Brief Definitive Report

CLONING AND CHARACTERIZATION OF A cDNA FOR MURINE MACROPHAGE INFLAMMATORY PROTEIN (MIP), A NOVEL MONOKINE WITH INFLAMMATORY AND CHEMOKINETIC PROPERTIES

By GEORGE DAVATELIS,* PATRICIA TEKAMP-OLSON,‡ STEPHEN D. WOLPE,* KATHY HERMSEN,‡ CHRISTINA LUEDKE,* CAROL GALLEGOS,‡ DORIS COIT,‡ JAMES MERRYWEATHER,‡ and ANTHONY CERAMI*

*From the *Laboratory of Medical Biochemistry, The Rockefeller University, New York, New York 10021; and ‡Chiron Corporation, Emeryville, California 94608*

Macrophages secrete a wide variety of proteins that mediate many aspects of acute and chronic inflammation (for review see reference 1). While some of these factors have been well characterized (e.g., IL-1 [2] and cachectin/TNF [3]), others remain poorly defined. Recently, we described the purification and characterization of a new monokine found in the culture medium of an LPS-stimulated mouse macrophage tumor cell line (RAW264.7) (4). This protein, termed macrophage inflammatory protein or MIP, has several properties indicative of an endogenous mediator of inflammation (e.g., neutrophil attraction and activation). Since MIP represents an important new addition to the family of activated macrophage products, it is important to investigate its structure and regulation on the molecular level. Here we describe the cloning and sequencing of the cDNA for murine MIP.

Materials and Methods

Construction of the cDNA Library. RAW264.7 cells were obtained from American Type Culture Collection (Rockville, MD) and grown in RPMI 1640 (Gibco Laboratories, Grand Island, NY) supplemented with 20 mM Hepes and 10% FCS (HyClone Laboratories, Logan, UT) until they reached confluency. The cells were then washed five times in HBSS (Gibco Laboratories) and the medium was replaced with serum-free RPMI supplemented with 1 μg/ml of LPS W (*Escherichia coli* 0127:B8, Difco Laboratories, Detroit, MI). The cells were incubated at 37°C for 2 h and total RNA was extracted by the addition of 6 M guanidinium thiocyanate (5). Poly(A)$^+$ RNA was then isolated by two cycles of oligo-dT-cellulose chromatography, essentially as described by Maniatis et al. (6). Double-stranded cDNA was prepared from the poly(A)$^+$ selected RNA as described by Gubler and Hoffman (7). After methylation of the internal Eco RI sites and addition of Eco RI linkers, the cDNA was inserted into the Eco RI sites of the bacteriophage λgt10 (8).

Construction of the Probe Pools. Oligonucleotide probe pools were synthesized as described by Warner et al. (9) against amino acids 22–30 of a partial NH_2-terminal sequence. This portion of the polypeptide was selected because of its lower degeneracy in the codon dictionary when compared with the remainder of the sequence. The resulting probe pools are two 512-fold degenerate pools of 26 nucleotides in length.

Screening the Library. Nitrocellulose filter lifts of a low-density plating (5×10^3 PFU/plate) of the library were hybridized using the synthetic probe pool that had been 5′-end labeled

```
AA#22                    #30
  GLNPHEILEVALASPTYRPHEGLUTHR      Protein

    A  T        T  T  T  A
  CA+TT+ATXGTXGA+TA+TT+GA+ACX      cDNA
    G  C        C  C  C  G

    T  A        A  A  A  T
  GT+AA+TAXCAXCT+AT+AA+CT+TGX      Coding
    C  G        G  G  G  C

     T  A  A  A        A
5'-GT+TC+AA+TA+TCXACXAT+AATTG-3'   Probe 1
     C  G  G  G        G

     T  A  A  A        A
5'-GT+TC+TT+TA+TCXACXAT+AACTG-3'   Probe 2
     C  G**G  G        G  *
```

FIGURE 1. 512-fold degeneracy pools for MIP. The asterisk below the base indicates a constant base change between the two probe pools.

with [^{32}P] ATP (New England Nuclear, Boston, MA). After the hybridization, the lifts were washed using the method of Wood et al. (10). After several rounds of screening, 18 recombinant phage clones were isolated and grown in bulk for DNA isolation.

DNA Sequence Analysis. The cDNA inserts to be analyzed were subcloned into the M13 phage vectors and DNA sequencing was performed by the dideoxy-chain termination method of Sanger et al. (11).

Blot Hybridization Analysis. Northern blot hybridization was performed according to the method of Lehrach et al. (12). Total RNA of LPS-stimulated and nonstimulated RAW 264.7 cells were electrophoresed through 1.2% agarose gels and transferred to nitrocellulose filters.

Primer Extension. The synthetic oligonucleotide primer was end labeled using γ-[^{32}p]ATP (3,000 Ci/mmol, Amersham Corp., Arlington Heights, IL) and the T4 polynucleotide kinase. The primer extension method was a modification of that described by Walker et al. (13).

Results and Discussion

To elucidate the molecular structure of murine MIP (MuMIP) a cDNA clone was isolated containing the sequence coding for MuMIP. As a first step, the mouse macrophage cell line RAW264.7 was stimulated with LPS. Since RAW264.7 cells have been shown to be a source of MIP protein after LPS stimulation, the MIP mRNA was expected to be highly reiterated in these cells 2 h after LPS induction of these cells. Poly(A)$^+$ RNA was prepared from total RNA by two cycles of oligo-dT-chromatography and a cDNA library was constructed in λgt10. The cloning efficiency was 10^6 clones/μg of poly(A)$^+$. The library was amplified and shown to contain >1,000 bp inserts in ~60% of recombination plaques. Nitrocellulose filter lifts of a low-density plating of the library were screened using two synthetic oligonucleotide pools that were based on a partial NH_2-terminal amino acid sequence of purified protein (4). Each pool consisted of a 512-fold degeneracy pool of 26 nucleotides in length (Fig. 1). After the initial library screening, positive plaques were streaked onto fresh bacterial lawns and a secondary screening was performed by differential plaque hybridization. Two replicate lifts of the secondary streaks were hybridized to either ^{32}P-labeled pool 1 or pool 2. Since the melting temperature (T_m)

```
AAGCAGCAGCGAGTACCAGTCCCTTTTCTGTTCTGCTGACAAGCTCACCCTCTGTCACCTG

         -23          -20
         Met Lys Val Ser Thr Thr Ala Leu Ala Val Leu Leu Cys
CTCAACATC ATG AAG GTC TCC ACC ACT GCC CTT GCT GTT CTT CTC TGT

                                                  1
Thr Met Thr Leu Cys Asn Gln Val Phe Ser Ala Pro Tyr Gly Ala Asp
ACC ATG ACA CTC TGC AAC CAA GTC TTC TCA GCG CCA TAT GGA GCT GAC

          10                                                  20
Thr Pro Thr Ala Cys Cys Phe Ser Tyr Ser Arg Lys Ile Pro Arg Gln
ACC CCG ACT GCC TGC TGC TTC TCC TAC AGC CGG AAG ATT CCA CGC CAA

                              30
Phe Ile Val Asp Tyr Phe Glu Thr Ser Ser Leu Cys Ser Gln Pro Gly
TTC ATC GTT GAG TAT TTT GAA ACC AGC AGC CTT TGC TCC CAG CCA GGT

    40                                                  50
Val Ile Phe Leu Thr Lys Arg Asn Arg Gln Ile Cys Ala Asp Ser Lys
GTC ATT TTC CTG ACT AAG AGA AAC CGG CAG ATC TGC GCT GAC TCC AAA

                         60                                   69
Glu Thr Trp Val Gln Glu Tyr Ile Thr Asp Leu Glu Leu Asn Ala OP
GAG ACC TGG GTC CAA GAA TAC ATC ACT GAC CTG GAA CTG ATT GCC TGA

GAGTCTTGGAGGCAGCGAGGAACCCCCCAAACCTCCATGGGTCCCGTGTAGAGCAGGGGCTTGAGC

CGAACATTCCTGCCACCTGCATAGCTCCATCTCCTATAAGCTGTTTGCTGCCAAGTAGCCACATCG

AGGGACTCTTCACTTGAAATTTTATTTAATTTAATCCTATTGGTTTAATACTATTTAATTTTGTAA

TTTATTTTATTGTCATACTTGTATTTGTGACTATTTATTCTGAAAGACTTCAGGACACGTTCCTCA

ACCCCCATCTCCCTCCCAGTTGTTCACACTGTTTGGTGACAGCTATTCTAGGTAGACATGATGACA

AAGTCATGAACTGACAATGTACAATAGATGCTTTGTTTATACCAGAGAAGTAATAAATATGCCCTT

TAACAAGTAAAAAA
```

FIGURE 2. The complete nucleotide sequence of the cDNA clone for MIP is shown. The underlined sequence indicates the complementary sequence of oligonucleotide used in the primer extension experiments. The predicted translated molecular weight is 10,346. The mature protein sequence, starting at position one, is 69 amino acids in length and has a predicted molecular weight of 7,889.

of DNA/DNA hybrids can be approximated by the empirical formula: T_m = 16.6(log[Na+]+0.41(%[G+C])+81.5−500/number of bp in homology, we could effectively eliminate one of the probe pools through the differential melting temperatures of the hybrids based on a 26-bp homology. By using the tetramethylammonium chloride washing technique of Wood et al. (10), which abolishes the preferential melting of A-T vs. G-C base pairs, the T_m becomes dependent simply on the length of the hybrid.

After several rounds of differential hybridization, probe pool 1 yielded 18 recombinant phage clones out of 10^4 screened that hybridized under maximally stringent conditions for MIP. All of the plaques were purified and DNA was prepared from each. The recombinant phage clone 52 appeared to contain the largest cDNA Eco RI insert of ~750 bp and was chosen for further characterization. The complete nucleotide sequence of cDNA clone 52, as well as 262 bp of 5′ sequence of another, partially overlapping clone, 32, have been determined and are shown in Fig. 2. The latter clone, which was isolated in a later screening, had a smaller Eco RI insert than clone 52, but a larger 5′-end fragment and therefore presumably less poly(A)$^+$ tail. The MIP nucleotide sequence of 763 bp predicts a single open reading frame starting at nucleotide 2. The mature protein sequence, starting at position one, is 69 amino acids in length and encodes the major sequence previously defined by NH_2-terminal analysis of the purified protein (4).

The first methionine present in the sequence is found at position −23. We postulate this to be the initiating methionine for the MIP precursor based on the following observations. Structural analysis of the putative presequence (−23 to −1) indicates that it has features characteristic of signal sequences (i.e., α-helix and a hydrophobic core[14]). The predicted initiating ATG has a purine at position −3, which has been shown by Kozak (15) to have a dominant effect on translation initiation efficiency. Furthermore, in a survey of the frequency of A,C,T,G around the translation start

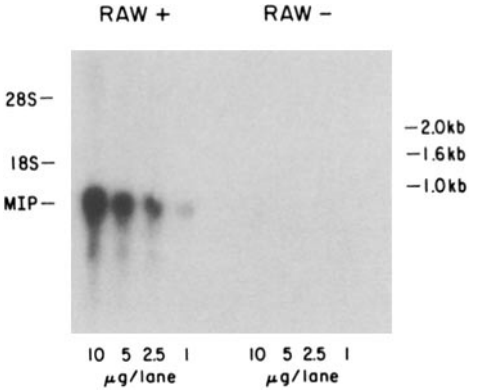

FIGURE 3. Northern blot of total RNA from RAW264.7 cells, a transformed murine macrophage cell line. The cells were stimulated with LPS at 2 μg/ml in serum-free medium for 6 h. Control cells were also given serum-free medium in the absence of LPS for 6 h. The amount of total RNA loaded in each lane is indicated on the figure.

site of 699 vertebrate mRNAs, 97% had a purine at position −3, 61% having an A at that position (16).

We have also performed a primer extension analysis to determine the amount of 5′ sequence lacking from the cDNA clone. A labeled oligonucleotide primer (Fig. 2) was hybridized to LPS-stimulated RAW264.7 poly(A)$^+$ RNA and elongated with reverse transcriptase. After hybridization, an extended primer of 98 ± 2 nucleotides was obtained (data not shown). After subtracting out the primer length of 25 nucleotides and the sequence 5′ to the primer we had previously determined (61 nucleotides), we can conclude that our known sequence is 10–14 nucleotides short of a full-length cDNA. While it is possible that an in-frame AUG is present in this unknown region, it seems highly unlikely given that only 14 of 346 sequenced vertebrate mRNAs have 5′ noncoding sequences of <19 nucleotides in length (16). We would thus estimate the 5′ untranslated sequence to be ~82 nucleotides, well within the 20–100 nucleotide length of most vertebrate 5′ noncoding sequenced to date.

The proposed preMIP is 92 amino acids in length. There are no Asn-X-Ser,Thr sites for *N*-linked glycosylation evident in the molecule. There are 7 cysteines, 3 in the presequence and 4 in the mature sequence. The codon usage of the putative pre-MIP agrees well with that determined for 66 other sequenced murine genes (17). The protein has no significant sequence similarity to any protein as defined to date by the dfast-p program homology search (18) of the Dayhoff protein data base. The DNA sequence was also compared against the GenBank genetic sequence data with a similar result.

In the 3′-untranslated region there is a single consensus polyadenylation site at bp 711–716. There are also 4 sequences that have only one mismatch to the cytokine consensus 3′-untranslated sequence defined by Caput et al. (19). The 3′-untranslated consensus cytokine sequence $(TATT)_n$ defined by Reeves et al. (20) is also present. When $n = 2$ and one mismatch is allowed, four of these sequences are found. There is an overlap in three of these between the sequence defined by Caput et al. (19) and that defined by Reeves et al. (20).

Since MIP is an inducible polypeptide, we have also studied the expression of murine MIP mRNA by Northern blot hybridization in RAW264.7 cells. As shown in Fig. 3, total RNA from LPS-induced cells exhibited a positive hybridization band with an estimated size of 800 bp, while total RNA from uninduced cells showed

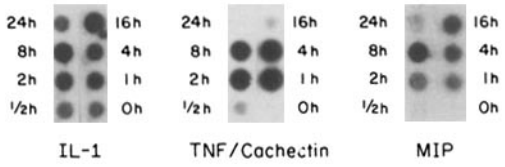

Figure 4. Time-course study of RAW264.7 mRNA induction by LPS. The time course is measured in hours. All three probes were plasmid cDNA clones labeled with α-[^{32}P] dCTP.

very little of a positive signal at that or any other size. In a time-course study on the induction of MuMIP mRNA with endotoxin in these same cells (Fig. 4), MuMIP mRNA exhibits detectable levels of mRNA within 1 h after LPS stimulation and peaks between 8–16 h after LPS stimulation. This time course is quite different from either MuTNF-α/cachectin or MuIL-1α mRNAs when their respective plasmid probes were hybridized to the same blot. Further analysis is currently underway to examine how MIP is regulated at the molecular level and how it relates to other known inflammatory mediators.

Summary

In the course of studies on cachectin/TNF being conducted in our laboratory, a novel macrophage product has been detected and characterized. Termed macrophage inflammatory protein or MIP, this protein appears to be an endogenous mediator of the inflammatory events induced by endotoxin. A cDNA cloned probe for this protein has been isolated from a λgt10 phage library prepared from poly(A)$^+$ RNA obtained of endotoxin-induced RAW264.7 cells. The sequence codes for a 92 amino acid-long polypeptide, of which 69 amino acids correspond to the mature product. The sequence predicts a molecular weight of 7,889 and structural analysis of the protein indicates a characteristic signal sequence α-helix and a hydrophobic core. Sequence data also confirm no sequence similarity to any other protein listed in the Dayhoff data base.

We would like to thank Dr. Pablo Valenzuela and Leslie Rall (Chiron Corp.) for their help with the cDNA library construction; and Dr. Donna Atherton for protein sequencing.

Received for publication 22 February 1988.

Note added in proof: We have recently become aware of an mRNA produced by human tonsillar lymphocytes in response to mitogen (21). While the sequence is not listed in the Dayhoff protein data base or the GenBank genetic sequence data base, there is a 75.3% amino acid sequence similarity with murine MIP. The relationship between these proteins is not known.

References

1. Nathan, C. F. 1987. Secretory products of macrophages. *J. Clin. Invest.* 79:319.
2. Beutler, B., and A. Cerami. 1986. Cachectin and tumour necrosis factor as two sides of the same biological coin. *Nature (Lond.).* 320:584.
3. Dinarello, C. A. 1985. An update on human interleukin 1: from molecular biology to clinical relevance. *J. Clin. Immunol.* 5:287.
4. Wolpe, S. D., G. Davatelis, B. Sherry, B. Beutler, D. G. Hesse, H. T. Nguyen, L. L. Moldawer, C. F. Nathan, S. F. Lowrey, and A. Cerami. 1987. Macrophages secrete a

novel heparin-binding protein with inflammatory and neutrophil chemokinetic properties. *J. Exp. Med.* 167:570.

5. Ullrich, A., J. Shine, J. Chirgwin, R. Pictet, E. Tischer, W. J. Rutter, and H. M. Goodman. 1977. Rat insulin genes: construction of plasmids containing the coding sequences. *Science (Wash. DC).* 196:1313.
6. Maniatis, T., E. F. Fritsch, and J. Sambrook. 1982. Molecular Cloning. A Laboratory Manual. Cold Spring Harbor Laboratory, Cold Spring Harbor, NY. 197.
7. Gubler, U., and B. J. Hoffman. 1983. A simple and very efficient method for generating cDNA libraries. *Gene (Amst.).* 25:263.
8. Huynh, T. V., R. A. Young, and R. W. Davis. 1985. DNA Cloning: A Practical Approach. D. M. Gloves, editor. IRL Press, Oxford. 49.
9. Warner, B. D., M. E. Warner, G. A. Karns, L. Ku, S. Brown-Shimer, and M. S. Urdea. 1984. Construction and evaluation of an instrument for the automated synthesis of oligodeoxyribonucleotides. *DNA (NY).* 3:401.
10. Wood, W. I., J. Gitschier, L. A. Laskey, and R. M. Lawn. 1985. Base composition-independent hybridization in tetramethylammonium chloride: a method for oligonucleotide screening of highly complex gene libraries. *Proc. Natl. Acad. Sci. USA.* 82:1585.
11. Sanger F., S. Nicklen, and R. Coulson. 1977. DNA sequencing with chain-terminating inhibitors. *Proc. Natl. Acad. Sci. USA.* 74:5463.
12. Lehrach, H., D. Diamond, J. M. Wozney, and H. Boedtker. 1977. RNA molecular weight determinations by gel electrophoresis under denaturing conditions, a critical reexamination. *Biochemistry.* 16:4743.
13. Walker, M. D., T. Edlund, A. M. Boulet, and W. J. Rutter. 1983. Cell-specific expression controlled by the 5'-flanking region of the insulin and chymotrypsin genes. *Nature (Lond.).* 306:557.
14. Perlman D., and H. O. Halvorson. 1983. A putative signal peptidase recognition site and sequence in eucaryotic and procaryotic signal peptides. *J. Mol. Biol.* 167:391.
15. Kozak, M. 1986. Point mutations define a sequence flanking the ATG initiator codon that modulates translation by eucaryotic ribosomes. *Cell.* 44:283.
16. Kozak, M. 1987. An analysis of 5'-noncoding sequences from 699 vertebrate messenger RNAs. *Nucleic Acids Res.* 15:8125.
17. Marayama, T., T. Gojobori, S. Aota, and T. Ikemura. 1986. Codon usage tabulated from the GenBank genetic sequence data. *Nucleic Acids Res.* 14(Suppl.):r151.
18. Lipman, D. J., and W. R. Pearson. 1985. Rapid and sensitive protein similarity searches. *Science (Wash. DC).* 227:1435.
19. Caput, D., B. Beutler, K. Hartog, R. Thayer, S. Brown-Shimer, and A. Cerami. 1986. Identification of a common nucleotide sequence in the 3'-untranslated region of mRNA molecules specifying inflammatory mediators. *Proc. Natl. Acad. Sci. USA.* 83:1670.
20. Reeves, R., A. G. Spies, M. S. Nissen, C. D. Buck, A. D. Weinberg, P. J. Barr, N. S. Magnuson, and J. A. Magnuson. 1986. Molecular cloning of a functional interleukin 2 cDNA. *Proc. Natl. Acad. Sci. USA.* 83:3228.
21. Obaru, K., M. Fukuda, S. Maeda, and K. Shimada. 1986. A cDNA clone used to study mRNA inducible in human tonsillar lymphocytes by a tumor promoter. *J. Biochem.* 99:885.