The A20 cDNA Induced by Tumor Necrosis Factor α Encodes a Novel Type of Zinc Finger Protein*

(Received for publication, March 27, 1990)

Anthony W. Opipari, Jr.‡, Mark S. Boguski§, and Vishva M. Dixit†

From the Department of Pathology, University of Michigan Medical School, Ann Arbor, Michigan 48109 and the National Center for Biotechnology Information, National Library of Medicine, National Institutes of Health, Bethesda, Maryland 20894

We have previously identified in endothelial cells a gene product designated A20, that is rapidly and profoundly induced by a stimulus in the absence of protein synthesis. One of the clones obtained, designated as A20, has a cognate mRNA approximately 4 kb in size which is rapidly (15 min) and transiently induced following TNF treatment. A20 was not identical to any known gene product based on sequence analysis of a partial length cDNA clone.

With this report we extend our characterization of the TNF-responsive A20 cDNA. Analysis of the full-length cDNA has revealed an open reading frame coding for a protein composed of 790 amino acid residues which is not identical to any other previously described protein. Within this putative protein two types of repeated elements are found including seven novel zinc finger motifs which may serve to define a new class of zinc finger proteins. Southern analysis supports the existence of other members of the class.

Tumor necrosis factor α (TNF) is a cytokine secreted predominantly by cells of the monocyte lineage following activation. The importance of TNF in the development and maintenance of inflammation is suggested by the wide range of proinflammatory activities attributed to TNF and the variety of disease states with which it has been associated (1).

Among the many tissue types which are responsive to TNF the vascular endothelium has a fundamental role in mediating proinflammatory activities of TNF. Positioned at the interface between the bloodstream and solid body tissues, the endothelium activity controls the flux of leukocytes into an inflammatory nidus. The endothelium has been shown to respond to TNF by developing procoagulant activities (2-4) and by actively increasing its adhesive interactions with circulating leukocytes through the expression of adhesive molecules such as intercellular adhesion molecule-1 (5), endothelial leukocyte adhesion molecule-1 (6), and vascular cell adhesion molecule-1 (7). Our earlier work has shown that neutrophil chemotactic factor and monocyte chemotactic factor are expressed as primary response genes in endothelial cells following TNF treatment (8). Each of these responses is likely to contribute to the proinflammatory actions of TNF.

In an effort to discover regulatory factors which function to initiate the cascade of endothelial responses to TNF, we previously used a strategy involving differential hybridization to identify several cDNA clones induced as primary response genes by TNF in human umbilical vein endothelial cells (8, 9). Primary response genes are those which are induced rapidly and profoundly by a stimulus in the absence of protein synthesis. In other systems it has been demonstrated that genes induced as part of this initial wave of expression encode regulatory proteins that are either transcriptional factors or paracrine factors. Thus, with this type of approach we expected to obtain clones which would encode, among other things, transcriptional factors regulating the expression of the many genes TNF is likely to influence. One of the clones obtained, designated as A20, has a cognate mRNA approximately 4 kb in size which is rapidly (15 min) and transiently induced following TNF treatment. A20 was not identical to any known gene product based on sequence analysis of a partial length cDNA clone.

With this report we extend our characterization of the TNF-responsive A20 cDNA. Analysis of the full-length cDNA has revealed an open reading frame coding for a protein composed of 790 amino acid residues which is not identical to any other previously described protein. Within this putative protein two types of repeated elements are found including seven novel zinc finger motifs which may serve to define a new class of zinc finger proteins. Southern analysis suggests that the A20 gene is a member of a large family of related genes.

**EXPERIMENTAL PROCEDURES**

**cdDNA Library Screening**—Poly(A)+ RNA from TNF/cycloheximide-treated human umbilical vein endothelial cells was used to construct cDNA libraries in the plasmid vector CDMS (10) and in λgt11 (8, 9). These libraries were screened as previously described by us (8, 9) using a 32P-labeled partial length A20 cDNA. The DNA probes were labeled with [32P]dCTP using either the random primer method (11) or by the polymerase chain reaction (12). When necessary, hybridizing cDNA inserts were isolated and subcloned into pGEM7zf (Promega Corp., Madison, WI).

**Sequence Determination and Computer Analysis**—Plasmid DNA prepared through polyethylene glycol precipitation (13) was sequenced in both directions by the dideoxynucleotide method with modified T7 DNA polymerase using oligonucleotide primers (8, 9). Homology searches for sequences related to A20 were done using the FASTA program (14) and the following data bases: GenBank release 63.0, NBRF/PIR release 24.0, EMBL release 22.0, Swiss Protein Database release 14.0, and DNA Database of Japan (2/15/90). In addition, the GenBank on-line service was used via Internet to access translated GenBank (new sequences) as of the time of this writing (6/5/90). Finally, a private data base of approximately 1200 fungal sequences was searched (Mark Goebl, University of Indiana). Detection and quantitation of internal repeats was performed using standard methods (reviewed in Ref. 15). Sequence alignments were carried out according to Lipman et al. (16).

**Southern Blot Analysis**—Human genomic DNA was prepared by standard methods (15) from the A431 cell line. 12-μg portions of DNA were digested overnight with various restriction enzymes (Boehringer Mannheim). Following electrophoresis on a 0.6% agarose gel, the DNA was transferred onto nitrocellulose and hybridized with a 32P-labeled probe corresponding to a specific region of the A20 cDNA. Hybridizations were performed at 42 °C in buffer containing 30% formamide, 5 × SSC, 10 × Denhardt's, 0.1% sodium dodecyl sulfate, 14705
**RESULTS AND DISCUSSION**

The A20 cDNA clone isolated previously represented approximately 1.5 kb of the predicted 4.0-kb full-length mRNA. To obtain a full-length cDNA clone, TNF-induced cDNA libraries were screened using the 32P-labeled 1.5-kb partial length A20 cDNA as a probe. Hybridizing clones were isolated and the cDNA insert sized by appropriate restriction mapping. The nucleotide and predicted protein sequence of A20 is displayed in Fig. 1. The length of the cloned sequence (4440 nucleotides) is in good agreement with the size predicted for a full-length clone by Northern analysis. An open reading frame of 2370 nucleotides is indicated. The initiator methionine is encoded beginning at nucleotide 67, and this sequence is in complete agreement with Kozak's translation initiation consensus sequence (17). The cDNA encodes a protein pre-
TNF Induces a Novel Zinc Finger Protein

dicted to contain 790 amino acids; contains a 5'-untranslated region of 66 nucleotides and a 3'-untranslated region of 2001 nucleotides. Within the 3'-untranslated region, four copies of the consensus polyadenylation signal AATAAA are found along with four copies of the sequence ATTATA which is present in the mRNA of various oncogenes and cytokines and confers message instability (18, 91). Data base searches (as described under "Experimental Procedures") using the A20 protein sequence gave uniformly negative results, i.e. there were no statistically significant matches to any known finger protein nor to any other proteins. However, intrasequence comparisons (see below) showed that A20 has a highly significant internally repetitive structure.

A20 was compared with itself using the RELATE program (20), and a score of 20.3 S.D. units was obtained. For comparative purposes, an S.D. score of ≥3.0 is generally indicative of significant internal repeats, and RELATE analysis of the Xenopus transcription factor IIIA gives an S.D. score of 21.9 under similar conditions. Furthermore, segment displacement frequencies computed by RELATE indicated that the repeat unit was approximately 43-46 residues in length. To visualize the organization of repeats within the A20 protein, a self-comparison matrix was computed (Fig. 2). This analysis reveals that greater than 50% of the A20 protein is composed of repeated elements and that the COOH-terminal half of the protein contains nine tandem copies of this element forming a continuous domain extending from approximately residue 386 to the carboxyl terminus. Alignment of the repetitive sequences demonstrates a striking conservation of cysteine residues (Fig. 3B). The cysteines in these repeats form a pattern similar to the Cys2/Cys2 finger motif found in the C class of zinc finger proteins (21). Within the putative A20 protein, this finger motif is present 6 times in the form of Cys-X2-Cys-X11-Cys-X3-Cys and once in the form of Cys-X2-Cys-X11-Cys-X2-Cys (Figs. 1 and 3B). Besides the nearly invariant occurrence and positioning of cysteines within these repeats, other residues are conserved to a lesser extent. This is illustrated in Fig. 3C by consensus sequences derived from three different cut-off frequencies.

A second unrelated repeat is also seen on the self-comparison matrix plot. Two copies of this element are found between

![Fig. 2. Comparison matrix analysis. Program CMPSEQ84 (26) was used to compute a self-comparison matrix with a window of 43 residues using the PAM 250 values (17) for scoring. A plotting threshold of >2.5 standard deviations above the mean for all scores was used. This threshold corresponds to a p value of <0.001 although the highest scoring diagonals achieved scores with p values of <10^{-11}. For comparative purposes, the highest scoring repeats in Xenopus transcription factor IIIA have scores with p values of <10^{-10} (27).](image-url)

![Fig. 3. Alignments of internal repeats and consensus of finger repeats. The alignments depicted in panels A and B were derived in the following manner. Start and stop positions for the tandem repeat blocks were estimated from the locations of diagonals on the self-comparison matrix in Fig. 2. Start and stop positions for individual repeat units were then estimated from segment displacements reported by the RELATE program, and individual repeats were aligned manually. Finally, to optimize the alignments, each individual repeat was treated as a separate sequence, and the entire set was simultaneously aligned using program MSA (16). Panel C illustrates the consensus sequences obtained from the repeats in panel B derived at three different cut-off frequencies. Amino acid residues are color-coded according to hydropathy index and charge: hydrophobic, green; acid/amide, red; basic, blue; small and neutral, black. Proline and cysteine are colored yellow and violet, respectively, to indicate their unique structural potentials.](image-url)
residues 286 and 356 (Fig. 3A). Furthermore, the blank left lower and right upper quadrants on the matrix analysis reveal that the NH2-terminal domain (residues 1–285) contains no repetitive sequences and is unrelated to the COOH-terminal half of the molecule. These structural distinctions are consistent with the possibility that separable functional domains exist within the A20 protein and provide a basis for future mutational analyses.

Comparison of the putative A20 finger protein with previously described C, finger proteins (21–25) reveals two unique and important distinctions. The first is that A20 contains multiple repeated motifs of the CysP/CysP type. Other previously described proteins of this class do not contain more than two nonequivalent fingers. The second distinction is the intercysteine spacing of XnXnXnXn. This is the first report of a CysP/CysP motif in which the finger loop domain is composed of 11 amino acid residues. Furthermore, because an absolute characteristic of C, finger proteins has been the presence of no more than two finger motifs, these findings suggest that A20 is a member of a novel class of finger proteins. This class would be defined by the presence of multiple related CysP/CysP finger motifs each containing a finger-loop domain of 11 amino acids.

The discovery of this novel type of finger motif suggests the possibility that the A20 gene is a member of a new family of zinc finger coding genes related to each other through transcriptional factor. Southern analysis suggests that other genes may exist which belong to this same class. Given the association between other zinc finger proteins and genetic regulation, it is plausible that A20 may function as a TNF-inducible transcriptional factor.

Acknowledgments—We thank Larry Holzman, Romy Marks, Karen O’Rourke, Vidya Sarma, and Fred Wolf for assistance and helpful discussions, and Peter Ward for support and encouragement.

REFERENCES
1. Beutler, B., and Cerami, A. (1989) Annu. Rev. Immunol. 7, 625–655
2. Bevilacqua, M. P., Prober, J. S., Majeau, G. R., Fiers, W., Cotran, R. S., and Giambri, M. A., Jr. (1989) Proc. Natl. Acad. Sci. U. S. A. 86, 4533–4537
3. Nawroth, P. P., and Stern, D. M. (1986), J. Exp. Med. 163, 740–746
4. Schleef, R. R., Bevilacqua, M. P., Sawdey, M., Giambri, M. A., Jr., and Loskutoff, D. J. (1988) J. Biol. Chem. 263, 5797–5803
5. Bevilacqua, M. P., Prober, J. S., Mendrick, D. L., Cotran, R. S., and Giambri, M. A., Jr. (1989) Proc. Natl. Acad. Sci. U. S. A. 84, 9238–9242
6. Prober, J. S., Giambri, M. A., Jr., Lapierre, L. A., Mendrick, D. L., Fier, W., Rothlein, R., and Springer, T. A. (1986) J. Immunol. 137, 1893–1898
7. Osborn, L., Hessen, C., Tizmad, R., Vassallo, C., Lukoweky, S., Chi-Rose, G., and Lobbe, R. (1989) Cell 59, 1203–1211
8. Diets, V. M., Green, S., Sarma, V., Holman, L. B., Wolf, F. W., O’Rourke, K. E., Ward, P. F., Prochownik, E. V., and Marks, R. M. (1989) J. Biol. Chem. 265, 2973–2978
9. Dutz, V. M., Marks, R. M., Sarma, V., and Prochownik, E. V. (1989) J. Biol. Chem. 264, 16005–16110
10. Seed, B. (1987) Nature 329, 840–842
11. Peinberg, A. P., and Vogelstein, B. (1983) Anul. Biochem. 132, 6–13
12. Schwalbe, D. B., and Sommer, S. S. (1989) Anul. Biochem. 177, 90–94
13. Sambrwock, J., Fritsch, K. F., and Mamaties, T. (1989) Molecular Cloning: A Laboratory Manual, 2nd Ed. pp. 1.40–1.41, 3.31–3.58, Cold Spring Harbor Laboratory, Cold Spring Harbor, NY
14. Pearson, W. R., and Lipman, D. J. (1988) Proc. Natl. Acad. Sci. U. S. A. 85, 2444–2448
15. Boguski, M. S., Freeman, M., Elshourbagy, N. A., Taylor, J. M., and Gordon, J. J. (1988) J. Lipid Res. 27, 1011–1034
16. Lipman, D. J., Altschul, S. F., and Kececioglu, J. D. (1989) Proc. Natl. Acad. Sci. U. S. A. 86, 4442–4447
17. Kozak, M. (1989) J. Cell Biol. 108, 229–241
18. Caput, D., Beutler, B., Hartog, K., Brown-Shriver, S., and Cerami, A. (1986) Proc. Natl. Acad. Sci. U. S. A. 83, 1670–1674
19. Shaw, G., and Kamen, R. (1988) Cell 56, 859–867
20. Dayhoff, M. O. (1975) Atlas of Protein Sequence and Structure, Vol. 5, Suppl. 3, pp. 345–355, 316, 313, 310, National Biomedical Research Foundation, Washington, DC
21. Evans, R. M., and Huggins, S. M. (1980) Cell 52, 1–3
22. Payre, F., and Vincent, A. (1988) FEBS Lett. 234, 245–250
23. Sundaram, N. W., Jr., and Barber, A. M. (1988) Anul. Clin. Lab. Sci. 18, 267–286
24. Wingender, E. (1988) Nucleic Acids Res. 16, 1879–1902
25. Tronier, C. D., Evans, T., Felsenfeld, G., and Boguski, M. S. (1990) Nature 343, 92–96
26. McLachlan, A. D. (1983) J. Mol. Biol. 169, 15–30
27. Miller, J., McLachlan, A. D., and Klug, A. (1985) EMBO J. 4, 1609–1614