J. Mol. Biol. (1993) 229, 1175–1183

# Expression of the Human Placental Folate Receptor Transcript is Regulated in Human Tissues

## Organization and Full Nucleotide Sequence of the Gene

**Stephanie T. Page, William C. Owen, Kirsten Price and Patrick C. Elwood†**

*Medicine Branch, National Cancer Institute*
*National Institutes of Health, Bethesda, MD 20892, U.S.A.*

*(Received 3 August 1992; accepted 21 October 1992)*

The primary structures of the human KB cell (FR-KB1) folate receptor (FR) and of a human placental (FR-P2) FR, proteins important in cellular accumulation of folates, have been deduced from cDNA sequences. Herein, we report a novel human FR cDNA (FR-P3) isolated from a placental library and the chromosomal organization of the human FR-P3 gene. Compared to the FR-P2 cDNA, the composite 1084 base-pair (bp) FR-P3 cDNA is homologous, but contains a unique 5′ terminus and sequence differences within the open reading frame (ORF) and at the exon I–II junction. Polymerase chain reaction and RNase protection assays demonstrate that the FR-P3 cDNA represents the major transcript, and suggest that the FR-P2 cDNA is encoded by an independent FR gene. The nucleotide sequences of two non-overlapping human genomic clones contain the FR-P3 gene, which spans 5148 bp, is composed of five exons, and is polymorphic relative to 5′ restriction sites. The transcript size (1084 bp) predicted from structural analysis of the FR-P3 gene correlates with the size (1100 bp) determined by Northern blots. Based on RNase protection assays, both FR-P3 and FR-KB1 transcripts are expressed in human fetal and adult tissues, and the abundance of each transcript varies among the tissues studied. These results indicate that the FR transcripts are products of independent, conserved genes; that neither FR gene is preferentially expressed during fetal development; and that specific FR transcripts are differentially expressed in human tissues, suggesting that transcription of each FR gene is regulated independently. The isolation of the FR-P3 gene will permit functional analysis of the *cis* and *trans* regulatory elements of the FR-P3 gene and the mechanisms involved in tissue-specific FR gene expression.

*Keywords:* folate; receptor; folate-binding; transport; gene family

## 1. Introduction

Human folate receptors (FR‡) comprise a heterogeneous family of single chain glycoproteins with related biochemical (Antony *et al.*, 1981, 1982; Antony, 1992; Elwood *et al.*, 1986, 1991; Henderson, 1990; Kane & Waxman, 1989; Luhrs *et al.*, 1987) and molecular (Elwood, 1989; Ratman *et al.*, 1989; Sadisavan & Rothenberg, 1989; Lacey *et al.*, 1989) properties. The membrane-associated form of human folate binding proteins, hereafter designated the folate receptor (FR), is a component of a high affinity folate transport system described in malignant tissue culture cells (Kane *et al.*, 1988; Antony *et al.*, 1985; Jansen *et al.*, 1990; Henderson & Strauss, 1990; Kamen & Capdevila, 1986; Kamen *et al.*, 1988) and normal human erythrocyte progenitors (Antony *et al.*, 1987).

To investigate the structure and function of human FRs, the amino acid sequences of two human FRs have been deduced from the nucleotide sequences of cDNAs isolated from human KB cell (Elwood, 1989; Sadisavan & Rothenberg, 1989), CaCo-2 cell (Lacey *et al.*, 1989) and placental (Elwood, 1989; Ratman *et al.*, 1989) libraries. As standard nomenclature has not been adopted, the

† Author to whom all correspondence should be addressed.

‡ Abbreviations used: FR, folate receptor; KB cells, human nasopharyngeal epidermoid carcinoma cells; FR-P2, reported placental FR (Ratman *et al.*, 1989); FR-KB1, KB cell FR; FR-P3, placental FR; UTR, untranslated region; ORF, open reading frame; bp, base-pair(s); kb, $10^3$ base-pairs; PCR, polymerase chain reaction.



0022-2836/93/041175–09 $08.00/0

© 1993 Academic Press Limited

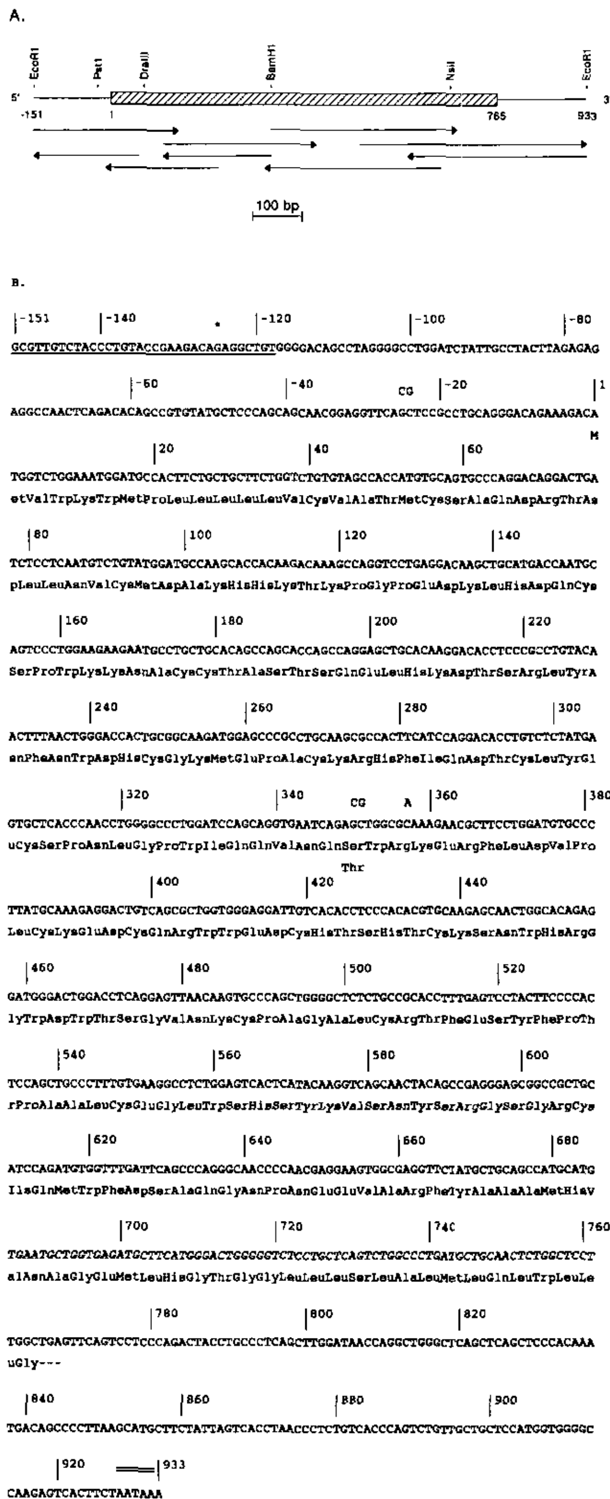


**Figure 1.** Primary structure of the placental folate receptor (FR-P3) cDNA. A. Partial restriction map and the sequencing strategy of the composite placental folate receptor cDNA (FR-P3). The arrows indicate the extent and orientation of sequence determined from each primer by the dideoxynucleotide chain-termination method (Sanger, 1977) using Sequenase and reagents supplied in the Sequenase Sequencing Kit (United States Biochemical Corp.). The hatched rectangle represents the ORF. B. The cDNA nucleotide sequence and predicted amino acid sequence of FR-P3. The 5′ terminus of the longest λGT11 cDNA (FR c15) is indicated by "*". The underlined 5′ terminal sequence and nucleotides at residues −25, −24, +350, +351 and +357 diverge from the reported FR-P2 cDNA. The termination codon of the ORF and the consensus polyadenylation signal are indicated by "– – –" and by a double overline, respectively. Placental folate receptor cDNAs were isolated from a placental λGT11 cDNA library (Clontech) ($n = 16$) by screening $9{\cdot}5 \times (10^5)$ recombinant phage with a radiolabeled oligonucleotide guessmer (Elwood, 1989) or by amplification using RACE (rapid amplification of cDNA ends) PCR ($n = 11$) (Frohman, 1990). The latter cDNAs were amplified from the placental cDNA library or from poly(A) tailed first strand cDNA synthesized from oligo(dT) primed placental total RNA (10 μg) using AMV reverse transcriptase and reagents supplied in a first Strand cDNA Synthesis Kit (Stratagene) as recommended by the vendor. PCR was performed using the reagents and *Taq*I polymerase supplied in a Perkin–Elmer Cetus GeneAmp PCR reagent kit as recommended by the vendor. The 5′ sense primers consisted of universal λGT11 primers (Promega) or oligo(dT); and the 3′ antisense cDNA-specific primers corresponded to sequence within the ORF of FR-P3 cDNA (PCR-1 (residue +38 to +60), 5′-TCTGTGTAGCCACCATGTGCAGT-3′; PCR-2 (residue +330 to +345), 5′-ATTCACCTGCTGGATC-3′). All oligonucleotides were synthesized on an Applied Biosystems Inc. DNA Synthesizer (model 392) or purchased from Synthecell. The authenticity of each PCR product was verified by restriction digests (e.g. *Pst*I) and by nucleotide sequence determination. The *Eco*RI cDNA inserts from λGT11 clones were subcloned into pGEM4Z plasmids (Promega), and the specific PCR fragments were subcloned into pCR-1000 (Invitrogen Corp.) exactly as recommended by the vendor. All nucleotide sequences were determined from both strands.

reported FR cDNAs are designated herein as FR-P2 (Ratman *et al.*, 1989) or FR-KB1 (Elwood, 1989), where the letters reflect the tissue ("P", placenta) or cellular ("KB", KB cells) sources from which their respective proteins and cDNAs were originally isolated and the number denotes independent cDNAs. Although the 5′ untranslated regions (UTRs) of the reported "KB cell" cDNA isoforms (Elwood, 1989; Lacey *et al.*, 1989; Campbell *et al.*, 1991; Coney *et al.*, 1991) are heterogeneous in length and sequence, their open reading frames (ORF) and 3′ UTRs are identical. Based on the heterogeneity of determined partial amino acid sequences between purified human placental and milk (Svendson *et al.*, 1984) FRs, Ratman *et al.* (1989) isolated a homologous "placental" FR cDNA encoding the second FR (FR-P2). Although their nucleotide sequences suggest that the FR cDNAs are products of different genes, the molecular basis of their sequence heterogeneity, the chromosomal organization of their gene(s), the normal tissue distribution of each FR gene transcript and its product, and the factors regulating transcript expression, have not been determined.

We have isolated a cDNA (FR c15) from a human placental λGT11 library that is homologous to, but distinct from, the FR-KB1 and FR-P2 cDNAs. A partial restriction map with the sequencing strategy, and the full length nucleotide sequence of a composite cDNA (designated FR-P3) (composed

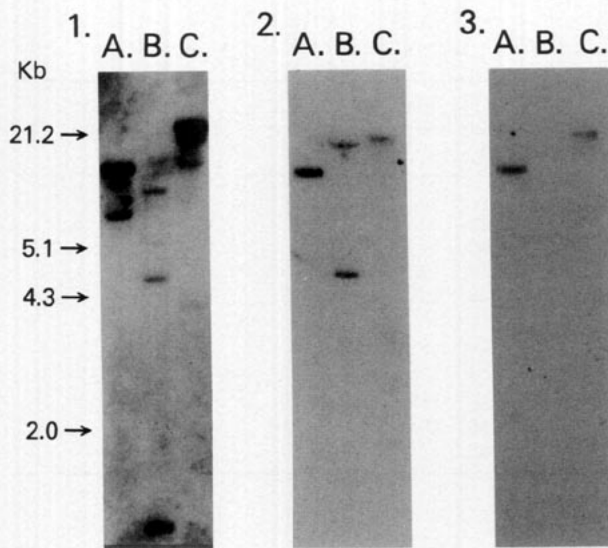


**Figure 2.** Southern analysis. Normal human genomic DNA (20 μg) was digested with *Eco*RI (lane A), *Bam*HI (lane B), or *Hin*dIII (lane C), resolved on a 0·8% agarose gel containing 0·5 × TBE, stained, photographed and transferred to S & S Nytran. The blots were prehybridized, hybridized with radiolabeled probes (0·5–2 × ($10^6$) cts/min per ml of hybridization solution) and washed (Elwood, 1989). Panel 1, the blot was probed with full length FR c15 cDNA radiolabeled by nick translation (Nick Translation Kit, Promega). Panel 2, a different blot was probed with 5′ *Eco*RI–*Dra*III and panel 3, 3′ *Eco*RI-*Nsi*I FR c15 restriction fragments radiolabeled to a specific activity ≥0·5 × ($10^6$) cts/min per μg using random oligonucleotide hexamers (Prime-a-Gene Kit, Promega).

of sequence determined from the longest λGT11 clone (FR c15) and from the 5′ terminus (26 bp) of the longest PCR-cloned cDNA (FR pcr1)) are shown in Figure 1A and B, respectively. The 5′ terminal sequence of FR-P3 was verified by determination of the cDNA sequence of 11 other PCR-cloned cDNAs. The 1084 bp cDNA contains a 151 bp 5′ untranslated region (UTR), a 765 bp open reading frame (ORF) initiated by a consensus translational start site (Kozak, 1987) and encoding a 255 residue protein, and a 168 bp 3′ UTR. The FR-P3 cDNA does not contain a poly(A) tail; however, the 3′ terminal six nucleotides (AATAAA) constitute a consensus polyadenylation signal which is 14 bp upstream from the polyA tail of FR-P2 cDNA.

The ultrastructure of the FR-P3 gene was determined by Southern analysis (Fig. 2) of normal human genomic DNA. The restriction enzymes were selected because the FR-P3 cDNA lacked internal *Eco*RI (lane A) and *Hin*dIII (lane C) sites and contained a single *Bam*HI (lane B) site. Figure 2, panel 1, contains the results of a Southern blot probed with radiolabeled full-length FR c15 cDNA. The multiplicity of bands observed in each lane resembles the Southern analysis of the FR-KB1 cell gene (Elwood, 1989) and suggests that the FR-P3 gene locus contains intronic restriction sites, that genomic DNA contains FR-P3 pseudogenes, or that the FR-P3 gene is a member of a gene family. In contrast, single bands (approx. 10 and >21 kb) are observed in lanes containing the *Eco*RI (lane A) and *Hin*dIII (lane C) digests, respectively, when a Southern blot was probed with 5′ and 3′ cDNA-specific FR c15 probes (Fig. 2, panels 2 and 3, respectively). Two *Bam*HI genomic fragments hybridize with the 5′ FR c15 probe (Fig. 2, panel 2, lane B) suggesting that the FR-P3 gene is polymorphic relative to 5′ *Bam*HI sites, that the gene contains an intronic *Bam*HI site, or that the genomic DNA was partially digested. The latter possibility is unlikely, as the 3′ FR c15 probe hybridizes with a single band and because the DNA appeared completely digested on ethidium bromide-stained gels. These results indicate that the FR-P3 gene is contained within 10 kb *Eco*RI and >21 kb *Hin*dIII genomic restriction fragments.

To determine the organization of the FR-P3 gene, we isolated 45 hybridization-positive clones from a normal human lymphocyte genomic *EMBL3* library. Based on restriction mapping and Southern blot analysis of purified recombinant *EMBL3* DNA, five non-overlapping genomic clones, ranging in length from 16·2 kb to 18·8 kb, were identified (Fig. 3A). Clones C1 and C2 contain single *Eco*RI (approx. 9 kb) and *Hin*dIII (>14·5 kb) restriction fragments, which hybridize strongly with FR-P3-specific probes. Furthermore, the restriction map of clone C1 is identical to that of C2, except for their 5′ restriction sites, suggesting that both clones contain the FR-P3 gene and represent polymorphisms.

The nucleotide sequence of the placental FR-P3 gene including 610 bp 5′ and 301 bp 3′ relative to the termini of FR-P3 (Fig. 1) was determined from genomic clone C1. The sequencing strategy, and chromosomal organization and partial restriction map of the gene are shown in Figure 3B. The complete FR-P3 gene sequence is available from EMBL under accession number X69516. The FR-P3 gene locus spans 5148 bp and is composed of five exons. The length of the exons and introns, and the nucleotide sequences of 5′ donor and 3′ acceptor splice sites of the FR-P3 gene are summarized in Table 1. The nucleotide sequence of putative exon I is identical to that of the 5′ terminus of FR-P3, and diverges at residue 34 of exon I from the reported 5′ terminal sequence of FR-P2. The nucleotides flanking residue 34 of exon I contained in C1 do not constitute consensus splice donor or acceptor sequences. Exon II contains the remaining 24 bp of the 5′ UTR of the FR-P3 cDNA and the translation initiation codon (ATG). Exon V contains the 3′ ORF sequence that encodes the carboxyl terminus of the placental FR and the 3′ UTR. The sequences flanking introns A, B, C and D are consensus splice site sequences (Lewin, 1990*a*), except for the 5′ donor splice sequence of intron A, which contains a "GC" variant rather than the more typical "GT" splice sequence. The "GC" splice sequence appears to be authentic as: (1) RNase protection assays demonstrate protection of the FR-P3 transcript contained in normal human tissues by riboprobes

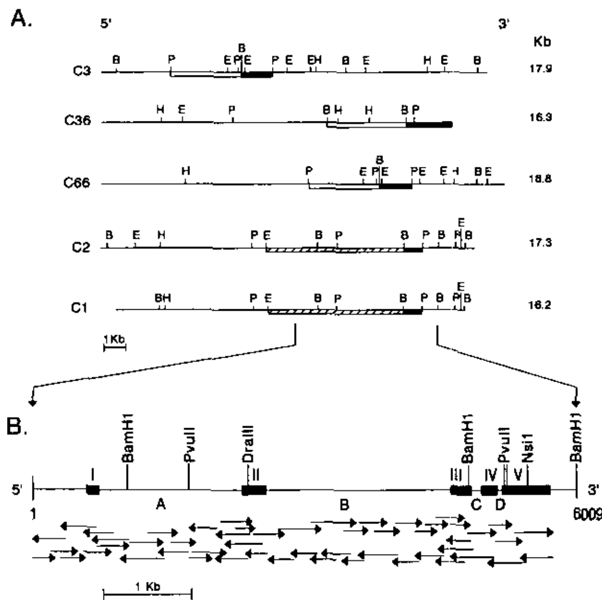


**Figure 3.** Organization and structure of the placental folate receptor (FR-P3) gene. A, The hybridization and restriction maps of 5 non-overlapping *EMBL3* clones homologous to the FR cDNAs. The restriction sites for *Eco*RI, *Bam*HI, *Hin*dIII and *Pvu*II are indicated by the "E", "B", "H" and "P", respectively. Restriction fragments hybridizing with 5′ FR-P3 and FR-KB1 probes are designated by the hatched (▨) and open (□) rectangles, respectively. The closed rectangle (■) represents restriction fragments that hybridize with 3′ FR cDNA probes. B, A partial restriction map, the sequencing strategy, and the chromosomal organization of the FR-P3 gene determined from clone C1. The arrows indicate the extent and orientation of sequence determined from each primer using the dideoxynucleotide chain-termination method of Sanger *et al.* (1977). Exons, represented by closed rectangles, are designated by roman numerals (I–V) and introns by capital letters (A–D). For these experiments, genomic clones ($n = 45$) were isolated from a normal human lymphocyte genomic *EMBL3* library (Clontech) by screening $3 \times (10^6)$ plaque-forming units with radiolabeled full length FR c32 (FR-KB1) cDNA (Benton & Davis, 1977) under stringent conditions. Plaque-purified recombinant *EMBL3* DNA was isolated from liquid or plate lysates using lambdasorb (Promega) as recommended by the vendor for restriction and hybridization mapping. FR-KB1 specific probes consisted of the 5′ *Eco*RI–*Hinc*II and 3′ *Eco*RI–*Pst*I restriction fragments of FR c32 (FR-KB1) (Elwood, 1989) radiolabeled using random oligonucleotide hexamers. FR-P3-specific probes are described in the legend of Figure 2. The specificities of each of the 5′ and 3′ cDNA restriction fragment probes in discriminating between the FR-P3 and FR-KB cDNAs were verified by Southern analysis.

corresponding to the sequence of exon I of genomic clone C1 (see below) and by riboprobes corresponding to cDNA sequence spanning the putative splice junction between exons I and II (data not shown); and (2) because the same "GC" splice site is present in two polymorphic genomic clones (C1 and C2) containing the FR-P3 gene. This "GC" variant is contained in other genes including platelet GPIIb (Heidenreich *et al.*, 1990), alpha crystallin (King & Piatigorsky, 1983), alpha globin (Dodgson & Engel, 1983) genes, and adenine phosphoribosyltransferase (Broderick *et al.*, 1987). Padgett *et al.* (1986) have also demonstrated proper splicing of the "GC" variant *in vitro*.

**Table 1**
*Exon/intron sizes of FR genomic clones*

| Clone | Size (bp) C1 | C3 | C36 | C66 | Exon/intron splice junctions |
|---|---|---|---|---|---|
| *Intron* | | | | | |
| A | 1645 | | | | CAG/gcaaga . . tag/C |
| B | 2130 | | | | CAA/gtacgg . . cag/T |
| C | 117 | 196 | 151 | | CAG/gtgagg . . cag/G |
| D | 158 | 105 | 149 | 105 | CAG/gtgagg . . cag/G |
| *Exon* | | | | | |
| I | 127 | | | | |
| II | 174 | | | | |
| III | 189 | 187 | 195 | | |
| IV | 136 | 134 | 141 | 134 | |
| V | 472 | | | | |

The nucleotide sequences from each exon and from introns C and D of clone C2 are identical to the corresponding sequence of clone C1, except for a single nucleotide substitution ("A" for "G") 319 bp upstream from exon I. The size of amplified introns A (1·6 kb) and B (2·0 kb) and the splice acceptor and donor sequences of C2 are identical to clone C1. Combined with their differences in restriction sites, these results indicate that C1 and C2 represent FR-P3 polymorphisms. Although it remains possible that the unique 5′ restriction sites near the *EMBL3* cloning sites of C1 and C2 represent cloning artifacts, this seems unlikely because the restriction maps of each independent copy of C1 ($n = 23$) and of C2 ($n = 10$) are identical to the maps shown in Figure 3A and because the sequences of C1 and C2 are different. Partial nucleotide sequences of genomic clones C3, C36 and C66 (Fig. 3A) indicate that these clones are homologous to the FR cDNAs, but do not encode FR-KB or FR-P transcripts. Based on the presence of introns, and in-frame termination codons, these clones appear to represent unprocessed FR pseudogenes. The sizes of the putative exons and introns, and their splice site sequences of clones C3, C36 and C66, are summarized in Table 1.

The FR-P3 gene nucleotide sequence upstream from the first nucleotide of putative exon I does not contain a consensus "TATA" or "CAAT" box within the first 100 nucleotides. Potential eukaryotic promoter sequences (Locker & Buzzard, 1990; Lewin, 1990*b*) are present 413 bp and 225 bp ("TATA" boxes); 314 bp ("CAAT" box); and 550 bp, 515 bp, 76 bp ("GC" boxes) upstream from the first nucleotide of exon I. Interestingly, the immediate upstream sequence (residues −72 to −23 relative to exon I) is purine rich (84%) and resembles the purine motif (Pu box) involved in kidney and lymphoid cell-specific enhancer activity (Weber *et al.*, 1983).

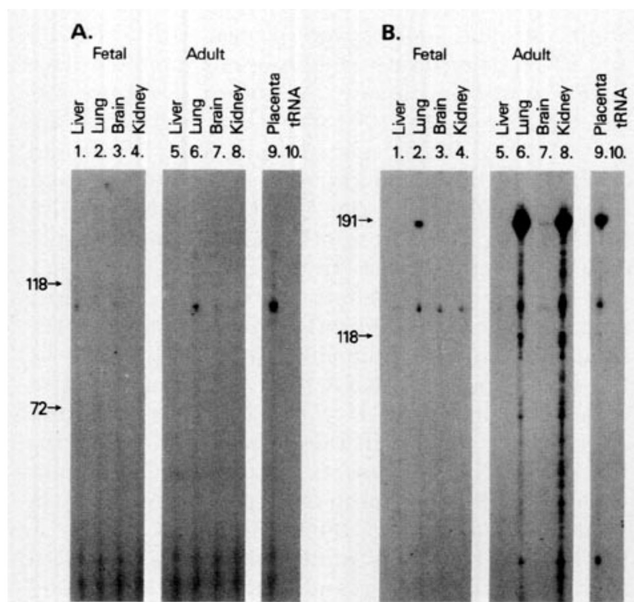


**Figure 4.** RNase protection assays: analysis of folate receptor transcripts in normal human fetal and adult tissue. The FR-P3 antisense riboprobe was synthesized from a pCR-1000 construct which contains a 449 bp insert corresponding to 5′ terminal 106 bp of putative exon I and 343 bp of upstream genomic sequence. The insert was amplified from clone C1, subcloned into pCR-1000, and sequenced to ensure the sequence fidelity of the construct. The FR-KB antisense riboprobe is complementary to residues 203 to 383 of FR-KB1 cDNA and was synthesized from the 5′ *Eco*RI–*Hin*cII restriction fragment of FR c10 (Elwood, 1989) subcloned into pGEM4Z. Riboprobes were synthesized with reagents supplied in an *in vitro* transcription kit (Promega) using a protocol provided by the vendor. RNase protection assays were performed as described (Ausubel *et al.*, 1989) on total RNA (25 μg) from adult tissues and placenta, and on poly(A)$^+$ mRNA (0·5 μg) from fetal tissues following hybridization with 5′ FR-P3 (A) or FR-KB1 (B) riboprobes. The tissue sources of RNA are indicated above each lane. For RNase protection assays and Northern analyses (see below), human RNA was extracted from cells (Elwood, 1989) or purchased from Clontech (poly(A)$^+$ mRNA from fetal liver, lung, kidney and brain; and total RNA from placenta or adult lung, liver, kidney, brain, testes, thymus, spleen, mammary gland, cerebellum, ovary and small intestinal mucosa). Protected fragments were resolved on a 6% polyacrylamide/8 M-urea sequencing gel in 0·5 × TBE at 60 W (constant power) until the tracking dye approached the bottom of the gel (approx. 2·5 h). Size markers included end-labeled *Hae*III-digested phiX174 DNA and a sequencing ladder.

We have mapped the transcription start site of the FR-P3 gene by three different methods including: (1) RACE PCR; (2) primer extension assays; and (3) RNase protection assays. The length of the FR-P3 5′ UTR was determined from the sizes of specific primer extended products and of specific cDNAs. The cDNAs were amplifed from the placental λGT11 cDNA library and poly(A)-tailed first strand human placental cDNA using cDNA-specific antisense primers (complementary to residues +38 to +60, and +330 to +345 of FR-P3) and universal λGT11 or oligo(dT) sense primers. The specificities of the amplified fragments were verified by restriction mapping and sequence determination. Primer extension assays of placental total RNA with an antisense primer complimenting residues +50 through +80 of FR-P3 cDNA yields a specific extended product of 228 bp (data not shown). Based on the sizes of amplified FR-P3 cDNAs and primer extended products, the FR-P3 transcripts contain a 5′ UTR that is approximately 148 bp to 150 bp in length.

The transcription start site of the FR-P3 gene was determined by RNase protection assays of selected human tissue RNA. A 449 bp antisense FR-P3 riboprobe complementing the 5′ terminal 106 bp of putative exon I and 343 bp of upstream sequence of the FR-P3 gene was utilized to define the 5′ boundary of exon I and to determine the spatial relationship between the transcriptional start site and the uptream TATA, CAAT and GC boxes. It is important to note that this riboprobe contains 73 nucleotides complementary to sequence reported in the 5′ UTR of FR-P2 (Ratman *et al.*, 1989). Figure 4A contains the results of RNase protection of the FR-P3 riboprobe following hybridization with various human RNAs. Placental RNA (lane 9) specifically protects two to four major fragments between 100 bp and 106 bp in length compared to control wheat-germ tRNA (lane 10). RNA from fetal liver (lane 1) and adult lung (lane 6) also protected 100 bp to 106 bp fragments; and faint signals are present in lanes 7 and 8 containing RNA from adult brain and kidney, respectively. Combined with the results of PCR and primer extension assays, these results indicate that exon I contains approximately 127 nucleotides. As functional "TATA" and "CAAT" boxes are classically (Lewin, 1990*c*) approximately 25 bp and 80 bp, respectively, upstream from the transcription start site, it is unlikely that the "TATA" and "CAAT" promoter sequences upstream from exon I (see above) are involved in expression of the FR-P3 gene. Furthermore, the stuttering at the transcription start site manifested by the four closely grouped protected fragments in the RNase protection assay is consistent with a "TATA"-less promoter (Lewin, 1990*b*).

Previous studies have shown that FR and FR mRNA levels are inversely related to folate concentrations in tissue culture cells (Kane *et al.*, 1988; Sadisavan & Rothenberg, 1989). Although FRs are widely distributed in mammalian tissues (Henderson, 1990; Kane & Waxman, 1989), the specific expression of each of the two FRs and the mechanism(s) by which folates interact during transcription or post-transcriptional processing, the effect of other factors on expression, and the elements involved in regulating the expression of each FR gene are unknown. The KB cell FR protein and/or transcript has been identified in cultured human cells (Elwood, 1989; Campbell *et al.*, 1991; Coney *et al.*, 1991) and in RNA from human tissues (Elwood, 1989; Weitman *et al.*, 1992). The placental

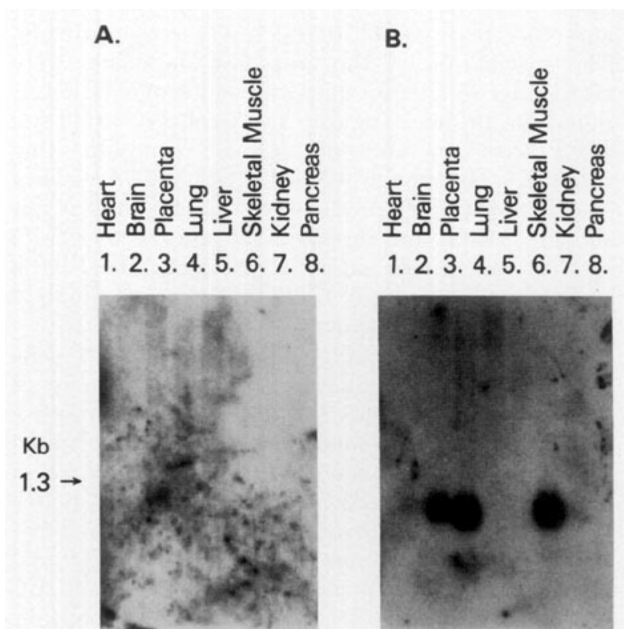


**Figure 5.** Northern analyses. The sources of poly(A)$^+$ RNA (2 $\mu$g) are indicated above each lane. The blot (Clontech) was hybridized with radiolabeled 5′ cDNA-specific restriction fragments of FR-P3 and FR-KB1 cDNAs (see Figs 2 and 3) in panels A and B, respectively. The autoradiographs were exposed for 10 days (panel A) and 3 days (panel B). Following transfer to S & S Nytran, the blot was prehybridized, hybridized and washed as described (Elwood, 1989). The integrity of the RNA and the efficiency of transfer was verified by hybridization of the blot with radiolabeled human actin (data not shown).

FR protein and transcript have been demonstrated only in human placenta, suggesting that the FR-P2 encoded FR may represent a fetal form of the human FR (Ratman *et al.*, 1989).

Utilizing cDNA-specific probes, we have studied expression of the FR-P3 and FR-KB gene transcripts in selected normal human fetal and adult tissues by Northern analysis (Fig. 5) and RNase protection assays (Fig. 4). The radiolabeled 5′ *Eco*RI–*Dra*III FR-P3 probe weakly hybridizes with a single 1100 bp mRNA species contained in human placental poly(A) mRNA (Fig. 5A, lane 3), but is not detectable in poly(A) mRNA from adult human tissues including heart, brain, lung, liver, skeletal muscle, kidney, pancreas, or from KB cell total RNA (data not shown). The size of the FR-P3 transcript agrees with studies of the structure of the FR-P3 gene. In contrast to the restricted FR-P3 tissue expression, the 5′ FR-KB1 cDNA-specific probe strongly hybridizes with an 1100 bp mRNA species contained in human lung, kidney, placenta, brain and pancreas (Fig. 5B). These differences in FR-P3 and FR-KB1 transcript expression as determined from Northern blots are consistent with the proposal that FR-P2 represents a fetal FR.

To determine if the FR genes are differentially expressed during human development, and to determine if the divergent 5′ terminal sequences of the FR-P3 and FR-P2 cDNAs result from alternative splicing of exon I, RNAs from corresponding human adult and fetal sources were probed with 5′ FR-P3 and FR-KB1 antisense riboprobes by more sensitive RNase protection assays. As shown in Figure 4A, specific 100 bp to 106 bp FR-P3 fragments are protected by RNA isolated from fetal (liver) and adult (lung, brain and kidney) tissues and placenta. Specific fragments of the 5′ FR-P3 probe are also protected by RNA from other adult human tissues including testes, retina, thymus, spleen, mammary gland, ovary, cerebellum and small intestinal mucosa (data not shown). We did not observe smaller protected fragments (e.g. 73 bp) of the FR-P3 riboprobe in RNA from the tissues studied indicating that exon I of the FR-P3 gene is not alternatively spliced in these tissues, and that the FR-P3 cDNA represents the most abundant placental FR transcript in these tissues. As shown in Figure 4B, a specific 180 bp fragment of the 5′ FR-KB1 riboprobe is protected by RNA isolated from human fetal (lung) and adult (lung, kidney and brain) tissues and placenta. The autoradiograph shown in Figure 4B was overexposed to demonstrate the faint 180 bp band protected by human adult brain RNA (lane 7). The signal intensities of the smaller protected bands (<180 bp) contained in lanes 6 and 8 are faint compared to the 180 bp fragment and most likely represent over-digestion with RNase, or hybridization with other homologous transcripts.

The results of Northern analysis and RNase protection assays indicate that the relative FR-P3 transcript abundance among tissues and the tissues expressing FR-P3 transcripts (Figs 4A and 5A) are independent of the FR-KB1 transcript abundance and tissue expression (Figs 4B and 5B). These differences are consistent with independent regulation of the FR-P3 and FR-KB1 genes. Furthermore, the variable abundance of each specific FR transcript among human tissues indicates that expression of the FR genes are non-co-ordinately regulated rather than constitutively expressed, presumably as a result of transcriptional regulation or, alternatively, of differences in transcript processing (e.g. splicing, nuclear export or stability). The absence of a "TATA" box in close proximity to the transcription start site of the FR-P3 gene resembles other recently described mammalian genes (e.g. terminal deoxynucleotide transferase (TdT) gene (Landau *et al.*, 1984), p53 gene (Bienz-Tadmor *et al.*, 1985), and the T cell receptor beta-chain gene (Anderson *et al.*, 1988)), which are not constitutively expressed by all cells, which contain "TATA"-less promoters, and which are regulated during cell differentiation or development, or exhibit restricted tissue expression. The presence of FR-P3 and FR-KB1 transcripts in normal fetal and adult tissue demonstrates that transcription of these genes is not related to fetal development.

To study the conservation of DNA homologous to the FRs, a Southern blot containing genomic DNA from a spectrum of eukaryotic sources was probed with radiolabeled FR-KB1 cDNA (Fig. 6). One or

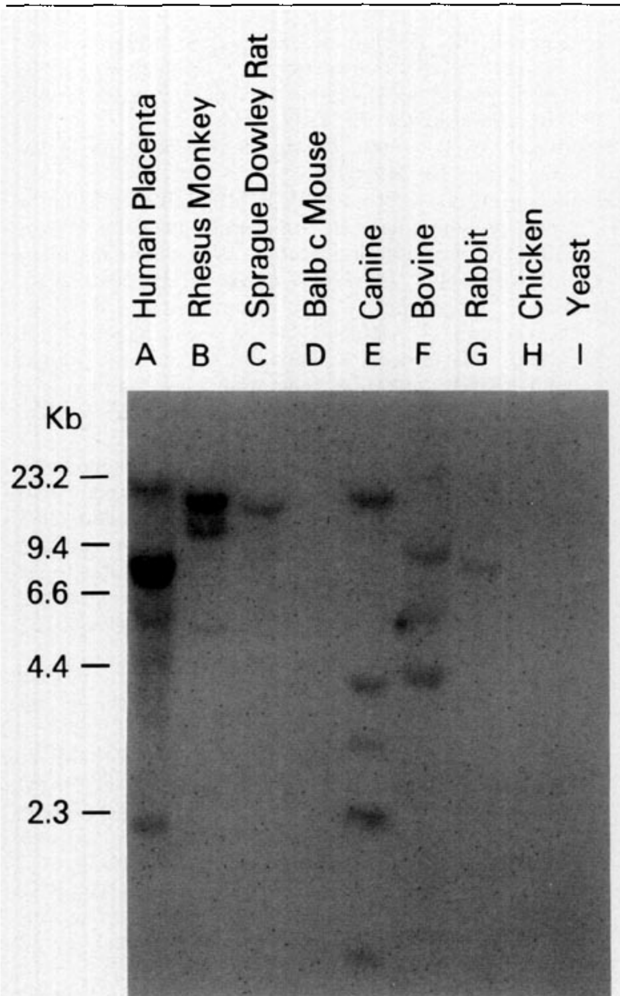


**Figure 6.** Southern analysis of eukaryotic DNA. Genomic DNAs (8 μg) from the sources designated above each lane were digested with *Eco*RI. The blot (Clontech) was hybridized with full-length FR-KB1 cDNA and washed under stringent conditions.

more bands are present in lanes containing DNA from each source, except for chicken (lane H) and yeast (lane I) DNA. Furthermore, under relaxed stringency, the lanes containing chicken and yeast DNA also contained signals. These results are consistent with conservation of DNA homologous to the human FR cDNA throughout eukaryotic evolution and are consistent with their role in folate homeostasis. Compared to the FR-KB1 cDNA, the full-length FR-P3 cDNA and its ORF are approximately 59% and 77% homologous, respectively. Moreover, the nucleotide sequences contained within exons III and IV of the FR-P3 gene are more highly conserved (80% and 85% homology, respectively) when compared to corresponding sequence of FR-KB1, suggesting a common and/or important functional role (e.g. ligand binding site) of the protein encoded by these exons (Lewin, 1990*e*). The homology (77%) between the ORFs of FR-P3 and FR-KB1 (Elwood, 1989) cDNAs and the similarity of chromosomal organization of the FR-P3 gene and FR-KB1 gene (Page & Elwood, 1990) suggest that the genes diverged approximately 2 ($10^8$) years ago (unit evolutionary period (UEP) = 10·4 ($10^6$) years (Lewin, 1990*d*)), presumably by reduplication.

Although homologous to the reported FR-P2 cDNA, the FR-P3 cDNA contains important differences, including a unique 33 bp 5′ terminal nucleotide sequence, and sequence differences involving the putative splice junction between exons I and II and within the ORF. We considered the possibilities that the sequence divergence at the 5′ terminus resulted from cloning or ligation artifacts, from alternative splicing of the first exon(s) as proposed for the FR-KB1 cell gene (Coney *et al.*, 1991) or from placental FR gene polymorphisms. The authenticity of the FR-P3 cDNA sequence (see Fig. 1B) was verified by sequence analysis of 11 other cDNA clones and two genomic clones, and by RNase protection assays with a riboprobe containing putative exon I and upstream genomic sequence. Alternative splicing of exon I is unlikely because the FR-P3 gene sequence at the residue of sequence divergence between the FR-P3 and FR-P2 cDNAs is not a consensus splice acceptor sequence, as a smaller fragment(s) (e.g. 73 bp) is not present in RNase protection assays using a probe complementary to putative exon I, and as we were unable to amplify specific FR-P3 cDNA fragments from genomic DNA, clones C1 and C2, or human placental cDNA using sense primers derived from the divergent 5′ FR-P2 sequence. The presence of FR-P3 gene polymorphisms and the internal sequence differences between the FR-P3 and FR-P2 cDNAs are most consistent with the possibility that the FR-P2 transcript is encoded by a polymorphic or independent FR-P gene. However, the prevalence of such a polymorphism or abundance of its respective transcript would be relatively low, given the results of RNase protection assays.

In this report, we describe a placental FR cDNA (FR-P3) and the chromosomal organization and full nucleotide sequence of its gene. Furthermore, we demonstrate that both FR-P3 and FR-KB1 transcripts are expressed in fetal and adult tissues, and that the abundance of FR-P3 and FR-KB1 transcripts are variable among the tissues studied. These data indicate that the human KB cell and the two placental FR transcripts are products of distinct, but closely related, members of a conserved gene family which are variably and independently expressed in normal human tissues. With the upstream sequence of the FR-P3 gene and the cDNA-specific restriction fragment probes, we can study the potential *cis* and *trans* acting factors that may be involved in tissue differential expression of the FR-P3 gene, clone the gene encoding the FR-KB1 cDNA, and advance our understanding of the role of the FRs in folate and antifolate accumulation and cytotoxicity.

## References

Anderson, S. J., Chow, H. S. & Loh, D. Y. (1988). A conserved sequence in the T-cell receptor B-chain promotor region. *Proc. Nat. Acad. Sci., U.S.A.* **85**, 3551–3554.

Antony, A. C. (1992). The biologic chemistry of folate receptors. *Blood*, **79**, 2807–2820.

Antony, A. C., Utley, C., Van Horne, K. C. & Kolhouse, J. F. (1981). Isolation and characterization of a folate receptor from human placenta. *J. Biol. Chem.* **256**, 9684–9692.

Antony, A. C., Utley, C., Marcell, P. D. & Kolhouse, J. F. (1982). Isolation, characterization and comparison of the solubilized particulate and soluble folate-binding proteins for human milk. *J. Biol. Chem.* **257**, 10081–10089.

Antony, A. C., Kane, M. A., Portillo, R. M., Elwood, P. C. & Kolhouse, J. F. (1985). Studies of the role of a particulate folate-binding protein in the uptake of 5-methyltetrahydrofolate by cultured human KB cells. *J. Biol. Chem.* **260**, 14911–14917.

Antony, A. C., Bruno, E., Briddell, R. A., Brandt, J. E., Verma, R. S. & Hoffman, R. (1987). Effect of perturbation of specific folate receptors during *in vitro* erythropoiesis. *J. Clin. Invest.* **80**, 1617–1623.

Ausubel, F. M., Brent, R., Kingston, R. E., Moore, D. D., Seidman, J. G., Smith, J. A. & Struhl, K. (1989). Ribonuclease protection assay. In *Current Protocols in Molecular Biology*, pp. 4.7.1–4.7.8, Greene Publishing and Wiley-Interscience, New York.

Benton, W. D. & Davis, R. W. (1977). Screening λgt recombinant clones by hybridization to single plaques *in situ*. *Science*, **196**, 180–182.

Bienz-Tadmor, B., Zakut-Houri, R., Libresco, S., Givol, D. & Oren, M. (1985). The 5′ region of the p53 gene: evolutionary conservation and evidence for a negative regulatory element. *EMBO J.* **4**, 3209–3213.

Broderick, T. P., Schaff, D. A., Bertino, A. M., Dush, M. K., Tischfield, J. A. & Stambrook, P. J. (1987). Comparative anatomy of the human APRT gene and enzyme: nucleotide sequence divergence and conservation of a nonrandom CpG dinucleotide arrangement. *Proc. Nat. Acad. Sci., U.S.A.* **84**, 3349–3353.

Campbell, I. G., Jones, T. A., Foulkes, W. D. & Trowsdale, J. (1991). Folate-binding protein is a marker for ovarian cancer. *Cancer Res.* **51**, 5329–5338.

Coney, L. R., Tomassetti, A., Carayannopoulos, L., Frasca, V., Kamen, B. A., Colnaghi, M. I. & Zurawski, V. R. (1991). Cloning of a tumor-associated antigen: MOv18 and MOv19 antibodies recognize a folate-binding protein. *Cancer Res.* **51**, 6125–6132.

Dodgson, J. B. & Engel, J. D. (1983). The nucleotide sequence of the adult chicken alpha-globin gene. *J. Biol. Chem.* **258**, 6165–7177.

Elwood, P. C. (1989). Molecular cloning and characterization of the human folate-binding protein cDNA from placenta and malignant tissue culture (KB) cells. *J. Biol. Chem.* **264**, 14893–14901.

Elwood, P. C., Kane, M. A., Portillo, R. M. & Kolhouse, J. F. (1986). The isolation, characterization and comparison of the membrane-associated and soluble folate-binding proteins from human KB cells. *J. Biol. Chem.* **261**, 15416–15423.

Elwood, P. C., Deutsch, J. C. & Kolhouse, J. F. (1991). The conversion of the human membrane-associated folate binding protein (folate receptor) to the soluble folate binding protein by a membrane-associated metalloprotease. *J. Biol. Chem.* **266**, 2346–2353.

Frohman, M. A. (1990). RACE: rapid amplification of cDNA ends. In *PCR Protocols, a Guide to Methods and Applications* (Innis, M. A., Gelfand, D. H., Sninsky, J. J. & White, T. J., eds), 1st edit., pp. 28–38, Academic Press, San Diego.

Heidenreich, R., Eisman, R., Surrey, S., Delgrosso, K., Bennett, J. S., Schwartz, E. & Poncz, M. (1990). Organization of the gene for platelet glycoprotein IIb. *Biochemistry*, **29**, 1232–1244.

Henderson, G. B. (1990). Folate binding proteins. *Annu. Rev. Nutr.* **10**, 319–335.

Henderson, G. B. & Strauss, B. P. (1990). Growth inhibition by homofolate in tumor cells utilizing a high-affinity folate binding protein as a means for folate internalization. *Biochem. Pharmacol.* **39**, 2019–2025.

Jansen, G., Westerhof, G. R., Jarmuszewski, M. J. A., Kathmann, I., Rijksen, G. & Schornagel, J. H. (1990). Methotrexate transport in variant human CCRF-CEM leukemia cells with elevated levels of reduced folate carrier. *J. Biol. Chem.* **265**, 18272–18277.

Kamen, B. A. & Capdevila, A. (1986). Receptor-mediated folate accumulation is regulated by the cellular folate content. *Proc. Nat. Acad. Sci., U.S.A.* **83**, 5983–5987.

Kamen, B. A., Wand, M., Streckfuss, A. J., Peryea, X. & Anderson, R. G. W. (1988). Delivery of folates to the cytoplasm of MA104 cells is mediated by a surface membrane receptor that recycles. *J. Biol. Chem.* **263**, 13602–13609.

Kane, M. A. & Waxman, S. (1989). Role of folate binding proteins in folate metabolism. *Lab. Invest.* **60**, 737–746.

Kane, M. A., Elwood, P. C., Portillo, R. M., Najfeld, V., Finley, A., Waxman, S. & Kolhouse, J. F. (1988). Influence on immunoreactive folate-binding proteins of extracellular folate concentration in cultured human cells. *J. Clin. Invest.* **81**, 1398–1406.

King, C. R. & Piatigorsky, J. (1983). Alternative RNA splicing of the murine alpha A-crystalline gene: protein coding information within an intron. *Cell*, **32**, 707–712.

Kozak, M. (1987). At least six nucleotides preceding the AUG initiator codon enhance translation in mammalian cells. *J. Mol. Biol.* **196**, 947–950.

Lacey, S. W., Sanders, J. M., Rothberg, K. G., Anderson, R. G. W. & Kamen, B. A. (1989). Complementary DNA for the folate binding protein correctly predicts anchoring to the membrane by glycosyl-phosphatidylinositol. *J. Clin. Invest.* **84**, 715–720.

Landau, N. R., St John, T. P., Weissman, I. L., Wolf, S. C., Silverstone, A. E. & Baltimore, D. (1984). Cloning of terminal transferase cDNA by antibody screening. *Proc. Nat. Acad. Sci., U.S.A.* **81**, 5836–5840.

Lewin, B. (1990*a*). Mechanisms of RNA splicing. In *Genes IV*, 4th edit., p. 597, Cell Press, Cambridge, MA.

Lewin, B. (1990*b*). Building the transcriptional complex. In *Genes IV*, 4th edit., pp. 545–550, Cell Press, Cambridge, MA.

Lewin, B. (1990*c*). Building the transcription complex. In *Genes IV*, 4th edit., p. 551, Cell Press, Cambridge, MA.

Lewin, B. (1990*d*). Structural genes evolve in families. In *Genes IV*, 4th edit., p. 505, Cell Press, Cambridge, MA.

Lewin, B. (1990*e*). The organization of interrupted genes. In *Genes IV*, 4th edit., p. 493, Cell Press, Cambridge, MA.

Locker, J. & Buzard, G. (1990). A dictionary of transcriptional control sequences. *J. DNA Seq. Mapp.* **1**, 3–11.

Luhrs, C. A., Pitiranggon, P., da Costa, M., Rothenberg, S. P., Slomiany, B. L., Brink, L., Tous, G. I. & Stein, S. (1987). Purified membrane and soluble folate

binding proteins from cultured KB cells have similar amino acid compositions and molecular weights but differ in fatty acylation. *Proc. Nat. Acad. Sci., U.S.A.* **84**, 6546–6549.

Padgett, R. A., Grabowski, P. J., Konarska, M. M., Seiler, S. & Sharp, P. A. (1986). Splicing of messenger RNA precursors. *Annu. Rev. Biochem.* **55**, 1119–1150.

Page, S. T. & Elwood, P. C. (1990). Human folate receptor gene family: initial characterization. *Blood,* **76**, 43a.

Ratnan, M., Marquardt, H., Duhring, J. L. & Freisheim, J. H. (1989). Homologous membrane folate binding proteins in human placenta: cloning and sequence of a cDNA. *Biochemistry*, **28**, 8249–8254.

Sadasivan, E. & Rothenberg, S. P. (1989). The complete amino acid sequence of a human folate binding protein from KB cells determined from the cDNA. *J. Biol. Chem.* **264**, 5806–5811.

Sanger, F., Nicklen, S. & Coulson, R. (1977). DNA sequencing with chain terminating inhibitors. *Proc. Nat. Acad. Sci., U.S.A.* **74**, 5463–5467.

Svendsen, I. B., Hansen, S. I., Holm, J. & Lyngbye, J. (1984). The complete amino acid sequence of the folate binding protein from cow's milk. *Carlsberg Res. Commun.* **49**, 123–131.

Weber, F., deVilliers, J. & Schaffner, W. (1983). An SV40 "enhancer trap" incorporates exogenous enhancers or generates enhancers from its own sequence. *Cell,* **36**, 983–992.

Weitman, S. D., Lark, R. H., Coney, L. R., Fort, D. W., Frasca, V., Zurawski, V. R. & Kamen, B. A. (1992). Distribution of the folate receptor GP38 in normal and malignant cell lines and tissues. *Cancer Res.* **52**, 3396–3401.

*Edited by J. Karn*