

Structure of the genes for porcine endometrial secreted and membrane folate binding proteins

J.L. Vallet*, T.P.L. Smith, T.S. Sonstegard, M. Heaton, S.C. Fahrenkrug

USDA, ARS, Roman L. Hruska U.S. Meat Animal Research Center, Clay Center, NE, USA

Received 14 March 2001; accepted 4 April 2001

Abstract

The endometrium of the pig produces two types of folate binding proteins (FBP) which, based on their sequences, are likely to be membrane (m) and secreted (s) forms. A clone containing both a gene coding for the sFBP cDNA and a gene coding for the mFBP was isolated from a yeast artificial chromosome (YAC) library. Each gene was subcloned and sequenced. The gene for sFBP spanned 4.4 kbp and included 5 exons. The mFBP gene spanned 7.0 kbp and also contained 5 exons. Structures of the genes were very similar for the last three exons, and this similarity was shared with other known FBP/folate receptor (FR) gene sequences. Unexpectedly, portions of introns 3 and 4 of both genes were highly homologous, suggesting the possibility that sequences within these introns served some as yet unknown function. In contrast, the structures of the 5' exons differed between the two genes and other known FBP/FR genes. Comparison of putative promoter regions for the two genes with promoter regions for human FBP/FR genes revealed significant sequence homology between sFBP and human γ FBP and between mFBP and human α FR. These regions of homology may play a role in control of transcription of each gene. © 2001 Elsevier Science Inc. All rights reserved.

1. Introduction

Folates are vitamins that participate in methyl transfer reactions and are essential for methionine and DNA synthesis [1]. Consequently, rapidly growing tissues, such as the erythron and the developing conceptus, have a high requirement for folate, and deficiencies lead to abnormal erythropoiesis [2] and birth defects [3]. In swine, folate transport to the developing conceptus is not well understood, but recent evidence suggests that two different

* Corresponding author. Tel.: +1-402-762-4187; fax: +1-402-762-4382.

E-mail address: vallet@email.marc.usda.gov (J.L. Vallet).

Table 1

Primers used for screening and PCR cloning of portions of the FBP genes

YAC screening	
Forward A	GAGTGCTCGCCCAACCTG
Reverse B	AGTAGAAGTCAAGCGGTGGC
SFBP cosmid screening	
5 sec FBP-F1	AAAGCTCAGACTGCACTGTCC
5 sec FBP-R1	CAGATTTCTGTTTCCCTTCC
mFBP primers	
5' cosmid screening	
5mem FBP-F1	GGATTCTGCTGCTTTTGAC
5mem FBP-R-1	TCATGCAGGCCATCTTCC
3' cosmid screening	
COSMFBP-F3	GCCTGGCCTCTCCAGTTC
COSMFBP-R2	TAGAGGCACTGACGAGCTG
PCR Cloning	
MEMGEN-5	GCGCTGATCTGGCAACTC
FBPRCE-R2	TTCCCACCAGTTCTGACAGTC

forms of folate binding protein (FBP), a secreted (s) and a membrane (m) form, likely play central roles in this process during early pregnancy [4,5].

The sFBP is a 30,000 Mr protein that appears as a diffuse band after SDS-PAGE and Coomassie blue staining [4]. It binds folic acid with high affinity, and increases from Day 11 to Day 13 of the cycle or pregnancy to reach $\mu\text{g/ml}$ concentrations within the intrauterine lumen [4,6]. This increase in sFBP within the intrauterine lumen occurs in the absence of obvious changes in the amount of mRNA present in endometrium [5]. Cloning and sequencing of the cDNA for this protein indicated that it was related to other FBP/folate receptor (FR) cDNAs previously characterized, but differs from the porcine FR described previously [7]. Heterogeneity of the 5' untranslated region (UTR) was also demonstrated, likely due to differential mRNA splicing [6].

A putative mFBP cDNA was also isolated from endometrium along with the cDNA for sFBP [6]. The membrane linkage of this protein was predicted due to the presence of an intact glycosphosphatidylinositol linkage site [5,8]. The mFBP cDNA shares sequence homology with the sFBP cDNA as well as with other known FBP/FR cDNAs. The mRNA for mFBP is present in a variety of tissues and expression increases dramatically from Day 15 to Day 24 in endometrium from pregnant pigs. It is also relatively highly expressed in Day 30 placental tissue. The 5' UTR of this cDNA is also heterogeneous, probably due to differential splicing [5]. The mFBP protein itself has not yet been characterized.

To obtain clues to the control of production of these proteins in the intrauterine environment of swine during pregnancy and to determine the basis of the heterogeneity of the 5' UTRs for each cDNA, we cloned and sequenced the genes corresponding to each cDNA.

2. Materials and methods

A yeast artificial chromosome (YAC) genomic library was screened by PCR using primers capable of amplifying either gene (Table 1). Screening yielded one positive YAC clone that

```

sFBP cDNA GACACTGCTTCCGGGTGGGCCTCCAGG.....AGGGCCGAGGC 38
          |||
sFBP gene GACACTGCTTCCGGGTGGGCCTCCAGGTGGGCCTCCAGGAGGGCCGAGGC 50
          |||
          39 AGAG|GAGCCTCTGCCTGTGGGTGAAGCACTGGCTGGCGAACTCCGGAAG 88
          |||
          51 AGAG|GAGCCTCTGCCTGTGGGTGAAGCACTGGCTGGCGAACTCCGGAAG 912
          |||
          89 GGAGGTCCGGAGAGGTGGTGCCTCCCCCGCAGCAAAGCTCAGACTGCAC 138
          |||
          913 GGAGGTCCGGAGAGGTGGTGCCTCCCCCGCAGCAAAGCTCAGACTGCAC 962
          |||
          139 TGTCCCTCAGGTGGCAGTGGTGTCTACCACTTGGCACAGACCTCCACGGG 188
          |||
          963 TGTCCCTCAGGTGGCAGTGGTGTCTACCACTTGGCACAGACCTCCACGGG 1012
          |||
          189 CCCTTCATCGCTTGGCTCCACTGTGCTGTGGGGTAAGCGGCGCGGGGAGG 238
          |||
          1013 CCCTTCATCGCTTGGCTCCACTGTGCTGTGGGGTAAGCGGCGCGGGGAGG 1062
          |||
          239 GACGACGATCTGGGCTTGGAAAGGAAACAGGAAATCTGGCCAAGAAGCTT 288
          |||
          1063 GACGACGATCTGGGCTTGGAAAGGAAACAGGAAATCTGGCCAAGAAGCTT 1112
          |||
          289 ACGGCAGCTTTCTGGCAGAAGTGGATCAACATGGCCTGGCGGCTGACGCT 338
          |||
          1113 ACGGCAGCTTTCTGGCAGAAGTGGATCAACATGGCCTGGCGGCTGACGCT 1162
          |||
          339 CTTCGTGCTCCTGGGTTTGGTGGCTGCTGTGGGGGGCGCCCGGGCCAAGT 388
          |||
          1163 CTTCGTGCTCCTGGGTTTGGTGGCTGCTGTGGGGGGCGCCCGGGCCAAGT 1212
          |||
          389 CGGACATGCTCAATGTCTGCATGGATGCCAAGCACCACAAGCCAAAGCCA 438
          |||
          1213 CGGACATGCTCAATGTCTGCATGGATGCCAAGCACCACAAGCCAAAGCCA 1262
          |||
          439 AGCCCCGAGGACAAGCTGCACGACCAG|TGCAGCCCTGGAGGAAGAACTC 488
          |||
          1263 AGCCCCGAGGACAAGCTGCACGACCAG|TGCAGCCCTGGAGGAAGAACTC 3296
          |||
          489 CTGCTGCTCAGTCAACACCAGCCTAGAAGCCATAAAGACATCTCCTACC 538
          |||
          3297 CTGCTGCTCGGTCAACACCAGCCTAGAAGCCATAAAGACATCTCCTACC 3346
          |||
          539 TGTACAGATTCAACTGGGACCACTGCGGCAAGATGGAGCCGGCCTGCAAG 588
          |||
          3347 TGTACAGATTCAACTGGGACCACTGCGGCAAGATGGAGCCGGCCTGCAAG 3396
          |||
          589 CGCCACTTCAATCAAGACACCTGTCTCTATGAGTGTGCGCCCAACCTGGG 638
          |||
          3397 CGCCACTTCAATCAAGACACCTGTCTCTATGAGTGTGCGCCCAACCTGGG 3446
          |||
          639 GCCCTGGATCCAGGAG|GTGAACCAGAAGTGCCGCGAGAGCGGATCCTGA 688
    
```

Fig. 1. The sFBP and mFBP cDNA sequences aligned with their respective gene sequences are illustrated. Mismatched nucleotides are indicated in bold. The start and stop codons are underlined. For sFBP, the cDNA sequence at 913 bp codes for a serine while the gene codes for an arginine residue. All other mismatches do not affect the predicted amino acid sequence.

was found to contain both genes upon PCR analysis using primers specific to the 5' end of each cDNA (Table 1; 5). The YAC incorporating the FBP genes was isolated by pulse field electrophoresis and then partially digested with *SauIIIa* to generate fragments of a suitable size for cosmid cloning. Fragments were gel isolated, ligated with the superCOS vector, and the resulting DNA was packaged and bacteria were infected according to the directions contained in the Gigapack III packaging kit (Stratagene, La Jolla, CA). Resultant colonies were screened with primers specific to the 5' ends of each gene. A positive clone for the sFBP gene containing the entire gene and a positive clone containing only the 5' end of the

```

|||||
3447 GCCCTGGATCCAGGAG|GTGAACCAGAAGTGGCGCAGAGAGCGGATCCCTGA 3692
|||||
689 ACGTGCCCTCTGCAAAGAGGACTGTCAGATCTGGTGGGAAGACTGCCGT 738
|||||
3693 ACGTGCCCTCTGCAAAGAGGACTGTCAGATCTGGTGGGAAGACTGCCGC 3742
|||||
739 ACCTCCTACACCTGCAAGAGCAACTGGCACAAAGGGCTGGAACCTGGACCTC 788
|||||
3743 ACCTCCTACACCTGCAAGAGCAACTGGCACAAAGGGCTGGAACCTGGACCTC 3792
|||||
789 AG|GGTATAACCAAGTGCCCAAGTGAAGCGCCGCTGCCACCGCTTCGACTTCT 838
|| ser
3793 AG|GGTATAACCAAGTGCCCAAGTGAAGCGCCGCTGCCACCGCTTCGACTTCT 3974
|||||
839 ACTTCCCCAGCCCGCTGCCCTGTGCAACAGATCTGGAGCCACTCCTTT 888
|||||
3975 ACTTCCCCAGCCCGCTGCCCTGTGCAACAGATCTGGAGCCACTCCTTT 4024
|||||
889 GAAGTCAGCAGCTACAGCCGGGCGAGCGCCGCTGCATCCAGATGTGGTT 938
|||||
4025 GAAGTCAGCAGCTACAGCCGGGCGAGCGCCGCTGCATCCAGATGTGGTT 4074
|||||
939 CGACCCGGCCAGGGCAACCCCAACGAGGCGGTGGCGAGATACTATGCAG 988
|||||
4075 CGACCCGGCCAGGGCAACCCCAACGAGGCGGTGGCGAGATACTATGCAG 4124
|||||
989 AGAATGGGGATGCTGGGGCCGTGGCCAGGGGATCGGGCCTCTCCTGACC 1038
|||||
4125 AGAATGGGGATGCTGGGGCCGTGGCCAGGGGATCGGGCCTCTCCTGACC 4174
|||||
1039 AACTTGACGGAGATGGTGAaaCACTGGGTClCCGGCTAAgCTGTTCCCCC 1088
|||||
4175 AACTTGACGGAGATGGTGAaaCACTGGGTClCCGGCTAAgCTGTTCCCCC 4224
|||||
1089 GCCGACCCTGCTTTCGCCCCACACCCCTGGGTACTCTCGGGTGGCC 1138
|||||
4225 GCCGACCCTGCTTTCGCCCCACACCCCTGGGTACTCTCGGGTGGCC 4274
|||||
1139 TCAGCACCCCGGTCATTGGCTCCTGATCTAAGATCCGATGGGGAGCCTCT 1188
|||||
4275 TCAGCACCCCGGTCATTGGCTCCTGATCTAAGATCCGATGGGGAGCCTCT 4324
|||||
1189 GATGGCCTCTTCCAATACAATATCCACGTG 1218
|||||
4325 GATGGCCTCTTCCAATACAATATCCACGTG 4354

```

Fig. 1. (Continued)

mFBP gene were thus obtained and resultant cosmid DNAs were purified and sequenced using automated sequencing (ABI 377, Perkin Elmer, Foster City, CA). The 3' end of the gene for mFBP was obtained by rescreening the YAC subclones using primers specific to the 3' end of the mFBP cDNA. The resulting positive cosmid clone contained only the 3' end of the gene. To obtain the intervening mFBP sequence, primers based on the previously obtained gene sequence and the mFBP cDNA (Table 1) were used to amplify a portion of the missing region using the YAC clone as template. PCR generated DNA was cloned into the PCRII vector according to the directions included with the kit and resultant colonies were screened using the same primers used for amplification. Positive colonies were completely sequenced in both directions. At least three positive colonies for each fragment were sequenced to reduce PCR generated errors. To obtain the remaining 3' portion of the mFBP gene, a BAC library was screened by hybridization using a probe specific to the 3' end of the mFBP gene [5]. A positive mFBP BAC clone was digested with PST1 and EcoR1 and a fragment containing the missing region of the mFBP gene was cloned into PBSIISK. The

```

mFBP cDNA 1 GATGAGGGAGTCCAGGAGTTCCAGCAAGCTCGACCTGCTTAACTCCCA 50
            |
mFBP gene 1 GATGAGGGAGTCCAGGAGTTCCAGCAAGCTCGACCTGCTTAACTCCCA 50
            |
51 GACGGTCACAGGATTTCAG 68
   |
51 GCCGGTCACAGGATTTCAG 68
   |
1  GGATTCCTGCTGCTTTTGACCACAGTCTCTTC 32
   |
3525 ATTCTGCTGCTTTTGACCACAGTCTCTTC 3554
   |
33  TGCAG|GACAAGCATGGCCCTTGGGAGAGCACGGCTGCTGCTCTTGGT 82
   |
3555 TGCAG|GACAAGCATGGCCCTTGGGAGAGCACGGCTGCTGCTCTTGGT 3604
   |
83  GTGTGTGGCTGTACATGGGCGGCCCGGCTGATCTCCTCAACATCTGCA 132
   |
3605 GTGTGTGGCTGTACATGGGCGGCCCGGCTGATCTCCTCAACATCTGCA 3654
   |
133 TGGACGCCAAGCACCACAAGACCAAGCCCGGCCGGAAGATGGCCTGCAT 182
   |
3655 TGGACGCCAAGCACCACAAGACCAAGCCAGGCCCGGAAGATGGCCTGCAT 3704
   |
183 GAGCAG|TGCAGCCCCTGGGAGATGAACGCCTGCTGCTCCGTCAACACCAG 232
   |
3705 GAGCAG|TGCAGCCCCTGGGAGATGAACGCCTGCTGCTCCGTCAACACCAG 5127
   |
233 CCAAGAAGCCATAAACGACATCTCTACCTGTACAAATTCAACTGGGAGC 282
   |
5128 CCAAGAAGCCATAAACGACATCTCTACCTGTACAAATTCAACTGGGAGC 5177
   |
283 ACTGCGGCAAGATGAAGCCGGCCTGCAAGCGCCACTTCATTCAAGACACC 332
   |
5178 ACTGCGGCAAGATGAAGCCGGCCTGCAAGCGCCACTTCATTCAAGACACC 5227
   |
333 TGTCTCTATGAGTGTCTGCCCAACCTGGGGCCCTGGATCCAGGAG|GTGAA 382
   |
5228 TGTCTCTATGAGTGTCTGCCCAACCTGGGGCCCTGGATCCAGGAG|GTGAA 6279
   |
383 CCAGAAGTGGCGCAGAGAGCGGATCCTGAACGTGCCCTCTGCAAAGAGG 432
   |
6280 CCAGAAGTGGCGCAGAGAGCGGATCCTGAACGTGCCCTCTGCAAAGAGG 6329
   |
433 ACTGTCAGAACTGGTGGGAAGACTGCCGCACCTCCTACACCTGCAAGAGC 482
   |
6330 ACTGTCAGAACTGGTGGGAAGACTGCCGCACCTCCTACACCTGCAAGAGC 6379
   |
483 AACTGGCAGGAGGCTGGAAGTGGAGCTCAG|GGTATAACCGTGCCCGC 532
   |
6380 AACTGGCAGGAGGCTGGAAGTGGAGCTCAG|GGTATAACCGTGCCCGC 6573
   |
533 GAACGCCGCTGCCACCCTTCGACTTCTACTTCCACGCTGCTGCC 582

```

Fig. 1. (Continued)

region containing the missing 3' portion of the mFBP gene was then sequenced using automated sequencing.

3. Results

Both the sFBP and mFBP genes were isolated from the same YAC clone, and sequencing of the cosmid containing the 5' end of the mFBP gene resulted in sequence matching the 3' end of the sFBP gene. These results indicate that the two genes must be very close (within ~50 kb of each other), with the sFBP gene 5' to the mFBP gene, a result which is similar to the human FBP/FR locus [9]. The sFBP and mFBP gene sequences obtained are shown

```

|||||
6574 GAACGCCGCTGCCACCCCTTCGACTTCTACTTCCCCACGCCCTGCTGCC 6623
583 TGTGCAGCCAGATCTGGAGCAACTCCTACAAACAAGCAACTACAGCCGG 632
|||||
6624 TGTGCAGCCAGATCTGGAGCAACTCCTACAAACAAGCAACTACAGCCGG 6673
633 GGCAGCGGCCGCTGCATCCAGATGTGGTTCGACCCGGAAACAGGGCAACCC 682
|||||
6674 GGCAGCGGCCGCTGCATCCAGATGTGGTTCGACCCGGAAACAGGGCAACCC 6723
683 CAACGAGGTGGTGGCGAGATACTACGCCAGATCATGAGTGGCGCTGGGC 732
|||||
6724 CAACGAGGTGGTGGCGAGATACTACGCCAGATCATGAGTGGCGCTGGGC 6773
733 TCTCCGAGGCCTGGCCTCTCCAGTTCGGCCTGGCCCTGACGCTGCTCTGG 782
|||||
6774 TCTCCGAGGCCTGGCCTCTCCAGTTCGGCCTGGCCCTGACGCTGCTCTGG 6823
783 CTGCTGAGCTGAGCTTCTGTCTTCGGAGAGCTGGACAGCCCTCCCCTGTT 832
|||||
6824 CTGCTGAGCTGAGCTTCTGTCTTCGGAGAGCTGGACAGCCCTCCCCTGTT 6873
833 CGGCCCCACAGCACCCAGCTCGTCAGTGCCTCAGTGGTGGTGGTGGTGGT 882
|||||
6874 CGGCCCCACAGCACCTAGCTCGTCAGTGCCTCGGTGGTGGTGGTGGTGGT 6923
883 GGTGGT...GGTGGCGCGGGGGGACTCTGAATAAACAGTCACCCC 926
|||||
6924 GGTGGTGGCGCGGTGGCGCGGGGGGACTCTGAATAAACAGTCACCCC 6973

927 AC 928
||
6974 AC 6975

```

Fig. 1. (Continued)

aligned with their corresponding cDNAs in Fig. 1. The genes were 99.2 and 99.1% identical with the previously determined cDNA sequences, respectively. Each gene consisted of 5 exons and 4 introns. The few mismatches that were present are likely to be the result of polymorphisms occurring between animals. Of the differences between the cDNA and gene sequences, most do not result in changes in the coding sequence. One exception is a C to A change (base 913 of the cDNA) which in the sFBP gene codes for an arginine instead of a serine. The sequences obtained for the sFBP and mFBP genes spanned 6.2 and 9.1 kbp, respectively, including 4.4 and 7.0 kbp, respectively, corresponding to each cDNA, and approximately 1.4 and 1.8 kbp, respectively, of the 5' proximal regions for each gene.

Fig. 2 compares the structure of the genes for porcine secreted and membrane folate binding proteins with each other and with the structures of other known FBP/FR genes from other species. The structures of all the FBP/FR genes were very similar for the last three exons of each gene; the sizes and the positions of splice junctions for the last three exons are similar in all genes. Both porcine genes contained multiple copies of a swine SINE [10] repeat element (Fig. 3). Furthermore, comparison of introns 3 and 4 of both genes indicated that both the last two introns contained regions of significant sequence homology between sFBP and mFBP (85–92%), while homology was much less between the porcine and human genes in this region (Fig. 3 and 4). In contrast to the 3' end of the gene, the sizes of the exons in the 5' end of the gene were more variable and sequence homology within the introns between these exons occurred between species but was FBP/FR form specific (see below). Table 2 indicates the sequences found at the splice junctions for sFBP and mFBP. These sequences match the consensus splice donor (NNG gt(a/g)agn) and acceptor (cag NNN)

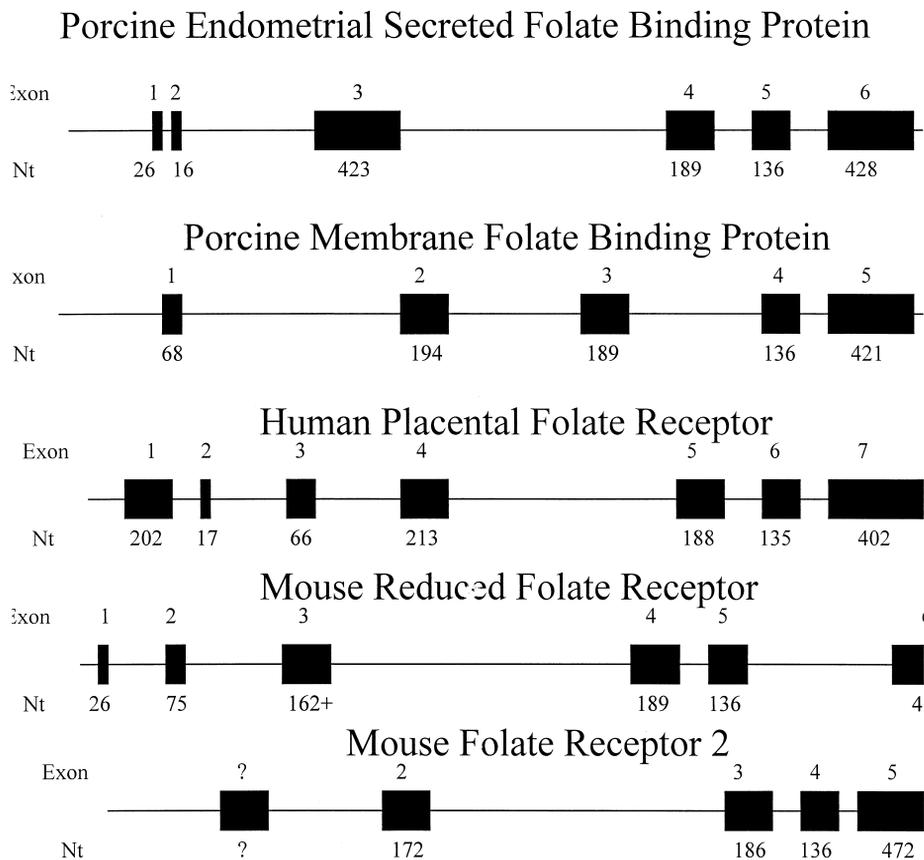


Fig. 2. Schematic diagrams of the porcine sFBP and mFBP genes are illustrated along with the structures of human α -folate receptor [13] and mouse folate receptor 1 [21] and 2 [22]. All gene structures are similar for the last three exons.

sequences that have been obtained for mammals [11]. Finally, the sequence beginning exon 1 of the sFBP gene and exon 2 of the mFBP gene share some homology with the consensus sequence (YAYTCYYY, Y = pyrimidine, sFBP has three mismatches, exon 2 of the mFBP gene has one mismatch) for initiator regions for transcription [12]. The beginning of exon 1 of the mFBP gene has no homology to this sequence.

Significant sequence homology was present in the 5' proximal regions and/or the first introns of the secreted and membrane FBP genes when they were compared with similar regions present in the human secreted (γ) and placental (α) forms of FBP/FR, respectively (Fig. 3, 5, 6). For sFBP, sequence homology was found both 5' and 3' of exon 1 of the sFBP gene, while the sequence of exon 1 itself did not display significant homology (Fig. 5). Curiously, intron 1 of the sFBP gene is homologous to the 5' proximal region of the h γ FBP gene (i.e., the region upstream of the first exon of h γ FBP). A swine SINE repeat element appears to have been inserted into the sFBP gene at or near the region homologous to the transcription start site of h γ FBP, possibly disrupting transcription from this site.

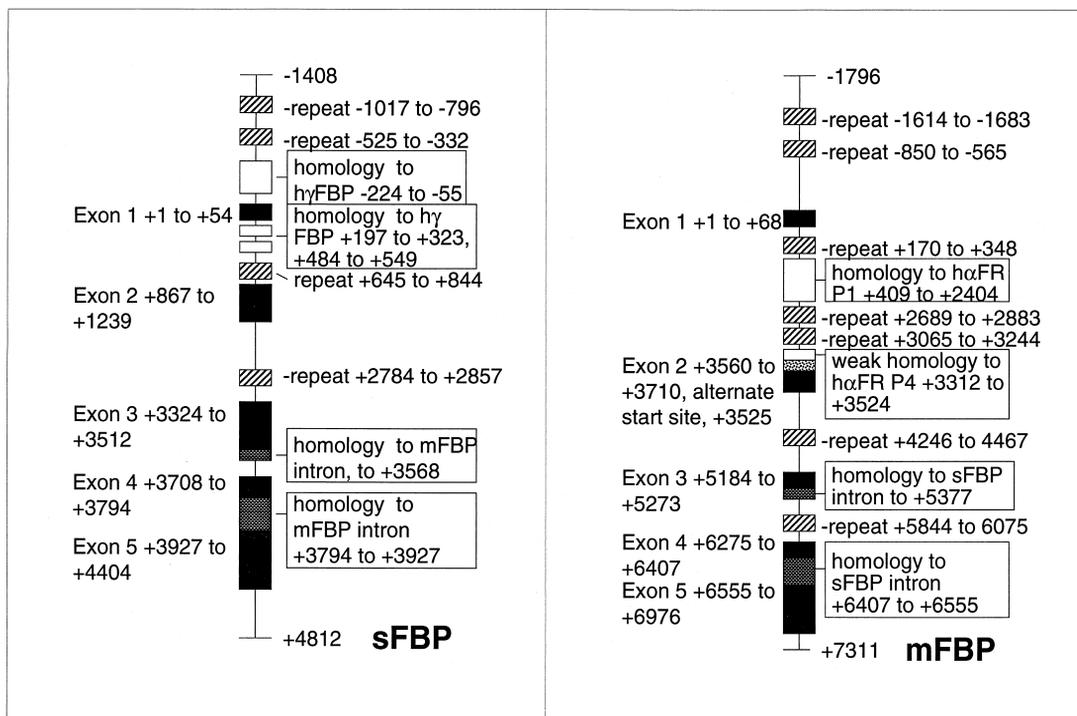


Fig. 3. Schematic diagrams of the sFBP and mFBP genes indicating the positions of exons (black boxes) repeat regions (hatched boxes), regions of homology with human genes (open boxes), and intronic regions of homology between the two porcine genes (dotted boxes).

For the mFBP gene, significant homology with hαFR was present in the first intron of mFBP only, and the region in the hαFR sequence that is homologous to mFBP includes both the P1 and P4 promoters previously identified for that gene [13]. There was no homology to the human αFR sequence found for the region 5' to exon 1 of the mFBP gene. These results suggest that at least two promoter regions are likely present in the mFBP gene, one corresponding to the P4 promoter region in the human gene, and one different from any that has been previously described. A comparison of sFBP and mFBP genes in the regions near the start of transcription showed that no homology between the two genes was present in this region, suggesting that the 5' proximal regions of each gene are specific to each different form of FBP.

4. Discussion

The complete nucleotide sequences for the porcine sFBP and mFBP genes provides an explanation for the different mRNAs for mFBP that have been obtained previously [5] and allows a comparison of these two genes with each other, and with the corresponding genes of humans. These comparisons suggest several hypotheses regarding FBP/FR genes gener-

Table 2

The splice donor and acceptor sites for the secreted and membrane folate binding protein genes are listed for each splice junction

Secreted FBP	Donor sequence	Pyrimidine stretch	Acceptor sequence
Exon 1 to Exon 2	GAG gtaagg	8 bp	cag GAGCCT
Differential splice*	GGG gtaagc	6 bp	cag AAGTGG
Exon 2 to Exon 3	CAG gtgagg	9 bp	cag TGCAGC
Exon 3 to Exon 4	GAG gtatag	10 bp	cag GTGAAC
Exon 4 to Exon 5	CAG gtgagg	8 bp	cag GGTATA
<i>Membrane FBP</i>			
Exon 1 to Exon 2	CAG gtatgg	8 bp	cag GACAAG
Exon 2 to Exon 3	CAG gtgggc	12 bp	cag TGCAGC
Exon 3 to Exon 4	GAG gtacag	10 bp	cag GTGAAC
Exon 4 to Exon 5	CAG gtgagg	8 bp	cag GGTATA

* Splicing variant reported in Vallet et al. (5).

ally, and about the control of each specific gene type. First, the heterogeneous 5' untranslated regions found previously for the mFBP cDNA are the result of initiation of transcription of the gene from two different initiation sites. Second, the remarkably high conservation of regions within the last two introns of the sFBP and mFBP genes suggests that these introns contain sequences that may influence some aspect of the function of both FBP genes. Third, the regions of homology with similar regions in the h α FR and h γ FBP genes near the 5' end of each porcine FBP gene suggests that these regions contain type specific sequences that control the function of each gene type.

It was previously reported that the 5' untranslated regions (UTR) of both the sFBP and mFBP mRNAs were heterogeneous [5]. For sFBP, the heterogeneity appeared to be due to the differential splicing of a region within the sFBP mRNA. The current results indicate that this region is contained within Exon 2 of the sFBP gene, lending support to the concept of differential splicing within Exon 2 versus an alternative exon 2. For mFBP, the heterogeneity in the 5' UTR of the mRNAs appears to be due to initiation of transcription from two different sites within the mFBP gene, which are separated by a large (3456 bp) intron. Sequence homology to both the P1 and P4 promoter regions for h α FR are contained within this intron and these homologous regions appear to be separated from one another by the insertion of two swine SINE repeat elements. One of the two different 5' UTRs obtained for mFBP mRNA appears to correspond well to initiation of transcription from the human P4 promoter (corresponding to FB4 cDNA; 13), despite the fact that this region of the gene was only weakly homologous to the P4 region of the hFR gene. Saikawa et al., [14] reported that the activity of the P4 promoter region was strongly influenced by a cluster of two Sp1 binding sites and a CAC/Sp1 binding site near the transcription initiation site for the KB4 mRNA. Sequence alignment of this region with the mFBP gene indicated that none of these sites are well conserved in the mFBP gene. However, signal scan analysis [15] of this region indicates the presence of a possible Sp1 binding site [16] and a possible CAC binding site [17,18] within the region, both are close to the position of their counterparts in the h α FR gene [14]. Whether these regions or other regions influence the rate of transcription from this site requires further study.

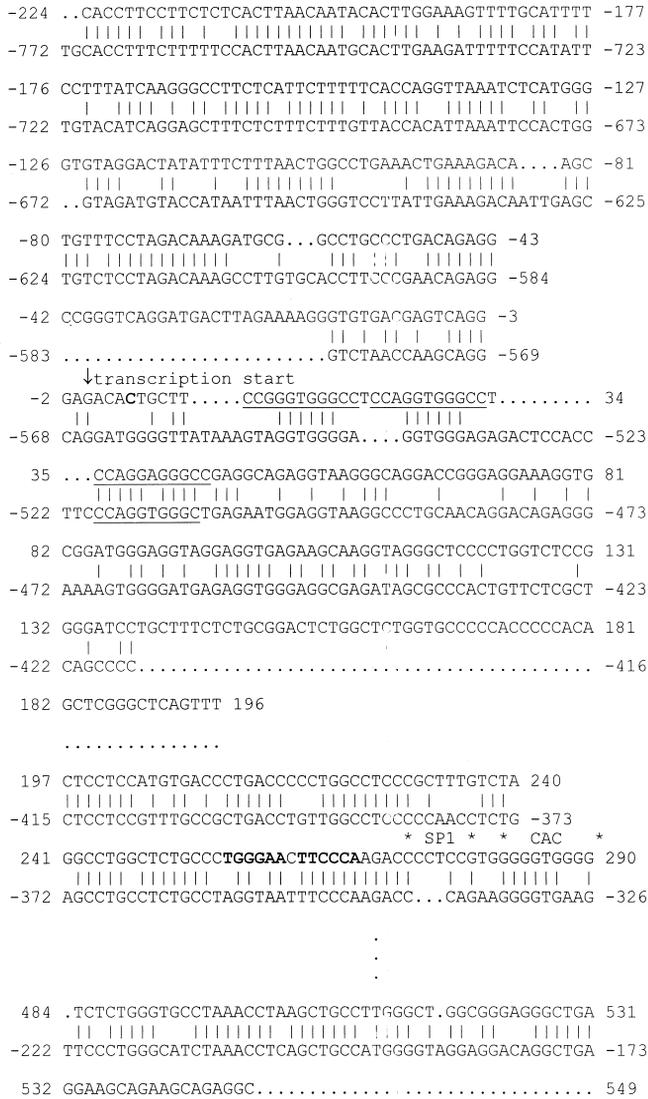


Fig. 5. The 5' regions of the sFBP and human γ FBP genes are aligned to indicate homologies between the two genes. Repeat sequences are underlined, palindromic sequences are in bold letters. Transcription start sites for each gene are indicated.

active in the endometrium of the pig. The mRNA corresponding to initiation of transcription from this promoter region may occur in other tissues. Elwood et al., [13] reported that the activity of this promoter is specific to kidney and cerebellum. Resolution of this question will require the cloning and sequencing of mFBP transcripts from a variety of tissues, to determine if the region corresponding to the P1 promoter is active in other tissues in swine. Surprisingly, Exon 1 of the mFBP gene is found 5' to the regions showing homology to the P1 and P4 h α FR promoter regions, suggesting that a third promoter region may lie in the


```

1671 CCAGGACCCCTTGGCTTTTCAGCACTGCACTGAGCTCTCAGTGCTTGAAGTC 1720
      ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
402 CCTGGGCTCTAGCCCTTCAGTCCAGAGCTGAGTTCTCAGCTCTTCTAGTC 451

1721 TGGGACCCGGGGGAATGTGCGTGTGTGTGCTGAGCGTTGG.GGCGGGG 1769
      ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
452 TGGGCCCCAAGG..TTGGGTGTGGGGTCATGATTGTTGGTGGGGAGGG 499

1770 GTCATAGCTGGACAAGGAAGGCAGGGAATGAC..... 1801
      ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
500 GTCACAGCTGGACTAAGACCTGAAGGTGAGACTAGGCAGGTGGGAAAGGA 549

1802 .....ATGCAGGCACAGATGCAGTGGAGGGGAGGAGAAAGCTCTAGA 1843
      ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
550 GCTTGACAGAGTGATGCTGCTCAAAAGGCAGGAAGAGAGCCTGGCTTCAG 599

1844 AAGCAGCCACAGCAAGGAAAACCACTGA.....GACCCGCGTGT 1882
      ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
600 AAGCAGCCACAGCAAGAGAGACTACTGACTGAACAGGTGGGCTCCACTGG 649

1883 G.....CTCCCCAGGGCTCTGGGCGG 1904
      |
650 GGGCTCCGAAAGGATTTTCTCAGCCCCATCCCCAGCACTGTGTGTTGG 699
      ↑transcription start

1905 TCACCCCTGTGAGAGCCTCCGCACAGCCAAGGTGCAGGGGACCGAAGGTG 1954
      ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
700 CCGCACCCATGAGAGCCTCAGCACTCTGAAGTGCAGGGGGCAAAAGGCCA 749

1955 AAGGAGCTCTGGCC.....GTGAGGGTCCCAAGTCCGACTTGGGCGG 1996
      ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
750 AAAGAGCTCTGGCCTGAACTTGGGTGGTCCCTACTGTGTGACTTGGGGCA 799

1997 CCCCCGCATGTGACCTGGGGCACGTCCCTCAACCCGGGCTGGTTGGCC 2046
      ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
800 TGGCCCTCATCTGTGCTGAAATGATTCCACAAGATTAAACTGGCTATCA 849

2047 AGCATTGCTGGGGGTTTCCCCTACATTTAATCCTCAGAAGCGGAAGCT 2096
      ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
850 TTTGTTGATTTCCCCTT...CTTACATTTAATCCTTGCAGGAGAAAGCT 896

2097 AAGCCTCAGGATGCTGTGAGTTCTTTTCCCCAAGGTCAAGCAGGAGGGA 2146
      ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
897 AAGCCTCAAGATAGTTTGTCTTCTTTCCCCAAGGCCAAGGAGAAAGGTG 946

2147 GAGTGTGAGGGCTGAGATCGCAAACGGGCGGATCTGGTACCGGCTGCTCT 2196
      ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
947 GA..GTGAGGGCTGGGGTCG.GGACAGGTGAACGGGAACCCCTGTGCTCT 993

2197 AAACCTGCTAGACTTGGTTTCCCCTAAGGAGCCTCATCTTGCATCACCTGG 2246
      ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
994 AAACAGTTAGGGTTTGTTCCTCCGCAAGTGAACCAAGGATCACCTGG 1043

2247 CATTGCC.GGGAGGACAGGCTTGT..GCTGTGCGCTGGAGATCAGTGAG 2293
      ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
1044 TATTCCCTGAGAGTACAGATTTCTCCGGCGTGGCCCTCAAGGTTAGTGAG 1093
      ↑transcription start

```

Fig. 6. (Continued)

gene [14]. Finally, there is a short region of homology between the sFBP and h γ FBP genes in the region 5' of exon 1. Further upstream are two repeat regions. Using signal scan analysis, there are very few potential binding factor sites within this upstream homologous region. The actual role of each of these regions in the control of transcription and/or translation awaits further work.

The high homology between regions in the 3rd and 4th intron of the two porcine FBP/FR genes suggests that these regions contain elements that control transcription or mRNA processing. Within the homologous region of the third exon, there are two completely conserved partial palindromic sites. Also, this region of the sFBP gene contains four

are similar to the other known FBP/FR genes. The sequence of the 5' proximal region and intron 1 of the gene for the secreted form of FBP is homologous to the putative promoter region of the human γ FBP gene, suggesting the possibility that this region may contain elements that control transcription of this gene. Likewise, similar homologous regions exist between the mFBP gene and human α FR. This homology coupled with the divergence in the 5' UTR of the mFBP gene strongly suggest that at least two promoter regions are present within this gene. Finally, highly conserved sequence homology within the third and fourth introns of both porcine genes suggests the possible presence of controlling elements within these introns.

Acknowledgments

Names are necessary to report factually on available data; however, the USDA neither guarantees nor warrants the standard of the product, and the use of the name by USDA implies no approval of the product to the exclusion of others that may also be suitable.

References

- [1] Shane B, Stokstad EL. Vitamin B12-folate interrelationships. *Annu Rev Nutr* 1985;5:115–41.
- [2] Babior BM. Erythrocyte disorders: anemias related to disturbance of DNA synthesis (megaloblastic anemias). In: Hematology, 4th ed. Williams WJ, Beutler E, Erslev AJ and Lichtman MA (eds.) McGraw-Hill, Inc., New York, NY, p. 453–81, 1990.
- [3] Thenen SW. Gestational and neonatal folate deficiency in rats. *Nutr Res* 1991;11:105–16.
- [4] Vallet JL, Christenson RK, Klemcke HG. Purification and characterization of intrauterine folate binding proteins from swine. *Biol Reprod* 1998;59:176–81.
- [5] Vallet JL, Smith TPL, Sonstegard T, Pearson PL, Christenson RK, Klemcke HG. Isolation of complementary deoxyribonucleic acids encoding putative secreted membrane-bound folate binding proteins from endometrium of swine. *Biol Reprod* 1999;61:372–9.
- [6] Vallet JL, Christenson RK, Klemcke HG. Technical Note: A radioimmunoassay for porcine intrauterine folate binding protein. *J Anim Sci* 1999;77:1236–40.
- [7] Van Hoozen CM, Ling E-H, Halsted CH. Folate binding protein: molecular characterization and transcript distribution in pig liver, kidney and jejunum. *Biochem J* 1996;319:725–9.
- [8] Moran P, Raab H, Kohr WJ, Caras IW. Glycophospholipid membrane anchor attachment, molecular analysis of the cleavage/attachment site. *J Biol Chem* 1991;266:1250–7.
- [9] Ragoussis J, Senger G, Trowsdale J, Campbell IG. Genomic organization of the human folate receptor genes on chromosome 11q13. *Genomics* 1992;14:423–30.
- [10] Singer DS, Parent LJ, Ehrlich R. Identification and DNA sequence of an interspersed repetitive DNA element in the genome of the miniature swine. *Nucleic Acids Res* 1987;15:2780.
- [11] Burge CB, Tuschl T, Sharp PA. Splicing of precursors to mRNAs by the spliceosomes. In: *The RNA World*, 2nd ed. Gesteland RF, Cech TR and Atkins JF (eds). Cold Spring Harbor Laboratory Press, Cold Spring Harbor, NY, p. 525–560, 1999.
- [12] Roeder RG. The complexities of eukaryotic transcription initiation: regulation of preinitiation complex assembly. *Trends Biol Sci* 1991;16:402–11.
- [13] Elwood PC, Nachmanoff K, Saikawa Y, Page ST, Pacheco P, Roberts S, Chung K-N. The divergent 5' termini of the α human folate receptor (hFR) mRNAs originate from two tissue-specific promoters and alternative splicing: Characterization of the α hFR gene structure. *Biochemistry* 1997;36:1467–78.

- [14] Saikawa Y, Price K, Hance KW, Chen TY, Elwood PC. Structural and functional analysis of the human KB cell folate receptor gene P4 promoter: cooperation of three clustered Sp1-binding sites with initiator region for basal promoter activity. *Biochemistry* 1995;34:9951–61.
- [15] Prestridge DS. Signal scan: A computer program that scans DNA sequences for eukaryotic transcriptional elements. *Comput Appl Biosci* 1991;7:203–6.
- [16] Kadonga JT, Jones KA, Tjian R. Promoter-specific activation of RNA polymerase II transcription by Sp1. *Trends Biol Sci* 1986;11:20–3.
- [17] Schule R, Muller M, Otsuka-Murakami H, Renkawitz. Cooperactivity of the glucocorticoid receptor and the CACCC-box binding factor. *Nature* 1988;332:87–90.
- [18] Yu C.-Y, Motamed K, Chen J, Bailey AD, Shen C-KJ. The CACC box upstream of human embryonic ϵ globin gene binds Sp1 and is a functional promoter element in vitro and in vivo. *J Biol Chem* 1991;266:8907–15.
- [19] Savagner P, Miyashita T, Yamada Y. Two silencers regulate the tissue-specific expression of the collagen II gene. *J Biol Chem* 1990;265:6669–74.
- [20] Wang H, Ross JF, Ratnam M. Structure and regulation of a polymorphic gene encoding folate receptor type γ/γ' . *Nuc Acids Res* 1998;26:2132–42.
- [21] Tolner B, Roy K, Sirotinak FM. Organization, structure and alternate splicing of the murine RFC-1 gene encoding a folate transporter. *Gene* 1997;189:1–7.
- [22] Brigle KE, Seither RL, Westin EH, Goldman ID. Increased expression and genomic organization of a folate-binding protein homologous to the human placental isoform in L1210 murine leukemia cell lines with a defective reduced folate carrier. *J Biol Chem* 1994;269:4267–72.