



Since January 2020 Elsevier has created a COVID-19 resource centre with free information in English and Mandarin on the novel coronavirus COVID-19. The COVID-19 resource centre is hosted on Elsevier Connect, the company's public news and information website.

Elsevier hereby grants permission to make all its COVID-19-related research that is available on the COVID-19 resource centre - including this research content - immediately available in PubMed Central and other publicly funded repositories, such as the WHO COVID database with rights for unrestricted research re-use and analyses in any form or by any means with acknowledgement of the original source. These permissions are granted for free by Elsevier for as long as the COVID-19 resource centre remains active.

GENE 07066

Expression of the *Bgp* gene and characterization of mouse colon biliary glycoprotein isoforms

(Carcinoembryonic antigen gene family member; reverse transcription and polymerase chain amplification; mouse hepatitis virus receptor; adhesion molecule)

Kimberly McCuaig^a, Madelaine Rosenberg^{a,b}, Patrick Nédellec^{a,c}, Claire Turbide^a and Nicole Beauchemin^{a,b,c}

^aMcGill Cancer Centre, ^bDepartment of Biochemistry, and ^cDepartment of Medicine, McGill University, Montreal, Québec, Canada

Received by: J.A. Engler: 9 October 1992; Revised/Accepted: 7 December/8 December 1992; Received at publishers: 1 February 1993

SUMMARY

The biliary glycoprotein (BGP)-encoding gene is a member of the human carcinoembryonic antigen (CEA) gene family. We have now cloned several mouse *Bgp* cDNAs from an outbred CD^R-1 mouse colon cDNA library, as well as by reverse transcription-PCR amplification of colon RNA. The distinguishing features of the deduced *Bgp* protein isoforms are found in the two divergent N-terminal domains, the highly conserved internal C2-set immunoglobulin domains, and an intracytoplasmic domain of either 10 or 73 amino acids (aa). The cDNA structures suggest that these mRNAs are produced through alternative splicing of a *Bgp* gene and the usage of multiple transcriptional terminators. The *Bgp* deduced aa sequences are highly homologous to several well characterized rat hepatocyte proteins such as the cell CAM105/ecto-ATPase/pp120/HA4 proteins. Oligodeoxyribonucleotide probes representing the various cDNA isoform domains revealed predominant transcripts of 1.8, 3.1 and 4.0 kb on Northern analyses of mouse colon RNA; some of these bands are actually composed of several co-migrating transcripts. The transcripts encoding the long intracytoplasmic-tailed *Bgp* proteins are expressed at one-tenth the relative abundance of the shorter-tailed species. We have previously demonstrated that several mouse *Bgp* cDNAs, when transfected into eukaryotic cells, express BGP proteins at the cell surface and function in vitro as cell adhesion molecules, much like their human and rat counterparts. The expression of the many *Bgp* isoforms at the surface of epithelial cells, such as colon, suggests that these proteins play a determinant role, through self- or heterologous contact, in renewal and/or differentiation of their epithelia.

INTRODUCTION

Carcinoembryonic antigen (CEA); (Gold and Freedman, 1965) is a human tumor marker used to evalu-

ate recurrences of gastrointestinal, breast and lung cancers (Shuster et al., 1980). The *CEA* gene family, composed of at least 22 genes clustered on human chromosome 19q13.1-q13.2, is divided into two subgroups

Correspondence to: Dr. N. Beauchemin, McGill Cancer Centre, McGill University, 3655 Drummond St., Montreal, H3G 1Y6, Québec, Canada. Tel. (514) 398-3541; Fax (514) 398-6769.

Abbreviations: aa, amino acid(s); BGP, biliary glycoprotein; *Bgp*, mouse biliary glycoprotein; *BgpX*, gene isoforms encoding mouse *Bgps* (replaces mmCGM to conform with mouse genome nomenclature; X is assigned by order of characterization); bp, base pair(s); CD, complement determining; CEA, carcinoembryonic antigen; CGM, CEA-related

gene family member; Cyt, intracytoplasmic tail; ICAM-1, intercellular adhesion molecule 1; Ig, immunoglobulin; kb, kilobase(s) or 1000 bp; MHV, mouse hepatitis virus; MHVR, MHV receptor; NCA, non-specific cross-reacting antigen; nt, nucleotide(s); oligo, oligodeoxyribonucleotide; ORF, open reading frame; PCR, polymerase chain reaction; PSG, pregnancy-specific glycoprotein; 5' or 3'UTR, 5' or 3' untranslated region; RIT, oligo specific for Cyt in the antisense orientation; RT, reverse transcription; TM, transmembrane.

based on sequence comparisons and expression patterns (Thompson et al., 1992; Khan et al., 1992): the CEA subgroup [*CEA*, *NCA*, *BGP*, *CGM6* and a number of genes with as yet undefined gene products (designated *CGMs*; Thompson et al., 1992)] and the pregnancy-specific (*PSG*) subgroup (Khan et al., 1992). The *CEA*-related gene products exhibit structural features of the immunoglobulin (Ig) supergene family with N-terminal domains resembling variable regions and internal domains homologous to C2-set constant domains (Williams and Barclay, 1988). *BGPs* (Svenberg, 1976; Svenberg et al., 1979; Hinoda et al., 1988; Barnett et al., 1989) are unique in this family in that all isoforms, produced through alternative splicing of a single gene (Barnett et al., 1989), bear either short (10 aa) or long (71 aa) intracytoplasmic tails (Cyt) (Hinoda et al., 1988; Barnett et al., 1989). In contrast, *CEA*, *NCA* and *CGM6* are linked to the membrane by a glycopospholipid anchor (Hefta et al., 1988; Kolbinger et al., 1989; Berling et al., 1990).

CEA, *NCA* and *BGP* are postulated to participate in intestinal tissue organization during embryonic development and function in vitro as intercellular adhesion molecules (Benchimol et al., 1989; Oikawa et al., 1989; Oikawa et al., 1992; Rojas et al., 1990; Zhou et al., 1990). The *CEA* and *NCA* glycoproteins are also involved in the bacterial recognition and in colonization (Leusch et al., 1991). *PSGs* are postulated to act as immunomodulators during pregnancy (Cerni et al., 1977).

To pursue functional investigations, we have begun characterization of the mouse *CEA*-related gene family. In previous work, we have shown that mouse *Bgp* proteins (the cDNAs had previously been called mmCGM1a and mmCGM2 and have been renamed *BgpA* and *BgpB*), expressed at the surface of transfectant cells, function in vitro as adhesion molecules (Turbide et al., 1991; McCuaig et al., 1992). Moreover, Dveksler et al. (1991) have demonstrated that one of these mouse *Bgp* proteins (MHVR with identical coding sequence to the deduced aa sequence of *BgpA*) acts as the mouse hepatitis virus (MHV) receptor (Dveksler et al., 1991).

Contrary to human *BGP* isoforms which contain the same Ig-like N-terminal domain (Barnett et al., 1989), the previously reported mouse *Bgp*-like cDNAs each encode a distinct N-terminal domain, and internal C2-set domains that are conserved but not identical (Turbide et al., 1991; McCuaig et al., 1992). This finding implied that either a *Bgp* gene subfamily exists, or that complex transcriptional mechanisms are operative to account for these *Bgp* cDNAs. It therefore became important to define the number of *Bgp* isoforms and to study their expression in mouse tissues for future functional studies. The data presented in this paper describes the characterization, the relative abundance and the differential expres-

sion of nine *Bgp* cDNA isoforms expressed in colon. The results suggest that a unique mouse *Bgp* gene undergoes alternative splicing, and that different polyadenylation signals are used to produce a number of related proteins with possible complementary functions. However, the two classes of related *Bgp* cDNAs cloned from the outbred CD-1 mouse may represent allelic variants of this *Bgp* gene (Dveksler et al., 1993).

RESULTS AND DISCUSSION

(a) Library cDNA clones

To characterize the mouse *CEA*-related gene family, we screened a CD-1 mouse colon cDNA library (Beauchemin et al., 1989). This yielded a significant number of cDNA clones (i.e., 392), some of which presented different restriction patterns or different cDNA sequences. To verify if these cDNA clones were representative of other mouse *CEA*-related genes, clones 23, 32, 37, 58 and 64 were completely sequenced while clone 132 was partially sequenced. As shown in Fig. 1, some of these clones (clones 23 and 132) overlapped perfectly within sequences encoding either the signal sequence, the N terminal or the A2 domain of the *BgpA* cDNA (formerly mmCGM1a) (McCuaig et al., 1992), while others (clones 23, 32, 37, 58 and 64) were aligned with DNA coding for either the leader, the N-terminal, the A2, the TM, the Cyt domains or the 3'UTR sequence of another published cDNA clone, *BgpB* (formerly mmCGM2) (Turbide et al., 1991). Clone 23 extended the published *BgpB* cDNA 3'UTR (Turbide et al., 1991) by 119 nt and ended in a poly(A) tail. Another cDNA clone (clone 58) encoded a longer Cyt (73 aa) than the previously reported 10 aa Cyt domain (Turbide et al., 1991; McCuaig et al., 1992); this new mouse Cyt domain, however, resembled closely the cDNA sequence of human *BGP* (Barnett et al., 1989) and of a rat ecto-ATPase cDNA (Lin and Guidotti, 1989). Clones 32 and 64 also encoded conserved but not identical A1 and B1 domains when compared to similar domains of the *BgpA* cDNA. cDNA clone 37 demonstrated high homology to the reported 3'UTR of the MHVR isolated from Balb/c mice (Dveksler et al., 1991), but extended this sequence by 199 nt to another poly(A) tail.

(b) RT-PCR amplification

The information from the cDNA sequences and computer-translated aa sequences compiled from the cDNA clones were indicative of multiple *Bgp* cDNA isoforms present in mouse colon. We therefore conducted a series of experiments using a RT-PCR amplification technology (Frohman et al., 1988) to identify the type of *Bgp* isoforms

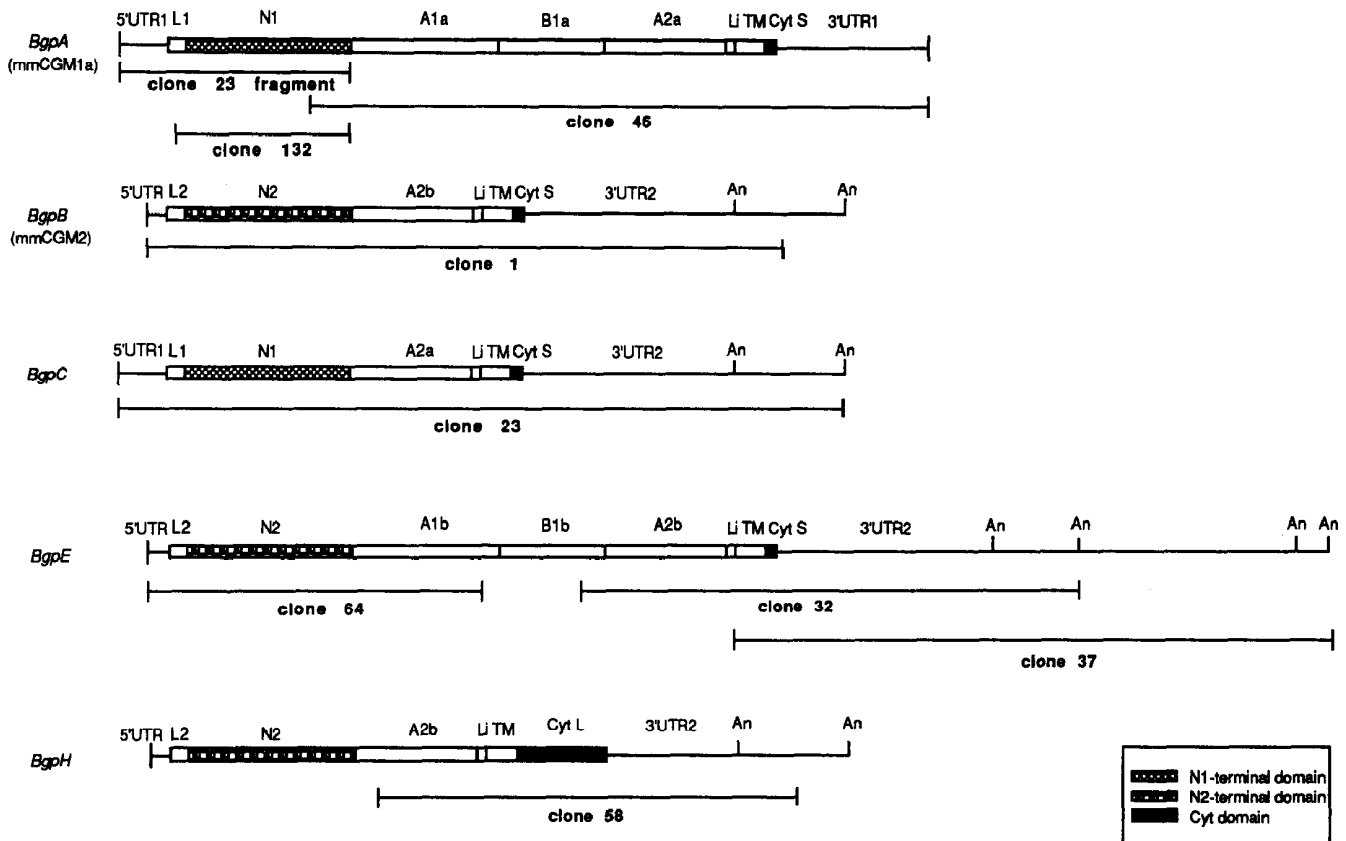


Fig. 1. Topology of library cDNA clones. *5'UTR1* and *5'UTR*, 5'-untranslated regions; L1 and L2, signal sequences; N1 and N2, N-terminal domains; A1a, A1b, B1a, B1b, A2a, A2b, C2-type Ig domains; Li, linker; TM, transmembrane domain; CytS, putative short Cyt domain; CytL, long Cyt domain; *3'UTR1* and *3'UTR2*, 3' untranslated regions; An, positions of poly(A) tails. **Methods:** cDNA was synthesized as previously described (Beauchemin et al., 1989). The CD-1 mouse colon cDNA library was screened with cDNA restriction fragments corresponding to the *EcoRI-SstI* and *AccI-EcoRI* restriction fragment of clone 46 (McCuaig et al., 1992) and with the oligo RIT (antisense: 5'-TGAGGGTTTGTGCTCTGTGAGATC) representing a region of the long Cyt conserved between human BGP (Barnett et al., 1989) and rat ecto-ATPase (Lin and Guidotti, 1989). *Bgp* cDNA restriction fragments were subcloned into unique sites of the BluescriptSK+ plasmid (Stratagene, La Jolla, CA) and overlapping DNA restriction fragments were sequenced on both strands with either the T7 or the T3 promoter primers or with internal primers by the dideoxy chain termination method (Sanger et al., 1977) using T7 DNA polymerase (Pharmacia, Montreal, Canada). Sequences were analyzed using the DNAsis, Prosis (Pharmacia) and the GCG sequence analysis system (Devereux et al., 1984) programs.

present in this tissue. Total RNA from CD-1 mouse colon was reverse-transcribed using either a $(dT)_{17}$ adaptor primer or KM5, a primer located in the *3'UTR* of *BgpB* and *BgpD* (see Fig. 3b). PCR amplifications were then performed on single-stranded cDNAs with combinations of primers located in the coding and non-coding regions and the products were cloned and sequenced (Fig. 2a). Since we had observed that major sequence differences were found between the N-terminal domains (designated N1 and N2; Turbide et al., 1991; McCuaig et al., 1992), divergent primers within these two N-terminal domains (46N1, R46N1, CGM2N and RCGM2N) were synthesized, tested at medium and high hybridization stringency and demonstrated non-cross-reactivity (data not shown). The coding sequence of the two A1 domains (A1a, A1b), B1 domains (B1a, B1b) and the two A2 domains (A2a, A2b) differed by only 4, 1 and 3 nt, respectively, (indicated in Fig. 3b). Common primers specific for each of these domains (KM7 and KM8 for A1 and 33–35 for A2) as

well as a primer specific to the region encoding the long Cyt (RIT) were synthesized.

The use of primers specific to the *5'UTR* of *BgpA* (KM2), and to the N-terminal domains (46N1 or RCGM2N) resulted in the amplification of a 360-bp fragment only when KM2 was combined with a primer from the N1-terminal domain (46N1). This result indicates that the *5'UTR* of the transcripts containing the N2 domain is different in the region of KM2 (Fig. 2a). Varying the Mg^{2+} concentrations of the amplification reaction for the KM2-RCGM2N primer pair did not produce a fragment. The identity of the KM2–46N1 fragment was confirmed by sequencing of the fragment.

The sequences of the cDNA clones suggested that splicing events occurred on the primary *Bgp* transcript(s), resulting in the deletion of the A1 and B1 domains (McCuaig et al., 1992), or in the inclusion of an exon to generate a long Cyt protein. To ascertain the numerous splicing possibilities downstream from the N-terminal

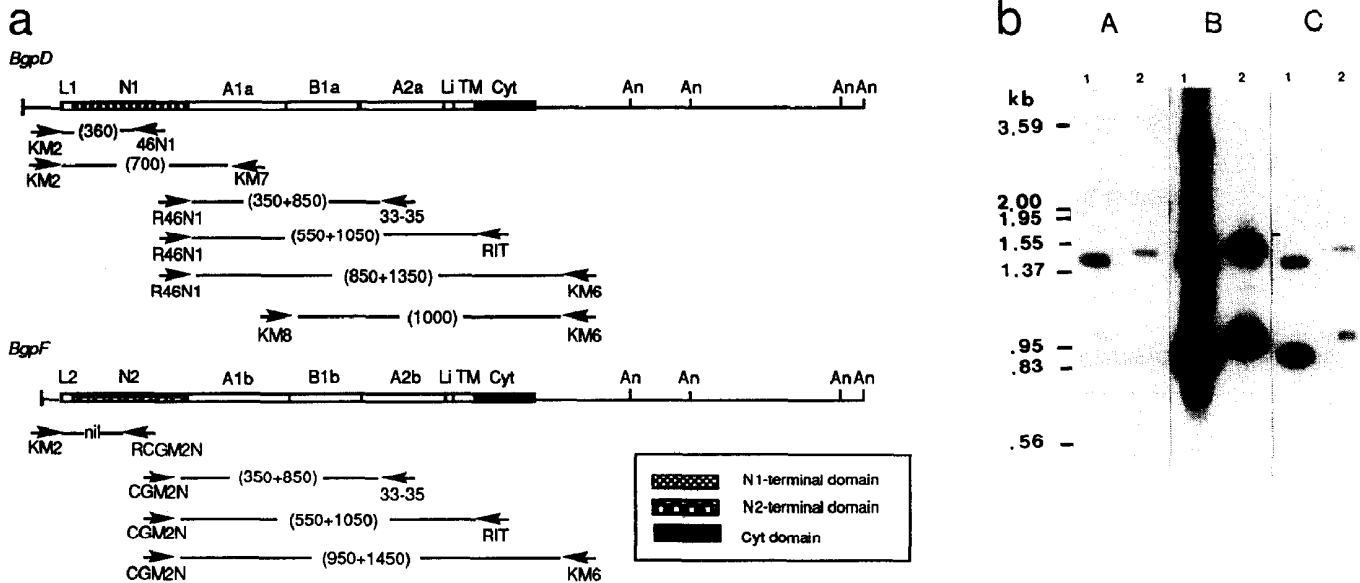


Fig. 2. PCR amplification reactions and hybridizations of amplified fragments. (a) PCR amplification reactions. The names and positions of the various oligo primers used in the PCR amplification reactions are indicated. The numbers in parentheses represent the approximate length of the fragments obtained from each reaction. (b) Hybridization of PCR fragments. PCR fragments were obtained from amplifications with oligos R46N1-KM6 (lane 1) and CGM2N-KM6 (lane 2). They were fractionated on a 1% agarose gel and hybridized with KM7 [³²P]oligo (panel A), with oligo 33-35 (B) and with oligo RIT (C). **Methods:** Total RNA from CD-1 mouse colon was prepared by guanidium isothiocyanate extraction and centrifugation (Beauchemin et al., 1989). RT was performed using AMV reverse transcriptase (Pharmacia, Montreal, Canada) and 10 μg of total mouse colon RNA as template essentially as described previously (McCuaig et al., 1992). The oligo primers used in this reaction were either a (dT)₁₇ containing restriction sites (5'-GACTCGAGTCGACGGTACCCT₁₇) or oligo KM5 (antisense: 5'-TTGATACCTCACTCTCAGCCA). The amplification reactions were incubated at an annealing temperature of 44°C for 2 min, at an elongation temperature of 72°C for 3 min and at a denaturing temperature of 94°C for 1 min for 40 cycles in a total volume of 100 μl containing 20 mM Tris-Cl pH 8.8 (at 24°C)/10 mM KCl/2 mM MgSO₄/10 mM ammonium sulfate/0.1% Triton X-100/0.1 μg BSA/0.2 mM of each dNTP/40 pmol of phosphorylated primers, using 2 units of Vent[®] DNA polymerase (New England Biolabs), to decrease possible proofreading errors (Frohman et al., 1988; Mattila et al., 1991). The oligo primers used were: KM2: sense 5'-CCAAGTCCCGACAAGTAGTG; RCGM2N: antisense 5'-GTCTTATTAGTGCCTGTTAC; CGM2N: sense 5'-GTAACAGGC-ACTAATAAGAC; 46N1: antisense 5'-CTTCATGGTGATCATTGG; R46N1: sense 5'-CCAAATGATCACCATGAAG; KM7: antisense 5'-GGGTCACCTTCGGTTGACACT; KM8: sense 5'-CCAGTGAGTGCAACCGAAG; 33-35: antisense 5'-CCGGCATCTTCCCTCTTAATAGGGTCTATTCTG; RIT: antisense 5'-TGAGGGTTTGTGCTCTGTGAGATC; KM6: antisense 5'-GGCTCCAGGATCCACCTTTTCTTC. The PCR amplification products were blunt-ended (Sambrook et al., 1989) with T4 DNA polymerase (Boehringer-Mannheim) to degrade single-stranded DNA resulting from non-symmetrical amplification (Innis et al., 1990), separated by electrophoresis and cloned into the *Sma*I site of the BluescriptSK + plasmid (Stratagene, La Jolla, CA, USA) for sequencing. A minimum of three independent clones were sequenced for each fragment resulting from every amplification reaction with different pairs of primers. Control samples omitting reverse transcription were included to verify that the RNA was not contaminated by genomic DNA. These fragments were also electrophoresed through 1% agarose gels and transferred to GeneScreen Plus membranes (NEN Dupont). Hybridizations were carried out as previously described (McCuaig et al., 1992) with 2 × 10⁶ cpm/ml of [³²P]oligos (Sambrook et al., 1989). The filters were washed to a final stringency corresponding to 5°C below the predicted melting temperature of each oligo.

domains, amplifications were carried out between N-terminal primers (R46N1 or CGM2N) and oligos corresponding to the A2 domain (33-35), the long Cyt (RIT) or one of the 3'UTRs (KM6). As is summarized in Fig. 2a, amplification between primers R46N1 and 33-35 or CGM2N and 33-35 resulted in the synthesis of two different-sized fragments in each case. Cloning and sequencing revealed that the shorter-sized fragments were derived from cDNAs where the N-terminal coding domains, N1 or N2, were followed by their respective A2 domains (N1 domain with A2a, N2 domain with A2b). None of the sequenced clones spliced either the N1 domain to the A2b domain, or the N2 domain to the A2a domain. Sequencing of the longer fragments cloned from each amplification demonstrated similar specificity; the

N1 domain only associated with the A1a and B1a domains followed by the A2a domain while the N2 domain was joined to the A1b, B1b and A2b domains.

Similarly, amplifications using the N-terminal domain and the long Cyt-specific primers produced two DNA fragments. Sequencing confirmed that the shortest ones carried a deletion of the A1 and B1 domains, while the longer fragments included these two domains. Amplifications conducted with primers located in the N-terminal coding domains and in the 3'UTR (KM6) confirmed the previous results. A greater number of clones (seven) representing these fragments were sequenced, one of which was shown to code for a long Cyt domain. Results from these amplifications were confirmed by Southern analyses. As can be seen in Fig. 2b,

the 1350 and 1450-bp fragments amplified by either pairs (R46N1-KM6 or CGM2N-KM6) hybridized to an oligo specific to the A1 domains (KM7) (Fig. 2b, panel A), while all the fragments hybridized to the 33–35 oligo defining the A2 domain (Fig. 2b, panel B). The bands hybridizing to oligo RIT (Fig. 2b, panel C) confirmed that the short-tailed cDNA isoforms have a counterpart encoding a long Cyt. These hybridizations also revealed a new 3.0-kb DNA fragment specifically amplified in primer combinations containing the N1 primer. This novel cDNA fragment will be the subject of future reports. This fragment is not due to contamination with genomic DNA since it was not present in PCR amplifications when the RT reaction was omitted (data not shown).

(c) The nt sequences of the cDNAs and the aa sequences

The structure of the deduced isoforms are depicted in Fig. 3a. A minimum of nine cDNA isoforms are generated from the mouse *Bgp* gene(s). Two of these cDNA isoforms (*BgpA* and *MHVR*) have an identical coding region but somewhat different 5' and 3'UTRs (the first 41 nt upstream from the ATG start codon are identical) (Dveksler et al., 1991; McCuaig et al., 1992). As was demonstrated by Southern analysis (Fig. 2b, panel C) and sequencing of the PCR products, cDNAs which lack a region encoding two C2-set domains such as *BgpB* and *BgpC* (clones 1 and 23 respectively) can also encode a long Cyt to generate the *BgpG* and *BgpH* cDNA isoforms. Likewise, the Southern and sequencing analyses indicate that the regions encoding the long Cyt as well as three internal repeats are present in the two longest cDNA isoforms designated *BgpD* and *BgpF* (Fig. 3b).

The *Bgp* isoforms exhibit two variants of the signal sequence: four-nt substitutions (Fig. 3b) lead to change in two aa (Fig. 3c). The leader region is followed by an N-terminal domain. There are two possible N-terminal domains (N1 and N2) which exhibit both common and divergent areas: the nt sequences encoding the first 37 aa are identical except for one nt substitution, while the last two thirds of these domains exhibit numerous nt differences, many of which are located in the third position of a codon (Fig. 3b and 3c). As has been noted in the human *CEA* gene (Schrewe et al., 1990), the last nt of the N-terminal domain exon is paired with the first two nt of the following exon to form the first codon of the next domain. Because of this splicing position, different aa are encoded by the various *Bgp* isoforms, depending on whether the N-terminal domain is followed by the A1 or the A2 domain. If there is an A1 domain, the first aa in A1 is Pro (*BgpA/D/E/F*) and in A2, Glu. If there is no A1 domain, the first aa in A2 is a Gln (*BgpB/C/G/H*) (Fig. 3c). There are two variants for each of the internal C2-set Ig domains; however, there are few nt changes

(Fig. 3b) and those that do occur lead to aa substitutions (Fig. 3c), and in one case, a silent mutation (indicated by arrow in Fig. 3c). A linker region (Li) joining the A2 domain to the TM region is predicted by hydropathy computer analysis.

Two Cyt domains are derived from essentially the same sequence of DNA. As reported for human BGP (Barnett et al., 1989), the mouse long Cyt is generated by the inclusion of a 53 bp exon (nts 1464–1517 in Fig. 3b) which shifts the ORF from what would otherwise encode a short cytoplasmic tail. The translated proteins with a short Cyt are produced by using the first in frame stop codon (nt 1533–1535) while proteins bearing a long Cyt are generated by the use a stop codon at nt 1667–1669 due to a the shift in reading frame. These stop codons are followed by a 3'UTR of 1.25 (CytL) or 1.39-kb (CytS) (Figs. 1 and 3).

The UTRs of these cDNAs are also variable. cDNA clone 23 contains a 103-bp 5'UTR. However, a longer 5'UTR has been demonstrated by primer extension analysis (P.N. and N.B., unpublished results). As indicated above, *BgpA* and *MHVR* share the first 41 nt upstream from the ATG start codon but diverge further upstream from this point, suggesting that there are a minimum of two 5'UTR exons, or that a more complicated exon pattern lies upstream from the signal sequence. Sequencing of cDNA library clones and PCR-amplified cDNA clones have indicated a polymorphic region in the 3'UTR (indicated at the bottom of Fig. 3b). Since these sequence differences occur in a highly repetitive region and are from outbred CD-1 mice, it is not known presently if this variable region is due to stuttering during the RT reactions, errors during PCR amplifications, or if it represents a bona fide genetic polymorphism. Further downstream, cDNA library clones terminate at three different positions: clone 23 exhibits the shortest 3'UTR (nt 2114 in Fig. 3b), the *MHVR* cDNA (Dveksler et al., 1991) ends at nt 2710 (Fig. 3b) and clone 37 contains a further 199 nt ending with a poly(A) tail at nt position 2922. Several poly(A) consensus signals (double underlines in Fig. 3b) are located upstream from the various poly(A) tails, only one of which fits the perfect consensus sequence (AATAAA) (Proudfoot and Brownlee, 1976).

The predicted aa sequence of the *Bgp* isoforms are presented in Fig. 3c and may be sorted into two classes, defined by their N-terminal domains. The first 37 aa of these domains are identical; however, in the following stretch of 19 aa, nine residues are not conserved; charged structures replace non-polar aa and vice versa (K→A, M→K, F→Q). The Cys residues of the C2-set domains, thought to be involved in intrachain disulfide bonding (Williams and Barclay, 1988), are well conserved in all isoforms (shading in Fig. 3c).

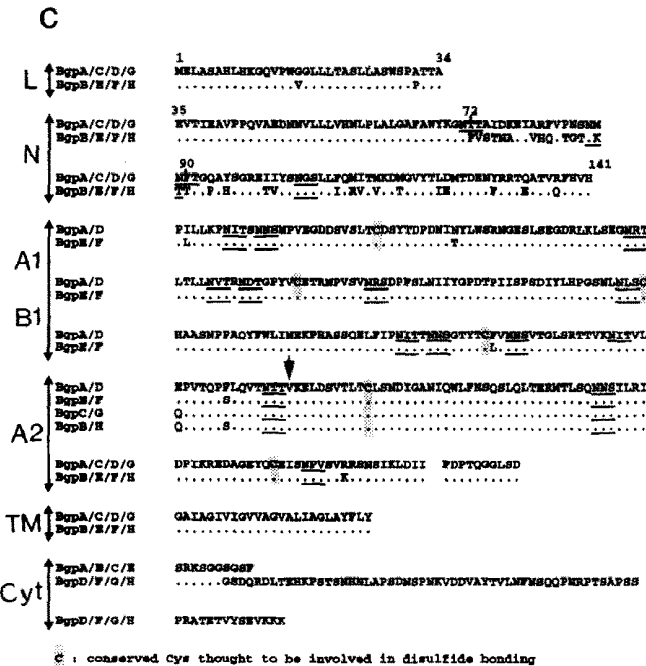
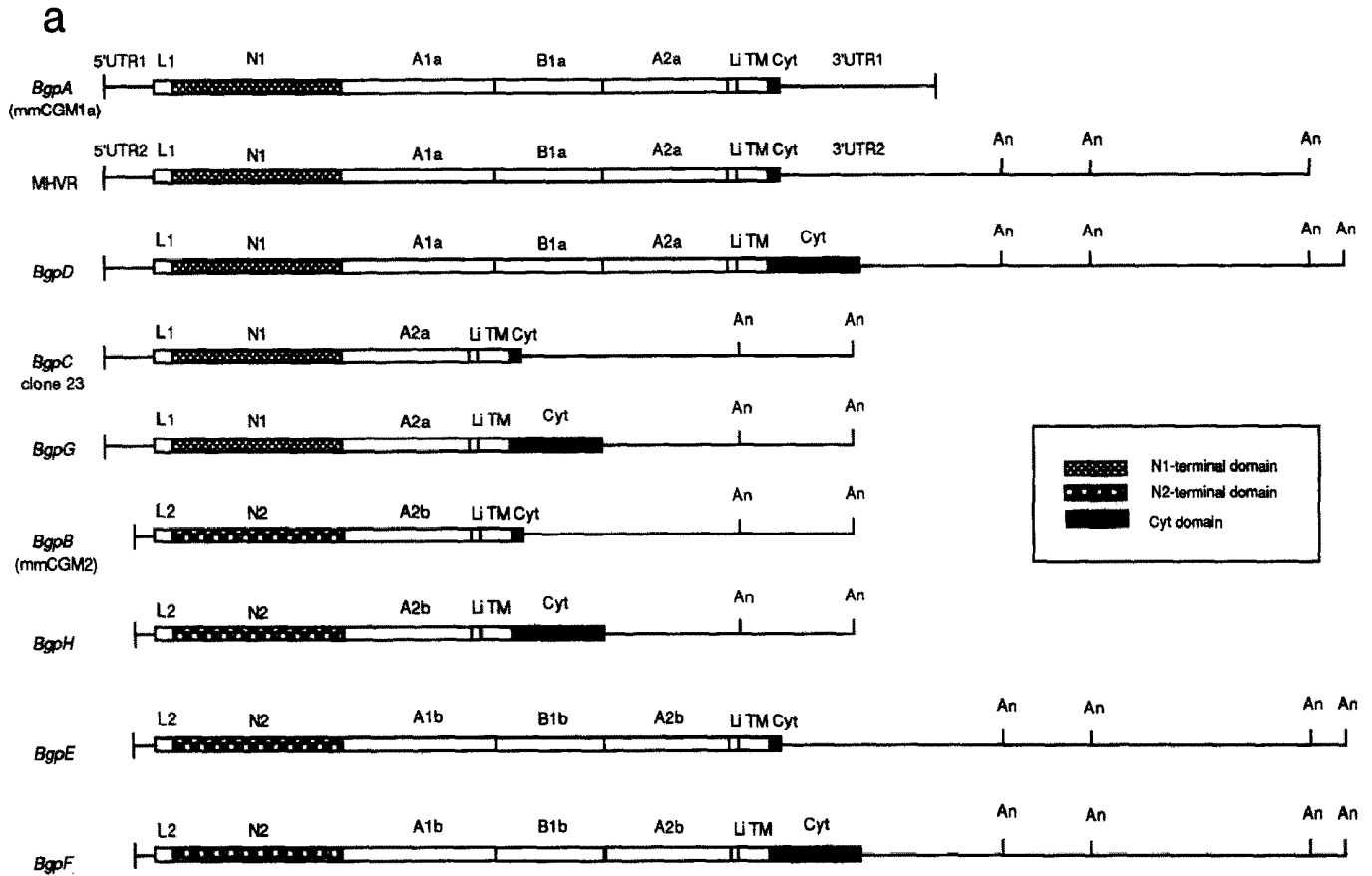
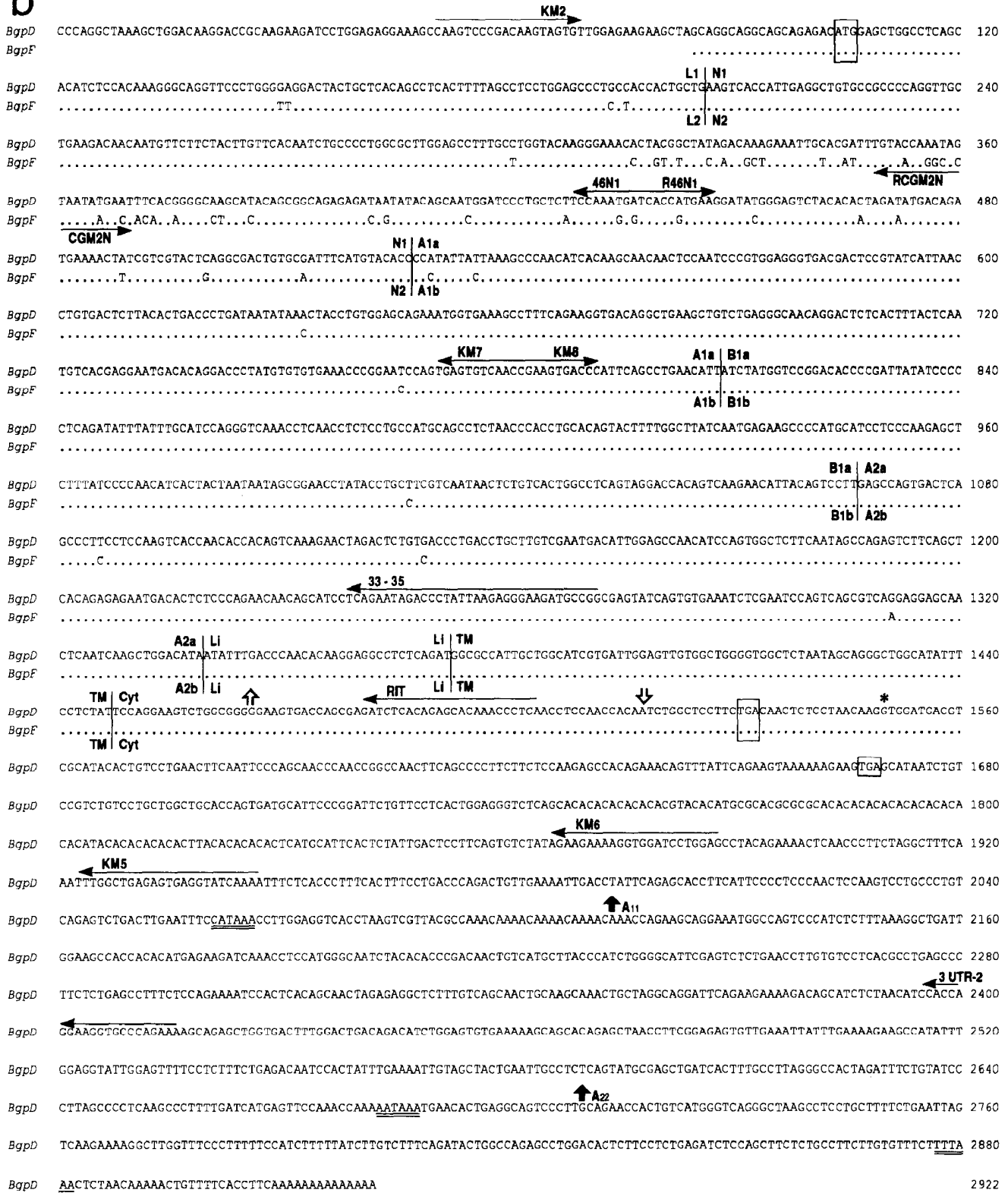


Fig. 3. Topology and cDNA sequences of mouse *Bgp* isoforms and deduced aa sequences. (a) Topology of *Bgp* isoforms. Structures of the various cDNA isoforms. Abbreviations used have been described in legend to Fig. 1, MHVR, mouse hepatitis virus receptor. (b) cDNA sequences. The nt sequences of the longest isoforms (*BgpD* and *BgpF*) are presented. Dots indicate identical nt. The sequences of the oligos used as primers or probes are indicated by lines above or below the nt sequences and their orientation by arrowheads on these lines. The ATG start and TGA stop codons are boxed. The domains are separated by lines and captions over and under the sequences. Open arrowheads delineate the 53 bp exon inserted to generate the long Cyt domain. The nt 1534–1550 are common to 3'UTR1 and 3'UTR2. The asterisk (at nt 1550) represents the position where the 3'UTR1 (McCuaig et al., 1992) replaces the 3'UTR2. Double underlines indicate putative or actual poly(A) consensus signals. ↑A₁₁ at position 2114 represents the poly(A) tail of cDNA clone 23. ↑A₂₂ at nt 2710 represents the poly(A) tail of the MHVR cDNA (Dveksler et al., 1991). The poly(A) tail at nt 2908 is that of cDNA clone 37. Sequence differences between nt 1747–1829 in the 3' UTR between various clones are indicated at the bottom of the figure. GenBank accession Nos.: *BgpC*, X67278; *BgpD*, X67279; *BgpE*, X67280; *BgpF*, X67281; *BgpG*, X67282; *BgpH*, X67283. (c) Deduced aa sequences of *Bgp*. The nt sequences were computer-translated and the deduced aa sequences were aligned. Double arrowheads indicate delineation of the domains. The protein sequences are written in the one letter aa code. Aa positions are indicated above the signal sequences and N-terminal domains. Glycosylation sites are underlined. Shaded C represent the conserved Cys residues. Arrow indicates position of the nt modification leading to a silent mutation. **Methods:** The full coding cDNAs for *BgpD*, *BgpE*, *BgpF*, *BgpG*, *BgpH* were reconstructed by ligating overlapping cDNA clones at various restriction sites. Reconstructed cDNA clones were completely resequenced to ensure accuracy of sequence.

b



clone 37: (CA) 5 CGTACACAT (GC) 2 AC (GC) 3 A (CA) 12 TA (CA) 5 CTT (AC) 5
 clone 58: (...) 5 (...) 2 (...) 3 11 - (...)
 clone 23: (...) 6 (...) 5 - (...) - (...) 14 - (...)
 PCR clones: (...) 6 (...) 4 - (...) - (...) 14 - (...)
 (R46N1-KM6)

LEGEND	
↓ ↑	53 bp exon
*	position of insertion of 3' UTR 1
↑ A _n	poly(A) tail

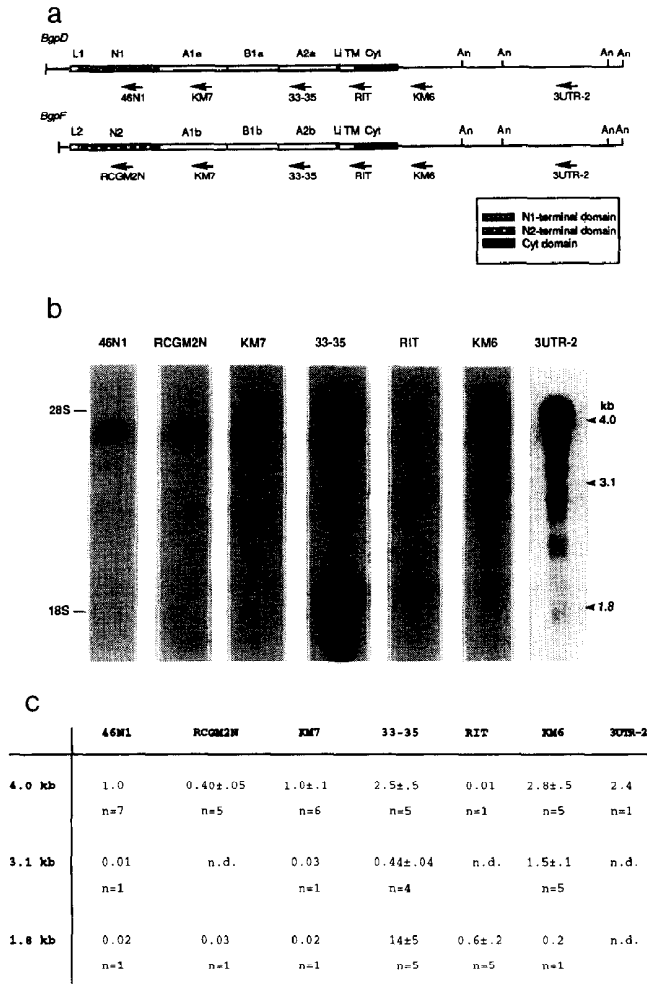


Fig. 4. Northern analyses and quantification. **(a)** Positions of oligos in the *BgpD* and *BgpF* isoforms. The positions of the oligos in the various regions of the cDNAs are indicated by arrows below the structures. **(b)** Northern analyses. Separation of mouse colon total RNA on a formaldehyde-agarose gel and hybridization with oligos. Positions of the 18S, 1.86-kb and 28S, 4.71 kb as well as the three major transcripts quantitated in **c** are indicated. **(c)** Quantification. Autoradiograms of different exposure times of the above Northern blots were scanned by laser densitometry using a Bio-Imager scanner (Millipore, Montreal, Canada). The 4.0-kb band revealed by hybridization to oligo 46N1 was weighed as 1.0. All other readings were computed proportionally. n.d., not determined. **Methods:** 20 μ g of total colon RNA, prepared by guanidium isothiocyanate extraction and centrifugation, was electrophoresed through a 2.2 M formaldehyde-1.5% agarose gel and transferred to GeneScreen Plus membranes. Hybridizations were carried out as previously described (McCuaig et al., 1992) with the [32 P]oligos (Sambrook et al., 1989) indicated above the autoradiograms. The filters were washed to a final stringency corresponding to 5°C below the predicted melting temperature of each oligo.

(d) Northern analysis and relative abundance of transcript isoforms

Northern analyses were performed on mouse colon transcripts with oligos representing different domains of the cDNAs (Fig. 4a and b) to quantify the relative abundance of the *Bgp* isoforms (Fig. 4c). Three major transcripts were identified with these probes. These were measured as 1.8, 3.1 and 4.0-kb relative to rRNA markers.

In comparison, human *BGP* restriction fragments have been shown to hybridize to 2.2 and 3.9-kb transcripts from a variety of cell lines (Hinoda et al., 1988; Barnett et al., 1989). The nt sequences of *BgpD* and *BgpF* isoforms reported in Fig. 3b are 2922 nt long, while the *BgpB* and *BgpC* or *BgpG* and *BgpH* isoforms are 1518 or 1569 nt long, respectively. The size discrepancy between the cDNA sequences and the transcript lengths measured by Northern analysis may be due to an additional 300 nt at the 5' end (P.N. and N.B., unpublished results). As well, the length of the poly(A) tails are thought to be 400-bp longer than the DNA sequences of the human *CEA* gene family members (Beauchemin et al., 1987). This adjustment would be consistent with a size of 3622-bp for the longer cDNAs. The existence of an even longer 3'UTR has not been ruled out. The 4.0-kb transcripts and a strong hybridization signal at 3.1-kb with a relative abundance of 1.5 are detected with oligo KM6; the 3.1-kb length is compatible with cDNAs expressing L-N-A2-TM-CytS or CytL carrying a complete 3'UTR region.

To quantify the number of transcripts co-migrating in the revealed bands, equal pmol of the labelled oligos were added to hybridization mixtures to ensure that the intensity of the bands would reflect the relative abundance of the transcripts. The most striking observation was that, although the cDNAs encoding the isoforms containing only two Ig domains (N and A2) could be readily cloned from a cDNA library or amplified by PCR, their abundance as 1.8-kb transcripts was minimal (Fig. 4c: 46N1, RCGM2N and KM7). The N1 and N2-terminal domains are preferentially expressed with three Ig C2-set constant domains as full-length transcripts. But, the transcripts encoding proteins with long Cyt are preferentially expressed without A1 and B1 domains and migrate as 1.8-kb rather than 4.0-kb mRNA. These are 1/10 to 1/3 as abundant as the short Cyt-encoding isoforms. As expected, hybridization with an oligo (3'UTR-2) located in the distal region of the 3' UTR detects preferentially 4.0 kb transcripts.

An oligo specific for the A2 domain (33-35) binds to both the 1.8 and 4.0-kb transcripts; however, quantification of the autoradiograms at different exposure times indicated exceedingly high values for this domain. Although the migration of these transcripts is similar to the 18S rRNA, the signal seen does not represent cross-hybridization to rRNA; when this oligo was used to probe non *Bgp*-expressing cell lines, no signal was encountered. However, this A2 domain is very conserved when compared to a similar domain expressed in the mouse *Psg* cDNAs (Kodelja et al., 1989; Rebstock et al., 1990; Rudert et al., 1992). Thus, the oligo 33-35 may be detecting some of these PSG-encoding transcripts.

(e) Conclusions

(1) The data presented in this report emphasize the multiplicity of the mouse *Bgp* isoforms and suggest mechanisms responsible for their generation. So far, nine similar, but not identical cDNA isoforms have been cloned from CD-1 mouse colon (Turbide et al., 1991; McCuaig et al., 1992; this report) and from Balb/c liver (Dveksler et al., 1991). Alternative splicing is active in the generation of isoforms exhibiting different Cyt domains as well as isoforms lacking internal A1 and B1 domains. Northern analysis performed on mouse colon RNA with specific oligo probes and quantification of the autoradiograms have confirmed the diversity of the *Bgp* isoforms; three major transcripts (1.8, 3.1 and 4.0 kb) encode these isoforms, albeit in different relative amounts. The relative abundance of the *Bgp* long Cyt-containing transcripts are approximately tenfold less abundant than the short Cyt-tailed isoforms. This finding may reflect on the function of the long-tailed BGP proteins since they contain consensus Tyr phosphorylation sites that have been conserved throughout evolution and are present in the mouse, the rat and the human BGP counterparts (Lin et al., 1989; Afar et al., 1992; Culic et al., 1992). Overexpression of these isoforms may be detrimental to cell growth and/or differentiation.

(2) In contrast to what has been described for human BGP isoforms (Barnett et al., 1989) there are, in fact, two classes of mouse *Bgp* isoforms. Sequence analyses performed on the cDNA isoforms from CD-1 mice have demonstrated that the N1-terminal domain is only associated with either the A1a or A2a domain, while the N2-terminal domain is joined to the A1b or A2b domain. These C2-set Ig domains differ from each other by only a few nt while the N-terminal domains exhibit a greater number of changes. The A2 domain is also a required structural feature of the mouse colon *Bgp* isoforms contrary to the human *BGPb* and *BGPd* isoforms (Barnett et al., 1989) which lack this domain as do *BGPg*, *BGP_h*, *BGP_x* and *BGP_y* (T. Barnett, personal communication). However, all isoforms examined in this report carry identical TM and long or short Cyt domains. The length of the 3'UTR varies but the sequence within remains invariable (except for a few nt modifications in a highly repetitive region), strongly suggesting that a single *Bgp* gene encompasses all exons.

However, the generation of isoforms bearing a few nt modifications could also be explained by allelic variation. The present characterization was done on the outbred CD-1 mouse. A recent report suggests that expression of the two classes of isoforms (defined by their N-terminal domains) are due to allelic variation of the same gene. Balb/c mice do not express the N2-containing cDNA isoforms and conversely, SJL/J mice lack the N1-bearing

isoform transcripts and proteins (Dveksler et al., 1993). The splicing versus allelic variation issue will be resolved with the characterization of the complete Balb/c *Bgp* gene (P.N. and N.B., in progress).

(3) The feature which renders the two isoform classes distinct lies in their N-terminal domains. The first 37 aa of these domains are identical, while the last third encloses many modifications mostly concentrated in a central core. Recently, Bates et al. have constructed a three-dimensional model for human CEA based on the NMR structure of the rat CD2 as well as X-ray crystal structures of human CD4 and Ig variable domains (Bates et al., 1992). The first two domains of CEA are predicted to have a rod-like appearance with numerous tightly packed β sheets and a few exposed loops on either side of the molecule. Interestingly, the most divergent aa residues (positions 72–90) within the mouse *Bgp* N-terminal domains would be localized in the C-C' loop which is predicted to be exposed. On the other hand, some of the conserved 37 aa are positioned within another exposed loop. These loops are the most likely targets for interactions with other cells or ligands. In fact, two described properties of these mouse glycoproteins are dependent upon such interactions; two protein products of the *Bgp* cDNA isoforms (*BgpA* and *BgpB*) can function in vitro as cell adhesion molecules (Turbide et al., 1991; McCuaig et al., 1992) and one (MHVR or *BgpA*) has also been shown to bind to MHV spike glycoprotein (Williams et al., 1990; Dveksler et al., 1991). Although other domains of the CEA/BGP proteins are necessary for intercellular adhesion, the N-terminal domain appears to be critical for this function (Oikawa et al., 1991). Similarly, other members of the Ig supergene family (CD4, ICAM-1, CD2, poliovirus receptor) require that their N-terminal domains enter into contact with either viral envelope glycoproteins (Greve et al., 1989; Mendelsohn et al., 1989; Staunton et al., 1989; White and Littman, 1989; Koike et al., 1990; Pollard et al., 1991) or their respective ligands to be functionally activated (Staunton et al., 1990; Koike et al., 1991; Register et al., 1991).

Human BGP mRNAs and proteins appear to be less abundant than CEA in colon but more abundant in liver (Hinoda et al., 1990; Drzeniek et al., 1991). Extensive expression analyses of these two protein entities have been hampered by the lack of specific anti-BGP antibodies which have only recently been developed (Drzeniek et al., 1991). However, the mouse *Bgp* mRNAs (Turbide et al., 1991; McCuaig et al., 1992) and proteins (M.R., P.N., S.J., C.T., D.F. and N.B., manuscript in preparation) are more abundant in colon than liver. Since a true CEA homolog or any of the phosphatidylinositol-linked CEA family members have yet to be identified in the mouse or rat models, it is tempting to speculate that the functions

associated with CEA in the human may represent a recent evolutionary event (Stanners et al., 1992) and that these functions may be ensured in the mouse by other *Bgp*-related genes (P.N. and N.B., in preparation). The significance of the concurrent expression of the many *Bgp* isoforms on the surface of mouse tissues such as colon, intestine, liver and uterus (Turbide et al., 1991; McCuaig et al., 1992) may be that these proteins play a determinant role, through self- or heterologous contact, in renewal and/or differentiation of their epithelia.

ACKNOWLEDGEMENTS

The authors wish to thank Dr. Annette Herscovics and Dr. Clifford Stanners for critical reading of the manuscript and Dr. André Veillette for the gift of the NIH3T3 cells. K.McC. is a recipient of a studentship from the Medical Research Council of Canada, M.R., from the Cancer Research Society Inc. and P.N., from the 'Institut Mérieux', Lyon, France, N.B. is a scholar of the 'Fonds de la Recherche en Santé du Québec'. This research is supported by the Medical Research Council of Canada and The Cancer Research Society Inc.

REFERENCES

- Afar, D.E.H., Stanners, C.P. and Bell, J.C.: Tyrosine phosphorylation of biliary glycoprotein, a cell adhesion molecule related to carcinoembryonic antigen. *Biochim. Biophys. Acta* 1134 (1992) 46–52.
- Barnett, T., Kretschmer, A., Austen, D.A., Goebel, S.J., Hart, J.T., Elting, J.J. and Kamarck, M.E.: Carcinoembryonic antigens: alternative splicing accounts for multiple mRNAs that code for novel members of the carcinoembryonic antigen family. *J. Cell Biol.* 108 (1989) 267–276.
- Bates, P.A., Luo, J. and Sternberg, M.J.E.: A predicted three-dimensional structure for the carcinoembryonic antigen (CEA). *FEBS Lett.* 301 (1992) 207–214.
- Beauchemin, N., Benchimol, S., Cournoyer, D., Fuks, A. and Stanners, C.P.: Isolation and characterization of full-length functional cDNA clones for human carcinoembryonic antigen. *Mol. Cell. Biol.* 7 (1987) 3221–3230.
- Beauchemin, N., Turbide, C., Afar, D., Bell, J., Raymond, M., Stanners, C.P. and Fuks, A.: A mouse analogue of the human carcinoembryonic antigen. *Cancer Res.* 49 (1989) 2017–2021.
- Benchimol, S., Fuks, A., Jothy, S., Beauchemin, N., Shirota, K. and Stanners, C.P.: Carcinoembryonic antigen, a human tumor marker, functions as an intercellular adhesion molecule. *Cell* 57 (1989) 27–34.
- Berling, B., Kolbinger, F., Grunert, F., Thompson, J.A., Brombacher, F., Buchegger, F., von Kleist, S., Zimmermann, W.: Cloning of a carcinoembryonic antigen gene family member expressed in leukocytes of chronic myeloid leukemia patients and bone marrow. *Cancer Res.* 50 (1989) 6534–6539.
- Cerni, C., Tatra, G. and Bohm, H.: Immunosuppression by human placenta lactogen (HPL) and the pregnancy-specific β 1-glycoprotein (SP1): inhibition of mitogen-induced lymphocyte transformation. *Arch. Gynäkol.* 223 (1977) 1–7.
- Culic, O., Huang, Q.H., Flanagan, D., Hixson, D. and Lin, S.H.: Molecular cloning and expression of a new rat liver cell-CAM105 isoform. *Biochem. J.* 285 (1992) 47–53.
- Devereux, J., Haerberli, P. and Smithies, O.: A comprehensive set of sequence analysis programs for the VAX. *Nucleic Acids Res.* 12 (1984) 387–395.
- Drzeniek, Z., Lamerz, R., Fenger, U., Wagener, C. and Haubeck, H.D.: Identification of membrane antigens in granulocytes and colonic carcinoma cells by a monoclonal antibody specific for biliary glycoprotein, a member of the carcinoembryonic antigen family. *Cancer Lett.* 56 (1991) 173–179.
- Dveksler, G.S., Pensiero, M.N., Cardellicchio, C.B., Williams, R.K., Jiang, G.S., Holmes, K.V. and Dieffenbach, C.W.: Cloning of the mouse hepatitis virus (MHV) receptor: expression in human and hamster cell lines confers susceptibility to MHV. *J. Virol.* 65 (1991) 6881–6891.
- Dveksler, G.S., Dieffenbach, C.W., Cardellicchio, C.B., McCuaig, K., Pensiero, M.N., Jiang, G.S., Beauchemin, N. and Holmes, K.V.: Several members of the mouse CEA-related glycoprotein family are functional receptors for murine coronavirus MHV-A59. *J. Virol.* 67 (1993) 1–8.
- Frohman, M.A., Dush, M.K. and Martin G.R.: Rapid production of full-length cDNAs from rare transcripts: amplification using a single gene-specific oligonucleotide primer. *Proc. Natl. Acad. Sci. USA* 85 (1988) 8998–9001.
- Greve, J.M., Davis, G., Meyer, A.M., Forte, C.P., Yost, S.C., Marlor, C.W., Kamarck, M.E. and McClelland, A.: The major human rhinovirus receptor is ICAM-1. *Cell* 56 (1989) 839–847.
- Gold, P. and Freedman, S.O.: Specific carcinoembryonic antigens of the human digestive system. *J. Exp. Med.* 122 (1965) 467–481.
- Hefta, S.A., Hefta, L.J., Lee, T.D., Paxton, R.J. and Shively, J.E.: Carcinoembryonic antigen is anchored to membranes by covalent attachment to a glycosylphosphatidylinositol moiety: identification of the ethanolamine linkage site. *Proc. Natl. Acad. Sci. USA* 85 (1988) 4648–4652.
- Hinoda, Y., Neumaier, M., Hefta, S.A., Drzeniek, Z., Wagener, C., Shively, L., Hefta, L.J.F., Shively, J.E. and Paxton, R.J.: Molecular cloning of a cDNA coding for biliary glycoprotein I: primary structure of a glycoprotein immunologically crossreactive with carcinoembryonic antigen. *Proc. Natl. Acad. Sci. USA* 85 (1988) 6959–6963.
- Hinoda, Y., Imai, K., Nakagawa, N., Ibayashi, Y., Nakano, T., Paxton, R.J., Shively, J.E. and Yachi, A.: Transcription of biliary glycoprotein I gene in malignant and non-malignant human liver tissues. *Int. J. Cancer* 45 (1990) 875–878.
- Khan, W. N., Teglund, S., Bremer, K. and Hammarström, S.: The pregnancy-specific glycoprotein family of the immunoglobulin superfamily: identification of new members and estimation of family size. *Genomics* 12 (1992) 780–787.
- Kodelja, V., Lucas, K., Barnert, S., von Kleist, S., Thompson, J.A. and Zimmermann, W.: Identification of a carcinoembryonic antigen gene family in the rat. *J. Biol. Chem.* 264 (1989) 6906–6912.
- Koike, S., Horie, H., Ise, I., Okitsu, A., Yoshida, M., Iisuka, N., Takeuchi, K., Takegami, T. and Nomoto, A.: The poliovirus receptor protein is produced both as membrane-bound and secreted forms. *EMBO J.* 9 (1990) 3217–3224.
- Koike, S., Ise, I. and Nomoto, A.: Functional domains of the poliovirus receptor. *Proc. Natl. Acad. Sci. USA* 88 (1991) 4104–4108.
- Kolbinger, F., Schwarz, K., Brombacher, F., von Kleist, S. and Grunert, F.: Expression of an NCA cDNA in NIH/3T3 cells yields a 110K glycoprotein, which is anchored into the membrane via glycosylphosphatidylinositol. *Biochem. Biophys. Res. Commun.* 161 (1989) 1126–1134.
- Leusch, H.G., Drzeniek, Z., Markos-Pusztai, Z. and Wagener, C.: Binding of *Escherichia coli* and *Salmonella* strains to members of the carcinoembryonic antigen family: differential binding inhibition

- by aromatic α -glycosides of mannose. *Inf. Immun.* 59 (1991) 2051–2057.
- Lin, S.H. and Guidotti, G.: Cloning and expression of a cDNA coding for rat liver plasma membrane ecto-ATPase: the primary structure of the ecto-ATPase is similar to that of the human biliary glycoprotein I. *J. Biol. Chem.* 264 (1989) 14408–14414.
- Mattila, P., Korpela, J., Tenkänen, T. and Pitkänen, K.: Fidelity of DNA synthesis by the *Thermococcus litoralis* DNA polymerase. An extremely heat stable enzyme with proofreading activity. *Nucleic Acids Res.* 19 (1991) 4967–4973.
- McCabe, P.C.: Production of single-stranded DNA by asymmetric PCR. In: Innis, M.A., Gelfand, D.H., Sninsky, J.I. and White, T.J. (Eds.), *PCR Protocols: A Guide to Methods and Applications*. Academic Press, San Francisco, CA, 1990, pp. 76–83.
- McCuaig, K., Turbide, C. and Beauchemin, N.: mmCGM1a: a mouse carcinoembryonic antigen gene family member, generated by alternative splicing, functions as an adhesion molecule. *Cell Growth Differ.* 3 (1992) 165–174.
- Mendelsohn, C.L., Wimmer, E. and Racaniello, V.R.: Cellular receptor for poliovirus: molecular cloning, nucleotide sequence, and expression of a new member of the immunoglobulin superfamily. *Cell* 56 (1989) 855–865.
- Oikawa, S., Inuzuka, C., Kuroki, M., Matsuoka, Y., Kosaki, G. and Nakazato, H.: Cell adhesion activity of non-specific cross-reacting antigen (NCA) and carcinoembryonic antigen (CEA) expressed on CHO cell surface: homophilic and heterophilic adhesion. *Biochem. Biophys. Res. Commun.* 164 (1989) 39–45.
- Oikawa, S., Inuzuka, C., Kuroki, M., Arakawa, F., Matsuoka, Y., Kosaki, G. and Nakazato, H.: A specific heterotypic cell adhesion activity between members of carcinoembryonic antigen family W272 and NCA, is mediated by N-domains. *J. Biol. Chem.* 266 (1991) 7995–8001.
- Oikawa, S., Kuroki, M., Matsuoka, Y., Kosaki, G. and Nakazato, H.: Homotypic and heterotypic Ca^{2+} -independent cell adhesion activities of biliary glycoprotein, a member of carcinoembryonic antigen family, expressed on CHO cell surface. *Biochem. Biophys. Res. Commun.* 186 (1992) 881–887.
- Pollard, S.R., Meier, W., Chow, P., Rosa, J.J. and Wiley, D.C.: CD4-binding regions of human immunodeficiency virus envelope glycoprotein gp120 defined by proteolytic digestion. *Proc. Natl. Acad. Sci. USA* 88 (1991) 11320–11324.
- Proudfoot, N.J. and Brownlee, G.G.: 3' Non-coding region sequences in eukaryotic messenger RNA. *Nature* 263 (1976) 211–214.
- Rebstock, S., Lucas, K., Thompson, J.A. and Zimmermann, W.: cDNA and gene analyses imply a novel structure for a rat carcinoembryonic antigen-related protein. *J. Biol. Chem.* 265 (1990) 7872–7879.
- Register, R.B., Uncapher, C.R., Naylor, A.M., Linerberger, D.W. and Colonna, R.J.: Human-murine chimeras of ICAM-1 identify amino acid residues critical for rhinovirus and antibody binding. *J. Virol.* 65 (1991) 6589–6596.
- Rojas, M., Fuks, A. and Stanners, C.P.: Biliary glycoprotein, a member of the immunoglobulin supergene family, functions in vitro as a Ca^{2+} -dependent intercellular adhesion molecule. *Cell Growth Differ.* 1 (1990) 527–533.
- Rudert, F., Saunders, A.M., Rebstock, S., Thompson, J.A. and Zimmermann, W.: Characterization of murine carcinoembryonic antigen gene family members. *Mammalian Genome* 3 (1992) 262–273.
- Sambrook, J., Fritsch, E.F. and Maniatis, T.: *Molecular Cloning. A Laboratory Manual*, 2nd Ed., Cold Spring Harbor Laboratory Press, Cold Spring Harbor, NY, 1989.
- Sanger, F., Nicklen, S. and Coulson, A.R.: DNA sequencing with chain-terminating inhibitors. *Proc. Natl. Acad. Sci. USA* 74 (1977) 5463–5467.
- Schrewe, H., Thompson, J., Bona, M., Hefta, L.J.F., Maruya, A., Hassauer, M., Shively, J.E., von Kleist, S. and Zimmermann, W.: Cloning of the complete gene for carcinoembryonic antigen: analysis of its promoter indicates a region conveying cell type-specific expression. *Mol. Cell. Biol.* 10 (1990) 2738–2748.
- Shuster, J.D., Thompson, M.P., Fuks, A. and Gold, P.: Immunologic approaches to diagnosis of malignancies. *Prog. Exp. Tumor Res.* 25 (1980) 89–139.
- Stanners, C.P., Rojas, M., Zhou, H., Fuks, A. and Beauchemin, N.: The CEA family: a system in transitional evolution? *Int. J. Biol. Markers* 7 (1992) 137–142.
- Staunton, D.E., Merluzzi, V.J., Rothlein, R., Barton, R., Marlin, S.D. and Springer, T.A.: A cell adhesion molecule, ICAM-1, is the major surface receptor for rhinoviruses. *Cell* 56 (1989) 849–853.
- Staunton, D.E., Dustin, M.L., Erickson, H.P. and Springer, T.A.: The arrangement of the immunoglobulin-like domains of ICAM-1 and the binding sites for LFA-1 and rhinovirus. *Cell* 61 (1990) 243–254.
- Svenberg, T.: Carcinoembryonic antigen-like substances of human bile: isolation and partial characterization. *Int. J. Cancer* 17 (1976) 588–596.
- Svenberg, T., Hammarström, S. and Hedin, A.: Purification and properties of biliary glycoprotein I (BGP 1). Immunochemical relationship to carcinoembryonic antigen. *Mol. Immunol.* 16 (1979) 245–252.
- Thompson, J., Zimmermann, W., Osthus-Bugat, P., Schleussner, C., Eades-Perner, A.M., Barnert, S., von Kleist, S., Willcocks, T., Craig, I., Tynan, K., Olsen, A. and Mohrenweiser, H.: Long-range chromosomal mapping of the carcinoembryonic antigen (CEA) gene family cluster. *Genomics* 12 (1992) 761–772.
- Turbide, C., Rojas, M., Stanners, C.P. and Beauchemin, N.: A mouse carcinoembryonic antigen gene family member is a calcium-dependent cell adhesion molecule. *J. Biol. Chem.* 266 (1991) 309–315.
- White, J.M. and Littman, D.R.: Viral receptors of the immunoglobulin superfamily. *Cell* 56 (1989) 725–728.
- Williams, A.F. and Barclay, A.N.: The immunoglobulin superfamily: domains for cell recognition. *Ann. Rev. Immunol.* 6 (1988) 381–405.
- Williams, R.K., Jiang, G.S., Snyder, S.W., Frana, M.F. and Holmes, K.V.: Purification of the 110-kilodalton glycoprotein receptor for mouse hepatitis virus (MHV)-A59 from mouse liver and identification of a nonfunctional homologous protein in MHV-resistant SJL/J mice. *J. Virol.* 64 (1990) 3817–3823.
- Zhou, H., Fuks, A. and Stanners, C.P.: Specificity of intercellular adhesion mediated by various members of the immunoglobulin supergene family. *Cell Growth Differ.* 1 (1990) 209–215.