (lane 2) or *Hin*dIII (lane 3). The DNA was transferred onto a membrane and was hybridized with *CYPEG2* gene (a and b), *CYPEG4* gene (c) or *CYPEG3* gene (d), and washed under conditions of low stringency (a) or under conditions of high stringency (c, d). The *Eco*RI/*Hin*dIII fragments of lambda phage DNA were used as a size calibration standard. (Lower panel) The cDNA restriction endonuclease sites are shown. The arrow indicates the coding region of each clone and its direction of transcription.

the carboxyl-terminal region in CYP71, CYPEG2, CYPEG3, and CYPEG4 (Fig. 2). It is known that this region contains sequences highly conserved among P450s, such as the heme-binding site [1]. In the amino-terminal region, other conserved sequences, such as amino acid residues 36–50 (proline-rich segment), and 120–150 were found.

The cDNA clones of *CYP75*, *CYPEG2*, *CYPEG3* and *CYPEG4* have different restriction maps (Fig. 3e). Therefore, mapping followed by Southern hybridization can be used to distinguish between each genomic clone. Southern analysis under conditions of high stringency showed that the *CYPEG2* gene hybridized to one or two bands (Fig. 3b). In contrast, under conditions of low stringency, the probe recognized numerous bands (Fig. 3a), including bands of the same mobility as were observed under conditions of high stringency (Fig. 3b–d). Each cDNA hybridized to one or two bands of different mobility, implying that each cDNA is encoded by a single gene or a small number of genes. Although the majority of the multiple bands could be correlated to the three cDNAs, there were residual bands which could not be detected under conditions of high stringency. Thus, it will be possible to isolate other cDNA(s) by rescreening the library. These results, in addition to the similarity in sequence indicate that these genes form a family in the eggplant, and they belong to the same family as *CYP71*. This is because genes in the same family share greater than 40% amino acid sequence homology [11].

To study the function(s) of these genes, the expression pattern of *CYPEG2* was investigated in the eggplant and petunia. We showed that the *CYP75* transcripts accumulated by switching irradiation from red to white in eggplant hypocotyl tissues, this was in agreement with results obtained for other flavonoid biosynthesis genes [9]. Northern analysis of mRNA isolated from hypocotyl tissues before and after white light irradiation indicates that the level of CYPEG2 transcripts remained almost constant during both sets of light conditions (Fig. 4a), implying that the *CYPEG2* gene does not respond to this change in irradiation.

Next, the *CYPEG2* gene transcripts in petunia plants were studied using mRNAs isolated from leaves and flower buds at various developmental stages. We could detect a weak signal during the intermediate stage of flower bud development, and a rapid accumulation during the late stage, but could not detect transcripts in flower buds during the early stage, or in young leaves (Fig. 4b). This presents a striking contrast with the *CYP75* gene, which showed maximal expression during

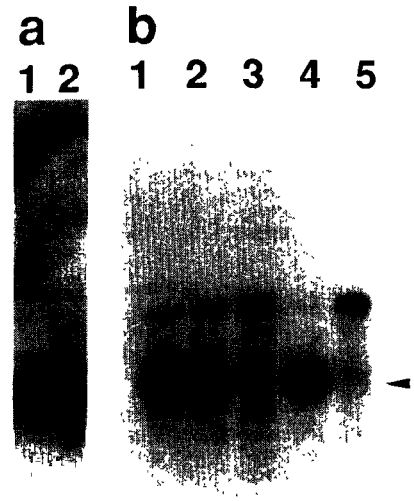

Fig. 4. Northern analysis of the *CYPEG2* gene in eggplant hypocotyl tissues and petunia tissues. Poly(A)$^+$ RNAs were prepared from non-induced (a, lane 1) and white light-induced (a, lane 2 and b, lane 4) hypocotyl tissues from eggplant, petunia flower buds of three developmental stages (b, lane 1, early; lane 2, intermediate; lane 3, late) and petunia leaves (b, lane 5). The mRNAs were transferred onto a membrane and were hybridized with the *CYPEG2* cDNA.

the intermediate stage of flower bud development, as is common with other flavonoid synthesis genes [9]. These hybridization results indicate that the P450s described here do not have a role in flavonoid metabolism. It is suggested that the physiological function of avocado CYP71 is the hydeoxylation of monoterpenoids [12]. Recently, Nebert and Nelson named CYPEG2, 3 and 4 as CYP71A4, A3 and A2, respectively (personal communication). In the case of eggplant *CYP71*s and the homologous petunia gene(s), we suggest that they may have a role in maturation, such as during flavor formation or other metabolite production specific to aging tissues. This is because the tissues expressing these P450 genes were fully elongated, and suffered from senescence when they were harvested. Another important point is that the expression of this gene family is not restricted to ripening fruits; it is possible that similar pathway(s) may function in different tissues, such as seedlings and flowers. This might provide us an important clue to determining the function of these genes. In order to study the function of these genes, we plan to introduce them into plants and study their expression.

Acknowledgements: We wish to thank Dr. T. Ohtani for valuable discussions and R. Kaneko for her help with experiments. In addition, we gratefully acknowledge Drs. D. Nebert and D. Nelson for naming the P450 cDNAs.

REFERENCES

[1] Nebert, D.W. and Gonzalez, F.J. (1987) Annu. Rev. Biochem. 56, 945–993.
[2] Bozak, K.R., Yu, H., Sirevag, R. and Christoffersen, R.E. (1990) Proc. Natl. Acad. Sci. USA 87, 3904–3908.
[3] Vetter, H.-P., Mangold, U., Schröder, G., Marner, F.-J., Werck-Reichhart, D. and Schröder, J. (1992) Plant Physiol. 100, 998–1007.
[4] Mizutani, M., Ward, E., DiMaio, J., Ohta, D., Ryals, J. and Sato, R. (1993) Biochem. Biophys. Res. Commun. 190, 875–880.
[5] Teutsch, H.G., Hansenfratz, M.P., Lesot, A., Stoltz, C., Garnier, J.-M., Jeltsch, J.-M., Durst, F. and Werck-Reichhart, D. (1993) Proc. Natl. Acad. Sci. USA 90, 4102–4106.
[6] West, C.A. (1980) in: Hydroxylase, Monooxygenase, and Cytochrome P-450 (Davies, D.D., Ed.), Biochemistry of Plants, Vol. 2, pp. 317–364, Academic Press, New York.
[7] Donaldson, R.P., Luster, D.G. (1991) Plant Physiol. 96, 669–674.
[8] Bozak, K.R., O'Keefe, D.P. and Christoffersen, R.E. (1992) Plant Physiol. 100, 1976–1981.
[9] Toguri, T., Umemoto, N., Kobayashi, O. and Ohtani, T. (1993) in: 2nd International Symposium on Cytochrome P450 of Microorganisms and Plants, p. L5, Tokyo.
[10] Sambrook, J., Fritsch, E.F. and Maniatis, T. (1989) in: Molecular Cloning; A Laboratory Manual, Cold Spring Harbor Laboratory, New York.
[11] Nelson, D.R., Kamataki, T., Waxman, D.J., Guengerich, F.P., Estabrook, R.W., Feyereisen, R., Gonzalez, F.J., Coon, M.J., Gunsalus, I.C., Gotoh, O., Okuda, K. and Nebert, D.W. (1993) DNA Cell Biol. 12, 1–51.
[12] Hallaham, D.L., Nurgent, J.H.A., Hallaham, B.J., Dawson, G.W., Smiley, D.W., West, J.M. and Wallsgrove, R.M. (1992) Plant Physiol. 98, 1290–1297.