Vol. 179, No. 1, 1991 BIOCHEMICAL AND BIOPHYSICAL RESEARCH COMMUNICATIONS
August 30, 1991 Pages 572-578

AMPHIREGULIN-ASSOCIATED PROTEIN: COMPLETE AMINO ACID SEQUENCE OF A PROTEIN PRODUCED BY THE 12-0-TETRADECANOYLPHORBOL-13-ACETATE-TREATED HUMAN BREAST ADENOCARCINOMA CELL LINE MCF-7

Mohammed Shoyab*, Vicki L. McDonald, Kenneth Dick, Brett Modrell, Najma Malik, and Gregory D. Plowman

Bristol-Myers Squibb Pharmaceutical Research Institute, 3005 First Avenue, Seattle, WA 98121

Received July 16, 1991

Amphiregulin-associated protein (ARAP) was purified from serum-free conditioned medium of MCF-7, human breast carcinoma cells, treated with 12-0- tetradecanoylphorbol-13-acetate (TPA). ARAP is a single-chain, extremely hydrophilic, heparin-binding protein. Its apparent molecular weight is approximately 21,500 as assessed by gel chromatography and approximately 15,500 as determined by polyacrylamide gel electrophoresis. The complete amino acid sequence of ARAP was determined. The larger form contains 123 amino acids, whereas a shorter form is missing two amino acids at the amino-terminal. ARAP contains 10 cysteines and 30 basic amino acids (23 lysines and 7 arginines). ARAP sequence has been found to be identical to protein encoded by human MK gene.

We previously reported the isolation of a novel growth modulatory glycoprotein, termed amphiregulin (AR), from serum-free conditioned medium (SFCM) of MCF-7 human breast carcinoma cells that were treated with 12-0- tetradecanoylphorbol-13-acetate (TAP)(1,2). AR inhibits growth of a number of human tumor cells and stimulates proliferation of human fibroblast, murine keratinocytes and other normal and tumor cells (1-3). The mature protein exists in two forms: the truncated form contains 78 amino acids, whereas a larger form contains 6 additional amino acids at the amino terminal end (3). AR is a member of the epidermal growth factor (EGF) family and binds to the EGF receptor, but not as well as EGF. The amino-terminal half of AR is extremely hydrophilic and contains two putative nuclear targeting signals (3,4). Isolation and characterization studies of cDNA and genomic clones for AR revealed that the mature proteins are derived from a

*To whom correspondence should be addressed.

Abbreviations: AR, amphiregulin; ARAP, amphiregulin-associated protein; TPA, 12-0-tetradecanoylphorbol-13-acetate; PAGE, polyacrylamide gel electrophoresis; MK, midkine; HB-GAM, heparin binding growth-associated molecule; RI-HB, retinoic acid induced-heparin binding; SFCM, serum-free conditioned medium.

0006-291X/91 $1.50

252 amino acid transmembrane precursor (4). AR binds to heparin and other sulfated glycosaminoglycans (GAG) which also modulate the biological activity of AR (5).

During the isolation of AR from the SFCM of TPA-treated MCF-7 cells, we found that a very hydrophilic protein copurified with AR and separated from it during the last step of purification (1). We called this protein AR-associated protein (ARAP). ARAP, like AR, is a basic protein and binds heparin. MCF-7 cells secrete AR and will also secrete ARAP when treated with TPA. Here we are reporting the complete amino acid sequence of ARAP which we find to be identical to the protein encoded by the human MK gene (6). These results indicated that ARAP belongs to a new family of heparin-binding proteins that exhibit neurotrophic activity. The members of this family include retinoic acid responsive gene product, murine MK (7-11); heparin-binding growth-associated molecule (HB-GAM) (12); pleiotrophin (13); heparin-binding neurotrophic factor (HBNF) (14,15); and retinoic acid-induced heparin-binding protein (RI-HB) (16).

MATERIALS AND METHODS

Purification of ARAP: ARAP was purified from SFCM of TPA-treated MCF-7 cells using the AR purification protocol (1). The gel-permeation HPLC, the last step of purification, separated ARAP from AR (1). The early eluting protein peak 1 was identified as ARAP, while peak 2 was characterized to be AR (Fig. 1). The homogeneous ARAP was used for all the chemical and biological studies reported here.

Amino Acid Sequence Determination: The amino acid sequence of human ARAP was determined by Edman degradation of reduced and S-pyridylethylated ARAP (SPE-ARAP) with various endopeptidases as described for AR (3).

Immunological Methods: The site specific antisera recognizing two peptides, which correspond to residues 1-18 (N-terminal peptide) and residues 106-123 (C-terminal peptide) of ARAP primary structure were raised in rabbits (Fig. 4A). For Western immunoblot analysis, proteins were subjected to reducing SDS-PAGE on 15% gel, transferred to nitrocellulose membrane, and bound antibody was detected using alkaline phosphatase-conjugated protein A as previously described (17). ELISA for the detection of AR and ARAP was performed using appropriate antibodies as described earlier (5).

Other Methods: Growth modulatory assays on various cell lines was performed as detailed earlier (1,3). Proteins were analyzed on SDS-polyacrylamide slab gel by the method of Laemmli and proteins were detected by silver staining (18,19).

RESULTS AND DISCUSSION

ARAP copurifies with AR during the first four chromatographic steps (1). ARAP, as with AR, elutes from analytical μ-Bondapak-C_{18} column at acetonitrile concentration of ~21% indicating the extreme hydrophilic nature of both proteins. The final step of purification, gel-permeation HPLC, resulted in the separation of ARAP and AR (Fig. 1). The early eluting peak 1 (Fig. 1A-C) was characterized to be ARAP by ELISA and Western blot analysis with anti-ARAP peptide antibodies and by direct amino-terminal sequencing of the peak 1 protein. The molecular weight of ARAP, determined by gel chromatography was ~21,500 (Fig. 1). Nearly 50 μg of pure ARAP was obtained from 4.5 liters of SFCM.

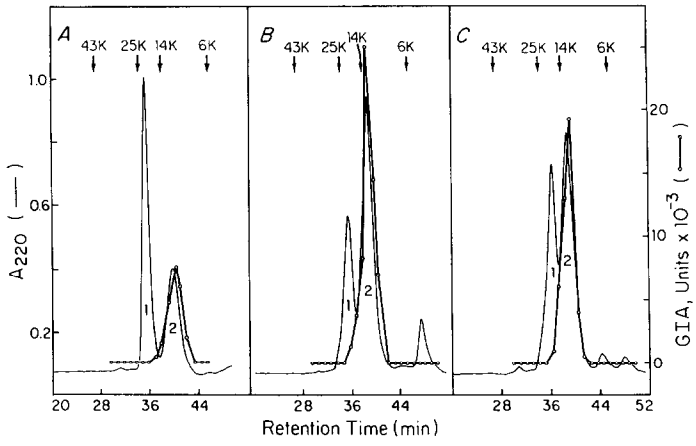

Fig. 1. Gel-permeation HPLC of fractions from Fig. 1C (Ref. 1). Experiments were performed as described (1). (A) HPLC of concentrated fraction 35. (B) HPLC of concentrated fraction 36. (C) HPLC of concentrated fraction 37.

Figure 2 shows an analysis of ARAP on 20% SDS-polyacrylamide gel under reducing and nonreducing conditions. ARAP migrated mainly as a single band with a molecular weight of ~15,500 under reducing condition and with a molecular weight of ~16,500 under nonreducing condition. These results indicate that ARAP, like AR, is a single-chain protein.

The Western blot analysis of crude extracts prepared with SFCM from TPA-treated or -untreated MCF-7 cells is shown in Fig. 3. The antibody to the amino terminal of the ARAP peptide detected only one major band (Mr ~15.5 kDa) in TPA-treated extract but not in the control extract. The molecular weight of this band was similar to that of ARAP (Fig. 3). The preimmune serum did not react with either extract nor with pure ARAP (Fig. 3, Lanes 1-3). We also assayed for ARAP in both crude extracts using ELISA. Again, ARAP immunoreactivity was detected only in the SFCM of TPA treated MCF-7 cells but not in the medium of control cells. These data indicate that ARAP is not secreted by MCF-7 cells without treatment with a stimulator such as TPA.

We fractionated the SFCM of TPA-treated MCF-7 cells on a column of heparin agarose (data not shown). AR as well as ARAP bound to the column and were eluted with NaCl solutions at median concentrations of 0.5 M and 0.9 M, respectively, suggesting that the binding to heparin is stronger with ARAP than with AR.

The amino acid sequence of human ARAP was determined with automated Edman degradation of S-pyridylethylated ARAP (SPE-ARAP) and of SPE-ARAP fragments generated from cleavage with *staphylococcus aureus* V8 protease and endoproteinase N-ASP (Fig. 4A). The amino-terminal analysis of SPE-ARAP revealed the presence of two sequences with one starting at residue 1, valine, and the other at residue 3, lysine (Fig. 4A). The yield of the larger form of ARAP was about 65% of the total ARAP. The larger form of ARAP thus contains two additional amino acids at the amino terminal of the truncated form of ARAP. The larger and truncated forms of ARAP are single chain polypeptides

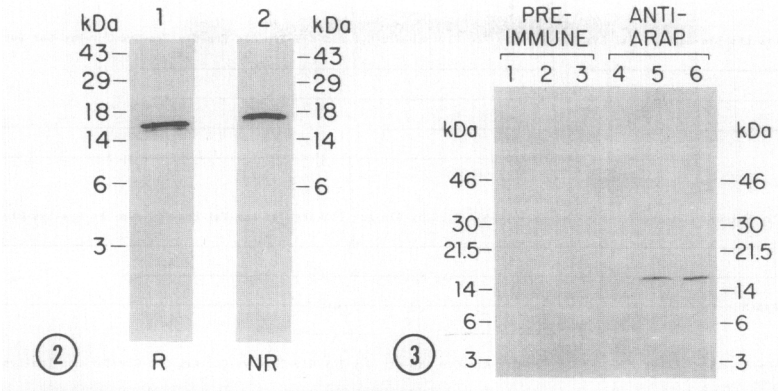

Fig. 2. SDS-polyacrylamide gel electrophoretic analysis of ARAP. Proteins were subjected to SDS-PAGE on 20% gel under reducing (R) and nonreducing conditions using a Bio-Rad mini-gel system. Molecular masses of protein markers are indicated: ovalbumin (43 kDa), carbonic anhydrose (29 kDa), β-lactoglobin (18 kDa), lysozyme (14 kDa), bovine trypsin inhibitor (6 kDa), and insulin subunit (3 kDa).

Fig. 3. Western immunoblot analysis. ~50 μg of crude extract prepared from SFCM of untreated cells (lanes 1 and 2) and from SFCM of TPA-treated MCF-7 cells (lanes 4 and 5), and 50 ng of ARAP (lanes 3 and 6) were subjected to SDS-PAGE under reducing conditions, electrotransferred to nitrocellulose, then probed with control IgG (lanes 1-3) or with anti-N-terminal ARAP IgG (lanes 4-6). Molecular masses of prestained markers (Amershams) are indicated: ovalbumin (46 kDa), carbonic aohydrase (30 kDa), trypsin inhibitor (21.5 kDa), lysozyme (14 kDa), aprotonion (6 kDa) and insulin subunit (3 kDa).

with a calculated molecular weight of 13,420 and 13,250, respectively. Both forms of ARAP have a similar carboxy-terminal (Fig. 4A).

Hydropathy profile of ARAP is shown in Fig. 4B. Data reveal that ARAP like AR is extremely hydrophilic. It is a very charged protein containing 23 lysines, 7 arginines, 5 arpartic acids and 6 glutamic acids, thus about one-third residues are charged amino acids in ARAP (Fig. 4A) The calculated isoelectric point of ARAP is 10.41. Similarly, unglycosylate AR has an isoelectric point of 10.1. The carboxy-terminal and amino-terminal of ARAP are very rich in lysine residues, and probably are responsible for the high affinity of ARAP for sulfated glycosaminoglycans. ARAP contain 10 cysteine residues, thus having potential for up to 5 intradisulfide bridges and being a very compact protein.

The sequence of ARAP was compared with all proteins in various data bases. The computer aided searches revealed that ARAP is identical to the protein encoded by human MK gene (Fig. 5). The search also showed that ARAP is about 86% homologous to murine MK protein, which contains an insertion of three additional residues between positions 11-13, and is about 50% homologous to pleiotrophin/HB-GAM (Fig. 5). Pleiotrophins from various mammalians species are highly conserved and exhibit similar neurite extension activity (13). The alignment of ARAP with MK and with

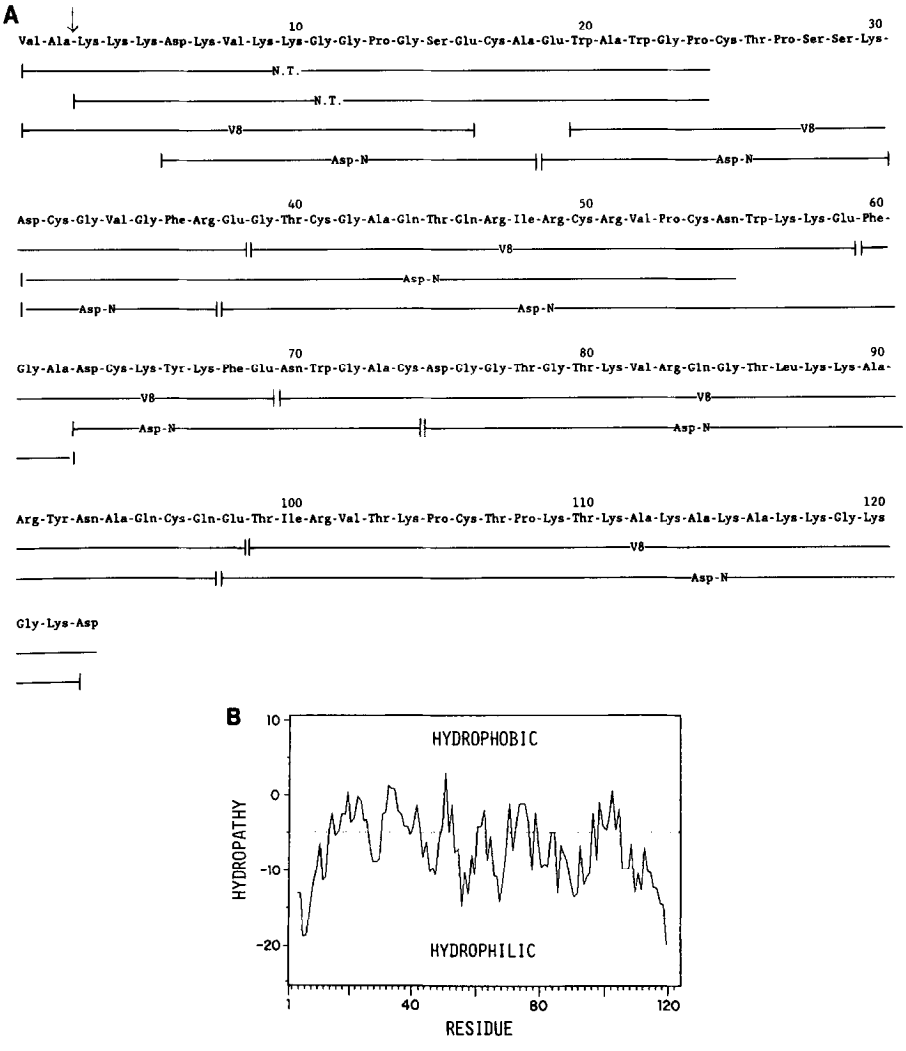

Fig. 4A. Amino acid sequence of ARAP and schematic outlines of the data supporting the sequence. The sequence of unfragmented SPE-ARAP is denoted by N.T. Peptides obtained by cleavage with endopeptidase-Glu, *staphylococcus aureus* V8 protease (V8) and with endoproteinase ASP-N (Asp-N) are indicated. The residues identified with Edman degradation are underlined. Vertical bar shows the beginning and ending of the fragments. Lines without two vertical bars show the incomplete sequences. Arrow indicates the start of truncated AR.

Fig. 4B. Hydropathy profile of ARAP based on the algorithm of Kyte and Doolittle (21).

pleiotrophin/HB-GAM indicates that all these protein belong to a superfamily. ARAP, like other family members, contains the hallmark 10 cysteine residues and conserves a spacing pattern, $CX_7CX_6C_{8\text{-}13}CX_8CX_3CX_9CX_9CX_{21}CX_9C$. ARAP/MK protein and pleiotrophin/HB-GAM are more homologous in the central portion of the molecules compared to the N-terminal or C-terminal domains of these two groups of proteins. RI-HB is probably a chicken homolog of ARAP/MK (16).

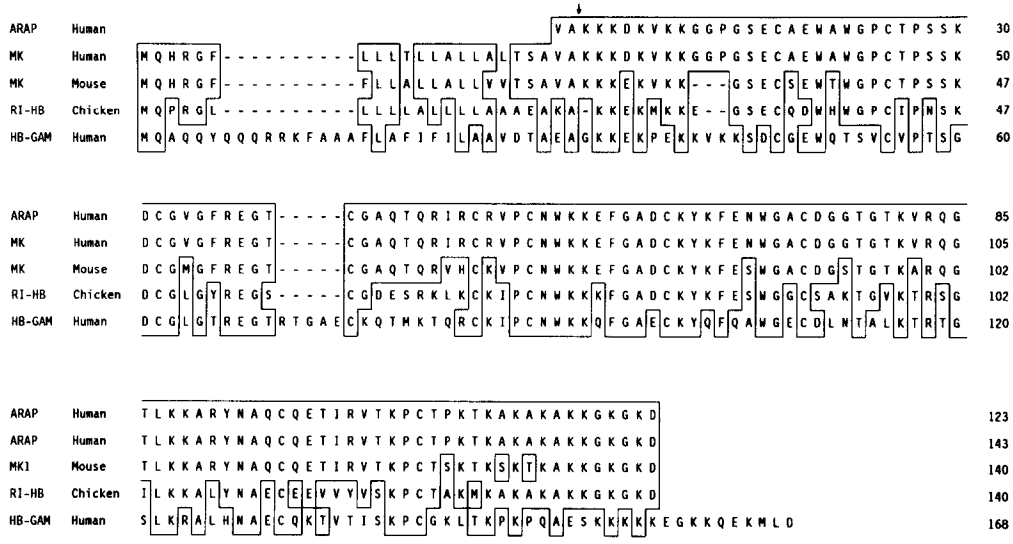

Fig. 5. Alignment of ARAP sequence with MK and pleiotrophin/HB-GAM related sequences. Amino acids are represented by standard one-letter symbols. Arrow indicates the start of truncated ARAP. Hyphens indicate gaps introduced to maximize homology. All sequences other than ARAP are deduced from the corresponding DNA.

The purified ARAP was evaluated for its potential growth modulatory activity on a number of tumor and normal cell lines that were tested previously with AR (1,3). ARAP, unlike AR, neither stimulated nor inhibited the proliferation of cell lines tested. ARAP and AR exhibit few common physical and chemical characteristics and appear to be functionally different proteins. ARAP was a very weak mitogen for rat pheochromocytoma PC-12 cells and it also slightly inhibited the differentiation of PC-12 cells that were induced by nerve growth factor (NGF) (data not shown). The murine recombinant MK protein has been reported to be a PC-12 mitogen and the human recombinant MK protein has been shown to be a weak mitogen for mouse 3T3 cells and a potent inducer of neurite extension in embryonic rat brain cells (10,20). Pleiotrophin/HB-GAM also exhibit neurite extension activity on embryonic nerve cells (12,13). Very little is known about the function of ARAP/MK proteins. It is interesting that murine MK gene is induced by retinoic acid and expressed in spatial and temporal fashion during embryogenesis and ontogenesis suggesting a putative role(s) in growth and differentiation (7-11). Additional work on the expression and regulation of the ARAP gene and protein in both physiological and pathological conditions as well as transgenic studies will help to unravel the function of ARAP. These studies may also provide clues to designing ARAP agonists and antagonists.

ACKNOWLEDGMENTS

We wish to thank Chris Kuepfer for her editorial assistance and Debby Baxter for her assistance in the preparation of this manuscript.

REFERENCES

1. Shoyab, M., McDonald, V.L., Bradley, J.G., and Todaro, G.J. (1988) Proc. Natl. Acad. Sci. USA 85, 6528-6532.
2. Shoyab, M. and Plowman, G.D. (1991) Methods Enzymol. 198, 213-221.
3. Shoyab, M., Plowman, G.D., McDonald, V.L., Bradley, J.G., and Todaro, G.J. (1989) Science 243, 1074-1076.
4. Plowman, G.D., Green, J.M., McDonald, V.L., Neubauer, M.G., Disteche, C.M., Todaro, G.J., and Shoyab, M. (1990) Mol. Cell. Biol. 10, 1969-1981.
5. Cook, P.W., Mattox, P.A., Keeble, W.E., Pittelkow, M.R., Plowman, G.D., Shoyab, M., Adelman, J.P., and Shipley, G.D. (1991) Mol. Cell. Biol. 11, 1547-2557.
6. Tsutsui, J., Uehara, K., Kadomatsu, K., Matsubara, S. and Muramatsu, T. (1991) Biochem. Biophys. Commun., 176, 792-797.
7. Kadomatsu, K., Tomomura, M., and Muramatsu, T. (1988) Biochem. Biophys. Res. Commun. 151, 1312-1318.
8. Tomomura, M., Kadomatsu, K., Matsubara, S. and Muramatsu, T. (1990) J. Biol. Chem. 265, 10765-10770.
9. Matsubara, S., Tomomura, M., Kadomatsu, K., and Muramatsu, T. (1990) J. Biol. Chem. 265, 9441-9443.
10. Tomomura, M., Kadomatsu, K., Nakamoto, M., Muramatsu, H., Kondoh, H., Imagawa, K. and Muramatsu, T. (1990) Biochem. Biophys. Res. Commun. 171, 603-609.
11. Kadomatsu, K., Huang, R.-P., Suganuma, T., Murata, F., and Muramatsu, T. (1990) J. Cell. Biol. 110, 607-616.
12. Merenmies, J. and Rauvala, H. (1990) J. Biol. Chem. 265, 16721-16724.
13. Li, Y.-S., Milner, P.G., Chauhan, A.K., Watson, M.A., Hoffman, R.M., Konder, C.M., Milbrandt, J., and Deuel, T.F. (1990) Science 250, 1690-1694.
14. Tezuka, K., Takeshita, S., Hakeda, Y., Kumegawa, M., Kikuno, R., and Hashimoto-Gotch, T. (1990) Biochem. Biophys. Res. Commun. 173, 246-251.
15. Kovesdi, I., Fairhurst, J.L., Kretchmer, P.J. and Bohlen, P. (1990) Biochem. Biophys. Res. Commun. 172, 850-854.
16. Raulais, D., Lagente-Chevallier, O., Guettet, C., Duprez, D., Coutois, Y., and Vigny, M. (1991) Biochem. Biophys. Res. Commun. 174, 708-715.
17. Malik, N., Kallestad, J.C., Gunderson, N.L., Austin, S.D., Neubauer, M.G. Ochs, V., Marquardt, H., Zarling, J.M., Shoyab, M., Wei, C., Linsley, P.S., and Rose, T.M. (1989) Mol. Cell. Biol. 9, 1847-1854.
18. Laemmli, U.K. (1970) Nature 227, 680-685.
19. Merril, C.R., Goldman, D., Sedman, S.A., and Ebert, M.H. (1981) Science 211, 1437-1439.
20. Muramatsu, H. and Muramatsu, T. (1991) Biochem. Biophys. Res. Commun. 177, 652-658.
21. Kyte, J. and Doolittle, R.F. (1982) J. Mol. Biol. 157, 105-132.