Molecular Cloning and Characterization of Alternatively Spliced Transcripts of the Mouse Neurofibromatosis 2 Gene

Tetsuo Hara, Albert B. Bianchi, Bernd R. Seizinger, and Nikolai Kley

Oncology Drug Discovery, Molecular Genetics and Cell Biology, Bristol-Myers Squibb Pharmaceutical Research Institute, Princeton, New Jersey 08543-4000

Abstract

The human neurofibromatosis 2 (NF2) gene has recently been isolated and predicted to encode a novel protein named merlin. Based on its high homology to the moesin-ezrin-radixin family of proteins, it may be involved in mediating interactions between the plasma membrane and the cytoskeleton. Here we report the isolation and characterization of multiple transcript isoforms of the mouse NF2 gene. The full length coding complementary DNA sequence of transcript isoform I is 1788 base pairs in length, shares 90% sequence identity with the human NF2 complementary DNA, and encodes a putative protein of 596 amino acids sharing 98% homology with the human merlin protein. Transcript isoforms II and III carry a 45- and 16-base pair insertion, respectively, at nucleotide 1740 at the 3′ end, generated by two different modes of alternative splicing; both insertions introduce premature termination codons. Thus, transcript isoforms II and III predict proteins of 591 and 584 amino acids with altered COOH-terminal sequences, respectively and predicted to encode a novel protein named merlin. Based on its high homology to the moesin-ezrin-radixin family of proteins (thus named merlin), merlin shares 65% homology at its NH2 terminus to moesin, ezrin, and radixin, suggesting that this region may localize it in the plasma membrane/cytoskeleton environment. Towards the COOH-terminus, the merlin protein displays only very weak homology to these proteins.

The human NF2 gene is widely expressed in different cell types (16). Moreover, it has recently been shown that mutations in NF2 gene transcripts are associated not only with hereditary but also sporadic vestibular schwannomas and meningiomas and tumor types seemingly unrelated to NF2 (breast carcinoma and melanoma). This indicates that the NF2 gene may constitute a tumor suppressor gene of more general importance in tumorogenesis. The isolation and characterization of the NF2 gene from other species will be important in the functional characterization of the merlin protein itself and for the establishment of cellular model systems to characterize growth regulatory properties of the NF2 gene.

Here we report the cloning and characterization of multiple transcript isoforms of the mouse NF2 gene. Our results indicate high conservation of the NF2 gene during evolution and suggest a possible role for the COOH-terminus in mouse merlin function.

Introduction

NF22 is inherited in an autosomal dominant manner and its morbidity is estimated as 1 in 50,000 people (1-4). The hallmark of this disorder is the development of bilateral vestibular schwannomas (acoustic neuromas) of the eighth cranial nerve (1-7). Other symptoms associated with NF2 include tumors of the central nervous system, such as meningiomas, gliomas, ependymomas, and spinal nerve root schwannomas (1-7). These tumors lead to serious neurological complications, particularly deafness, vertigo, gait disturbance, and paralysis within the first decades of life. Juvenile posterior capsular lens opacities are also characteristic of NF2 patients (8).

On the basis of genetic linkage studies and tumor deletion studies (familial and hereditary), it has been proposed that loss or inactivation of a tumor suppressor gene in the NF2 locus in chromosome 22q leads to tumorigenesis (9-15). Physical mapping and chromosome walking strategies recently culminated in the identification of a candidate NF2 gene during evolution and suggest a possible role for the COOH-terminus in mouse merlin function.

Materials and Methods

Cloning of Mouse NF2 cDNAs. A mouse brain oligodeoxynucleotidylic acid and random-primed cDNA library in AZAPII (Stratagene, La Jolla, CA) was used for the analysis. The library was screened by plaque hybridization to Ti0 filter-immobilized cDNA phage clones at a density of 5 x 105 plaque-forming units/150 mm filter. A 32P-labeled (19) PCR fragment representing nearly the entire coding region of the human NF2 cDNA (nucleotides 219 to 2031, as numbered in Ref. 16) was used as a probe.

Prehybridization was performed at 65°C for 3 h in 4 x standard sodium phosphate EDTA (1 x SSPF = 0.15 M NaCl, 0.010 M NaH2PO4, 0.001 M EDTA-Na2, pH 7.4), 6% polyethylene glycol 8000, 0.5% sodium dodecyl sulfate, 2 x Denhardt’s solution, and 100 mg/ml of denatured, sheared salmon sperm DNA. Hybridization was then carried out at 65°C in the same solution containing radiolabeled probe. Washes were performed under high stringency. Filters were exposed overnight to Kodak X-AR films at -80°C with intensifying screens. After 3 rounds of screening, cDNAs were rescued as pBluescript SK-phagemids by in vivo excision using the R408 helper phage.

Nucleotide Sequence Analysis. The nucleotide sequence of recombinant clones was determined by the dideoxy chain termination method using Sequenase version 2.0 (U.S. Biochemical Corp., Cleveland, OH, Ref. 20). Both strands were sequenced and analyzed using MacVector 4.1 sequence analysis software (International Biotechnologies, Inc., New Haven, CT).

RT-PCR and Expression Studies. Total cellular RNA (2 mg) from murine brain, heart, liver, lung, spleen, and testis was reverse transcribed using SuperScript RNase H reverse transcriptase ( Gibco BRL, Grand Island, NY)

Received 8/10/83; accepted 11/24/93.

The costs of publication of this article were defrayed in part by the payment of page charges. This article must therefore be hereby marked advertisement in accordance with 18 U.S.C. Section 1734 solely to indicate this fact.

1 To whom requests for reprints should be addressed.

2 The abbreviations used are: NF2, neurofibromatosis type 2; RT-PCR, reverse transcriptase-polymerase chain reaction; cDNA, complementary DNA; II-3′, antisense primer 5′-GAGGACTCATAATGCAGTAGGTCC-3′.
and the specific primer 3AS1 (5'-TCTTCAGAGAGAGCCAGGCTGC-3') complementary to the 3' end of the mouse NF2 mRNA. Reverse transcribed cDNA template was then subjected to two rounds of PCR. The first round of PCR was carried out with primers 3AS1 and 5 m9 (5'-GGTGGATGACCTG-GAGGAA-3'), extending from nucleotide 1576 with respect to the full length mouse NF2 cDNA; Fig. 1), and the second round (nested PCR) was performed using primers 3AS1 and 5AS1 (5'-ACACAGGAGAAGCTCAGGACAG-3'), extending from nucleotide 1684). Amplification was carried out on a Gene-Amp 9600 machine (Perkin Elmer Cetus, Norwalk, CT) in a total volume of 50 μl containing 10 μl Tris-HCl (pH 8.3), 50 mM KCl, 1.5 mM MgCl2, 0.25 mM of each dNTP and 2.5 units of Taq DNA polymerase (Boehringer Mannheim, Indianapolis, IN). The reactions were denatured at 94°C for 5 min, followed by 32 cycles of denaturing at 94°C for 15 sec, annealing at 58°C for 15 s, and extension at 72°C for 75 s. Final extension was at 72°C for 3 min. Five μl of the PCR reaction were analyzed by electrophoresis in a 3% agarose gel, 0.25 μl of each dNTP and 2.5 units of Taq DNA polymerase (Boehringer Mannheim, Indianapolis, IN). The reactions were denatured at 94°C for 5 min, followed by 32 cycles of denaturing at 94°C for 15 sec, annealing at 58°C for 15 s, and extension at 72°C for 75 s. Final extension was at 72°C for 3 min. Five μl of the PCR reaction were analyzed by electrophoresis in a 3% agarose gel.

**Isolation of Mouse NF2 Genomic Sequences.** In order to obtain exon-intron boundary information, mouse genomic DNA harboring the 3' end of the mouse NF2 gene was analyzed. PCR was used to amplify mouse genomic DNA across the region from nucleotide 1684 (primer 5AS1, numbering as in Fig. 1) to nucleotide 331 (primer 3AS1, numbering as in Fig. 1), in order to determine the exon-intron boundaries. Northern and Southern blot analysis. Filter-immobilized polyadenylated RNA (2 μg/lane) from multiple murine tissues (Clontech) and DNA (8 μg/lane) from different species were hybridized with a 32P-labeled (19) mouse or human NF2 cDNA probe, respectively, according to the manufacturer’s instructions. Blots were exposed to Kodak X-AR films at -80°C with intensifying screens for 2 h to 3 days.

**Results and Discussion.**

The search for the genetic defect underlying NF2 has recently culminated in the identification of a candidate gene shown to be mutated in hereditary vestibular schwannomas and meningiomas from unrelated NF2 families (16, 17). Further mutational analysis indicates that mutations in the NF2 gene are not only associated with NF2-related CNS neoplasms but also with other seemingly unrelated tumor types, suggesting a more widespread involvement of this gene in tumorigenesis. Here we report the isolation of the murine NF2 cDNA homologue, and the characterization of multiple transcript isoforms predicted to encode mouse merlin protein variants with altered structure and expression. A mouse brain cDNA library (in λZAPIII) was screened with a 32P-labeled DNA probe spanning nearly the entire human NF2 gene coding region. Twelve positive plagues were analyzed in more detail and helper phage rescued phagedenms (pBluescript) tested for insert size after the third round of screening. Four clones with inserts of 2.0 kilobases, 0.8 kilobases, 3.5 kilobases, and 2.0 kilobases in size turned out to be overlapping clones that cover the entire coding region of the mouse NF2 gene. The complete nucleotide and deduced amino acid sequences of the NF2 cDNA, as derived from sequence analysis of both strands of these four clones, are shown in Fig. 1. The coding region consists of 1788 nucleotides, encoding a predicted protein of 596 amino acids with a calculated molecular mass of 69-70 kilodaltons, as compared to 595 amino acids for the human protein (Fig. 1). The one residue difference is based on the presence of a three base insertion (CCG) at nucleotide 1710, identifying a proline residue at this site. Sequence analysis of all isolated cDNA clones and RT-PCR products derived from mouse brain RNA revealed the presence of this proline, indicating that this residue is indeed specific to the predicted mouse merlin protein. This was confirmed by sequencing of genomic DNA, which also revealed that the 3-base pair insertion does not represent a distinct exon in the mouse (see Fig. 3). The overall homology of mouse and human NF2 cDNAs is 90% at the nucleotide level (a total of 172 substitutions) and 98% at the deduced amino acid sequence of the mouse NF2 cDNA.
level (a total of 10 substitutions, 7 of which constitute conservative changes and 2 of which represent semiconservative changes). These substitutions are more clustered toward the COOH-terminus of the predicted protein (Fig. 1).

Analysis of several different cDNA clones revealed an additional transcript isoform of the mouse NF2 gene. Clone λZ8 showed an insertion of 16 base pairs at nucleotide 1740 (Fig. 2). Interestingly, at the same position we recently detected a 45 base pair insertion in an alternative transcript isoform of the human NF2 gene. In order to investigate whether this alternative transcript, presumably arising from alternative splicing, is also expressed in murine tissue, we performed RT-PCR on mouse brain RNA using oligonucleotide primers flanking the insertion site at nucleotide 1740. Two PCR products were identified by agarose gel electrophoresis. Sequencing of the largest PCR product revealed the presence of a 45-base pair insertion at nucleotide 1740, which is identical in sequence to that described for the human NF2 transcript except for a 1-base pair difference (Fig. 2, right, see arrow), which, however, does not translate into a difference in amino acid sequence between the two species. In concordance with the numbering system used for the human transcript variants, we designated the mouse transcript with the 45-base pair insertion as isoform II and that with the 16-base pair insertion as isoform III (Fig. 2). The 45-base pair insertion in transcript isoform II encodes 11 amino acids in frame and introduces a premature stop codon, predicting a protein of 591 amino acids. The 16-base pair insertion in transcript isoform III is entirely different. It encodes four amino acids in frame before introducing a premature stop codon and thus encodes a putative protein of 584 amino acids in length (Fig. 2).

Interestingly, two distinct modes of splicing appear to be involved in the generation of isoforms I, II, and III. Fig. 3 shows a schematic description of the mouse NF2 gene structure in this region, as determined by sequence analysis of mouse genomic DNA (genomic PCR fragments and mouse genomic phage clones, see “Materials and Methods”). Interestingly, in genomic DNA the 16-base pair insertion sequence present in transcript isoform III is contiguous with and represents an extension of the upstream exon. Thus no distinct intron separates the 5' constitutive exon from the alternative exon of 16 base pairs. In contrast, sequencing beyond the 16 nucleotides revealed an intronic sequence (Fig. 3). Similarly, sequencing with primer II-3' revealed intronic sequences at the 5'-boundary of the 45-base pair sequence found in isoform II cDNA. Thus, the 45-base pair sequence represents a distinct exon and is separated from the 3' end of the

Fig. 2. Alternative splice variants of NF2 gene transcripts. Nucleotide sequence and deduced amino acid sequence of transcript isoform III (left) (16-base pair insertion at nucleotide 1740) and transcript isoform II (right) (45-base pair insertion at nucleotide 1740).
using primers 5AS1 and II-3') and an isolated genomic phage clone.

isoform II, that is generated as a variant of isoform I by the inclusion
compared to that in isoform III; and (b) a casette exon, transcript

16-base pair insertion sequence by ~1.4 kilobases. The different
splicing pathways are shown by the diagonal lines. At the mechanistic
level, these appear to include: (a) alternative 5'-donor sites, transcript
isoforms I and II, that are generated by usage of an alternative 5'-
donor site within a particular exon, resulting in excision of an intron
of different length and a reduction in exon size of 16 base pairs, as
compared to that in isoform III; and (b) a casette exon, transcript
isoform II, that is generated as a variant of isoform I by the inclusion
of a distinct alternative exon of 45 base pairs in length. Such splice
mechanisms have been previously described. Competing splice sites
are common in viral transcription units (Ela; SV40 T/t units; Refs. 21
and 22) and have been described for the Drosophila ultrabithorax gene
(23). Cassette exons, for instance, are frequently used in the troponin-T
gene (24).

Hydropathy and structural analysis of the different murine NF2
gene transcripts by means of Kyte-Doolittle analysis (25) predicts that
isoforms II and III encode proteins with higher hydropicity and lack
\( \alpha \)-helical structures at their COOH-termini (data not shown).
Such differences could signify alter functions of the putative alternative
mouse merlin proteins. Alternative splicing is a frequent mechanism
used to generate multiple protein isoforms of different structure, often
varying only in specific domains and thereby allowing for the fine
modulation of protein function. In particular, heterogeneity due to
alternative splicing has been reported to modify effector-target inter-
actions (24). In the case of the erythrocyte protein 4.1, a member of
the moesin-ezrin-radixin family of proteins, alternative splicing at the
COOH-terminus has been suggested to possibly regulate its binding to
the cytoskeletal protein spectrin and thus the mechanical integrity of
the RBC membrane (26).

As indicated by the RT-PCR analysis of different mouse tissue
RNAs using primers flanking the insertion site at nucleotide 1740
(Fig. 4, top), a variable pattern of expression is observed for the
various transcript isoforms. The identity of all three PCR-amplified
DNA fragments was confirmed by subcloning and sequencing. Iso-
form II is more abundantly expressed than isoform I in tissues such as
brain, heart, liver, and lung. In contrast, in spleen and testis, isoform
I is the predominant expressed species with minor expression of
isoform II. This indicates differences in splice site usage and may
depend on the presence of specific, transacting factors. Weak but
detectable expression of isoform III was observed only in spleen and
testis. Although originally identified as an isolated cDNA clone from
a mouse brain cDNA library, we could not detect any amplification of
isoform III from mouse brain RNA. This may reflect a very low or
absent expression of this isoform in adult mouse brain as compared to
newborn mouse brain from which the cDNA library was derived. In
spleen and testis, isoform I is the predominant species. It is possible
that in these tissues, besides the possible presence of specific, trans-
acting factor(s), this type of splice site selection allows a somewhat
more efficient usage of the alternative 5' donor site. Splice junction 1
shows a higher sequence conservation than splice junction 2 to the
strict consensus splice site sequence (C/A)AG/GURAGU (see Fig. 3).
Although the most critical nucleotides are the last two and the first two
of the exon and intron, respectively, flanking sequences do affect
efficiency of splice site usage. It is possible that such a difference in
sequence (see Fig. 3) dictates the less efficient formation of isoform
III. Alternatively, its formation may be restricted to a particular cell
type and/or developmental stage. Further studies will be required to
address this issue.

The overall pattern of expression of the NF2 gene was determined
by Northern blot analysis as shown in Fig. 4, bottom. The mouse NF2
gene is widely expressed in tissues such as heart, brain, spleen, lung,
liver, skeletal muscle, kidney, and testis. A single, predominantly
hybridizing species of ~6 kilobases was detected in all tissues as
compared to the 7.5-kilobase, 4.4-kilobase, and 2.6-kilobase species
detected in human tissues (16), with highest expression in heart, brain
and testis (it should be noted that the signal above the 6-kilobase band
present in the lane showing lung tissue represents a nonspecific back-
ground signal). A faint band around ~4.4 kilobases is detectable in
several lanes. The nature of this band is unclear and may also repres-
ent cross hybridization to other members of the mer-family of pro-
teins. The nature of the different mRNA species in humans also needs
to be clarified. In summary, the expression of the mouse NF2 gene is
also clearly not restricted to the central nervous system, and it is likely
to play a more general role in cellular growth control.

The COOH-terminus appears likely to be an important domain of
the merlin protein, possibly associated with its growth inhibitory
functions. Many mutations identified in human tumors introduce stop
codons or reading frame shifts, resulting in predicted proteins with
truncated and structurally altered COOH-termini (16, 17). Mutations
have also been detected within the most COOH-terminal 50 amino
acids of the predicted human merlin protein. The different mouse
transcript isoforms described in this report originate from nucleotide
NF2 cDNA AND TISSUE EXPRESSION OF MULTIPLE ISOFORMS

Fig. 4. RT-PCR analysis and expression of NF2 gene transcript isoforms. A, RT-PCR was performed using primers 5AS1 and 3AS1 flanking the alternative splice site (nucleotide 1740). DNA markers and three PCR products representing isoforms I, II, and III are shown. B, mouse multiple tissue Northern blot. Molecular size standards in kilobases (Kb) are indicated.

insertions within this COOH-terminal region and may thus encode functional changes in the protein. That the alternative isoform II is also perfectly conserved between mouse and human homologues further suggests that these different forms, arising from alternative splicing, bear functional relevance. A high degree of conservation of the NF2 gene may also be shared with species other than mouse. Using a human NF2 cDNA probe, strong hybridization on Southern blots is detected to DNA from species such as rat, dog, cow, rabbit, and chicken (data not shown). Thus, the NF2 gene appears to be highly conserved during evolution, which may reflect a low tolerance for changes in amino acid sequence as related to function of the protein product.

The high degree (98%) of homology between the mouse and human NF2 gene products could mean that these forms may be used interchangeably in functional studies, i.e., human versions of the protein may be used reliably in mouse model systems. Furthermore, the availability of characterized human mutants and cloned alternative forms of mouse NF2 gene transcript homologues should now serve as important tools in the characterization of the biochemical and functional properties of the NF2 tumor suppressor gene and its encoded product.

References


Molecular Cloning and Characterization of Alternatively Spliced Transcripts of the Mouse Neurofibromatosis 2 Gene

Tetsuo Hara, Albert B. Bianchi, Bernd R Seizinger, et al.