Plant Physiol. (1992) 100, 1613–1614
0032-0889/92/100/1613/02/$01.00/0

Received for publication June 20, 1992
Accepted July 20, 1992

Plant Gene Register

Nucleotide Sequence of a cDNA Encoding 3-Hydroxy-3-Methylglutaryl Coenzyme A Reductase from *Catharanthus roseus*[1]

Ignacio E. Maldonado-Mendoza[2], Ronald J. Burnett, and Craig L. Nessler*
Department of Biology, Texas A&M University, College Station, Texas 77843–3258

HMGR[3] (EC 1.1.1.34) performs the final enzymic step in mevalonate biosynthesis. In plants, mevalonate serves as a precursor for a number of important primary metabolites including ubiquinone, plastoquinone, carotenoids, ABA, and gibberellins. Additionally, mevalonate is used in the synthesis of a wide variety of secondary compounds such as phytoalexins, rubber, and alkaloids. We are interested in the role of HMGR in the synthesis of secologanin, the monoterpene component of *Catharanthus roseus* indole alkaloids.

C. roseus has been extensively studied because of the pharmaceutical importance of its antitumor bisindole alkaloids vincristine and vinblastine. Biosynthesis of these compounds begins with the conversion of tryptophan to tryptamine, catalyzed by tryptophan decarboxylase (5), followed by the condensation of tryptamine with secologanin to form strictosidine by strictosidine synthase (10). *C. roseus* cDNA clones for tryptophan decarboxylase (4) and strictosidine synthase (7) have been isolated and expressed in tobacco (6, 9). With the HMGR cDNA described here, it will now be possible to study the regulation of an enzyme in indole alkaloid biosynthesis that is not exclusively dedicated to this pathway but is also a key enzyme of primary metabolism.

Table I. *Characteristics of C. roseus HMGR cDNA pHMC3*

Organism
  Catharanthus roseus (L.) G. Don cv Little Delicata.
Location in Genome:
  Nuclear genome.
Gene, Function, and Pathway:
  HMGR (EC 1.1.1.34); reduction of 3-hydroxy-3-methylglutaryl CoA to mevalonate in terpene biosynthetic pathway.
Techniques:
  PCR of HMGR active site region from *C. roseus* cDNA with oligonucleotide primers with the sequences: 5′-GATGCTATGGGGATGAATATG-3′, and 5′-ACCTCGATAGATGGCATAGT-3′; cDNA library screening; double-stranded plasmid sequencing of both strands.
Method of Identification:
  Amino acid sequence identity with *A. thaliana* (71.1%) and *H. brasiliensis* (76.6%) HMGRs.
Expression Characteristics:
  Abundant 2.7-kb transcripts in all tissues.
Features of cDNA Structure:
  Open reading frame of 1803 bp. Translation start at nucleotide 126 and stop at nucleotide 1929. (G + C) Content: 45.4%.
Structural Features of the Protein:
  Deduced amino acid sequence of 601 residues; M_r 64,106; predicted isoelectric point 6.13; two potential transmembrane domains toward the N terminus; "PEST" cleavage sequences at amino acids 112 to 131 and 167 to 185; potential asparagine-linked glycosylation sites (N-X-S/T) at residues 86, 146, 149, 350, and 596.
Genbank/EMBL Accession No.:
  M96068.

A 444-bp HMGR fragment was generated from *C. roseus* cDNA by PCR using oligonucleotide primers (Table I) corresponding to two conserved regions within the active site of published plant and animal HMGR sequences. Because the predicted amino acid sequence of the PCR fragment showed 87% identity with the *Arabidopsis thaliana* HMGR (2), the fragment was used as a probe to screen a *C. roseus* seedling cDNA library. The complete nucleotide and derived amino acid sequence of a full-length *C. roseus* HMGR cDNA, pHMC3, is shown in Figure 1. The cDNA is 2636 bp, with an open reading frame encoding a protein of 601 amino acids having a calculated M_r of 64,106. The hydropathy profile of the predicted *C. roseus* HMGR indicates the presence of two hydrophobic domains located at residues 41 to 57 and 72 to 104, each of which is sufficient to span a membrane bilayer. Two N-terminal membrane spanning domains appear to be typical of other plant HMGRs (2, 3).

As in yeast (1), *A. thaliana* (2), and *Hevea brasiliensis* (3), "PEST" cleavage sequences are present in *C. roseus* HMGR (residues 112–131 and 167–185 in Fig. 1). PEST sequences are present in many short-lived proteins (8), although the half-life of HMGR has not yet been determined in plants. Amino acid sequence comparisons of *C. roseus* HMGR with sequences from plant, animal, and fungus HMGRs indicate

---

[1] This work was supported by the U.S. Department of Agriculture (90–37262–5375) and the Texas Advanced Technology Program. I.E.M.M. was supported by a short-term United Nations Educational, Scientific and Cultural Organization fellowship in plant biotechnology.

[2] Permanent address: Centro de Investigacion Cientifica de Yucatan, Division de Biologia Vegetal, Apartado Postal 87, Merida, Yucatan, Mexico.

[3] Abbreviations: HMGR, 3-hydroxy-3-methylglutaryl-CoA reductase; PCR, polymerase chain reaction.

```
   1  CCACAACCACCGTCACCGCTTCCGTCCTCTCCCCTCCTAATCTCCGCCACTGGCCACCCC

  61  ATTCTTTCTCTTTCCTCCACTACCACCCCCACCCTTCTCACTCTAGACCTTTGTGCAAAC

 121  AGAAAATGGACTCTCGCCGGCGATCACCAACTGTGACGGCAAAGGCTGCTGCCGGTGAAC
            M  D  S  R  R  R  S  P  T  V  T  A  K  A  A  A  G  E  L   19
 181  TGCCATTGGCACCTCACGAAGGCCAAAATCAACAGCCTTCCATTCCTAGGTCTTCCGATG
         P  L  A  P  H  E  G  Q  N  Q  Q  P  S  I  P  R  S  S  D  V   39
 241  TTTTACCTTTGCCTCTGTATCTCGCAAATGGGGTTTTCTTTACCCTGTTTTTCTCTGTTA
         L  P  L  P  L  Y  L  A  N  G  V  F  F  T  L  F  F  S  V  M   59
 301  TGTACTTTCTTCTGACCAGGTGGCGTGAGAAGATCCGAAATGCTACTCCTCTCCATGTCG
         Y  F  L  L  T  R  W  R  E  K  I  R  N  A  T  P  L  H  V  V   79
 361  TCACTCTTTCTGAGCTTGCCGCCTTGGCCTCATTAATCGCTTCTGTGATTTATCTTGTGA
         T  L  S  E  L  A  A  L  A  S  L  I  A  S  V  I  Y  L  V  S   99
 421  GTTTCTTTGGGCTTGACTTTGTTCAGTCCTTAATTTACAAGCCTAATAATGAAGGTTGGG
         F  F  G  L  D  F  V  Q  S  L  I  Y  K  P  N  N  E  G  W  E  119
 481  AAATTGAAGAAGAAATTCTTATGGTGGAAGATAGCCGCAACGGTACTAATTGTACTACCC
         I  E  E  E  I  L  M  V  E  D  S  R  N  G  T  N  C  T  T  L  139
 541  TAGGTTGCGCTGTCCCTCCACCATCTGTTCCCAAAATTGCCCCAGTCGTTCCTCAACAAC
         G  C  A  V  P  P  P  S  V  P  K  I  A  P  V  V  P  Q  Q  P  159
 601  CTTCTAAAATGGTCATCATAGAGAAGCCTGCCCCTCTCATTACCCCACAAAATTCAGAGG
         S  K  M  V  I  I  E  K  P  A  P  L  I  T  P  Q  N  S  E  E  179
 661  AAGATGAGGATATTATCAAAGCGGTGGTTGCTGGAAAAATTCCTTCGTATTCTCTCGAGT
         D  E  D  I  I  K  A  V  V  A  G  K  I  P  S  Y  S  L  E  S  199
 721  CAAAACTGGGGGACTGCAAGAGAGCTGCTGGAATTCGTAGGGAAGCGCTTCAGAGGATCA
         K  L  G  D  C  K  R  A  A  G  I  R  R  E  A  L  Q  R  I  T  219
 781  CAGGGAAGTCATTGGAGGGGTTACCTCTTGAGGGTTTTGATTATGCATCTATTTTGGGTC
         G  K  S  L  E  G  L  P  L  E  G  F  D  Y  A  S  I  L  G  Q  239
 841  AGTGCTGTGAGATGCCGGTGGGTTATGTACAGTTGCCGGTTGGAATAGCTGGGCCTTTGT
         C  C  E  M  P  V  G  Y  V  Q  L  P  V  G  I  A  G  P  L  L  259
 901  TGCTTGATGGGAGAGAATATATGTTGCCAATGGCTACTACAGAAGGGTGTTTAGTGGCTA
        L  D  G  R  E  Y  M  L  P  M  A  T  T  E  G  C  L  V  A  S  279
 961  GTACTAACAGGGGTTGTAAAGCTATTCTTGCATCTGGTGGGGCAAATAGCGTTTTGCTCA
         T  N  R  G  C  K  A  I  L  A  S  G  G  A  N  S  V  L  L  R  299
1021  GAGATGGAATGACTAGGGCTCCTGTTGTGAGGTTTGGAACTGCAAAAAGAGCCGCTGAGC
         D  G  M  T  R  A  P  V  V  R  F  G  T  A  K  R  A  A  E  L  319
1081  TCAAGTTCTACATGGAGGATACCCAGAATTTTGAGACTATATCTGTAGTTTTCAACAAAT
         K  F  Y  M  E  D  T  Q  N  F  E  T  I  S  V  V  F  N  K  S  339
1141  CGAGCAGATTTGCCAAATTGCAAAGCGTCCAATGTGCAATTGCAGGGAAGAATTTGTACA
         S  R  F  A  K  L  Q  S  V  Q  C  A  I  A  G  K  N  L  Y  I  359
1201  TAAGGTTTAGCTGCAGTACTGGTGATGCAATGGGGATGAACATGGTCTCAAAGGGCGTAC
         R  F  S  C  S  T  G  D  A  M  G  M  N  M  V  S  K  G  V  Q  379
1261  AAAATGTTCTTGAATTTCTCCAGACTGATTATCCTGATATGGATGTCCTTGGGATTTCTG
         N  V  L  E  F  L  Q  T  D  Y  P  D  M  D  V  L  G  I  S  G  399
1321  GAAATTTCTGTGCTGACAAGAAACCAGCAGCAGTCAACTGGATTGAAGGAAGAGGTAAAT
         N  F  C  A  D  K  K  P  A  A  V  N  W  I  E  G  R  G  K  S  419
1381  CTGTAGTTTGTGAGGCTATCATAAAGGAAGAGATTGTAAAGACGGTATTGAAGACAGAAG
         V  V  C  E  A  I  I  K  E  E  I  V  K  T  V  L  K  T  E  V  439
1441  TTGCGGCCCTTATTGAGCTTAACATGGTTAAAAATCTCGCTGGATCTGCAATTGCTGGAG
         A  A  L  I  E  L  N  M  V  K  N  L  A  G  S  A  I  A  G  A  459
1501  CCCTTGGTGGATTCAATGCTCATGCCAGCAATATAGTCTCTGCCATATTTATTGCCACCG

         L  G  G  F  N  A  H  A  S  N  I  V  S  A  I  F  I  A  T  G  479
1561  GCCAAGATCCTGCTCAGAATGTGGAGAGCTCTCAGTGTATCACTATGATGGAGGCTGTCA
         Q  D  P  A  Q  N  V  E  S  S  Q  C  I  T  M  M  E  A  V  N  499
1621  ATGATGGCAAAGATCTTCATATTTCTGTCACCATGCCTTCTATTGAGGTGGGCACAGTTG
         D  G  K  D  L  H  I  S  V  T  M  P  S  I  E  V  G  T  V  G  519
1681  GTGGAGGTACTCAACTTGCATCTCAGTCAGCATGTTTGAACTTGCTTGGAGTAAAAGGTG
         G  G  T  Q  L  A  S  Q  S  A  C  L  N  L  L  G  V  K  G  A  539
1741  CCAGCAAAGACTCACCTGGTGCAAATTCAAGGCTCTTGGCTACCATTGTTGCTGGTTCAG
         S  K  D  S  P  G  A  N  S  R  L  L  A  T  I  V  A  G  S  V  559
1801  TTCTTGCAGGTGAGCTTTCCCTCATGTCTGCCATTTCTGCAGGCCAACTCGTTAGGAGCC
         L  A  G  E  L  S  L  M  S  A  I  S  A  G  Q  L  V  R  S  H  579
1861  ACATGAAATACAACCGTTCTAGCAAGGATATAACCAACATAGCCTCCTCTCAGTTAGAGA
         M  K  Y  N  R  S  S  K  D  I  T  N  I  A  S  S  Q  L  E  S  599
1921  GTGATAGTTAAATACAATCCCTTCCCCCTCCCCCCCTTAAACCCCACCTTGTTAGTTAGT
         D  S  *                                                      601
1981  ATAGGCCATAGAAGAAATCCCCAAAAAAAAAAAAAAAAAAAAAAGGTTATGAAGTTAATAAA

2041  CAAATCAAGAAAGGGGTGGTGAATAAAAAAAATATATAAAACCCCCAAAAAAAAAAAAGG

2101  AAAGAGAATAGGGTGTATTGGTGGTGAATATGATATGTGAGAAGGAGAAATGAAGGAAGA

2161  AACAGACAAAGGCAAACTACATCTGTTGCAAGAAGGGAGTCATGTGATGCTGCTCTACCT

2221  CTCTTTCTCTCTCCTCTTTCTCTCTCTATTTCTTGAATGAAGTTGTCCTGAAGTCATGGG

2281  GAACAGTAGCTGGCTGGCTGGCTGGGTGCACCGTGGAGAAAGAAGAATCTGTTTGTTTAA

2341  TTATTTGTATCTTTTGTAATGTCATTTTGGCATCAAGAAATGGAAAGAGAAGAGGCAGTT

2401  AAGGTTAGAGTTTGGTGGCATTCATGTGCAAAGGAAGATTTTACTCTTAATTTTTAGTAT

2461  GTTTTTAATATAATCTTCTTCTTCACTGTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTG

2521  GTTTGTTGTTAACAGCAGTGATTAGTTAGTAATTAAGACAGAGCATTTCAAAAACATGAG

2581  ACTCTGATCTGTCTTACAAGTTTGTGTAAATGGAAAAAATGTTTTATTCTTCAAAA
```

Figure 1. Nucleotide and deduced amino acid sequence of *C. roseus* HMGR cDNA (pHMC3). PCR primer sites are underlined in the nucleotide sequence. Potential PEST sequences are underlined in the amino acid sequence.

a high degree of similarity in the C-terminal region, which contains the catalytic site.

ACKNOWLEDGMENTS

We wish to thank Dr. Vincenzo De Luca (Jardin Botanique de Montreal, Montreal, Canada) for providing the *C. roseus* cDNA library and Dr. Martin Rechsteiner (Department of Biochemistry, University of Utah, Salt Lake City, UT) for sending us his PESTFIND program. DNA sequencing was performed by the staff of the Advanced DNA Technologies Laboratory, Texas A&M University.

LITERATURE CITED

1. **Basson ME, Thorsness M, Finer-Moore J, Stroud RM, Rine J** (1988) Structural and funtional conservation between yeast and human 3-hydroxy-3-methylglutaryl coenzyme A reductases, the rate-limiting enzyme of sterol biosynthesis. Mol Cell Biol **8:** 3797–3808
2. **Caelles C, Ferrer A, Bacells L, Hegardt FG, Boronat A** (1989) Isolation and structural characterization of a cDNA encoding *Arabidopsis thaliana* 3-hydroxyl-3-methylglutaryl coenzyme A reductase. Plant Mol Biol **13:** 672–678
3. **Chye ML, Kush A, Tan CT, Chua NH** (1991) Characterization of cDNA and genomic clones encoding 3-hydroxyl-3-methylglutaryl CoA reductase from *Hevea brasiliensis*. Plant Mol Biol **16:** 567–577
4. **De Luca V, Marineau C, Brisson N** (1989) Molecular cloning and analysis of cDNA encoding a plant tryptophan decarboxylase: comparison with animal dopa decarboxylases. Proc Natl Acad Sci USA **86:** 2582–2586
5. **Luckner M** (1984) Secondary Metabolism in Microorganisms, Plants and Animals, Ed 2. Springer-Verlag, New York
6. **McKnight TD, Bergey DR, Burnett RJ, Nessler CL** (1991) Expression of enzymatically-active and correctly-targeted strictosidine synthase in transgenic tobacco plants. Planta **18:** 148–152
7. **McKnight TD, Rossner CA , Devagupta R, Scott AI, Nessler CL** (1990) Nucleotide sequence of a cDNA encoding the vacuolar protein strictosidine synthase from *Catharanthus roseus*. Nucleic Acids Res **16:** 4939
8. **Rogers S, Rodney W, Rechsteiner M** (1986) Amino acid sequences common to rapidly degraded proteins: the PEST hypothesis. Science **234:** 364–368
9. **Songstad DD, De Luca V, Brisson N, Kurz WGW, Nessler CL** (1990) High levels of tryptamine accumulation in transgenic tobacco expressing tryptophan decarboxylase. Plant Physiol **94:** 1410–1413
10. **Stockigt J, Zenk MH** (1977) Isovincoside (strictosidine), the key intermediate in the enzymatic formation of indole alkaloids. FEBS Lett **79:** 233–237