Isolation and sequencing of the cDNA of a novel cytochrome P450 from rat oesophagus

Nicola Brookman-Amissah², Alan G. Mackay¹ and Peter F. Swann

Department of Biochemistry and Molecular Biology and ¹Department of Surgery, University College London, London WC1E 6BT, UK

²To whom correspondence should be addressed
Email: ucbcnbr@ucl.ac.uk

Reverse transcription (RT) and PCR amplification of cytochrome P450 mRNAs

Introduction

Oesophageal cancer is the seventh most common cancer in the world. The majority of these tumours are squamous cell carcinomas, although there is a significant and rising incidence of adenocarcinoma. Identification of the cause of this cancer is very important because oesophageal cancer is incurable. In the UK, 77% of victims are dead within a year of diagnosis and the 5 year survival is only 7% (2). There is more than a 100-fold difference in the incidence of oesophageal cancer in different parts of the world and there have been rapid changes in incidence with time. For example, the incidence among males in Singapore fell by 3.6% per annum between 1970 and 1980 (3). These geographical and temporal variations suggest that oesophageal cancer is caused by exposure to carcinogens. Furthermore, the spectrum of p53 mutations in the squamous cell carcinomas suggests that oesophageal cancer is caused by exposure to carcinogens. Oesophageal cancer can be produced in the rat by administration of certain N-nitrosodialkylamines (6). Nitrosamines cause cancer in those tissues that contain an enzyme system to activate them. It is generally assumed that the enzyme system that carries out this activation is a cytochrome P450 and, since the original work of Labuc and Archer (7), the metabolism of nitrosamines by oesophageal microsomes, sometimes augmented by parallel studies on microsomes from liver and nasal epithelium, has been reported from many laboratories. The authors of these various studies, excellently summarized in two recent papers (8,9), have suggested that N-nitrosamines which are carcinogenic to the rat oesophagus, however, the actual P450 has not yet been unequivocally identified (9,10).

There have been two very important observations about oesophageal nitrosamine metabolism. First, the amount of methylation of DNA by N-nitrosomethylbenzylamine and N-nitrosomethyl-n-butylamine is greater in the oesophagus than in the liver (11,12), despite the fact that the total P450 content of the oesophageal epithelium is considerably less than 10% of that of the liver. Second, N-nitrosomethyl-(2-methylbutyl)amine, one of the nitrosamines that induce oesophageal cancer, produces methylation of the DNA of the rat oesophagus but not of the DNA of rat liver (13). This indicates that the oesophagus of rats contains an enzyme system capable of activating nitrosamines which is not expressed in the liver.

In the work reported here we have cloned and sequenced from the rat oesophagus the cDNA of a novel P450. The deduced amino acid sequence shows that it is from the CYP2B sub-family and it has now been given the systematic name CYP2B21. Using a combination of 5'- and 3'-RACE (rapid amplification of cDNA ends) and library screening, the cDNA was amplified and sequenced. The cDNA sequence (GenBank accession no. AF159245) covers the whole of the coding region and the whole of the 3'-untranslated region (UTR), but only 17 nt of the 5'-UTR. The DNA sequence has strong similarity to those of CYP2B1 and CYP2B2, with the derived amino acid sequence being 84 and 83% identical, respectively. The ease with which this cDNA was found in the cDNA library suggests that the actual P450 has not yet been unequivocally identified (9,10).

Materials and methods

Abbreviations: RACE, rapid amplification of cDNA ends; R; reverse transcription; UTR, untranslated region.

© Oxford University Press
The cDNA sequence of the cytochrome P450 identified in the oesophagus. The start codon is at position 18, the stop codon at 1500 and the polyadenylation signal at 1852. This sequence was determined from three independent cDNA library clones. The sequence between 964 and 1310 was confirmed by sequencing a fragment amplified by RT–PCR, the sequence between 759 and 1122 by nested 5′/H11032-RACE and the sequence between 1140 and 1894 by nested 3′/H11032-RACE. The forward primer used for 3′/H11032-RACE shown in Figure 3A began at 960 and the reverse, 5′/H11032-CCV AARCAGWTYCGCTTTCC; for CYP2B, forward, RAPID AMPLIFICATION OF CDNA ENDS (RACE) and nested RACE 5′/H11032-GGHTTCCTGCTCATGCTCAA, and reverse, 5′/H11032-CCARCACACACDCGYTTTCC. For CYP2E, forward, 5′/H11032-GGGCTCCTGATYCTCATGAA, and using oligo(dT)–Dynabeads (Dynal, UK) according to the manufacturer’s instructions. This RNA was reverse transcribed with Superscript II reverse transcriptase (Life Technologies) from an oligo(dT) primer according to the manufacturer’s instructions. Double-stranded DNA was made using a Marathon cDNA synthesis kit (Clontech) according to the manufacturer’s instructions. Sequencing. The procedures described above were also carried out with total RNA prepared from rat liver. Primers used by Shen et al. (21) to amplify a proof-reading DNA polymerase rTth (Perkin Elmer) using the gene-specific fragment of a novel P450 cDNA from mouse embryo fibroblast cells were also used. The downstream primer was designed to anneal to the sequence TACCCCCATGTTGCAGAG-3′, respectively, and the adapter primer AP1 (5′/H11032-AGGACCCAGG; Clontech) to re-amplify the 5′- and 3′-RACE PCR amplifications were expected to amplify between 321 and 336 bp.

Rapid amplification of cDNA ends (RACE) and nested RACE
Oesophageal poly(A) RNA was obtained from oesophageal total RNA using oligo(dT)–Dynabeads (Dynal, UK) according to the manufacturer’s instructions. This RNA was reverse transcribed with Superscript II reverse transcriptase (Life Technologies) from an oligo(dT) primer according to the manufacturer’s instructions. Double-stranded DNA was made using a Marathon cDNA synthesis kit (Clontech) according to the manufacturer’s instructions. Double-stranded RNA was used as a template for nested RACE PCR amplifications using the nested gene-specific primers 5′-GATAGGTCAATCGGGACAAG-3′ and 5′-AGGGCTCGAGCGGC; Clontech) to re-amplify the 5′- and 3′-RACE PCR amplifications were expected to amplify between 321 and 336 bp.

1  GGCACGAGTCT CAGGACCATTGATCCAGTGT TCGCTCTCT TTTGTGCTCTT
51  CCTCATGACCT GTCTTGACTACTTTTGTGACCAAGGGAA ACGGATATGG
101  CCCATCTTCCA CAGGACCTCT CCTCCCTGCCC CCTCTTGCGG AACGTTTTGTC
151  AAATGACGAC AGAGGTCCTC TCAAGTCTCTT ATATTTCGTTCT CTCGTGACAAA
201  TATGTTGAGAT TGTTGACATG CAACTTGGGA CAAAGGCAAGC TACGTCATTG
251  GTATGGAGCC GAGACCATAA GGAAAGCTCT GCCGACACATT CTCAGGCTGTT
301  TCTCTGGCGG GGAACAGGTT GCTGTGTTTC AGCCATATTAC AGACGACTAT
351  GGTATGATCT TCTGAAATGG GAAAGTTCTAG ACTGCTTCCT GCAGATTCTC
401  TCTAGGCACC ATGAGAAGCT TTGGAGATGG AAGCAGAGTT GTCGAGAGGA
451  GGAATAAGGA GGAAGCCCAAA TGGTCTGTGG AGGAACTGAA GAAATAACAG
501  GGAAGCCCAAC TGAAACCCAC CCACACAATTAC CAGTCCATTAC GCAGCCACAT
551  CATCTACACC ATTTCTTCTT GAGAAGTTTT CTGACTACA CACCAGCAAG
601  TCTTCTGCACT GCTGACCTTG TCTCAGTAAC CTCTTGCTAC CTGTCCTCTC
651  TTCTCCAGGC AGTTGTTTTG GCTTTCTACT GGGTCTGCTA AGTACTTCTC
701  TTGGACACCC GAGACCATCT CCACAAACAT CAAAGAACAT TTCAACTCTC
751  TTGGCCATTT TGTTGAGAGCA GACAGGCAAC TAGAGACCCC CAGACTCTCA
801  CGAGACTTCGA TAGAACACTT CTTCTCTAGA ATGGAGAGGA AGAAATTAAC
851  CCACCACACA GAATTACATG ACGAAGCTAC ATGATGCTGTT CTGTCCTCTC
901  TCTTCTCACA GGAAGACGTTT TTTTCTTGG TGAAGAGCT CCAGGCAATG
951  AGITGTTCTCT TTCTTCTACCA CACCACCTCC AGACACTTCTC ATGTCGACAG
1001 TGATCGAGTTG ATGGCTCACA AGCGGCTCCT ACCCTATCTG AGCCGATTC
1051 AATGACCATC AACTGAGCT GTCATCCATG AGATTCAGAG ATTTTCAGAT
1101 CCTTGTCCCGA TTGTTGCCGC ACACAGTCTG ACAGACAGAA CATGTGTTCCG
1151 AGGTATTCTG TCTTCCAAAG ACATTTGAGGT GTACCCCATC CTGGATTCAC
1201 CACCTCATGA TCCACAGTAC TTTGAGACCC GACAGCCTAC CACTCTTGAG
1251 CACCTCCTGG ATCCCAATTG GGCGATGAG AGAAGATAGG TTCTTCTGCC
1301 TTTTCTCACA GGAAGACGTTT TTTTCTTGG TGAAGAGCT CCAGGCAATG
1351 AGITGTTCTCT TTCTTCTACCA CACCACCTCC AGACACTTCTC ATGTCGACAG
1401 CCTGTTCTCT CTAGACGAGTC TAGACTCACT CCAAGGAGGA GTCGGTTCTGC
1451 CAAATATCCG CCAACATACCA AGATCTGCTT TTATCTGC AGATGGCTCT
1501 GAGGAGAACT GTTGAGACAA CTTTACGTAT CTATTGAGA TAAGTCCTCC
1551 AGAGAGAAGCA GTCTGAAATCC AGGCCAGGG TCAGTGCCTA CACCCCTACT
1601 CCTACTGCGAG CTCTCATATC CCGATGTGTT GTCTCCTCC TGTGTAATGG
1651 AGTTGAGAATT CAACTACGCT TCTTTCTTGG TGTAGACAA GACAGCTCTG
1701 AGTTGACTTG TCACCCTGTA ATACACTCTT AATGCCTTCC AACCAGCCAA
1751 GACCTCTACCT ATACCTCTCT CAGACATCT GTGTTATGCT ATTTACTCCT
1801 GTATAAGGTGT TGGATTTATGT GCTCACAATG TGCTCATGTA TGTCGCTTTG
1851 AAATACACAG AACACCGAGG TGTGTGAAAG AAAAAAAA AAAAAA

Fig. 1. The cDNA sequence of the cytochrome P450 identified in the oesophagus. The start codon is at position 18, the stop codon at 1500 and the polyadenylation signal at 1852. This sequence was determined from three independent cDNA library clones. The sequence between 964 and 1310 was confirmed by sequencing a fragment amplified by RT–PCR, the sequence between 759 and 1122 by nested 5′-RACE and the sequence between 1140 and 1894 by nested 3′-RACE. The forward primer used for 3′-RACE shown in Figure 3A began at 960 and the reverse, 5′-RACE primer began at 1251.
Preparation and screening of rat oesophageal cDNA library

Fig. 2. Amplification by RT–PCR of 350 bp of P450s of the 2A (lanes a and b), 2B (lanes c and d) and 2E (lanes e and f) sub-families from oesophageal (lanes a, c and e) and liver (lanes b, d and f) total RNA preparations. Note that only the CYP2B primers amplified anything from the oesophageal cDNA. The smallest molecular size standards are (from the bottom) 100, 200, 300 and 400 bp.

Fig. 3. (A) RACE. The 5′-RACE product is ~1150 bp and the 3′-RACE product ~1000 bp. (B) Northern blot of a total RNA preparation from the oesophageal (lanes a, c and e) and liver (lanes b, d and f) total RNA preparations. Note that only the CYP2B primers amplified anything from the oesophageal cDNA. The smallest molecular size standards are (from the bottom) 1.35, 2.37, 4.40, 7.46 and 9.49 kb.

were: 94°C for 1 min; 5 cycles of 94°C for 30 s, 72°C for 4 min; 5 cycles of 94°C for 30 s, 70°C for 4 min; 25 cycles of 94°C for 20 s, 68°C for 4 min; a final extension at 72°C for 10 min.

Preparation and screening of rat oesophageal cDNA library

A rat oesophageal cDNA library was prepared using a Zap-cDNA Synthesis Kit and Gigapack III Gold Packaging Kit (Stratagene, UK) according to the manufacturer’s instructions. The library was characterized as containing 1.2×10⁸ independent clones whose average insert size was ~800 bp. 500 000 plaques of an amplified phage stock of the library were plated on twenty 15 cm plates and screened by conventional library screening techniques (22). The 350 bp DNA fragment from oesophageal RNA obtained by RT–PCR using the CYP2B-specific primers as described above was random primer labelled with [³²P]dCTP and hybridized to plaque lifts (Hybond-N; Amersham) overnight in 50% formamide, 5× SSC, 5× Denhardt’s solution, 1% SDS at 42°C. Lifts were washed twice for 30 min at 68°C in each of the following solutions: 2× SSC, 0.1% SDS; 0.1× SSC, 0.1% SDS. The lifts were then rinsed in 2× SSC and autoradiographed. Positive plaques were corona and subjected to secondary screens. The positive clones were excised using ExAssist Interference-Resistant Helper Phage with SOLR Strain (Stratagene) according to the manufacturer’s instructions. The clones were grown and the plasmid DNA isolated (Maxiprep; Qiagen) and sequenced.

Results
cDNA fragments were amplified from total RNA preparations from rat oesophagus and liver by RT–PCR. In the first experiments a degenerate PCR primer pair containing inosine was used. These primers had been designed by Shen et al. (21) to amplify all P450s belonging to families 1, 2 and 3 and used by them to identify a fragment of a novel P450 from mouse embryonic fibroblast cells. However, PCR with these primers did not amplify DNA from either the liver or the oesophageal cDNA. Therefore, three different sets of PCR primers were designed to amplify an equivalent 350 bp fragment of CYPs of the 2A, 2B and 2C sub-families. The positions of the primers used to detect the presence of members of the CYP2B sub-family are given in Figure 1. The CYP2A-specific primers amplified a 350 bp cDNA fragment from the liver cDNA, which was observed as a very faint band, but no product was obtained from the oesophageal cDNA amplification (Figure 2). The CYP2B-specific primers amplified a 350 bp cDNA fragment from both oesophageal and liver cDNA (Figure 2). The band amplified in liver was more intense than that amplified from the oesophagus. Amplification with CYP2E-specific primers did not give a discernible product from oesophageal cDNA, but amplified a distinct 350 bp fragment from liver cDNA.

The CYP2B PCR product from the liver and from the oesophagus and the CYP2E PCR product from the liver were cloned into the pPCR-Script cloning vector and the cloned cDNA fragments were sequenced. Only one CYP2B clone amplified from the liver RNA was sequenced. This clone was found to have come from CYP2B2. Several different clones were sequenced from the fragment amplified from oesophageal cDNA by the CYP2B-specific primers. All of these had exactly the same sequence. The sequence had no exact matches in the database, but the closest match was to CYP2B12, a P450 sub-family 2B DNA fragment were used for RACE. The 5′-RACE product is ~1150 bp and the 3′-RACE product ~1000 bp. These internal primers were used so that the exact sequence of the two parts of the sequencing overlaps by 291 bp this would imply that the cDNA was ~1.9 kb. Consistent with this, a northern blot of total oesophageal RNA probed with the 350 bp fragment amplified by PCR as described above showed the presence of a RNA of ~1960 nt (Figure 3). The 3′-end was easily cloned into the multiple cloning site of pPCR-Script, but several attempts to clone the 5′-end into the same vector were unsuccessful.
Nested RACE–PCR was therefore carried out. This gave a more intense band from the 5' -end and this product could be cloned. More product was obtained from the 3' -end after nested RACE–PCR than after the initial RACE–PCR, so this product was also cloned. Cloned plasmids containing the 3'-nested RACE–PCR product consistently contained inserts of the expected size and the inserts in four clones were sequenced without difficulty. The four sequences were identical. However, plasmids containing the 5'-nested RACE–PCR product contained inserts of varying sizes, ranging from 500 to 1000 bp. Sequencing of four clones containing the longest inserts showed that the cloned DNA fragments were overlapping parts of the same sequence, but sequencing always resulted in bad data towards both ends of the clones.

To overcome this problem an oesophageal cDNA library was constructed as described in Materials and methods and 0.5×10^6 clones were screened using as probe the 350 bp oesophageal CYP2B obtained by RT–PCR as described above. Nine positive clones were obtained. On secondary screening, between two and five positive clones were isolated from each of the nine clones. Plasmid DNA from these clones was purified. Digestion with EcoRI and XhoI of the DNA from clones 1 and 7 gave three DNA fragments, the linearized nested RACE–PCR product consistently contained inserts of the expected size and the inserts in four clones were sequenced without difficulty. The four sequences were identical. However, plasmids containing the 5'-nested RACE–PCR product contained inserts of varying sizes, ranging from 500 to 1000 bp. Sequencing of four clones containing the longest inserts showed that the cloned DNA fragments were overlapping parts of the same sequence, but sequencing always resulted in bad data towards both ends of the clones.

Fig. 4. Comparison of the derived amino acid sequence of the newly discovered oesophageal P450, CYP2B21, aligned (31) with the sequence of other members of the CYP2B sub-family, CYP2B1 (35), CYP2B2 (32,33), CYP2B12 (28) and CYP2B3 (34). Blank areas in the amino acid sequences of CYP2B1, CYP2B2, CYP2B12 and CYP2B3 indicate that the amino acids are identical to those of the oesophageal protein. The C-terminal amino acid of all sequences has been put in to define the end of the peptide chain. The hyphen at position 21 indicates that CYP2B12 has one more amino acid than the other proteins at this point: and it should be noted that CYP2B21 has 494 amino acids.
aligned with the 3′-end of the library clones (Figure 1). The sequences obtained were compared with sequences in the DNA database using a BLAST search. No identical sequence was found, but the best 50 matches in this search were all to P450s of the 2B sub-family. The best matches were to CYP2B1, CYP 2B2 and CYP 2B12, with the derived amino acid sequence of the new P450 having, respectively, 84, 83 and 77% identity to their sequences (Figure 4).

Discussion

This paper describes the cloning and sequencing from rat oesophagus of an essentially full-length cDNA coding for a cytochrome P450. The sequence was assembled from PCR amplification of a short central sequence, nested 5′-RACE, nested 3′-RACE and sequencing of clones from an oesophageal cDNA library (Figure 1). In each case more than one independently derived clone was sequenced to avoid error. The final sequence contains the complete 3′-untranslated region (UTR) but only 17 nt of the 5′-UTR. The sequence translates to a protein of 494 amino acids. There are very close similarities between this amino acid sequence and the sequences of three cytochrome P450s of the 2B sub-family, CYP 2B1, CYP 2B2 and CYP 2B12 (Figure 4).

The amino acid sequence has all the features of a functional P450. There is the characteristic membrane targeting and anchoring sequence at the N-terminus with an acidic residue as the second amino acid, closely followed by a sequence of 16 mainly hydrophobic amino acids, then a positively charged Arg at position 21 and a His at position 29, followed immediately by a proline-rich region (Pro30, Pro31, Pro33, Pro35 and Pro38) (24–26). The part of the molecule immediately after the first hydrophobic region is less basic than in CYP2B1 and 2B2, which have Arg at position 21, Lys at 25 and Arg at 27 and 34, and also slightly less basic than CYP2B12, which has a Lys at position 26, two His at 28 and 30 and an Arg at 34. The Ser of the phosphorylation site (Ser128) is in the same sequence, LRRFSL, as the analogous Ser in CYP2B1 and 2B2. There are three conserved Cys at positions 152, 179 and 436. The last of these is the Cys of the heming binding site. The sequence surrounding this Cys, GKRVCL, differs by only a single amino acid (Val instead of Ile) from the analogous sequence in CYP2B1, 2B2 and 2B12. The DNA sequence around the start codon is TTCAGGACCATGG. This is only slightly different from the consensus Kozak sequence (27) and differs by only a single base, the 5′-terminal T, from the sequences around the start codons of CYP2B1 and CYP2B12, both of which are ACCAGGACCATGG (28).

This new P450 has not yet been expressed so we do not yet know whether it will metabolically activate nitrosamines. However, the previously published experiments showing that CYP2B1 or 2B2 will metabolically activate the oesophageal-selective methylalkylnitrosamines (8,14,16–19,29,30) suggest that this may be the elusive nitrosamine-activating P450 of the oesophagus. The oesophagus contains very little P450 but it can activate these nitrosamines very well, which suggests that the nitrosamine-metabolizing enzyme must be the major P450 in the oesophagus. Thus the relative ease with which we were able to clone the cDNA, suggesting that it must be one of the highly expressed P450s, also supports the view that this new P450 is the nitrosamine-metabolizing enzyme of the oesophagus. We have begun experiments to express this enzyme in vitro to see whether this view is correct.

Acknowledgements

We are very grateful to University College for a studentship for N.B.-A. and to the Cancer Research Campaign for their generous support.

References


Discussion

This paper describes the cloning and sequencing from rat oesophagus of an essentially full-length cDNA coding for a cytochrome P450. The sequence was assembled from PCR amplification of a short central sequence, nested 5′-RACE, nested 3′-RACE and sequencing of clones from an oesophageal cDNA library (Figure 1). In each case more than one independently derived clone was sequenced to avoid error. The final sequence contains the complete 3′-untranslated region (UTR) but only 17 nt of the 5′-UTR. The sequence translates to a protein of 494 amino acids. There are very close similarities between this amino acid sequence and the sequences of three cytochrome P450s of the 2B sub-family, CYP 2B1, CYP 2B2 and CYP 2B12 (Figure 4).

The amino acid sequence has all the features of a functional P450. There is the characteristic membrane targeting and anchoring sequence at the N-terminus with an acidic residue as the second amino acid, closely followed by a sequence of 16 mainly hydrophobic amino acids, then a positively charged Arg at position 21 and a His at position 29, followed immediately by a proline-rich region (Pro30, Pro31, Pro33, Pro35 and Pro38) (24–26). The part of the molecule immediately after the first hydrophobic region is less basic than in CYP2B1 and 2B2, which have Arg at position 21, Lys at 25 and Arg at 27 and 34, and also slightly less basic than CYP2B12, which has a Lys at position 26, two His at 28 and 30 and an Arg at 34. The Ser of the phosphorylation site (Ser128) is in the same sequence, LRRFSL, as the analogous Ser in CYP2B1 and 2B2. There are three conserved Cys at positions 152, 179 and 436. The last of these is the Cys of the heming binding site. The sequence surrounding this Cys, GKRVCL, differs by only a single amino acid (Val instead of Ile) from the analogous sequence in CYP2B1, 2B2 and 2B12. The DNA sequence around the start codon is TTCAGGACCATGG. This is only slightly different from the consensus Kozak sequence (27) and differs by only a single base, the 5′-terminal T, from the sequences around the start codons of CYP2B1 and CYP2B12, both of which are ACCAGGACCATGG (28).

This new P450 has not yet been expressed so we do not yet know whether it will metabolically activate nitrosamines. However, the previously published experiments showing that CYP2B1 or 2B2 will metabolically activate the oesophageal-selective methylalkylnitrosamines (8,14,16–19,29,30) suggest that this may be the elusive nitrosamine-activating P450 of the oesophagus. The oesophagus contains very little P450 but it can activate these nitrosamines very well, which suggests that the nitrosamine-metabolizing enzyme must be the major P450 in the oesophagus. Thus the relative ease with which we were able to clone the cDNA, suggesting that it must be one of the highly expressed P450s, also supports the view that this new P450 is the nitrosamine-metabolizing enzyme of the oesophagus. We have begun experiments to express this enzyme in vitro to see whether this view is correct.

Acknowledgements

We are very grateful to University College for a studentship for N.B.-A. and to the Cancer Research Campaign for their generous support.

References


Received July 10, 2000; revised September 26, 2000; accepted September 28, 2000