Vol. 192, No. 1, 1993 BIOCHEMICAL AND BIOPHYSICAL RESEARCH COMMUNICATIONS
April 15, 1993 Pages 223-230

HuMIG: A NEW HUMAN MEMBER OF THE CHEMOKINE FAMILY OF CYTOKINES

Joshua M. Farber

Division of Infectious Diseases, Department of Medicine
Johns Hopkins University School of Medicine
Baltimore, MD 21205

Received March 2, 1993

SUMMARY Characterization of lymphokine-activated genes in mouse macrophages led to the identification previously of Mu*mig*, an interferon-γ-inducible murine gene that encodes a member of the chemokine family of cytokines. The Mu*mig* cDNA probe was used to screen a cDNA library prepared from cultures of the THP-1 human monocytic cell line that had been treated with interferon-γ. This led to the identification of Hu*mig*, a new human member of the chemokine gene family. Hu*mig* is induced in THP-1 cells and in peripheral blood mononuclear cells by interferon-γ but not by interferon-α or by lipopolysaccharide. Analysis of mouse and human genomic DNAs suggested that the Mu*mig* and Hu*mig* genes are true mouse-human homologues. The Hu*mig* mRNA encodes a predicted secreted HuMig protein of 103 residues, M_r 11,725.

Studies of mine to identify genes and proteins induced as part of the macrophage response to lymphokines led to the isolation of cDNAs encoding two mouse members of the chemokine family of cytokines - Crg-2 (1) and Mig (2), both thus far of undescribed functions. The chemokines are small, inducible, secreted proteins that can function as chemotactic factors for a variety of cells, as immune cell activators, and as positive and negative regulators of cell growth (reviewed in 3). The most striking feature in the primary structure of the chemokines is the presence of four invariant cysteine residues.

Crg-2 is likely the mouse homologue of the human IP-10, an interferon-inducible protein of undescribed function, and the *crg*-2 gene could be induced by IFN's α, β and γ (1). In contrast, *mig* was found to be inducible in macrophage-like cells only by IFN-γ (2) and presumably mediates a subset of IFN-γ's biological activities. No homologue of the mouse *mig*, herein referred to as Mu*mig*, had been described previously. To pursue investigations into the regulation of the *mig* gene(s) and the functions of the Mig protein(s), particularly in regard to the possible roles for Mig in human biology, studies were done in order to identify and characterize a human counterpart to Mu*mig*. As described below, a cDNA has been isolated corresponding to a human gene designated Hu*mig*, the presumed human homologue of Mu*mig*, and encoding a new IFN-γ-inducible member of the chemokine family.

Abbreviations used in this paper: IFN, interferon; PBMC, peripheral blood mononuclear cells; LPS, lipopolysaccharide; EU, endotoxin units; TNF, tumor necrosis factor.

0006-291X/93 $4.00

Materials and Methods

Cell Culture, PBMC, IFNs and LPS. THP-1 cells were obtained from the Americn Type Culture Collection and grown in RPMI 1640 supplemented with 10% fetal bovine serum. For treating with interferons and other agents, the cells were suspended at approximately 1 x 10^6/ ml.

Human PBMC were prepared from citrate-anti-coagulated blood by centrifugation using Percoll (Pharmacia LKB) according to the supplier's specification. The banded mononuclear cells were washed twice in RPMI 1640 and resuspended in RPMI 1640 with 10% fetal bovine serum. In order to enrich for monocytes, 2.4 x 10^5 PBMC were plated per 10 cm gelatin-coated dish by the method of Jones et al (4). After 2 h at 37° C, the plates were washed gently three times and then treated as noted below.

Recombinant human IFN-γ was either from Collaborative Research or from Genentech (Actimmune), or was provided as a kind gift from Daniela Novick, Weizmann Institute of Science. Recombinant human IFN-α was either from Hoffman-LaRoche (Roferon-A) or was provided as a kind gift by Charles Weissmann, University of Zurich. LPS was from *E.coli* 0127.B8 (Difco Laboratories).

Following the treatments of the THP-1 cells and PBMC, media from the control and IFN-γ-treated cells were frozen and assayed subsequently for LPS contamination using a chromogenic *limulus* amebocyte lysate test (BioWhittaker). For the experiment shown with THP-1 cells, these LPS levels were between 5 and 6 EU/ml (about 0.5 ng/ml). In other similar experiments with THP-1 cells that showed the same responses to IFN-γ, LPS levels were < 0.25 EU/ml, and for the PBMC experiments shown below, levels were also < 0.25 EU/ml.

Construction and Screening of the cDNA Library. Using Fast Track Version 2.1 (Invitrogen), poly(A)+ RNA was prepared from 10^8 THP-1 cells cultured for 9 h at 5 x 10^5 cells/ml with 1000 U/ml of IFN-γ. Oligo-dT-primed cDNA was synthesized and a cDNA library was prepared in the λZAPII vector (Stratagene) and packaged according to the vendor's protocol with minor modifications. Screening using Mu*mig* cDNA probe, radiolabelled by nick-translation, was done by hybridizing in 1 M NaCl at 65° C with the most stringent wash done in 0.5 M NaCl at 65° C as described (5). In vivo excision of cDNA inserts in the pBluescript SK-phagemid (Stratagene) from within the λZAP vector was done according to the vendor's protocols.

Sequencing of Hu*mig* cDNAs. Using deletions made using exonuclease III as well as Hu*mig*-specific oligonucleotide primers, both strands of Hu*mig* cDNA clone H-1-3 were sequenced completely. Sequencing of cDNA H-1-3 and other cDNA clones was done by the dideoxy method (6) using reagents and protocols supplied by United States Biochemical Corp.

Northern Blot Analysis. Agarose/formaldehyde gel electrophoresis using 1.2% agarose gels was done according to the procedure of Goldberg (7), and prehybridization, hybridization and washing were done as described (1). ^{32}P-labeled cDNA probes were made using random primers (8) and with reagents and protocols provided by Amersham. The IP-10 cDNA clone was isolated from the THP-1 library described above by using an oligonucleotide probe based on published sequences (9). The human TNF-α cDNA probe was from λ clone 42-4 (10) and was a gift from Genentech. The mouse aldolase A cDNA clone was a gift from A. Levy, L. Sanders and D. Nathans (Johns Hopkins University). Autoradiography using Kodak XAR film was done with an intensifying screen at -70° C.

Southern Blot Analysis. The Southern filters containing genomic DNA were the kind gift of Anthony Lanahan, Johns Hopkins University. Twenty μg of digested human DNA were loaded per lane, separated on a 1% agarose gel and transferred to supported nitrocellulose as described (11). The Mu*mig* and Hu*mig* probes were prepared by the random primer method using cDNA's from which the poly(A) tail had been deleted. Prehybridization and hybridization were done overnight and for 2.5 days, respectively, at 42° C in 25% formamide, 0.8 M NaCl, 0.05 M NaH_2PO_4, 6 mM EDTA, 5 x Denhardt's reagent and 100 μg/ml salmon sperm DNA as described (12). The blots were washed twice briefly at room temperature in 2 x SSC (1x=150 mM sodium chloride, 15 mM sodium citrate) plus 0.1% SDS and then for 1 h at 42° C in 1 x SSC plus 0.1% SDS as described (12), followed by autoradiography.

Results and Discussion

A cDNA library was prepared using the λ ZAPII vector from cultures of the THP-1 human monocytic cell line that had been treated for 9 h with IFN-γ and the library was screened at

moderate stringency with a nearly-full-length Mu*mig* cDNA probe. Six related but independent cDNA clones were isolated. The longest was H-1-3 of approximately 2.5 kb, and the sequence of this cDNA clone is shown in Fig. 1. Additional H-1-3 related cDNA clones were subsequently isolated and using these as well as the initial collection of cDNAs, more than 1900 of the 2545 base pairs in H-1-3 were independently confirmed and no mismatches with H-1-3 were found. cDNA clone H-1-3 contained a 42-residue poly(A) tail, not shown in Fig. 1, at its 3' end with a consensus polyadenylylation sequence appropriately located from nucleotides 2525-2530. Since the mRNA detected using an H-1-3 probe (see below) is approximately 3 kb long, taking into account the mRNA's poly(A) tail, cDNA clone H-1-3 is likely to be close to full-length. The single longest open reading frame beginning with a methionine contains 125 amino acids and is designated HuMig.

The extent of the sequence, if any, missing from the 5' end of the H-1-3 cDNA has not been determined experimentally, and the position of the initiator methionine has been assigned based on the comparison with the Mu*mig* cDNA sequence. The open reading frame extends from nucleotides 40 to 414, with a termination codon at positions 415-417, followed by 2128 bases of presumed 3' non-translated sequence. The 125-residue HuMig protein is predicted to have a molecular weight (M_r) of 14,019. The amino terminal sequence has basic residues followed by a highly hydrophobic region characteristic of a signal peptide, and the most favorable site of cleavage as predicted by the rules of von Heijne (13) is C-terminal to Gly 22, leaving a protein of 103 residues, M_r 11,725. The absence of a second hydrophobic segment is consistent with HuMig being secreted and not membrane-bound. A prominent feature of the predicted HuMig sequence is an extremely basic C-terminal region that is characteristic of the chemokines, and HuMig has a calculated pI of >11.0.

A comparison between the predicted sequence for the MuMig and the HuMig proteins as shown in Fig. 2A demonstrates that allowing for a few single residue gaps, HuMig matches MuMig at 87 of 125 positions including the four invariant cysteines found in members of the chemokine family. The first and second cysteines of HuMig are separated by a single amino acid, as is the case for MuMig and other members of the so-called CXC subgroup of the chemokine family, a subgroup including platelet factor 4, IP-10 and IL-8 among others. The predicted site for N-linked glycosylation in the MuMig protein (Asn 58) is not conserved in HuMig.

The Hu*mig* mRNA (see below) is approximately 3 kilobases versus approximately 1.6 kilobases for the Mu*mig* mRNA (2). The Hu*mig* cDNA contains a 3' non-translated region that is 1320 base pairs longer than the corresponding region in Mu*mig*. Comparison between the Hu*mig* cDNA and the approximately 7 kilobases of the Mu*mig* gene that have been sequenced (Wright, T., S. Willenberger, and J. Farber, manuscript in preparation) showed that sequences corresponding to the 3' end of the Hu*mig* cDNA are present in the mouse gene 3' to the 3' end of the mouse mRNA (data not shown). Therefore, differences in polyadenylylation, presumably through the acquisition of mutations, most likely accounts for the difference in length between the mature mouse and human transcripts. While the segment of the mouse cDNA immediately upstream of the Mu*mig* polyadenylylation sequence can be readily aligned with the Hu*mig* cDNA, placing the site of polyadenylylation in the mouse at approximately position 1335 in the human, the

```
      10         20         30         40         50         60         70         80         90        100        110
      *  *       *  *       *  *       *  *       *  *       *  *       *  *       *  *       *  *       *  *       *  *
ATCCAATACA GGAGTGACTT GGAACTCCAT TCTATCACT ATG AAG AAA AGT GGT GTT CTT TTC CTC TTG GGC ATC ATC TTG CTG GTT CTG ATT GGA GTG CAA GGA ACC CCA GTA
TAGGTTATGT CCTCACTGAA CCTTGAGGTA AGATAGTGA TAC TTC TTT TCA CCA CAA GAA AAG GAG AAC CCG TAG TAG AAC GAC CAA GAC TAA CCT CAC GTT CCT TGG GGT CAT
                                        Met Lys Lys Ser Gly Val Leu Phe Leu Leu Gly Ile Ile Leu Leu Val Leu Ile Gly Val Gln Gly Thr Pro Val

     120        130        140        150        160        170        180        190        200        210        220
      *  *       *  *       *  *       *  *       *  *       *  *       *  *       *  *       *  *       *  *       *  *
GTG AGA AAG GGT CGC TGT TCC TGC ATC AGC ACC AAC CAA GGG ACT ATC CAC CTA CAA TCC TTG AAA GAC CTT AAA CAA TTT GCC CCA AGC CCT TCC TGC GAG AAA ATT
CAC TCT TTC CCA GCG ACA AGG ACG TAG TCG TGG TTG GTT CCC TGA TAG GTG GAT GTT AGG AAC TTT CTG GAA TTT GTT AAA CGG GGT TCG GGA AGG ACG CTC TTT TAA
Val Arg Lys Gly Arg Cys Ser Cys Ile Ser Thr Asn Gln Gly Thr Ile His Leu Gln Ser Leu Lys Asp Leu Lys Gln Phe Ala Pro Ser Pro Ser Cys Glu Lys Ile

     230        240        250        260        270        280        290        300        310        320        330
      *  *       *  *       *  *       *  *       *  *       *  *       *  *       *  *       *  *       *  *       *  *
GAA ATC ATT GCT ACA CTG AAG AAT GGA GTT CAA ACA TGT CTA AAC CCA GAT TCA GCA GAT GTG AAG GAA CTG ATT AAA AAG TGG GAG AAA CAG GTC AGC CAA AAG AAA
CTT TAG TAA CGA TGT GAC TTC TTA CCT CAA GTT TGT ACA GAT TTG GGT CTA AGT CGT CTA CAC TTC CTT GAC TAA TTT TTC ACC CTC TTT GTC CAG TCG GTT TTC TTT
Glu Ile Ile Ala Thr Leu Lys Asn Gly Val Gln Thr Cys Leu Asn Pro Asp Ser Ala Asp Val Lys Glu Leu Ile Lys Lys Trp Glu Lys Gln Val Ser Gln Lys Lys

     340        350        360        370        380        390        400        410        420        430        440
      *  *       *  *       *  *       *  *       *  *       *  *       *  *       *  *       *  *       *  *       *  *
AAG CAA AAG AAT GGG AAA AAA CAT CAA AAA AAG AAA GTT CTG AAA GTT CGA AAA TCT CAA CGT TCT CGT CAA AAG AAG ACT ACA TAA GAG ACCACTTCAC CAATAAGTAT
TTC GTT TTC TTA CCC TTT TTT GTA GTT TTT TTC TTT CAA GAC TTT CAA GCT TTT AGA GTT GCA AGA GCA GTT TTC TTC TGA TGT ATT CTC TGGTGAAGTG GTTATTCATA
Lys Gln Lys Asn Gly Lys Lys His Gln Lys Lys Lys Val Leu Lys Val Arg Lys Ser Gln Arg Ser Arg Gln Lys Lys Thr Thr ***

     450        460        470        480        490        500        510        520        530        540        550        560        570
      *  *       *  *       *  *       *  *       *  *       *  *       *  *       *  *       *  *       *  *       *  *       *  *
TCTGTGTTAA AAATGTTCTA TTTTAATTAT ACCGCTATCA TTCCAAAGGA GGATGGCATA TAATACAAAG GCTTATTAAT TTGACTAGAA AATTTAAAAC ATTACTCTGA AATTGTAACT AAAGTTAGAA
AGACACAATT TTTACAAGAT AAAATTAATA TGGCGATAGT AAGGTTTCCT CCTACCGTAT ATTATGTTTC CGAATAATTA AACTGATCTT TTAAATTTTG TAATGAGACT TTAACATTGA TTTCAATCTT

     580        590        600        610        620        630        640        650        660        670        680        690        700
      *  *       *  *       *  *       *  *       *  *       *  *       *  *       *  *       *  *       *  *       *  *       *  *
AGTTGATTTT AAGAATCCAA ACGTTAAGAA TTGTTAAAGG CTATGATTGT CTTTGTTCTT CTACCACCCA CCAGTTGAAT TTCATCATGC TTAAGGCCAT GATTTTAGCA ATACCCATGT CTACACAGAT
TCAACTAAAA TTCTTAGGTT TGCAATTCTT AACAATTTCC GATACTAACA GAAACAAGAA GATGGTGGGT GGTCAACTTA AAGTAGTACG AATTCCGGTA CTAAAATCGT TATGGGTACA GATGTGTCTA

     710        720        730        740        750        760        770        780        790        800        810        820        830
      *  *       *  *       *  *       *  *       *  *       *  *       *  *       *  *       *  *       *  *       *  *       *  *
GTTCACCCAA CCACATCCCA CTCACAACAG CTGCCTGGAA GAGCAGCCCT AGGCTTCCAC GTACTGCAGC CTCCAGAGAG TATCTGAGGC ACATGTCAGC AAGTCCTAAG CCTGTTAGCA TGCTGGTGAG
CAAGTGGGTT GGTGTAGGGT GAGTGTTGTC GACGGACCTT CTCGTCGGGA TCCGAAGGTG CATGACGTCG GAGGTCTCTC ATAGACTCCG TGTACAGTCG TTCAGGATTC GGACAATCGT ACGACCACTC

     840        850        860        870        880        890        900        910        920        930        940        950        960
      *  *       *  *       *  *       *  *       *  *       *  *       *  *       *  *       *  *       *  *       *  *       *  *
CCAAGCAGTT TGAAATTGAG CTGGACCTCA CCAAGCTGCT GTGGCCATCA ACCTCTGTAT TTGAATCAGC CTACAGGGCT CACACACAAT GTGTCTGAGA GATTCATGCT GATTGTTATT GGGTATCACC
GGTTCGTCAA ACTTTAACTC GACCTGGAGT GGTTCGACGA CACCGGTAGT TGGAGACATA AACTTAGTCG GATGTCCGGA GTGTGTGTTA CACAGACTCT CTAAGTACGA CTAACAATAA CCCATAGTGG

     970        980        990       1000       1010       1020       1030       1040       1050       1060       1070       1080       1090
      *  *       *  *       *  *       *  *       *  *       *  *       *  *       *  *       *  *       *  *       *  *       *  *
ACTGCAGATC ACCAGTGTGT GGCTTTCAGA GCCTCCTTTC TGGCTTTGGA AGCCATGTGA TTCCATCTTG CCCGCTCAGG CTGACCACTT TATTTCTTTT TGTTCCCCTT TGCTTCATTC AAGTCAGCTC
TGACCTCTAG TGGTCACACA CCGAAAGTCT CGGAGGAAAG ACCGAAACCT TCGGTACACT AAGGTAGAAC GGGCGAGTCC GACTGGTGAA ATAAAGAAAA ACAAGGGGAA ACGAAGTAAG TTCAGTCGAG

     1100       1110       1120       1130       1140       1150       1160       1170       1180       1190       1200       1210       1220
      *  *       *  *       *  *       *  *       *  *       *  *       *  *       *  *       *  *       *  *       *  *       *  *
TTCTCCATCC TACCACAATG CAGTGCCTTT CTTCTCTCCA GTGCACCTGT CATATGCTCT GATTTATCTG AGTCAACTCC TTTCTCATCT TGTCCCCAAC ACCCCACAGA AGTGCTTTCT TCTCCCAATT
AAGAGGTAGG ATGGTGTTAC GTCACGGAAA GAAGAGAGGT CACGTGGACA GTATACGAGA CTAAATAGAC TCAGTTGAGG AAAGAGTAGA ACAGGGGTTG TGGGGTGTCT TCACGAAAGA AGAGGGTTAA

     1230       1240       1250       1260       1270       1280       1290       1300       1310       1320       1330       1340       1350
      *  *       *  *       *  *       *  *       *  *       *  *       *  *       *  *       *  *       *  *       *  *       *  *
CATCCTCACT CAGTCCAGCT TAGTTCAAGT CCTGCCTCTT AAATAAACCT TTTTGGACAC ACAAATTATC TTAAAACTCC TGTTTCACTT GGTTCAGTAC CACATGGGTG AACACTCAAT GGTTAACTAA
GTAGGAGTGA GTCAGGTCGA ATCAAGTTCA GGACGGAGAA TTTATTTGGA AAAACCTGTG TGTTTAATAG AATTTTGAGG ACAAAGTGAA CCAAGTCATG GTGTACCCAC TTGTGAGTTA CCAATTGATT

     1360       1370       1380       1390       1400       1410       1420       1430       1440       1450       1460       1470       1480
      *  *       *  *       *  *       *  *       *  *       *  *       *  *       *  *       *  *       *  *       *  *       *  *
TTCTTGGGTG TTTATCCTAT CTCTCCAACC AGATTGTCAG CTCCTTGAGG GCAAGAGCCA CAGTATATTT CCCTGTTTCT TCCACAGTGC CTAATAATAC TGTGGAACTA GGTTTTAATA ATTTTTTAAT
AAGAACCCAC AAATAGGATA GAGAGGTTGG TCTAACAGTC GAGGAACTCC CGTTCTCGGT GTCATATAAA GGGACAAAGA AGGTGTCACG GATTATTATG ACACCTTGAT CCAAAATTAT TAAAAAATTA

     1490       1500       1510       1520       1530       1540       1550       1560       1570       1580       1590       1600       1610
      *  *       *  *       *  *       *  *       *  *       *  *       *  *       *  *       *  *       *  *       *  *       *  *
TGATGTTGTT ATGGGCAGGA TGGCAACCAG ACCATTGTCT CAGAGCAGGT GCTGGCTCTT TCCTGGCTAC TCCATGTTGG CTAGCCTCTG GTAACCTCTT ACTTATTATC TTCAGGACAC TCACTACAGG
ACTACAACAA TACCCGTCCT ACCGTTGGTC TGGTAACAGA GTCTCGTCCA CGACCGAGAA AGGACCGATG AGGTACAACC GATCGGAGAC CATTGGAGAA TGAATAATAG AAGTCCTGTG AGTGATGTCC

     1620       1630       1640       1650       1660       1670       1680       1690       1700       1710       1720       1730       1740
      *  *       *  *       *  *       *  *       *  *       *  *       *  *       *  *       *  *       *  *       *  *       *  *
GACCAGGGAT GATGCAACAT CCTTGTCTTT TTATGACAGG ATGTTTGCTC AGCTTCTCCA ACAATAAGAA GCACGTGGTA AAACACTTGC GGATATTCTG GACTGTTTTT AAAAAATATA CAGTTTACCG
CTGGTCCCTA CTACGTTGTA GGAACAGAAA AATACTGTCC TACAAACGAG TCGAAGAGGT TGTTATTCTT CGTGCACCAT TTTGTGAACG CCTATAAGAC CTGACAAAAA TTTTTTATAT GTCAAATGGC

     1750       1760       1770       1780       1790       1800       1810       1820       1830       1840       1850       1860       1870
      *  *       *  *       *  *       *  *       *  *       *  *       *  *       *  *       *  *       *  *       *  *       *  *
AAAATCATAT AATCTTACAA TGAAAAGGAC TTTATAGATC AGCCAGTGAC CAACCTTTTC CCAACCATAC AAAAATTCCT TTTCCCGAAG GAAAAGGGCT TTCTCAATAA GCCTCAGCTT TCTAAGATCT
TTTTAGTATA TTAGAATGTT ACTTTTCCTG AAATATCTAG TCGGTCACTG GTTGGAAAAG GGTTGGTATG TTTTTAAGGA AAAGGCTTC CTTTTCCCGA AAGAGTTATT CGGAGTCGAA AGATTCTAGA

     1880       1890       1900       1910       1920       1930       1940       1950       1960       1970       1980       1990       2000
      *  *       *  *       *  *       *  *       *  *       *  *       *  *       *  *       *  *       *  *       *  *       *  *
AACAAGATAG CCACCGAGAT CCTTATCGAA ACTCATTTTA GGCAAATATG AGTTTTATTG TCCGTTTACT TGTTTCAGAG TTTGTATTGT GATTATCAAT TACCACACCA TCTCCCATGA AGAAAGGAAA
TTGTTCTATC GGTGGCTCTA GGAATAGCTT TGAGTAAAAT CCGTTTATAC TCAAAATAAC AGGCAAATGA ACAAAGTCTC AAACATAACA CTAATAGTTA ATGGTGTGGT AGAGGGTACT TCTTTCCCTT

     2010       2020       2030       2040       2050       2060       2070       2080       2090       2100       2110       2120       2130
      *  *       *  *       *  *       *  *       *  *       *  *       *  *       *  *       *  *       *  *       *  *       *  *
CGGTGAAGTA CTAAGCGCTA GAGGAAGCAG CCAAGTCGGT TAGTGGAAGC ATGATTGGTG CCCAGTTAGC CTCTGCAGGA TGTGGAAACC TCCTTCCAGG GGAGGTTCAG TGAATTGTGT AGGAGAGGTT
GCCACTTCAT GATTCGCGAT CTCCTTCGTC GGTTCAGCCA ATCACCTTCG TACTAACCAC GGGTCAATCG GAGACGTCCT ACACCTTTGG AGGAAGGTCC CCTCCAAGTC ACTTAACACA TCCTCTCCAA

     2140       2150       2160       2170       2180       2190       2200       2210       2220       2230       2240       2250       2260
      *  *       *  *       *  *       *  *       *  *       *  *       *  *       *  *       *  *       *  *       *  *       *  *
GTCTGTGGCC AGAATTTAAA CCTATACTCA CTTTCCCAAA TTGAATCACT GCTCACACTG CTGATGATTT AGAGTGCTGT CCGGTGGAGA TCCCACCCGA ACGTCTTATC TAATCATGAA ACTCCCTAGT
CAGACACCGG TCTTAAATTT GGATATGAGT GAAAGGGTTT AACTTAGTGA CGAGTGTGAC GACTACTAAA TCTCACGACA GGCCACCTCT AGGGTGGGCT TGCAGAATAG ATTAGTACTT TGAGGGATCA

     2270       2280       2290       2300       2310       2320       2330       2340       2350       2360       2370       2380       2390
      *  *       *  *       *  *       *  *       *  *       *  *       *  *       *  *       *  *       *  *       *  *       *  *
TCCTTCATGT AACTTCCCTG AAAAATCTAA GTGTTTCATA AATTTGAGAG TCTGTGACCC ACTTACCTTG CATCTCACAG GTAGACAGTA TATAACTAAC AACCAAAGAC TACATATTGT CACTGACACA
AGGAAGTACA TTGAAGGGAC TTTTTAGATT CACAAAGTAT TTAAACTCTC AGACACTGGG TGAATGGAAC GTAGAGTGTC CATCTGTCAT ATATTGATTG TTGGTTTCTG ATGTATAACA GTGACTGTGT

     2400       2410       2420       2430       2440       2450       2460       2470       2480       2490       2500       2510       2520
      *  *       *  *       *  *       *  *       *  *       *  *       *  *       *  *       *  *       *  *       *  *       *  *
CACGTTATAA TCATTTATCA TATATATACA TACATGCATA CACTCTCAAA GCAAATAATT TTTCACTTCA AAACAGTATT GACTTCTATA CCTTGTAATT TGAAATATTT TCTTTGTTAA AATAGAATGG
GTGCAATATT AGTAAATAGT ATATATATGT ATGTACGTAT GTGAGAGTTT CGTTTATTAA AAAGTGAAGT TTTGTCATAA CTGAACATAT GGAACATTAA ACTTTATAAA AGAAACAATT TTATCTTACC

     2530       2540
      *  *       *  *
TATCAATAAA TAGACCATTA ATCAG
ATAGTTATTT ATCTGGTAAT TAGTC
```

Figure 1. The Hu*mig* cDNA sequence and inferred amino acid sequence of HuMig. The nucleotide sequence is that of the single cDNA clone H-1-3. Both strands of cDNA H-1-3 were sequenced completely. The 42-residue poly(A) tail of clone H-1-3 is not shown. The inferred amino acid sequence is from the longest open reading frame that begins with a methionine, and the position of the initiator methionine was assigned based on a comparison with the sequence of the Mu*mig* cDNA.

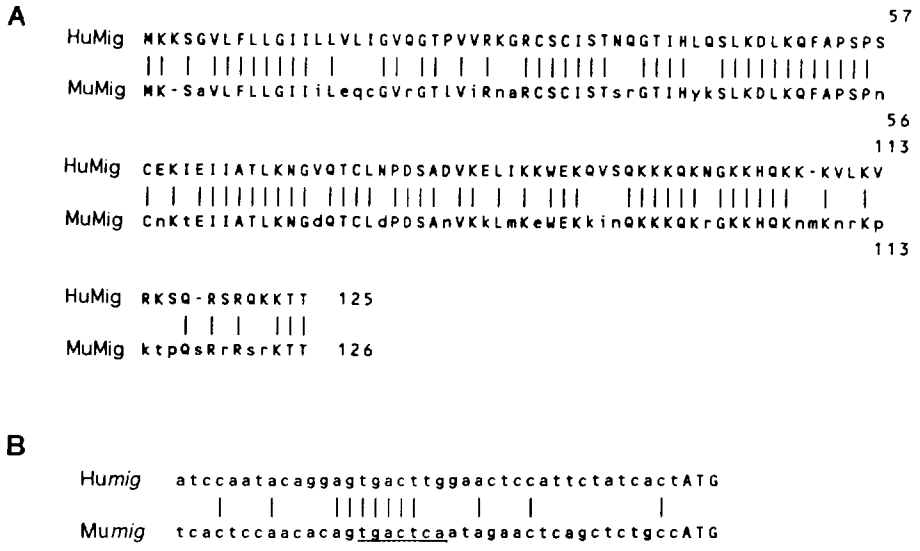

Figure 2. (A) Comparison of the inferred amino acid sequences of the HuMig and MuMig proteins. Numbers refer to the positions of the amino acid residues at the ends of the lines. Matching amino acids are connected by vertical lines and mismatches are noted by lower case letters in the MuMig sequence. (B) Comparison of 5' non-translated sequences of Mu*mig* and Hu*mig*. The sequences are aligned according to the ATG's (capitalized) corresponding to the codons for the presumed initiator methionines, and extend 5' as far as the end of the sequence available for Hu*mig*. Matches between the sequences are marked with vertical lines and the AP-1 site in Mu*mig* is underlined.

AGTAAA Mu*mig* polyadenylylation sequence (2) is not conserved in the human sequence. As regards the Hu*mig* and Mu*mig* 5' non-translated regions, Fig. 2B shows a comparison of these sequences aligned with respect to the presumed initiator ATG codons. There is notable conservation of the palindromic sequence AGTGACT. Part of this sequence, TGACT, corresponds to the five, 5' nucleotides in an AP-1 site (14), with a complete AP-1 site present in the Mu*mig* but not in the Hu*mig* sequence. The characteristics and strict conservation of the AGTGACT sequence provide presumptive evidence that the sequence serves as a recognition site for a DNA binding protein, although its presence in the *mig* mRNAs raises other possibilities as well. Except for the center base pair, the *mig* sequence matches a sequence found in the core of the X box (AGTAACT) in the promoter of the class II DQα-gene (15). The X boxes, identifiable as conserved, similar but non-identical elements in the class II MHC genes, are critical for induction of these genes by IFN–γ (16). In at least one experimental system, sequences 3' to the start site for transcription of the Mu*mig* gene were not required to confer responsiveness to IFN-γ on a heterologous promoter (17). Nonetheless, sequences outside the 5' flanking regions of the *mig* genes might be important in the overall response to IFN-γ, such as the sustained elevation in the levels of Mu*mig* mRNA seen with continued treatment with IFN-γ (18). To investigate further the relationship between the Hu*mig* and the Mu*mig* genes, Southern filters prepared using restricted human DNA were hybridized both with the Hu*mig* cDNA probe and with the Mu*mig* cDNA probe under conditions of reduced stringency. The Hu*mig* probe hybridized with a single band in each sample of digested human DNA (Fig. 3), suggesting that Hu*mig* is a single copy

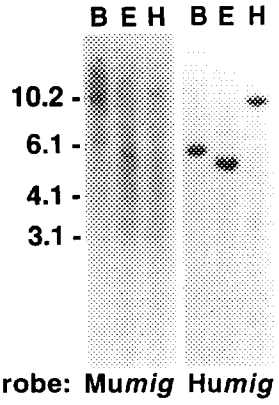

Figure 3. Human DNA probed with Mu*mig* and Hu*mig*. Human genomic DNA was digested with BamHI or EcoRI or Hind III and 20 μg were loaded per lane, separated on a 1% agarose gel and transferred to supported nitrocellulose as described (11). The filter was cut, and using either a Mu*mig* or Hu*mig* cDNA probe, portions of the filter were hybridized and washed under conditions of reduced stringency as described in Materials and Methods. The autoradiograms were aligned according to the sample wells of the gel. The sizes and positions of DNA markers in kilobase pairs are noted on the left. Exposures were for 20 h.

gene. The same Hu*mig* bands were detected using the mouse Mu*mig* probe, suggesting that the human gene most closely related to Mu*mig* is in fact Hu*mig*.

The induction of the Hu*mig* gene was evaluated in the THP-1 cell line and in human, monocyte-enriched PBMC. Fig. 4 demonstrates that the Hu*mig* mRNA, approximately 3 kb long as compared to RNA standards (not shown) is inducible dramatically in the THP-1 human monocytic cells after 8 h of treatment with IFN-γ but not significantly by IFN-α or by LPS, and that induction by IFN-γ is not prevented by the protein synthesis inhibitor cycloheximide. The results as shown using IP-10 and TNF-α cDNA probes demonstrated that the absence of induction of Hu*mig* was not due to the cells' failure to respond to IFN-α or LPS. The aldolase mRNA functioned as a constitutively expressed and non-induced control. Fig. 5 demonstrates similar results with human monocyte-enriched PBMC, showing induction in response to IFN-γ without significant induction by IFN-α or by LPS. The electrophoretic mobilities of the Hu*mig* mRNA species from THP-1 cells and from PBMC were indistinguishable.

The selective induction of Hu*mig* by IFN-γ suggests that Hu*mig* may mediate a portion of the physiological and/or pharmacological activities of IFN-γ such as anti-viral or anti-growth effects or the activation of macrophages for the production of reactive oxygen species important for the killing of tumor cells and intracellular pathogens (19). Experiments using reverse transcriptase-polymerase chain reaction to analyze the expression of cytokine mRNAs in the brains of human immunodeficiency virus-infected and control patients have shown, not surprisingly, that there is a positive correlation between the levels of Hu*mig* and IFN-γ mRNAs (20). Additional studies to evaluate the expression of Hu*mig* in conditions where IFN-γ is known or suspected to be involved will help direct investigations into the activities of the HuMig cytokine, investigations aimed at determining the role of HuMig in human biology.

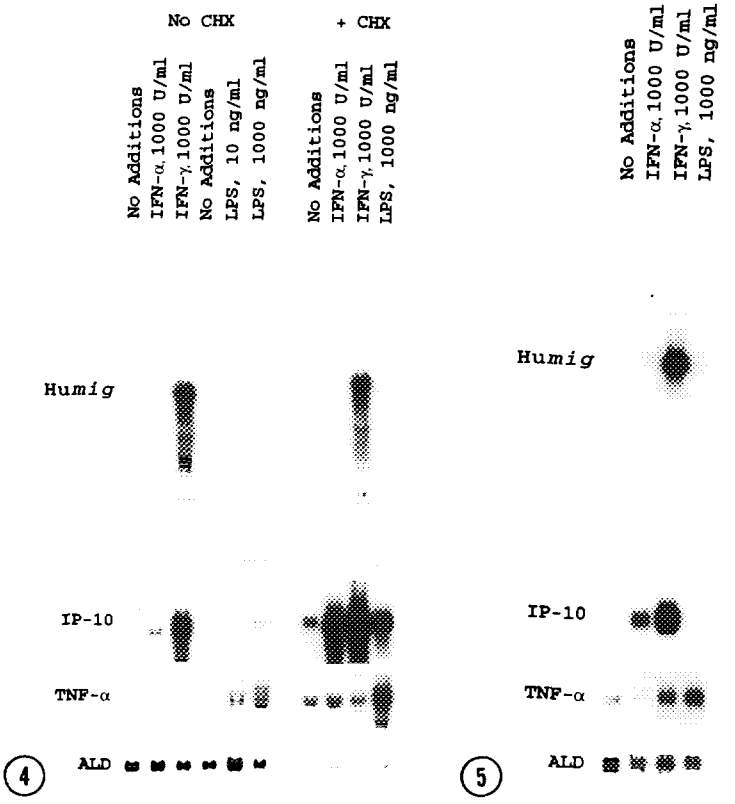

Figure 4. Hu*mig* mRNA in THP-1 cells treated with IFN-α and -γ and LPS. Total RNA was prepared from THP-1 cells that had been treated for 8 h as noted and 20 μg were loaded per lane, fractionated in a 1.2% agarose/formaldehyde gel, transferred to nitrocellulose and hybridized sequentially to cDNA probes for Hu*mig*, IP-10, TNF-α, and aldolase A (ALD). Cycloheximide (CHX) was used at a concentration of 50 μg/ml and was added 15 minutes before the other treatments. The IFN and LPS experiments were done separately and a No Additions control without CHX is shown for each experiment. The single No Additions control with CHX was done as part of the IFN experiment. The blot was stripped between hybridizations according to the manufacturer's specifications (Schleicher & Schuell). The size of the Hu*mig* mRNA was estimated at 3 kilobases on the basis of RNA markers (BRL) not shown. The Hu*mig* signal is from a 12 h exposure.

Figure 5. Hu*mig* mRNA in monocyte-enriched human PBMC treated with IFN-α and -γ and LPS. Total RNA was prepared from monocyte-enriched PBMC that had been treated for 6 h as noted and 3 μg were loaded per lane and analyzed as described in the legend to Figure 4. The cells were prepared using gelatin-coated plates according to the method of Jones et al (4). The mobility of the Hu*mig* mRNA from PBMC was indistinguishable from the mobility of the Hu*mig* mRNA from THP-1 cells. The Hu*mig* signal is from a 24 h exposure.

Acknowledgments

Thanks are due for extraordinary support to Daniel Nathans, Hyun Shin, Chaim Kahana and Thomas Moench.

I am indebted as well to Leonidas Koniaris, Anthony Lanahan, Eran Rom, Se-Jin Lee, Keith Peden, Padmavathy Vanguri, Steve Wesselingh, Timothy Wright, Diane Griffin, and James Kadonaga. Special thanks go to Christine Moss for work on the figures and the manuscript.

The work was supported by grants CA-48059 and CA-52001 from the National Cancer Institute.

References

1. Vanguri, P. and Farber, J.M. (1990) J. Biol. Chem. 265:15049-15057.
2. Farber, J.M. (1990) Proc. Natl. Acad. Sci. USA 87:5238-5242.
3. Oppenheim, J. J., Zachariae, C. O. C., Mukaida, N., and Matsushima, K. (1991) Ann. Rev. Immunol. 9:617-648.
4. Jones, B. M., Nicholson, J. K. A., Holman, R. C. and Hubbard, M. (1989) J. Immunol. Meth. 125:41-47.
5. Peden, K.W.C., Mounts, P. and Hayward, G. S. (1982) Cell 31:71-80.
6. Sanger, F., Nicklen, S. and Coulson, A. R. (1977) Proc. Natl. Acad. Sci. USA 74:5463-5467.
7. Goldberg, D. A. (1980) Proc. Natl. Acad. Sci. USA 77:5794-5798.
8. Feinberg, A. P., and Vogelstein, B. (1983) Anal. Biochem. 132:6-13.
9. Luster, A. D., Unkeless, J. C. and Ravetch, J. V. (1985) Nature (Lond.). 315:672-676.
10. Pennica, D., Nedwin, G. E., Hayflick, J. S., Seeburg, P. H., Derynck, R., Palladino, M. A., Kohr, W. J., Aggarwal, B. B., and Goeddel, D. V. (1984) Nature (Lond.). 312:724-729.
11. Lanahan, A., Williams, J.B., Sanders, L.K. and Nathans, D. (1992) Mol. and Cell. Biol. 12:3919-3927.
12. Phillips, M. A., Coffino, P. and Wang, C.C. J. Biol. Chem. (1987) 262:8721-8727.
13. von Heijne, G. (1986) Nucl. Acids. Res. 14:4683-4690.
14. Lee, W., Mitchell, P. and Tjian, R. (1987) Cell 49:741-752.
15. Ono, S. J., Liou, H-C., Davidon, R., Strominger, J. L. and Glimcher, L. H. (1991) Proc. Natl. Acad. Sci. USA. 88:4309-4312.
16. Yang, Z., Sugawara, M., Ponath, P. D., Wessendorf, L., Banerji, J., Li, Y. and Strominger, J. L. (1990) Proc. Natl. Acad. Sci. USA. 87:9226-9230.
17. Wright, T. M., and Farber, J. M. (1991) J. Exp. Med. 173:417-422.
18. Farber, J. M. (1992) Mol. and Cell. Biol. 12:1535-1545.
19. Nathan, C. F., Prendergast, T. J., Wiebe, M. E., Stanley, E. R., Platzer, E., Remold, H. G., Welte, K., Rubin, B. Y. and Murray, H. W. (1984) J. Exp. Med. 160:600-605.
20. Wesselingh, S.C., Powers, C.C., Glass, J., Tyor, W.R., McArthur, J.C., Farber, J.M., Griffin, J.W., and Griffin, D.E. (1993) Annals Neurol., in press.