Vol. 159, No. 1, 1989 BIOCHEMICAL AND BIOPHYSICAL RESEARCH COMMUNICATIONS
February 28, 1989 Pages 249-255

CLONING AND SEQUENCING OF THE cDNA FOR HUMAN MONOCYTE CHEMOTACTIC AND ACTIVATING FACTOR (MCAF)

Yasuji Furutani[1]*, Hideki Nomura[1], Mitsue Notake[1], Yoshihiro Oyamada[1], Toshikazu Fukui[1], Masaaki Yamada[1], Christian G. Larsen[2], Joost J. Oppenheim[2] and Kouji Matsushima[2]

[1] Research Laboratories, Dainippon Pharmaceutical Co. Ltd., Enoki-cho 33-94, Suita/Osaka 564, Japan

[2] Laboratory of Molecular Immunoregulation, Biological Response Modifiers Program, Division of Cancer Treatment, National Cancer Institute, Frederick, Maryland 21701-1013

Received January 20, 1989

cDNA clones having a nucleotide sequence encoding a human monocyte chemotactic and activating factor (MCAF) were isolated and sequenced. The amino acid sequence deduced from the nucleotide sequence reveals the primary structure of the MCAF precursor to be composed of a putative signal peptide sequence of 23 amino acid residues and a mature MCAF sequence of 76 amino acid residues. The amino acid sequence of MCAF showed 25-55% homology with other members of an inducible cytokine family, including macrophage inflammatory protein and some putative polypeptide mediators known as JE, LD78, RANTES and TCA-3. This suggests that MCAF is a member of family of factors involved in immune and inflammatory responses.

During the course of an inflammation, factor(s) seem to be generated that are specifically chemotactic for various leukocytes that infiltrate inflammatory sites.

Recently, we achieved the molecular cloning of one of such factor, namely neutrophil chemotactic factor (NCF) by isolating cDNA from the lipopolysaccharide (LPS)-stimulated human monocyte cDNA bank (1) and subsequently succeeded in producing and purifying of this factor in large quantity by recombinant DNA technology (2).

Furthermore, we have also purified and characterized a chemotactic factor for monocytes isolated from the conditioned media from a human promyelocytic cell line, THP-1 (3). Amino acid sequencing of a cyanogen bromide (CNBr) cleaved fragment of this factor disclosed a partial amino acid sequence, indicating that this factor is a previously unidentified novel peptide. It has also been shown that this factor possesses not only monocyte chemotactic fac-

* To whom correspondence should be addressed.

0006-291X/89 $1.50

tor (MCF) activity but also monocyte activating factor (MAF) activity, such as augmentation of cytostatic effect of monocytes on several type of human tumor cells (3).

To further investigate and elucidate the functional activities and the molecular structure of this factor, we have proceeded to isolate the cDNA clone for this novel cytokine. Here, we describe the cloning and characterization of the human monocyte chemotactic and activating factor (MCAF) cDNA.

MATERIALS AND METHODS

Human promyelocytic cell lines, THP-1 and HL-60, were first stimulated with 500 ng/ml phorbol-12-myristate-13-acetate (PMA) (P-L Biochemicals, Milwaukee, WI), and with both 500 ng/ml PMA and 1 µg/ml retinoic acid (Sigma, St. Louis, MO), respectively, for two days and were subsequently induced with 10 µg/ml E. coli LPS (Difco, Detroit, MI) in the presence of 1 µg/ml cycloheximide (Sigma) for 6 h.

Total RNA was extracted (4) from cells and poly(A)$^+$RNA was isolated by oligo(dT)-cellulose chromatography (5). Prior to synthesizing cDNA, poly(A)$^+$RNA was denatured with 10 mM methylmercuric hydroxide to increase the yield of full length cDNA as described (6). Double-stranded cDNA was constructed by the method of Gubler and Hoffman (7) and inserted into the PstI site of pBR322 by homopolymer tailing (8). Two replicas of the colony library fixed onto nitrocellulose filters (9) were hybridized separately with two different sets of ^{32}P-end-labelled tetradecamer probes (see text) in 4xSSC at 36°C for 40 h. Clones hybridized with both probes were selected and plasmid DNAs were isolated. The cDNA inserts were digested with appropriate restriction enzymes and subcloned into a pUC18 or pUC19 vector. The nucleotide sequences of these subclones were determined by the dideoxy chain termination method (10, 11) using universal primers or synthetic primers complementary to the determined cDNA sequences. Oligonucleotide probes and primers were synthesized by the phosphoramidite method (12).

Northern blot hybridization was carried out according to the method of Thomas (13) except that hybridization with oligonucleotide probe was performed as described (14).

RESULTS AND DISCUSSIONS

One of the most effective approaches to cDNA cloning is to use a synthetic oligonucleotide probe synthesized based on a defined amino acid sequence for screening cDNA library. The availability of a partial amino acid sequence of human MCAF (3) enabled us to isolate the corresponding cDNA clones by this strategy.

As a first step, two groups (I and II) of two pools of tetradecamers were synthesized based on the partial amino acid sequence of MCAF (Fig. 1). The presence of MCAF mRNA in our mRNA preparations from two distinct human promyelocytic cell lines, THP-1 and HL-60, was evaluated by Northern blot hybridization with the group II oligonucleotide probe. Despite the fact that the partial amino acid sequence was determined from a CNBr cleaved fragment of MCAF purified from the conditioned media from THP-1 cells, an mRNA species of about 700 nucleotides that hybridized with the group II oligonucleotide probe was only detected in HL-60 mRNA, but not in THP-1 mRNA (data not shown). This

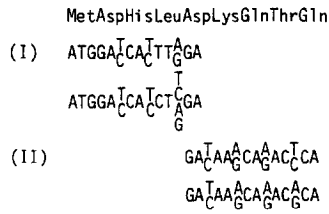


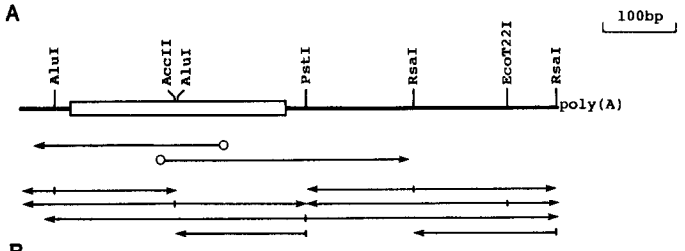
Figure 1. Synthetic oligonucleotide probes used for cloning MCAF cDNA (I and II) and for Northern blot hybridization analysis (II). Two groups (I and II) of two pools of tetradecamers represent all possible coding sequences that correspond to the indicated amino acids in the partial MCAF sequence (3).

indicated that the MCAF mRNA content in our mRNA preparation from THP-1 cells, if present at all, was extremely low. We, therefore, used HL-60 mRNA to construct the cDNA library.

Two replicas of a colony library of about 36,000 transformants constructed from the induced HL-60 mRNA were grown on nitrocellulose filters and hybridized separately with the group I or group II oligonucleotide probe. From this first screening, 35 clones specifically hybridizing with both probes were selected and recombinant plasmid DNA was isolated from each clone. Restriction enzyme mapping analysis of these clones showed that all of them contained a cDNA copy derived from a single mRNA species. From the 35 clones, three clones were selected for further analysis. Two of them, pHMCF7 and pHMCF29, contained a longer cDNA insert and another clone, pHMCF25, only contained the cDNA insert giving additional RsaI fragments in restriction analysis.

The complete nucleotide sequence of the cDNA insert of pHMCF7 was determined by dideoxy chain termination method according to the strategy as indicated in Fig. 2A and some additional nucleotide sequencing were also performed for the cDNA insert of pHMCF29 to assure the nucleotide sequence obtained and for that of pHMCF25 to define the site of the nucleotide difference which might cause generation of an additional RsaI site.

Figure 2B shows the complete nucleotide sequence of human MCAF cDNA. The nucleotide sequence was determined for both strands of the cDNA and that of the whole protein-coding region was determined with all three clones. The sequence of nucleotide residues 259-297 corresponded precisely with the amino acid sequence determined for the CNBr cleaved fragment of human MCAF (amino acid residues 87-99). Using this reading frame, a single large protein-coding sequence was identified extending from ATG initiation codon (nucleotide residues 1-3) to a termination codon TGA (nucleotide residues 298-300) which is preceded by the defined partial amino acid sequence of MCAF. The assigned ATG start codon is the first ATG triplet appearing in the cDNA sequence and the nucleotide sequence near this start codon (underlined in the figure) agrees well with the predicted consensus sequence for the translational start site in eukaryotic mRNAs (15).



B

```
                                                        AACCGAGAG   -61
GCTGAGACTAACCCAGAAACATCCAATTCTCAAACTGAAGCTCGCACTCTCGCCTCCAGC    -1

MetLysValSerAlaAlaLeuLeuCysLeuLeuLeuIleAlaAlaThrPheIleProGln   (20)
ATGAAAGTCTCTGCCGCCCTTCTGTGCCTGCTGCTCATAGCAGCCACCTTCATTCCCCAA    60

GlyLeuAlaGlnProAspAlaIleAsnAlaProValThrCysCysTyrAsnPheThrAsn   (40)
GGGCTCGCTCAGCCAGATGCAATCAATGCCCCAGTCACCTGCTGTTATAACTTCACCAAT   120
                                  C

ArgLysIleSerValGlnArgLeuAlaSerTyrArgArgIleThrSerSerLysCysPro   (60)
AGGAAGATCTCAGTGCAGAGGCTCGCGAGCTATAGAAGAATCACCAGCAGCAAGTGTCCC   180

                                  (Thr)
LysGluAlaValIlePheLysThrIleValAlaLysGluIleCysAlaAspProLysGln   (80)
AAAGAAGCTGTGATCTTCAAGACCATTGTGGCCAAGGAGATCTGTGCTGACCCCAAGCAG   240
                                  A

LysTrpValGlnAspSerMetAspHisLeuAspLysGlnThrGlnThrProLysThr      (99)
AAGTGGGTTCAGGATTCCATGGACCACCTGGACAAGCAAACCCAAACTCCGAAGACTTGA   300

ACACTCACTCCACAACCCAAGAATCTGCAGCTAACTTATTTTCCCCTAGCTTTCCCCAGA   360

CACCCTGTTTTATTTTATTATAATGAATTTTGTTTGTTGATGTGAAACATTATGCCTTAA   420

GTAATGTTAATTCTTATTTAAGTTATTGATGTTTTAAGTTTATCTTTCATGGTACTAGTG   480

TTTTTTAGATACAGAGACTTGGGGAAATTGCTTTTCCTCTTGAACCACAGTTCTACCCCT   540

GGGATGTTTTGAGGGTCTTTGCAAGAATCATTAATACAAAGAATTTTTTTTAACATTCCA   600

ATGCATTGCTAAAATATTATTGTGGAAATGAATATTTTGTAACTATTACACCAAATAAAT   660

(2) ATATTTTTGTAC                                                672 (3)
```

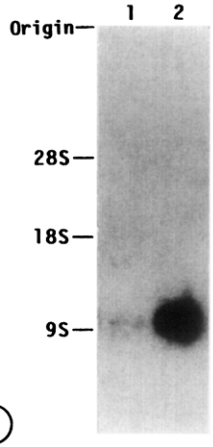


Figure 2. Nucleotide sequence of cDNA for human MCAF and deduced amino acid sequence. (A) Strategy for sequencing cDNA encoding human MCAF. The restriction map displays only relevant restriction endonuclease sites used for subcloning and sequencing. Arrows indicate the direction and extent of the sequencing. The open circles at the end of arrows indicate the positions of specific sequencing primers used. The protein-coding region is shown as a box. (B) Nucleotide sequence of the human MCAF cDNA and the deduced amino acid sequence. The nucleotide sequence was determined from those of the cDNA inserts of the clones, pHMCF7, pHMCF25 and pHMCF29. The deduced amino acid sequence is displayed above the nucleotide sequence. The predicted site for signal peptide cleavage is shown by the arrow. Two nucleotide differences observed in the nucleotide sequence of pHMCF25 are C for T at position 105 and A for G at position 226. The nucleotide difference at position 226 results in the replacement of Ala at position 76 by Thr. The nucleotide sequence similar to the consensus translational start sequence and the putative polyadenylation signal are underlined. The 3'-terminal sequence presented is followed by a poly(A) tract which is not included in the sequence.

Figure 3. Northern blot hybridization analysis of human MCAF mRNA. Ten µg of poly(A)$^+$RNA from uninduced (lane 1) and induced (lane 2) HL-60 cells were electrophoresed through 1.1% agarose gel and transferred to a nitrocellulose filter. The blot was hybridized with the ^{32}P-labelled PstI fragment of pHMCF7 which contains the entire protein-coding region. Globin mRNA (9S) was used as a size marker.

Two nucleotide differences were observed in the nucleotide sequence of the cDNA insert of pHMCF25 (see Fig. 2 legend), one of which at position 226 results in the formation of an additional RsaI recognition site and therefore in an amino acid replacement at position 76 (Thr for Ala). These observed nucleotide differences may be due to polymorphism (16, 17).

In the 3'-untranslated region, a polyadenylation signal (18), AATAAA (nucleotide residues 654-659, underlined in the figure) is located 14 nucleotides upstream from the poly(A) tract. Northern blot hybridization analysis demonstrated a hybridization positive band of about 700 nucleotides (see Fig. 3), indicating that the cDNA sequence shown in Fig. 2 represents a full length copy of the mRNA.

The amino acid sequence deduced from the nucleotide sequence reveals the primary structure of human MCAF precursor. The sequence of the first 23 amino acid residues present in the precursor molecule is presumably eliminated during secretion (19) following the concept predicted for a signal peptide cleavage site (20). After the proposed processing to remove the putative signal peptide sequence from the precursor of MCAF, a mature MCAF sequence consisting of 76 amino acid residues starting with Gln residue at position 24 is predicted. The amino acid sequence of 13 amino acid residues determined biochemically from a CNBr cleavage fragment of human MCAF is located at the carboxy terminus. The calculated molecular weight of mature MCAF thus predicted is 8,681 daltons which is considerably smaller than the value (about 15 kilodaltons) obtained from SDS-PAGE analysis of a purified natural MCAF (3). Since the predicted mature MCAF sequence contains a potential site for N-glycosylation (Asn-Phe-Thr, amino acid residues 37-39) (21), it is possible that the authentic human MCAF is a glycosylated protein. Precise reasons for this discrepancy between the estimated molecular size for purified natural MCAF by SDS-PAGE analysis and the calculated molecular size for predicted mature form of MCAF deduced from nucleotide sequence remain to be established.

To examine the MCAF mRNA expression, poly(A)$^+$RNA from induced and uninduced HL-60 cells were analyzed by Northern blot hybridization. As shown in Fig. 3, a dark hybridization-positive band of about 700 nucleotides was detected in mRNA from induced HL-60 cells, in contrast with a faint band detected in mRNA from uninduced HL-60 cells, demonstrating that the level of the expression of MCAF mRNA increased dramatically in response to LPS-stimulation. Kinetic studies of induction of MCAF mRNA with LPS or other cytokines are now in progress.

A comparison of the amino acid sequence of mature MCAF with other known proteins clearly demonstrated that there are a number of homologous polypeptides that are members of a family of proteins which are believed to be produced by activated T lymphocytes or monocytes. Figure 4 shows the align-

Figure 4. Alignment of the amino acid sequence of human MCAF (a) with those of the JE (b), LD78 (c), murine MIP (d), RANTES (e) and TCA-3 (f). Only the part of the amino acid sequence of JE (residues 24-99) is aligned. Putative signal peptide sequences are not included in the sequences. The positions occupied by the identical or chemically similar residues are marked with an asterisk or a dot, respectively. The positions of four cysteine residues conserved are shown by shadows.

ment of the amino acid sequence of mature MCAF with that of JE (22), LD78 (23), murine macrophage inflammatory protein (MIP) (24), RANTES (25) and TCA-3 (26). The degree of homology among these polypeptides, which is determined from the percentage of amino acid positions being occupied by identical matching amino acids, is 55%, 36%, 36%, 28% and 25% for the MCAF/JE, /LD78, /MIP, /RANTES and /TCA-3, respectively. Four cysteine residues are completely conserved at the same site in all six molecules (shadowed in the figure), suggesting that they are functionally important to the three dimensional structure by forming disulfide bridges. Our previous work concerning the cloning of cDNA for monocyte derived NCF (1) showed that NCF also belongs to another closely related family of inducible proteins, in which four cysteine residues are also perfectly conserved and located at a somewhat different position from that observed for the MCAF family. Significant amino acid sequence homology is also found between these two families, especially between the two chemotactic factors, MCAF and NCF (21%). These observations imply that the MCAF family and NCF family are functionally and evolutionally related.

The cloning and sequencing of cDNA for MCAF has enabled us to elucidate the complete primary structure of the factor. This will provide useful information for further studies of the biological and biochemical properties of MCAF and will provide means of isolating the corresponding genomic clone(s).

ACKNOWLEDGMENTS

Thanks are due to Drs. Masahisa Hashimoto, Michiko Yamayoshi, Mayumi Ohue and Hitoshi Kawashima for their helpful discussions.

REFERENCES

1. Matsushima, K., Morishita, K., Yoshimura, T., Lavu, S., Kobayashi, Y., Lew, W., Appella, E., Kung, H. F., Leonard, E. J., and Oppenheim, J. J. (1988) J. Exp. Med. 167, 1883-1893.
2. Furuta, R., Yamagishi, J., Kotani, H., Sakamoto, F., Fukui, T., Matsui, Y., Sohmura, Y., Yamada, M., Yoshimura, T., Larsen, C. G., Oppenheim, J. J., and Matsushima, K. (submitted)
3. Matsushima, K., Larsen, C. G., DuBois, G. C., and Oppenheim, J. J. (submitted)

4. Chirgwin, J. M., Przybyla, A. E., MacDonald, R. J., and Rutter, W. J. (1979) Biochemistry 18, 5294-5299.
5. Aviv, H., and Leder, P. (1972) Proc. Natl. Acad. Sci. U. S. A. 69, 1408-1412.
6. Maniatis, T., Fritsch, E. F., and Sambrook, J. (1982) Molecular Cloning: A Laboratory Manual, p. 231. Cold Spring Harbor Laboratory Press, Cold Spring Harbor.
7. Gubler, U., and Hoffman, B. J. (1983) Gene 25, 263-269.
8. Nelson, T., and Brutlag, D. (1979) Methods Enzymol. 68, 41-50.
9. Hanahan, D., and Meselson, M. (1980) Gene 10, 63-67.
10. Sanger, F., Nicklen, S., and Coulson, A. R. (1977) Proc. Natl. Acad. Sci. U. S. A. 74, 5463-5467.
11. Messing, J. (1983) Methods Enzymol. 101, 20-78.
12. Giles, J. G. (1985) American Biotechnology Laboratory 3, 22-27.
13. Thomas, P. S. (1983) Methods Enzymol. 100, 255-266.
14. Comb, M., Herbert, E., and Crea, R. (1982) Proc. Natl. Acad. Sci. U. S. A. 79, 360-364.
15. Kozak, M. (1984) Nucleic Acids Res. 12, 857-872.
16. Jeffreys, A. J. (1979) Cell 18, 1-10.
17. Slightom, J. L., Blechi, A. E., and Smithies, O. (1980) Cell 21, 627-638.
18. Proudfoot, N. J., and Brownlee, G. G. (1976) Nature 263, 211-214.
19. Blobel, G., and Dobberstein, B. (1975) J. Cell Biol. 67, 852-862.
20. Von Heijne, G. (1983) Eur. J. Biochem. 133, 17-21.
21. Winzler, R. J. (1973) In Hormonal Proteins and Peptides (C. H. Li, Ed.), pp. 1-15. Academic Press, New York.
22. Rollins, B. J., Morrison, E. D., and Stiles, C. D. (1988) Proc. Natl. Acad. Sci. U. S. A. 85, 3738-3742.
23. Obaru, K., Fukuda, M., Maeda, S., and Shimada, K. (1986) J. Biochem. 99, 885-894.
24. Davatelis, G., Tekamp-Olson, P., Wolpe, S. D., Hermsen, K., Luedke, C., Gallegos, C., Coit, D., Merryweather, J., and Cerami A. (1988) J. Exp. Med. 167, 1939-1944.
25. Schall, T. J., Jongstra, J., Dyer, B. J., Jorgensen, J., Clayberger, C., Davis, M. M., and Krensky, A. M. (1988) J. Immunol. 141, 1018-1025.
26. Burd, P. R., Freeman, G. J., Wilson, S. D., Berman, M., DeKruyff, R., Billings, P. R., and Dorf, M. E. (1987) J. Immunol. 139, 3126-3131.