dlk, a Putative Mammalian Homeotic Gene Differentially Expressed in Small Cell Lung Carcinoma and Neuroendocrine Tumor Cell Line*

(Received for publication, October 19, 1992)

Jorge Labordat, Edward A. Sausville†, Thomas Hoffman†, and Vicente Notario‡

From the †Center for Biologies Evaluation and Research, Food and Drug Administration, Bethesda, Maryland 20892 and the ‡Division of Medical Oncology, Department of Medicine, and the §Lombardi Cancer Center, Department of Radiation Medicine, Georgetown University, Washington, District of Columbia 20007

Gastrin releasing peptide is mitogenic for mouse Swiss 3T3 fibroblasts and certain human small cell lung carcinoma (SCLC) cells but not for mouse Balb/c 3T3 fibroblasts. To identify new molecules associated with the gastrin releasing peptide-responsive phenotype, clones isolated from a differential cDNA library between Swiss and Balb/c 3T3 fibroblasts were used to screen for their expression in human SCLC cell lines. Using this approach, we have isolated and characterized human and mouse cDNA clones encoding a novel protein. This protein is a putative transmembrane protein belonging to the epidermal growth factor-like superfamily. In vitro transcription and translation studies detect a 42-kDa protein, in agreement with the size predicted from the translated cDNA sequence. This protein (termed Delta-like or dlk) is highly homologous to invertebrate homeotic proteins, including Delta, and Notch, the products of neurogenic loci involved in normal neural differentiation in Drosophila. dlk is expressed in tumors with neuroendocrine features, such as neuroblastoma, pheochromocytoma, and a subset of SCLC cell lines. However, its expression in normal tissues is restricted to the adrenal gland and placenta. These data suggest that dlk may be involved in neuroendocrine differentiation and, because of its cellular location and restricted expression in normal tissues, it may be a potential therapeutic target in neuroendocrine tumors, particularly SCLC.

Small cell lung carcinoma (SCLC) is a tumor of neuroendocrine origin that accounts for approximately 25% of all human lung cancers. Several cell lines of SCLC have been established, which differ in their expression of several neuroendocrine markers (1, 2). In particular, some SCLC cell lines produce gastrin releasing peptide (GRP), which interacts with a specific G-protein-coupled receptor on the cell membrane (3). GRP has been reported to be a mitogen for normal human epithelial lung cells, for mouse Swiss 3T3 fibroblasts, and for some SCLC (4). For these SCLC cells, GRP seems to act as an autocrine growth factor (4). In contrast, GRP is not mitogenic for some other SCLC cell lines and mouse Balb/c 3T3 fibroblasts.

The precise genetic differences between GRP-responsive and unresponsive cells are unknown. The immediate purpose of this investigation was the identification of new genes associated with the GRP-responsive phenotype in order to improve our understanding of the autocrine mechanism of GRP action in SCLC. This, in addition, could lead to the development of new therapeutic approaches for some SCLC. Even if the molecules identified were not directly involved with the GRP-response mechanism, their characterization still could provide new information about the differences in gene expression in SCLC associated with the differentiation state of these tumors. Furthermore, these molecules might constitute new and useful markers for the detection and staging of SCLC.

MATERIALS AND METHODS

Library Screening and Hybridization Techniques—The differential library of Swiss 3T3 compared with Balb/c 3T3 fibroblasts was constructed as explained in detail previously (5). RNA isolation, electrophoresis, Northern blots, and hybridization techniques were performed according to standard protocols (6). The cDNA probes were labeled with [35S]dCTP (Amersham Corp.) by the random primer method (7). The mouse Swiss 3T3 fibroblast AZAPII cDNA library used to obtain a mouse clone of dlk was purchased from Stratagene (La Jolla, CA). Screening procedures and plasmid rescue of positive λ clones were performed following the manufacturer's protocol. Those plasmids were sequenced with Sequenase (United States Biochemical Corp.) by the chain termination method following the manufacturer's protocol, using [α-32P]dATP (Amersham Corp.) as labeling agent.

To obtain a human dlk clone, a λgt10 human adrenal gland cDNA library was purchased from Clontech (Palo Alto, CA) and the screening procedure was performed following the manufacturer's protocol. Positive λ clones were subcloned into pGEMZ (Promega, Madison, WI) and sequenced as described above.

The study of the presence of dlk gene in different species was done on a dried agarose gel wafer of DNA digested with EcoRI (Clontech). The hybridization with a mouse dlk probe was performed following the manufacturer's protocol. At the same time, 5 µg of Drosophila and Xenopus DNA were digested with EcoRI, and electrophoresed in a 1% agarose gel. The gel was then dried and hybridized following Clontech procedures.

In Vitro Translation Studies—dlk mRNA was selected by hybridization of 2 µg of poly(A)²/Swiss 3T3 RNA with 5 µg of microcellulose-immobilized denatured dlk cDNA. The RNA bound was eluted by boiling. RNA was also prepared in vitro from two different full-length mouse dlk cDNAs cloned in pGEMZ by using the Riboprobe Gemini System II (Promega), following the manufacturer's protocol. Those RNAs were then used for in vitro translation assays. The reticulocyte lysate used for those assays was purchased from Du Pont-New England Nuclear. The procedure was performed following the manufacturer's protocol. Labeled proteins were analyzed in a 12% polyacrylamide gel followed by fluorography.

Sequence Analysis—Most DNA and protein sequence analyses were performed with the software package PC/Genie (8). The open reading
frames were found by using the methods of Fickett (9) and Shepherd (10). The transmembrane domain of dlk was found with the program RAOARGOS (11). The signal peptide was analyzed with the program PSIGNAL, according to Von Heijne (12). The alignment of the EGF-like repeats of Fig. 3 was done with the program CLUSTAL. The sites of potential biological importance were analyzed with the program PROSITE. The homology searches in the DNA and protein sequence data bases were performed with the program FASTA (13). This program was created by the University of Wisconsin Genetic Computer Group (UWGCG). The program was used to search in the following data bases: GenBank release 69.0; EMBL release 26.0, NBRF-protein release 26.0, and Swissprot release 18.0.

To study the statistical significance of the homologies found, we used the method of Needleman and Wunsch (14). In this method, the optimal alignment score between two proteins is compared with the statistical distribution of a number of random alignments, 100 in our case. An optimal alignment score of more than 5 standard deviations to the right of the mean is considered significant, particularly when no functional or structural relationship between the proteins compared is known.

RESULTS AND DISCUSSION

To identify new molecules associated with the gastrin-releasing peptide (GRP) responsive phenotype, we searched for cDNAs, which, in addition to being differentially expressed between Swiss (responsive) and Balb/c (unresponsive) 3T3 fibroblasts, were expressed in GRP-responsive SCLC cell lines. This approach was based on the assumption that genes products related to the GRP-responsive phenotype should be missing from Balb/c and unresponsive SCLC cell lines but present in Swiss 3T3 fibroblasts and certain responsive SCLC cell lines. A differential library was constructed that enriched for clones expressed in Swiss 3T3 but not in Balb/c 3T3 fibroblasts (3). A partial length clone (150 nucleotides long) isolated from this differential library hybridized with a 1.6-kb mRNA showing a pattern of expression compatible with the two screening requirements. To obtain a full-length cDNA clone, we used this partial length clone to screen a commercial oligo(dT)-primed cDNA library of Swiss 3T3 fibroblasts in the XZAPII vector. Several clones with inserts around 1.6 kilobase pairs were obtained and sequenced.

Nucleotide sequence analysis of the cDNAs defined an open reading frame of 1155 nucleotides. Both Fickett’s (9) and Shepherd’s (10) methods classified this open reading frame as coding. This open reading frame encodes a putative protein (dlk) of 385 amino acids with a molecular mass of 41,320 daltons. To confirm the existence of the protein predicted from the dlk cDNA sequence, we performed in vitro translation assays from mouse dlk mRNA using a rabbit reticulocyte lysate system. dlk mRNA was selected by hybridization of poly(A)’ RNA from Swiss 3T3 fibroblasts with denatured full-length dlk cDNA immobilized on nitrocellulose filters. Mouse dlk RNA was also prepared by in vitro transcription from two independent full-length cDNA clones. These three RNAs were used as templates for in vitro translation. A protein of around 42 kDa was present in all three samples (Fig. 1), in agreement with the molecular mass of dlk predicted from its cDNA sequence.

Structural analysis revealed that dlk is a transmembrane protein which contains six EGF-like repeats in the extracellular domain, and a short intracellular domain. A signal peptide is also present at the amino terminus (12) (Fig. 2). Based on these structural features, dlk appears to be a new member of the family of EGF-like homeotic proteins, which are implicated in differentiation decisions related to the commitment of a given class of cells to become part of a fully differentiated structure (15–17). In fact, a computer search within the Swissprot and NBRF protein sequence data bases (13) showed that there was a high degree of homology between dlk and the proteins encoded by several homeotic genes, including Delta (18), Serrate (19), and Notch (20, 21) of D. melanogaster, lin-12 (22) and glp1 (23) of C. elegans, and uEGF1 (24) of the sea urchin. A high degree of homology with the Xenopus and human homologues of Notch, Xotch (25), and TAN-1 (26), respectively, was also found. The degree of homology varied between the individual proteins, but regions of maximum homology showed that more than 33% of the amino acids were identical with those of our protein. Considering conserved amino acid substitutions, the homology rose to around 75%. Statistical analysis showed that the homologies found are highly significant (14). Some of the more representative alignment scores (see “Materials and Methods”) are: Delta: 20.2, Serrate: 19.7, TAN-I: 16.2, Notch: 14.6, Xotch: 13.6, Drosophila laminin β2: 6.3, mouse laminin β2: 4.1, human laminin β2: 2.8. These homologies indicate that the function of dlk is not directly related with the GRP response. However, as discussed below, dlk expression seems to be another feature of the neuroendocrine phenotype in a fashion similar to GRP receptors or other neuroendocrine markers.

To reveal possible homologies of other regions of dlk not containing EGF repeats, FASTA searches within Swissprot and NBRF data bases were performed with a partial dlk protein sequence devoid of the EGF-like regions. Using the method of Needleman and Wunsch (14), no significant homologies were found.

A search of DNA nucleotide sequence homologies in GenBank and EMBL data bases (13), identified a 81.2% sequence identity with pg2, a human cDNA cloned from an adrenal gland cDNA library (27). pg2 is a gene expressed in neuroendocrine tumors whose expression in normal tissues seems limited to the adrenal gland and to specific differentiation stages in the development of neuroendocrine tissue (28, 29). However, the putative protein coded for by pg2 (27) did not appear in the list of proteins showing the greatest degree of homology with the dlk protein. In fact, the pg2 protein contains 286 amino acids (30,000 daltons) with no EGF-like repeats, no signal peptide, and no transmembrane domains. To clarify the relationship between mouse dlk and human pg2, we isolated and characterized cDNAs from a rat 10 cDNA library of human adrenal gland, a tissue expressing both pg2 and dlk. Mouse dlk cDNA was used as a probe. Even under low stringency conditions in the screening experiments, we were unable to isolate cDNAs that coded for proteins with
Fig. 2. Alignment of mouse and human dlk. Identical amino acids are shown by the character |. Similar amino acids are indicated by A. Amino acids said to be "similar" are: A, S, T; D, E; N, Q; R, K; I, L, M; V; F, Y, W. The proteins show 86.2% identity and 90.1% similarity in their amino acid sequence. Potential biologically significant sites found in the data base PROSITE (18) are indicated by asterisks (*). The proteins also share six repetitive EGF-like sequences highly homologous to those found in the invertebrate neuregulin proteins.

Fig. 3. Alignment of the consensus sequence of the EGF-like repeats of dlk with the consensus sequences of the same repeats from several homeotic genes. To obtain the dlk EGF-like repeat consensus sequence, an alignment of the 12 EGF-like repeats of dlk from human and mouse was performed using the program CLUSTAL (18). This consensus sequence was aligned with the consensus sequences of several invertebrate homeotic genes, obtained in the same way, as well as with mouse EGF.

Fig. 4. Dried agarose gel wafer of DNA from different species digested with EcoRI and hybridized with mouse dlk. A gel wafer was purchased from Clontech, and the hybridization was performed following the manufacturer's protocol. Five pg of Drosophila and Xenopus DNA were digested with EcoRI and run in a 1% agarose gel in TBE buffer, and the gel was treated following Clontech procedures. The hybridizing bands are, for the most part, of a similar size which is consistent with a gene size of 25–30 kb.

Fig. 5. Northern blot hybridization of human dlk. In A, 20 pg of total RNA or 2 pg of poly(A)+ RNA from SCLC cell lines and Swiss and Balb/c 3T3 fibroblasts were run in a 1% agarose gel and then blotted on a nitrocellulose filter. A 1.6-kb band corresponding to dlk can be observed only in the SCLC cell lines NCI-N592, NCI-H69, and NCI-H510, as well as in Swiss 3T3 fibroblasts. In B, a human multiple tissue Northern blot (Clontech), containing 2 pg of poly(A)+ of each tissue, was used for hybridization, following the manufacturer's specifications. Only placenta shows high levels of expression; the rest of the tissues studied are negative. However, despite the manufacturer's assertion that equal amounts of RNA were loaded per lane, actin binding shows that this is not the case. This variation certainly affects the relative sensitivity for each tissue, particularly for kidney, which shows the lowest level of actin binding.

two cDNAs may belong to the same gene superfamily but their protein products show significant structural and, presumably, functional divergence. dlk shows a high degree of homology with the EGF-like proteins of Drosophila and other invertebrates that are involved in the differentiation of different tissues and structures (18–24). Fig. 3 shows the alignment of mouse or human dlk EGF-like repeat consensus sequence with consensus sequences of the EGF repeats of several proteins; residues well conserved among homeotic genes are also conserved in dlk,
confirming dlk as a new member of the family of EGF-like homeotic proteins. The amino acid sequence and structure of the EGF-like repeats, as well as the overall structure of dlk, are more closely related to the products encoded by invertebrate homeotic genes than to other vertebrate non-homeotic EGF-like proteins such as the EGF-precursor, transforming growth factor α, the α, β1, and β2 chains of laminin, coagulation factors, or complement proteins, previously thought to be the mammalian counterparts of the invertebrate homeotic genes (18, 20–22, 30). For that reason, it was interesting to study the species distribution of the dlk gene. Fig. 4 shows that the dlk gene is present in species ranging from birds to human. However, despite its structural homology with invertebrate proteins, dlk is absent from invertebrates and low vertebrates. To our knowledge, this is the first protein homologous to invertebrate homeotic products that is exclusively present in higher animals. This suggests that dlk may possess a function specific to this class of animals for which the conservation of the invertebrate homeotic EGF-like repeats and protein structure is important. In this regard, it is noteworthy that proteins belonging to the EGF-like superfamilies are more closely related to the products encoded by invertebrate proteins, such as the EGF-precursor, transforming growth factor α, the α, β1, and β2 chains of laminin, coagulation factors, or complement proteins, previously thought to be the mammalian counterparts of the invertebrate homeotic genes, which is important. In this regard, it is of noteworthy that proteins belonging to the EGF-like superfamilies are implicated in protein-protein interactions (31). Furthermore, it has been described that the interactions between specific EGF-repeats between homeotic proteins play a role in the transduction of differentiation signals (32, 33). This suggests that, to fulfill its function, dlk may interact with other hitherto unknown proteins through its EGF-like repeats.

Expression of dlk can be detected by Northern blot or RNAase protection analysis in the SCLC cell lines NCI-H510, NCI-H69, and NCI-N592 (Fig. 5A), in SCLC NCI-H146, in the human neuroblastoma SK-N-SH, and in the rat pheochromocytoma PC-12 cell lines (data not shown). The data of dlk expression in the different SCLC cell lines shows no correlation between dlk expression and GRP responsiveness, despite this being the initial criterion for selection of differentially expressed genes in a limited number of SCLC cell lines. The Ewing’s sarcoma cell lines SK-B-S-1, A4573, and TC106, the breast cancer cell lines MCF7, MDA 231, MDA 453, and MDA 468, the prostate cancer cell lines PC3 and LNCAP, and the monocytic cell line U937 do not express dlk (data not shown). Mouse Swiss 3T3 fibroblast RNA also showed a high degree of expression of dlk, but Balb/c 3T3 fibroblast RNA is negative for dlk expression. In normal tissues of hamster (data not shown) and human origin, dlk expression can be detected exclusively in the placenta and adrenal gland, from which it was cloned (Fig. 5B). These data, together with the homology of dlk with homeotic proteins, suggest that dlk may play a role in the differentiation of the neuroendocrine phenotype. SCLC and neuroblastoma are the only tumors known to express dlk, and this expression is probably associated with some differentiation stages, at least in the case of SCLC. This offers the potential prognostic use of dlk expression for the staging of SCLC and neuroblastomas. Moreover, considering that dlk is a transmembrane protein, the expression of which in normal tissues is restricted to the adrenal gland, dlk may be a readily accessible target for antibody imaging or therapy of SCLC and neuroblastoma tumors.

Acknowledgments—We are grateful to Dr. James F. Battey for help at the beginning of this project, as well as for the review of the manuscript. We also appreciate the editorial comments of Dr. John Finlayson.

REFERENCES

1. Carney, D. N., Gazdar, A. F., Bepler, G., Guccion, J. C., Marangos, P. J., Moody, T. W., Zweig, M. H. & Minna, J. D. (1986) Cancer Res. 46, 2913
2. Gazdar, A. F., Carney, D. N., Nau, M. M. & Minna J. D. (1985) Cancer Res. 45, 5924
3. Betley, J. F., Way, J. M., Corjese, M. H., Shapira, H., Kusano, K., Harkins, R., Wu, J. M., Sistrery, T. Mann, E. & Feldman, R. I. (1991) Proc. Natl. Acad. Sci. U. S. A. 88, 265
4. Cuttitta, F., Carney, D. N., Mulshine, J., Moody, T. W., Fedorke, J., Fischer, A. & Minna, J. D. (1985) Nature 316, 825
5. Timblin, C., Betley, J. F. & Kubel, W. M. (1990) Nucleic Acids Res. 18, 1587
6. Davis, L. G. Diben, M. D. & Betley, J. F. (1986) Basic Methods in Molecular Biology, Elsevier Science Publishing, New York
7. Ausbel, F. M., Brent, R., Kingdom, T. E., Moore, D. D., Seidman, J. G., Smith, J. A. & Struhl, K. (1991) Current Protocols in Molecular Biology, pp. 3.5.8–3.5.10, John Wiley & Sons, New York
8. Bairoch, A. (1990) Ph.D. thesis, University of Geneva, Geneva, Switzerland
9. Fickett, J. W. (1982) Nucleic Acids Res. 10, 5953
10. Shepherd, J. C. W. (1990) Methods Enzymol. 188, 150
11. Rao M. K. & Argos, P. (1986) Biochim. Biophys. Acta 869, 197
12. Von Heijne, G. (1986) Nucleic Acids Res. 14, 4683
13. Pearson W. R. & Lipman, D. J. (1988) Proc. Natl. Acad. Sci. U. S. A. 85, 2444
14. Needleman S. B. & Wunsch, C. D. (1970) J. Mol. Biol. 48, 443
15. Corbin, V., Michaelson, A. M., Almey, S. M., Neal, V., Alcami, E., Maniatis, T. & Young, M. W. (1983) Cell 37, 311
16. Campos-Ortega, J. A. & Knust, E. (1990) Eur. J. Biochem. 190, 1
17. Kopczynski C. C., Fetchel, K., Koo, P. J., Shepherd, S. B., Bauer, K. A. & Musakivitch, M. A. T. (1990) Mol. Reprod. Dev. 27, 26
18. Kopczynski, C. C., Alton, A. K., Fetchel, K., Koo, P. J. & Musakivitch, M. A. T. (1986) Genes & Dev. 2, 1723
19. Fleming, R. J., Scottilc, T. N., Diederich, R. J. & Artavanis-Tsakonas, S. (1990) Genes & Dev. 4, 2136
20. Wharton, K. A., Johansen, K. M., Xu, T. & Artavanis-Tsakonas, S. (1985) Cell 43, 567
21. Kidd, S., Kelley, M. R. & Young, M. W. (1986) Mol. Cell. Biol. 6, 3094
22. Greenwald, I. (1985) Cell 38, 553
23. Yochem, J. & Greenwald, I. (1988) Cell 58, 553
24. Hrusch, D. A., Andrews, M. E. & Raff, R. A. (1987) Science 237, 1487
25. Coffman, C., Harris, W. & Kintner, C. (1990) Science 249, 1438
26. Elisen, L. W., Bird, J., West, D. C., Soreng, A. L., Reynolds, T. C., Smith, S. D. & Sklar, J. (1991) Cell 66, 949
27. Helman, L. J., Sack, N., Pion, S. E. & Israel, M. A. (1990) Nucleic Acids Res. 18, 688
28. Helman, L. J., Thiele, C. J., Linchen, W. M., Nelkin, B. D., Baylin, S. B. & Israel, M. A. (1987) Proc. Natl. Acad. Sci. U. S. A. 84, 2346
29. Cooper, M. J., Hutchins, G. M., Cohen, P. S., Helman, L. J., Menzies, R. J. & Israel, M. A. (1990) Cell Growth Different 1, 149
30. Musakivitch M. A. T. & Hoffmann, F. M. (1986) Curr. Top. Dev. Biol. 24, 283
31. Furie, B. & Furie, B. C. (1988) Cell 53, 505
32. Fehon, R. G., Koo, P. J., Deyso, R. T., Regan, C. L., Xu, T., Musakivitch, M. A. T. & Artavanis-Tsakonas, S. (1990) Cell 61, 523
33. Reboy, I., Fleming, R. J., Fehon, R. G., Cherbas, L., Cherbas, P. & Artavanis-Tsakonas, S. (1991) Cell 67, 657

3820

New Homeotic-like Gene in SCLC and Neuroendocrine Tumors