Vol. 176, No. 1, 1991 BIOCHEMICAL AND BIOPHYSICAL RESEARCH COMMUNICATIONS
April 15, 1991 Pages 247-254

The WDNM1 gene product is a novel member of the 'four-disulphide core' family of proteins[1]

T.N. Dear[2] and R.F. Kefford

Department of Medicine, University of Sydney, Westmead Hospital, Westmead, NSW, 2145, Australia

Received February 25, 1991

It has been previously demonstrated that the*WDNM1* gene is downregulated in metastatic rat mammary adenocarcinomas. Detailed analysis of the structure of *WDNM1* mRNA reveals the presence of two transcripts generated by alternative splicing of a single exon. The unspliced version possesses an in-frame termination codon which encodes a predicted protein of 18 amino acids. The spliced version encodes a predicted protein of 60 amino acids which exhibits strong homology to a family of proteins possessing a conserved arrangement of cysteine residues. This family includes several proteinase inhibitors suggesting that *WDNM1* could encode a product with proteinase inhibiting capacity. Possible modulation in the level of this protein product could be related to overall proteinase activity of specific tumour cells which, in turn, could influence the invasive and metastatic potential of such cell populations. © 1991 Academic Press, Inc.

While major advances have been made in the understanding of the molecular basis of cellular growth and the oncogene/tumour supressor gene systems which deregulate growth in cancer cells (1), comparatively little is known of the mechanisms by which a minority of cells in an established solid tumour metastasise. It has been postulated that the failure of feedback mechanisms that safeguard against metastasis may be responsible, in part, for the evolution of metastatic subpopulations (2). Consequently, several groups have focussed attention on the identification of molecules that are down-regulated during the progression towards metastasis. The *WDNM1* gene was first recognised in a nonmetastatic rat

[1]The nucleic acid sequence reported in this paper is available from the EMBL database accession no. X13309.

[2]Corresponding author.

0006-291X/91 $1.50
Copyright © 1991 by Academic Press, Inc.
All rights of reproduction in any form reserved.

mammary adenocarcinoma cell line DMBA-8. RNA levels for this gene were significantly reduced in highly metastatic derivatives of this line (3). Additionally, northern blot analysis suggested the presence of more than one transcript hybridizing to *WDNM1* cDNA (3). In the present investigation, we have used the sensitive technique of PCR to characterize two related *WDNM1* transcripts. These two transcripts encode predicted proteins of different lengths and with different carboxyl-terminal sequences. The full-length *WDNM1* polypeptide is related to a family of cysteine-rich proteins of which several proteinase inhibitors are known members.

MATERIALS AND METHODS

PCR. 5 μg of total cytoplasmic RNA was reverse transcribed into cDNA in a 20 μl reaction volume containing 50 mM Tris-HCl, pH 8.3, 150 mM KCl, 10 mM $MgCl_2$, 5 mM DTT, 0.5 mM dNTPs, 25 μg/ml oligo(dT), 1,000 units/ml AMV reverse transcriptase by incubation at 42^0 for 1 hour. DNA amplification was performed with a 0.5 μl aliquot of the cDNA in a 50 μl reaction volume using *Taq* DNA polymerase (Promega Biotec) under conditions recommended by the manufacturer. Amplified products were separated by electrophoresis in a 2% agarose gel and transferred to Hybond-N (Amersham) according to the manufacturers instructions.

Sequence analysis. Amplified PCR products were cloned into pBluescript. Sequencing was performed on single-stranded rescued DNA using the dideoxy chain-termination method with Sequenase (USB). Protein sequences were aligned using the FASTA algorithm.

RESULTS

PCR analysis. To further investigate the *WDNM1* mRNAs, PCR was used to amplify *WDNM1* cytoplasmic transcripts from rat thymus. As shown in Figure 1, two amplified transcripts are present; a major one of 389 bp and a second minor species of 469 bp. Although thymus was used as a source of RNA for this experiment, both transcripts were also present in PCR products of cytoplasmic RNA from spleen, kidney, brain, liver, lung and the cell line Rat-2 (data not shown). This suggests that a low level of *WDNM1* transcription occurs in most tissues. The genomic PCR product was approximately 2.2 kb in size (Figure 1) indicating that the amplified products were not the result of genomic DNA contamination.

Sequence analysis. Both amplified products were cloned and sequenced. The 389 bp fragment corresponded exactly to the

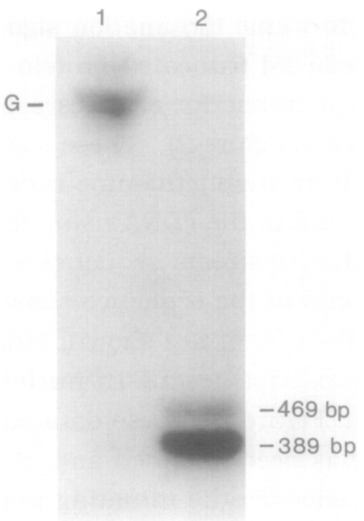

Figure 1. PCR analysis of *WDNM1* mRNA from rat thymus. PCR reactions were performed using primers corresponding to nucleotides 1-20 and 370-389 of the cDNA sequence in Figure 2. Amplified products were separated by agarose gel electrophoresis, transferred to nylon membrane and hybridized with a [^{32}P]-labelled oligomer corresponding to nucleotides 21-40 in the cDNA sequence. The 389 and 469 base pair fragments are indicated. The 2.2 kb product amplified from genomic DNA is also indicated (G). Lanes are: 1 - rat kidney genomic DNA; 2 -rat thymus total RNA.

previously published partial sequence of *WDNM1* (3). The larger product, however, possessed an 80 bp insertion which is most likely the result of alternative splicing as the concensus splice acceptor-donor nucleotide pair (GT...AG) is present at the ends of the insertion (Figure 2). The retention of this 'exon' in the mature

```
                                        GACAGCCTCAGTCTTGCTTCTGGTGGCTTTGATCGCCGTGGGA 43
ATGAACATTACCTATGCTCTGTTTTCTCCCACAAGTAAGTGTATTAAGGGCATCTGAGTTGAAGAGAGAGGGG 116
 M  N  I  T  Y  A  L  F  S  P  T  s  k  c  i  k  g  i  *

AATCTTCAAAGCTCCTACCACCTTAGGATGCGCTATTACAGAATTAGAAAAACCTGGAAAGTGTCCCAAGAAT 189
                                           K  L  E  K  P  G  K  C  P  K  N

CCCCCAAGAAGTATTGGCACTTGTGTTGAATTATGCTCAGGAGATCAATCGTGCCCCAACATACAGAAGTGCT 262
 P  P  R  S  I  G  T  C  V  E  L  C  S  G  D  Q  S  C  P  N  I  Q  K  C

GTTCCAATGGCTGTGGTCATGTTTGCAAATCTCCTGTCTTTTAAGTTACTAACAGCGATGTGGAATATGGATT 335
C  S  N  G  C  G  H  V  C  K  S  P  V  F  *

TGATCTTCATAAGCACGACTGATGGCCAGCCCCAGAAGATTTCTTCTGAATCCCTAGAGCCCATGCTTGGCTC 408
CTCCCTTGTCCTAGAATTGCATCCTTGGAAAAGGAAGATCTATACTGTGATGACAGCTTCCTAATGTGTTTGT 481
GTCCACAAATAAACTCTCCTTAGCATTCA$_n$                                          509
```

Figure 2. Nucleotide and deduced amino acid sequences of the alternative *WDNM1* transcripts. The spliced region is underlined. The altered C-terminal sequence of the unspliced version is indicated in lower case letters. Asterisks denote termination codons.

mRNA introduces an in-frame termination signal at nucleotides 98-100 resulting in a predicted truncated protein of only 18 amino acids, while the spliced transcript encodes a predicted protein of 60 amino acids in length (Figure 2). These sizes are based on translation initiation from the methionine codon at nucleotides 44-46. Cloning of the 5' end of the cDNA using the RACE protocol (4) did not yield any further upstream sequence with the largest clones terminating at the 5' end of the sequence shown. In addition, primer extension analysis (data not shown) indicates that the 5' end of the *WDNM1* transcript lies 13-15 nucleotides upstream of the sequence shown in Figure 2. These data suggest that a near-full-length sequence has been obtained and, therefore, that the ATG at nucleotides 44-46 encodes the initiating methionine.

The 60 amino acid predicted protein sequence exhibits striking similarity to a family of proteins that includes both the rat and mouse whey acidic proteins (5,6) and the proteinase inhibitors human antileukoprotease (7) and red sea turtle proteinase inhibitor, also known as chelonianin (8). The predicted amino acid sequence of *WDNM1* exhibits 45.2% identity with sea turtle chelonianin in a 42 amino acid overlap (Figure 3) and 42.0% identity with rat whey protein in a 50 amino acid overlap. If conservative amino acid substitutions are also considered, the identity increases to 78.6% for chelonianin and 70% for rat whey protein.

The region of similarity is found within a characteristic pattern of cysteine residues in the protein. This pattern of cysteines has been designated the 'four-disulphide core' (FDC) motif and is presumed to be important for folding of the protein (9). Other proteins which share this motif include snake venom neurotoxins, wheat germ agglutinin and neurophysin (6). The high level of identity of *WDNM1* with some of the proteins containing this motif suggest that *WDNM1* is a member of the FDC family.

```
                10        20        30        40        50        60
WDNM1   MNITYALFSPTKLEKPGKCPKNPPRSIGTCVELCSGDQSCPNIQKCCSNGCGHVCKSPVF
                 :.:: ::::.. ..:  :.. :..: .: . :::: .::::..:
CHE     KTKAECVRACRPPERPGVCPKTSGPGI--CLHGCDSDSDCKEGQKCCFDGCGYICLTVAP
        50        60        70        80        90        100
```

Figure 3. Comparison of *WDNM1* predicted amino acid sequence with that of the red sea turtle proteinase inhibitor chelonianin (CHE, Ref. 8). Gaps have been introduced to maximize homology. Identity of amino acids is indicated by colons and conservative substitutions by single dots.

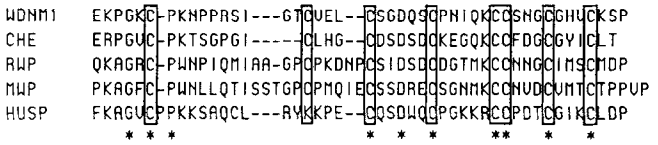

Figure 4. Alignment of the predicted protein sequence of *WDNM1* with members of the 'four-disulphide core' family of proteins. The alignment of *WDNM1* is made to the second cysteine-rich domain of red sea turtle chelonianin (CHE, Ref. 8), rat whey acidic protein (RWP, Ref. 5), mouse whey acidic protein (MWP, Ref. 6) and human antileukoprotease (HUSP, Ref. 7). The cysteine residues of the four-disulphide core motif are boxed. Fully conserved residues are indicated by asterisks below the sequence.

The arrangement of cysteines in rat and mouse whey acidic proteins and human antileukoprotease reveals a two domain protein structure. These domains are of similar length and are homologous to each other. It has been proposed that, in the case of the acid-stable proteinase inhibitor antileukoprotease, the first domain may be part of an antitryptic site while the second domain defines an antichymotryptic site (7). Furthermore, it has been previously demonstrated that the second domain of the red sea turtle proteinase inhibitor chelonianin possesses a subtilisin-inhibiting activity (8). It appears that only the second domain is present in *WDNM1*. An alignment of domain II of some of these proteins with the predicted protein for *WDNM1* is shown in Figure 4.

DISCUSSION

The data presented suggest that alternative splicing is involved in the production of distinct *WDNM1* mRNAs. A variety of genes demonstrate alternative splicing which results in amino acid insertions and deletions (10,11), or frame shifts which alter the C-terminal sequence of the protein and result in premature termination (12,13,14). Such alternative splicing may be involved in developmental regulation of gene expression (15). The unspliced *WDNM1* transcript encodes a predicted polypeptide of only 18 amino acids in length. It is unknown whether such a protein would be biologically active. The unspliced transcript could act as part of a mechanism to regulate the overall level of functional *WDNM1* protein by influencing translational activity of the spliced transcript within the cell.

The predicted protein of 60 amino acids demonstrated striking homology to the 'four-disulphide core' family of proteins which includes the whey acidic proteins and various proteinase inhibitors. The function of whey acidic protein remains to be

elucidated but the structural similarity of *WDNM1* to the proteinase inhibitors chelonianin and antileukoprotease suggests that *WDNM1* may possess proteinase-inhibiting activity. Metastasising tumour cells possess the capacity to invade host tissues by proteolytic degradation of the basement membrane and a strong correlation has been demonstrated between the metastatic potential of tumour cells and their level of activity of various proteinases (16). Genes associated with proteinase activity have been implicated in the malignant phenotype. Increased mRNA levels in malignant tumour populations have been reported for the proteinases transin (17,18), urokinase (19), cathepsin L (20) and collagenase IV (21). Increased proteinase activity can be achieved by either increasing proteinase levels or, alternatively, reducing the level of an inhibitor. Inactivation of the tissue inhibitor of metalloproteinases (TIMP) gene in murine fibroblasts leads to an increase in metastatic potential (22). Assuming that *WDNM1* functions as an inhibitor of a proteinase required for breakdown of the extracellular matrix, then loss of *WDNM1* could result in increased proteinase activity and, ultimately, enhanced metastatic ability. The metastatic cell lines cloned from the DMBA-8 rat mammary adenocarcinoma cell line possess enhanced levels of secreted urokinase activity (23) and an inhibitor of urokinase has been identified from nonmetastatic DMBA-8 cells (24). This suggests that increases in proteinase activity are required for metastasis to occur in this tumour cell line.

WDNM1 was the first of two genes which have been associated with the nonmetastatic phenotype in DMBA-8 cells. The second gene, initially termed *WDNM2* (25), encodes NADP(H):menadione oxidoreductase; an enzyme involved in the reduction of various quinones and redox dyes within the cell. The expression of a third gene, pGM21, has been associated with increased metastatic potential in DMBA-8 cells (26). However, lack of substantial sequence information on this gene severely limits any predictions that can be made of its function. The identification of the *WDNM1* gene as a member of a family which includes proteinase inhibitors suggests a possible mechanism of action for this gene in maintaining the nonmetastatic phenotype.

Another gene associated with low metastatic potential, *nm23*, encodes a subunit of a protein with nucleoside diphosphate (NDP) kinase activity (27,28)). This suggests that *nm23* may play a role in processes to which NDP kinases contribute, such as microtubule assembly or signal transduction (29). Therefore, *WDNM1*, *WDNM2* and *nm23* may contribute to the metastatic phenotype through

quite distinct pathways or, possibly, through different steps of a single complex pathway.

ACKNOWLEDGMENT

T.N. Dear is the recipient of a travelling fellowship from the Medical Foundation of the University of Sydney.

REFERENCES

1. Weinberg, R.A. (1989) Biochem. 28, 8263-8269.
2. Schirrmacher, V. (1985) Adv. Cancer Res. 43, 1-73.
3. Dear, T.N., Ramshaw, I.A. and Kefford, R.F. (1988) Cancer Res., 48, 5203-5209.
4. Frohman, M.A., Dush, M.K. and Martin, G.R. (1988) Proc. natl. Acad. Sci. USA 85, 8998-9002.
5. Dandekar, A.M., Robinson, E.A., Appella, E. and Qasba, P.-K. (1982) Proc. natl. Acad. Sci. USA 79, 3987-3991.
6. Hennighausen, L.G. and Sippel, A.E. (1982) Nucl. Acids Res. 10, 2677-2684.
7. Seemuller, U., Arnhold, M., Fritz, H., Wiedenmann, K., Machleidt, W., Heinzel, R., Appelhans, H., Gassen, H.-G. and Lottspeich, F. (1986) FEBS Lett. 199, 43-48.
8. Kato, I. and Tominaga, N. (1979) Fed. Proc. 38, 832 (Abstract)
9. Hennighausen, L.G. and Sippel, A.E. (1982) Nucl. Acids Res. 10, 3733-3744.
10. Bordereaux, D., Fichelson, S., Tambourin, P. and Gisselbrecht, S. (1990) Oncogene 5, 925-927.
11. Mosthaf, L., Grako, K., Dull, T.J., Coussens, L., Ullrich, A. and McClain, D.A. (1990) EMBO J. 9, 2409-2413.
12. LaRosa, G.J. and Gudas, L.J. (1988) Mol. Cell. Biol. 8, 3906-3917.
13. Hayzer, D.J. and Iynedjian, P.B. (1990) :Biochem. J. 270, 261-263.
14. Caras, I.W., Davitz, M.A., Rhee, L., Weddel, G., Martin, D.W. Jr. and Nussenxweig, V. (1987) Nature 325, 545-549.
15. Bond, R.W., Wyborski, R.J. and Gottlieb, D.I. (1990) Proc. natl. Acad. Sci. USA 87, 8771-8775.
16. Nicolson, G.L. (1988) Biochim. Biophys. Acta 948, 175-224.
17. Matrisian, L.M., Bowden, G.T., Krieg, P., Furstenberger, G., Briand, J.-P., Leroy, P. and Breathnach, R. (1986)Proc. natl. Acad. Sci. USA 83, 9413-9417.
18. Ostrwoski, L.E., Finch, J., Krieg, P., Matrisian, L., Patskan, G., O'Connell, J.F., Phillips, J., Slaga, T.J., Breathnach, R. and Bowden, G.T. (1988) Mol. Carcinog. 1, 13-19.
19. Sappino, A.P., Busso, N., Belin, D. and Vassalli, J.D. (1987) Cancer Res. 47, 4043-4046.
20. Denhardt, D.T., Greenberg, A.H., Egan, S.E., Hamilton, R.E. and Wright, J.A. (1987) Oncogene 2, 55-59.
21. Bonfil, R.D., Reddel, R.R., Ura, H., Reich, R., Fridman, R., Harris, C.C. and Klein-Szanto, A.J.P. (1989) J. Natl. Cancer Inst. 81, 587-594.
22. Khokha, R., Waterhouse, P., Yagel, S., Lala, P.K., Overall, C.M., Norton, G. and Denhardt, D.T. (1989) Science 243, 947-950.

23. Ramshaw, I.A. and Badenoch-Jones, P. (1985) Cancer Metastasis Rev. 4, 195-208.
24. Grant, A., Ramshaw, I.A., Badenoch-Jones, P., Eichner, R.D. and Hunt, N.H. (1986) Eur. J. Biochem. 154, 635-641.
25. Dear, T.N., McDonald, D.A. and Kefford, R.F. (1989) Cancer Res. 49, 5323-5328.
26. Phillips, S.M., Bendall, A.J. and Ramshaw, I.A. (1990) J. Natl. Cancer Inst. 82, 199-203.
27. Wallet, V., Mutzel, R., Heike, T., Barzu, O., Wurster, B., Veron, M. and Lacombe, M.-L. (1990) J. Natl. Cancer Inst. 82, 1199-1202.
28. Biggs, J., Hersperger, E., Steeg, P.S., Liotta, L.A. and Shearn, A. (1990) Cell 63, 933-940.
29. Liotta, L.A. and Steeg, P.S. (1990) J. Natl. Cancer Inst. 82, 1170-1172.