The DNA Rearrangement That Generates the *TRK-T3* Oncogene Involves a Novel Gene on Chromosome 3 Whose Product Has a Potential Coiled-Coil Domain

ANGELA GRECO,^{1*} CRISTINA MARIANI,^{1†} CLAUDIA MIRANDA,¹ ANDREI LUPAS,²
SONIA PAGLIARDINI,¹ MAURO POMATI,³ AND MARCO A. PIEROTTI¹

*Division of Experimental Oncology A, Istituto Nazionale Tumori,¹ and Servizio di Ematologia,
Istituto di Scienze Mediche, Università di Milano, Ospedale Maggiore IRCCS,³ Milan,
Italy, and Max-Planck-Institut für Biochemie, D-82152 Martinsried, Germany²*

Received 24 July 1995/Accepted 10 August 1995

Oncogenic rearrangements of the *NTRK1* gene (also designated *TRK4*), encoding one of the receptors for the nerve growth factor, are frequently detected in thyroid carcinomas. Such rearrangements fuse the *NTRK1* tyrosine kinase domain to 5'-end sequences belonging to different genes. In previously reported studies we have demonstrated that *NTRK1* oncogenic activation involves two genes, *TPM3* and *TPR*, both localized similarly to the receptor tyrosine kinase, on the q arm of chromosome 1. Here we report the characterization of a novel *NTRK1*-derived thyroid oncogene, named *TRK-T3*. A cDNA clone, capable of transforming activity, was isolated from a transformant cell line. Sequence analysis revealed that *TRK-T3* contains 1,412 nucleotides of *NTRK1* preceded by 598 nucleotides belonging to a novel gene that we have named *TFG* (*TRK*-fused gene). The *TRK-T3* amino acid sequence displays, within the TFG region, a coiled-coil motif that could endow the oncoprotein with the capability to form complexes. The *TRK-T3* oncogene encodes a 68-kDa cytoplasmic protein reacting with *NTRK1*-specific antibodies. By sedimentation gradient experiments the *TRK-T3* oncoprotein was shown to form, in vivo, multimeric complexes, most likely trimers or tetramers. The *TFG* gene is ubiquitously expressed and is located on chromosome 3. The breakpoint producing the *TRK-T3* oncogene occurs within exons of both the *TFG* gene and the *NTRK1* gene and produces a chimeric exon that undergoes alternative splicing. Molecular analysis of the *NTRK1* rearranged fragments indicated that the chromosomal rearrangement is reciprocal and balanced and involves loss of a few nucleotides of germ line sequences.

Chromosomal rearrangements producing chimeric oncogenes are frequently associated with human cancer, and several lines of evidence suggest that they are involved in the pathogenesis of their respective tumors (31).

Among solid tumors, papillary thyroid carcinoma provides a unique model of a frequent generation of chimeric oncogenes. In this type of neoplasia, in fact, two proto-oncogenes, *RET* and *NTRK1* (also named *TRK4*), have been found rearranged in about 50% of the samples assayed (5). *RET* and *NTRK1* both encode transmembrane receptor tyrosine kinases (RTKs), whose expression is highly tissue specific, being restricted to specific components of the peripheral nervous system (1, 38, 40). The common mechanism of activation is represented by somatic rearrangements juxtaposing their tyrosine kinase (TK) domain to 5'-end sequences derived from unrelated loci and producing chimeric oncogenes whose products display a constitutive and ectopic TK enzymatic activity (29).

The *NTRK1* gene encodes one of the receptors for the nerve growth factor (NGF) (15, 16) and is located on the q arm of chromosome 1 (24, 26). *NTRK1* was originally detected as an oncogene in a human colon carcinoma; this activated version of the gene was generated by a chromosomal rearrangement fusing the *NTRK1* TK domain with sequences of a tropomyosin gene, *TPM3*, and was designated the "TRK oncogene" (21).

Afterwards, by DNA transfection/focus formation assay, we have detected the activation of *NTRK1*-derived oncogenes in 8 of 52 cases of human papillary thyroid carcinoma (5). The subsequent molecular analysis revealed that in these tumors the oncogenic rearrangements involve at least three different genes. In three cases the activating 5' portion is provided by a tropomyosin gene (39) as in the original *TRK* oncogene isolated from a colon carcinoma (21). In three other cases the *TPR* (translocated promoter region) gene, first identified as part of the *MET* oncogene (27), is involved, and it is responsible for the generation of two different oncogenes, named *TRK-T1* and *TRK-T2* (12). In the remaining two cases, *NTRK1* activation was due to recombination with genes different from both the tropomyosin gene and *TPR*.

In this article we report the molecular cloning and the characterization of one of the latter oncogenes, hereafter designated *TRK-T3*, and the analysis of the relative breakpoint region. A full-length *TRK-T3* cDNA clone was isolated from an NIH 3T3 primary transformant and found to display transforming activity. Sequence analysis has revealed that it contains 1,412 nucleotides of the *NTRK1* gene preceded by 598 nucleotides belonging to a novel gene that we named *TFG* (*TRK*-fused gene). We have established that the *TFG* gene is ubiquitously expressed and that it is localized on chromosome 3. The *TRK-T3* cDNA clone encodes a 68-kDa cytoplasmic protein reacting with *NTRK1* antibodies and displaying, within the *TFG*-derived sequence, a coiled-coil domain most likely providing the capability to form complexes. We show that *TRK-T3* oncoprotein can form, in vivo, aggregates involving three or four molecules.

* Corresponding author. Mailing address: Division of Experimental Oncology A, Istituto Nazionale Tumori, Via G. Venezian 1, 20133 Milano, Italy.

† Present address: Laboratorio di Patologia Molecolare, Divisione di Ematologia, Ospedale San Gerardo, Monza, Italy.

MATERIALS AND METHODS

Cell culture and transfection. Mouse NIH 3T3 fibroblasts and human HeLa cells were maintained in Dulbecco's modified Eagle's medium supplemented with 10% calf serum and 10% fetal calf serum, respectively. CaPO₄-mediated DNA transfection was performed as previously described (5). NIH 3T3 cells (2.5×10^5 /10-cm plate) were transfected with 20 or 100 ng of plasmid DNA coprecipitated with 20 μ g of mouse DNA. Foci were fixed or isolated for further studies 2 weeks later. HeLa cells (10^6 /10-cm plate) were transfected by using 20 μ g of plasmid DNA plus 10 μ g of high-molecular-weight carrier DNA. Cells were harvested 2 days after transfection and assayed for the presence of TRK-T3 oncoprotein by Western blot (immunoblot).

Density gradient sedimentation. NF874 cells were lysed in CSK buffer (100 mM KCl, 10 mM PIPES [piperazine-*N,N'*-bis(2-ethanesulfonic acid), pH 7.0], 0.5% Nonidet P-40, 300 mM sucrose, 3 mM MgCl₂) containing 10 mM iodacetamide. Lysates were layered over 4.5-ml gradients of 5-to-20% sucrose containing 100 mM NaCl, 20 mM Tris (pH 7.5), and 0.1% Nonidet P-40. Gradients were centrifuged in a Beckman SW55 rotor at 39,000 rpm for 18 h at 4°C. Fractions of 150 μ l were collected dropwise from the bottom of the gradient and assayed for the presence of TRK-T3 protein by Western blot.

Western blot analysis. Cell lysates were prepared as described elsewhere (6). After immunoprecipitation with an affinity-purified NTRK1 antiserum, the protein samples were electrophoresed on a sodium dodecyl sulfate (SDS)-polyacrylamide gel (8.5%), transferred to nitrocellulose filters, and immunoblotted with the anti-NTRK1 or antiphosphotyrosine antiserum (UBI). Immunoreactive bands were visualized by using horseradish peroxidase-conjugated secondary antiserum and enhanced chemiluminescence (Amersham).

Nucleic acid extraction and blot analysis. DNA and RNA extraction and Southern and Northern (RNA) blot analyses were performed by standard procedures (36). Poly(A)⁺ RNA was purified by using the Quick Prep mRNA purification kit (Pharmacia) as suggested by the manufacturer. The different *NTRK1*-specific probes, kindly provided by M. Barbacid, comprised plasmid pLM6, containing the full-length cDNA (22), and plasmids pDM56 and pDM17, containing the 5' and the 3' moieties of the gene, respectively (10).

Genomic cloning. Fifty micrograms of high-molecular-weight DNA from the T75 tumor and NF874 NIH 3T3 transformant cell line were digested with the *EcoRI* restriction enzyme and fractionated by electrophoresis on a 0.8% agarose gel. DNA of the appropriate molecular weight was electroeluted and inserted into the EMBL4 phage vector. Library screening was performed by standard procedures (36).

cDNA isolation. cDNA synthesis was performed with oligo(dT)-primed NF874 transformant poly(A)⁺ RNA by using the cDNA Synthesis System (GIBCO-BRL). cDNA cloning was done according to standard procedures (36). The full-length *TRK-T3* cDNA was isolated by PCR amplification as described below.

Chromosomal localization. An informative panel of human-rodent somatic cell hybrids (24, 32) were analyzed by PCR for the presence of human *TFG* sequences. The following primers were used: primer 6, 5'-CTATAATTCAACA TCAAAGCT-3'; primer 7, 5'-AAAAGAGAATTGTGTGAAAATT-3'; primer 8, 5'-GAAACTATTGAAATAACTG-3'; and primer 5, 5'-GGTTTGTATAA CTTTCACATA-3'. The PCR conditions are described below.

Immunofluorescence microscopy. Cells were fixed with 4% paraformaldehyde in phosphate-buffered saline (PBS) for 20 min. After being washed with PBS, cells were permeabilized with 0.2% Triton X-100 in PBS for 10 min, washed twice with PBS, and treated with the first antibody (anti-NTRK1, 10 μ g/ml in PBS) overnight in a moist chamber at 4°C. After three washes with PBS, the samples were permeabilized as described above, washed twice with PBS, and incubated with fluorescein-labeled goat anti-rabbit immunoglobulins (Mely, Springfield, Va.) for 30 min at room temperature. Cells were analyzed by confocal laser scanning microscopy using a Bio-Rad MRC 600 equipped with a krypton-argon laser and mounted on a Nikon Optiphot-2 microscope. Since in this study a fluorescein isothiocyanate secondary antibody was employed, a single wavelength of 488 nm was used. The images were collected with a 60 $\times$ 1.4 NA objective and then elaborated by using CoMOS (Confocal Microscope Operating Software) (Bio-Rad; version 2.05).

PCR amplification. Full-length *TRK-T3* cDNA was amplified from the cDNA synthesis products by using primers 1 (5'-AACATCCTGGAGTCCACCAT-3', nucleotides 1 to 20 in Fig. 2A) and 2 (5'-TTTTTTTCAAGGGATAATAAA-3', nucleotides 1988 to 2010 in Fig. 2A). Reactions were performed in a final volume of 100 μ l. In particular, 1/15 of the cDNA synthesis product was mixed with H₂O, 1 $\times$ *Taq* polymerase buffer, and 1.5 mM MgCl₂; denatured for 2 min at 99°C; and chilled on ice. After addition of 200 μ M deoxynucleoside triphosphates (dNTPs) and 20 pM each primer, samples were heated at 85°C for 3 min, and then 2 U of *Taq* polymerase was added. After a 2-min denaturation at 92°C, 30 cycles of PCR were carried out (94°C for 20 s, 52°C for 20 s, 72°C for 100 s) and were followed by a final extension at 72°C for 5 min. Fifteen microliters of the PCR product was electrophoresed on a 2% agarose gel and visualized by ethidium bromide staining.

PCR amplification for other purposes was performed in a final volume of 100 μ l containing the template (1 μ g of high-molecular-weight DNA or an aliquot of the reverse transcriptase reaction mixture), 1 $\times$ *Taq* polymerase buffer (Perkin-Elmer), 2 U of *Taq* polymerase (Perkin-Elmer), 45 pM each primer, 200 mM dNTPs, and 2 mM MgCl₂. After an initial denaturation at 95°C for 4 min, 30

HeLa T75 M76

FIG. 1. Detection of *NTRK1* rearranged sequences in the T75 thyroid tumor and its related metastasis (M76) by Southern Blot hybridization to an *NTRK1*-specific probe. DNAs were digested with the *EcoRI* restriction endonuclease, electrophoresed on a 0.8% agarose gel, transferred to a GeneScreen membrane, and hybridized to a probe containing the full-length *NTRK1* cDNA (plasmid pLM6). Human DNA from HeLa cells is shown as a control. Tumor-specific *NTRK1* rearranged fragments are indicated (arrowheads).

cycles of PCR were carried out (94°C for 30 s, 56°C for 30 s, 72°C for 1 min), and were followed by a final extension at 72°C for 5 min. Fifteen microliters of the PCR product was electrophoresed on a 4% NuSieve gel and visualized by ethidium bromide staining. For amplification of the *TRK-T3* alternatively spliced exon, primer 3, 5'-AACAGTCTACTCAGGTAT-3', and primer 4, 5'-AACTT GTTTCACCGTCCAC-3', were used. For amplification of the *TFG* genomic region spanning the breakpoint, primer 5, 5'-GGTTTGTATAA CTTTCACATA-3', and primer 6, 5'-CTATAATTCAACATCAAAGCT-3', were used.

Cloning of PCR products. Twenty-microliter samples of PCR products were ligated with the pCRII (Invitrogen) or pGEM-T (Promega) plasmid vector under conditions suggested by the supplier.

DNA sequencing and sequence analysis. Genomic and cDNA fragments inserted into the pGEM4 plasmid vector and PCR-amplified fragments cloned into the pCRII or pGEM-T vector were subjected to nucleotide sequencing by the dideoxy chain termination method (37) using the Sequenase kit (U.S. Biochemical). A computer-assisted homology search of the EMBL data bank was performed. The search for predicted coiled-coil domains was performed by using COILS version 2.0, a program that evaluates the coiled-coil propensity of a sequence by the statistical significance of its similarity to a coiled-coil sequence profile (20). TFG was compared both to a profile derived from myosins, keratins, and tropomyosins and to a profile derived from a larger database that also included other intermediate filaments, kinesins, and desmosomal proteins. The final analysis was performed with a scanning window of 21 residues since the length of the potential coiled-coil segment was found to be shorter than four heptads. The top scores with the two profiles were 1.59 and 1.65, corresponding to probabilities of 84 and 97%, respectively.

RESULTS

Detection of a novel *NTRK1* tumor-specific rearrangement in a human papillary thyroid carcinoma. The evidence that tumor T75 and a related lymphonodal metastasis (M76) carry a tumor-specific rearrangement of the *NTRK1* gene was previously reported (5). By Southern blot analysis a *NTRK1* rearranged fragment was detected in tumor DNA but not in DNA from peripheral blood lymphocytes of the same patient (5). A Southern blot analysis of *EcoRI*-restricted DNAs from the primary tumor and from its metastasis hybridized with the full-length *NTRK1* cDNA probe (pLM6 plasmid) (22) is shown in Fig. 1. In the neoplastic specimens, besides the two germ line fragments of 12 and 20 kb present in control DNA from HeLa cells, the probe detects two novel rearranged *EcoRI* fragments of approximately 14 and 18 kb. Hybridization with different portions of the *NTRK1* cDNA revealed that the 18-kb fragment contains the TK domain, while the 14-kb fragment contains the 5' portion of the *NTRK1* gene (data not shown). The 18-kb fragment is part of the rearranged oncogene, being present in NIH 3T3 foci derived from T75 DNA (data not shown).

In the attempt to identify the sequences activating the *NTRK1* gene in tumor T75, we hybridized Southern and North-

ern blots of the NF874 primary transformant cell line with probes specific for the tropomyosin and *TPR* genes, already detected as activating sequences of the *TRK* (21) and *TRK-T1* (12) oncogenes, respectively. No signals were observed in these analyses (data not shown), thus indicating that tumor T75 contains a novel *NTRK1* oncogene, which we named *TRK-T3*.

Molecular cloning and sequence analysis of *TRK-T3* cDNA. We have previously reported that the *TRK-T3* oncogene is transcribed into a 2-kb mRNA (12) and encodes a protein of approximately 68 kDa (6). In an effort to elucidate the structure of the *TRK-T3* oncogene, we isolated the specific cDNA. A cDNA library was constructed in the λ gt10 phage vector by using oligo(dT)-primed poly(A)⁺ RNA from the NF874 primary transformant cell line. The screening with the pDM17 probe, containing the *NTRK1* 3' moiety (10), identified five positive clones bearing inserts of approximately 1.9 kb; the latter were subcloned into the pGM4 plasmid vector and sequenced. Sequence analysis revealed that all the cDNA clones contained the 3' moiety of the *NTRK1* gene, including the TK domain, preceded by novel sequences. A difference around the breakpoint region was noted, and it will be discussed later. The clones were slightly different in length at the 5' end: the shortest ones (pGT3.1, -2, and -3), starting at nucleotide 221 (Fig. 2A), did not contain an ATG initiation codon in frame with the *NTRK1* translation; the longest ones (pGT3.4 and 5), starting at nucleotide 92 in Fig. 2A, revealed the presence of an ATG codon (nucleotide 133; Fig. 2A) flanked by sequences partially matching the Kozak consensus (18). Translation starting from this triplet should produce a protein of 552 amino acids (61.5 kDa), smaller than that observed in Western blot analysis (68 kDa) (6). In spite of this difference, we proceeded to assess whether the longest cDNA clones would contain a functional *TRK-T3* cDNA, considering that the predicted molecular weight could be increased by posttranslational modifications. Therefore, one of the *TRK-T3* cDNA clones described above was inserted into the pRC/CMV expression vector (Invitrogen) in the sense and antisense orientations with respect to the cytomegalovirus promoter. The resulting plasmids (pT3E5A and -B) were transiently transfected into HeLa cells (see Materials and Methods), and protein cell extracts were assayed for the presence of TRK-T3 oncoprotein by Western blot analysis. No TRK-T3 protein was detected in extracts from cells transfected with pT3E5A and -B (data not shown), thus indicating the inability of the proposed ATG to function as an initiation codon.

To provide more information about the 5' end of the *TRK-T3* cDNA, we isolated a genomic clone containing the corresponding region of the activating gene (unpublished results). Sequence analysis revealed the presence of an ATG codon, 114 bp upstream of the previous one (nucleotide 19; Fig. 2A). This triplet represents most likely the true translation starting site for two reasons: (i) its flanking sequences perfectly conform to the Kozak consensus (18) and (ii) it is preceded by an in-frame TAG termination codon separated by 18 nucleotides that do not match consensus splicing acceptor sequences. In order to isolate *TRK-T3* cDNA clones including this ATG, cDNA synthesis using oligo(dT)-primed NF874 mRNA was performed. *TRK-T3* cDNA was then amplified by PCR using as primers oligonucleotides derived from the genomic region upstream to the ATG codon (primer 1, nucleotides 1 to 20 in Fig. 2A) and from the 3' end of the *TRK-T3* oncogene, including the poly(A) tail (primer 2, nucleotides 1988 to 2010 in Fig. 2A). Subsequently, it was inserted into the PCRII plasmid vector to produce plasmid PCRT3.

The nucleotide sequence of the *TRK-T3* cDNA contained in plasmid PCRT3 is shown in Fig. 2A (EMBL data library ac-

cession number X85960). It consists of 2,010 nucleotides, 1,412 of which belong to the *NTRK1* 3' moiety, including the transmembrane and the TK domains. With respect to the *NTRK1* cDNA, two nucleotide differences were observed. They are coincident with those reported for *TRK-T1* (12) and do not affect the amino acid sequence; most likely, they represent a silent polymorphism. In clone PCRT3 the breakpoint occurred at nucleotide 1261 of the *NTRK1* cDNA (22), similarly to the *TRK* (21) and *TRK-T2* oncogenes (10) but differently from *TRK-T1* (12). The 598 bp joined to *NTRK1* does not match any sequence in the EMBL data bank, thus indicating that it belongs to a novel gene. We propose to name it *TFG*.

The *TRK-T3* sequence contains an open reading frame coding for a protein of 592 amino acids, 193 of which are encoded by the *TFG* gene and 398 of which are encoded by the *NTRK1* gene; the Asp in position 194 is encoded by a new codon (GAC) created by the rearrangement, as judged by comparing the *TRK-T3* cDNA with the genomic sequences of *TGF* (unpublished results) and *NTRK1* (10). The predicted molecular size of the TRK-T3 protein is 66 kDa, compatible with that determined by Western blot analysis (68 kDa; Fig. 3) (6).

Biological activity assay of *TRK-T3* cDNA. The *TRK-T3* cDNA contained in clone PCRT3 was inserted into the pRC/CMV expression plasmid in the sense and antisense orientations with respect to the cytomegalovirus promoter present in the vector. The resulting plasmids, pT3E19A and -B, respectively, were transiently transfected into HeLa cells (see Materials and Methods). Western blot analysis with *NTRK1* antibodies revealed the presence of the 68-kDa TRK-T3 protein in extracts from cells transfected with pT3E19A (Fig. 3A, lane 2), thus indicating that it contains a functional clone of the *TRK-T3* oncogene. As previously reported (6), the TRK-T3 oncoprotein is constitutively phosphorylated on tyrosine (Fig. 3B). To test for their transforming activity, the expression plasmids were transfected into NIH 3T3 cells; plasmid pT3E19A was able to induce transformed foci (2×10^3 foci per μ g of DNA). As expected, the foci expressed the TRK-T3 oncoprotein (Fig. 3A, lane 3). HeLa cells transfected by an antisense *TRK-T3* cDNA did not contain detectable TRK-T3 oncoprotein (Fig. 3A, lane 1, and B).

***TRK-T3* encodes a cytoplasmic protein.** In order to determine the cellular localization of the *TRK-T3* oncoprotein, we have performed the experiments reported in Fig. 4. NIH 3T3 cell lines transformed by the *TRK-T3* oncogene (Fig. 4A) and by the original *TRK* oncogene (Fig. 4B), whose product has been already reported as cytoplasmic (25), were subjected to immunofluorescence with the *NTRK1* antiserum, as described in Materials and Methods. The antiserum recognizes a cytoplasmic protein in cells expressing the *TRK-T3* and *TRK* oncogenes but not in untransfected NIH 3T3 cells (Fig. 4C). No reaction was detected in *TRK-T3*-expressing cells treated with the secondary antibody alone (Fig. 4D).

Alternative splicing produces two different *TRK-T3* mRNAs. As anticipated above, a fraction of *TRK-T3* cDNA clones were identical in structure to the clone PCRT3 (Fig. 2A), while the remaining ones showed an insertion of 78 nucleotides at the *TFG/NTRK1* junction (Fig. 2B). Of the latter, 9 nucleotides belong to the *TFG* gene and 69 nucleotides belong to the *NTRK1* gene. This second class of cDNA clones shows the same junction as the genomic clone (see below), which occurs within an exon of the *NTRK1* gene, at nucleotide 1192 of the cDNA (10). To rule out the possibility that the two types of clones could be the product of a cloning artifact, reverse transcriptase PCR with RNA from the NF874 primary transformant was performed. Primers used for amplification correspond to nucleotides 482 to 500 and 717 to 746 of *TRK-T3*

A

```

aaaaccacttttatcagctcttctctagagttgtatatatagAACATCCTGGAGTCCACCATGAACGGACAGTTGGATCTAAGTGGGAAG 48
***MetAsnGlyGlnLeuAspLeuSerGlyLys
CTAATCGTCAAAGCTCAACTGGGGAGGATATTCGGCGAATTCCTATTCAATGAAGATATTACTTATGTAATTAGTGCTAATGATG 138
LeuIleValLysAlaGlnLeuGlyGluAspIleArgArgIleProIleHisAsnGluAspIleThrTyrAspGluLeuValLeuMetMet
CAACGAGTTTTTCAGAGGAAACTTCTGAGTAATGATGAAGTAACAATAAAGATGAAGATGGAGATCTTATAACAATTTTTGAT 228
GlnArgValPheArgGlyLysLeuLeuSerAsnAspGluValThrIleLysTyrLysAspGluAspGlyAspLeuIleThrIlePheAsp
AGTTCTGACCTTTCTTTGCAATTCAGTGCAGTAGGATACTGAACTGACATTATTTGTTAATGGCCAGCAAGACCTTTGAATCAAGT 318
SerSerAspLeuSerPheAlaIleGlnCysSerArgIleLeuLysLeuThrLeuPheValAsnGlyGlnProArgProLeuGluSerSer
CAGGTGAAATATCTCCGTCGAGAACTGATAGAACTCGAAATAAAGTGAATCGTTTATTGGATAGCTTGGAAACCACTGGAGAACCGGA 408
GlnValLysTyrLeuArgArgGluLeuIleGluLeuArgAsnLysValAsnArgLeuLeuAspSerLeuGluProProGlyGluProGly
CCTTCCACCAATATTCCTGAAATGATCTGTGGATGGTAGGGAAGAAAGTCTGCTTCTGATTCTTCTGGAAACAGTCTACTCAGGTT 498
ProSerThrAsnIleProGluAsnAspThrValAspGlyArgGluGlyLysSerAlaSerAspSerSerGlyLysGlnSerThrGlnVal
ATGGCAGCAAGTATGTCTGCTTTTATCTCTTTAAAAACCAAGATGAAATCAATAAAATGTTATGTCAGCGTTTGGCTTAACAGATGAT 588
MetAlaIleSerMetSerAlaPheAspProLeuLysAsnGlnAspGluIleAsnLysAsnValMetSerAlaPheGlyLeuThrAspAsp
TFG
CAGGTTTCACTACACAGCAGTCTGTGAGACCCGGTGGAGAGAAAGGACGAAACACCTTTTGGGGTCTCGGTGGCTGTGGCGCTGGCC 678
GlnValSerAspThrAsnSerThrSerGlyAspProValGluLysLysAspGluThrProPheGlyValSerValAlaValGlyLeuAla
TM
GTCTTTGCCTGCCTTCTCTTCTACGCTGCTCTTGTGCTCAACAAATGTGGACGGAGAAACAGTTTGGGATCAACCGCCCGCTGTC 768
ValPheAlaCysLeuPheLeuSerThrLeuLeuLeuValLeuAsnLysCysGlyArgArgAsnLysPheGlyIleAsnArgProAlaVal
CTGGCTCCAGAGGATGGGCTGCCATGTCCTGCATTTTCATGACATTGGGTGGCAGCTCCCTGCCCCACCGAGGGCAAGGCTCTGGG 858
LeuAlaProGluAspGlyLeuAlaMetSerLeuHisPheMetThrLeuGlyGlySerLeuSerProThrGluGlyLysGlySer
TK
CTCAAGGCCACATCATCGAACCACAACTACTTCAGTGATGCTGTGTTCCACCATCAAGCGCCGGGACATCGTGCTCAAGTGGGAG 948
LeuGlnGlyHisIleIleGluAsnProGlnTyrPheSerAspAlaCysValHisHisIleLysArgArgAspIleValLeuLysTrpGlu
CTGGGGAGGGCGCCTTTGGGAAGGCTCTCTTGTGAGTCCCAACCTCTGCTGAGCAGGACAAAGATGCTGGTGGCTGTCAAGGCA 1038
LeuGlyGluGlyAlaPheGlyLysValPheLeuAlaGluSerHisAsnLeuLeuProGluGlnAspLysMetLeuValAlaValLysAla
CTGAAGGAGCGCTCCGAGAGTGCTCGGAGGACTTCCAACGTGAGGCTGAGCTGCTCACCATGCTGACGACCAAGCAGCATCGTGCGCTTC 1128
LeuLysGluAlaSerGluSerAlaArgGlnAspPheGlnArgGluAlaGluLeuThrMetLeuGlnHisGlnHisIleValArgPhe
TTCGGCGTCTGCACCGAGGGCGCCCTGCTCATGGTCTTTGAGTATATGCGGCACGGGGACCTCAACCGCTTCTCCGATCCCATGGA 1218
PheGlyValCysThrGluGlyArgProLeuLeuMetValPheGluTyrMetArgHisGlyAspLeuAsnArgPheLeuArgSerHisGly
CCTGATGCCAAGCTGCTGGCTGGTGGGAGGATGTGGCTCCAGGCCCTGGGTCTGGGGCAGCTGCTGGCTGTGGCTAGCCAGGTCGCT 1308
ProAspAlaLysLeuLeuAlaGlyGlyGluAspValAlaProGlyProLeuGlyLeuGlyGlnLeuLeuAlaValAlaSerGlnVal
GCGGGGATGGTGTACCTGGCGGTGTCATTTTGCACCGGGGACCTGGCCACGCACTGTCTAGTGGGCCAGGACTGGTGGTCAAG 1398
AlaGlyMetValTyrLeuAlaGlyLeuHisPheValHisArgAspLeuAlaThrArgAsnCysLeuValGlyGlnGlyLeuValValLys
ATTGGTGATTTTGGCATGAGCAGGGATATCTACAGCAGGACTATTACCGTGGGAGGCCGACCATGCTGCCATTCGCTGGATGCCG 1488
IleGlyAspPheHisAlaSerAlaSerIleTyrSerThrAspTyrTyrArgValGlyGlyArgThrMetLeuProIleArgTrpMetPro
CCGAGAGCATCTGTACCTAAGTTCACACCGAGAGCGAGCTGTGGAGCTTCGGCGTGGTCTGGGAGATCTTACCTACGGCAAG 1578
ProGluSerIleLeuTyrArgLysPheThrThrGluSerAspValTrpSerPheGlyValValLeuTrpGluIlePheThrTyrGlyLys
CAGCCCTGGTACCACTCTCCAACACGGAGGCAATCGACTGCATCAGCAGGGAGCTGAGTTGGAGCGGCCACGTCCTGCCACCAAG 1668
GlnProTrpTyrGlnLeuSerAsnThrGluAlaIleAspCysIleThrGlnGlyArgGluLeuGluArgProArgAlaCysProProGlu
GTCTACGCCATCATGCGGGGCTGCTGGCAGCGGGAGCCCAAGCAAGCATCAAGGATGTGACGCCCGGCTGCAAGCCCTGGCC 1758
ValTyrAlaIleMetArgGlyCysTrpGlnArgGluProGlnGlnArgHisSerIleLysAspValHisAlaArgLeuGlnAlaLeuAla
CAGGCACCTCCTGTCTACCTGGATGCTCTGGGCTAGGGGGCCGCGGAGTGGTGTAGCCGAATCTGGGGCTGCCCTCA 1848
GlnAlaProProValTyrLeuAspValLeuGly***
GCATCCCCATAGCTCCAGCAGCCCCAGGGTGATCTCGAAGTATCTAATTCGCCCTCAGCATGTGGGAAGGACAGGTGGGGCTGGGA 1938
GTAGAGGATGTTCTCTCTCTAGGCAAGGTCCTCATAGCAATTATATTATTATCCCTTGAAAAA 2010

```

B

```

TGF ← NTRK1
oligo 12
GTTTCAGACTTTCATGCTCCGGCTCCATCATGGCTCCCTTCATGAGCAACCTTTTCGAGTTCAACCCGAGGAGCCCATCCCTGACACTAAC 93
ValSerAspPheHisAlaSerAlaSerIleMetAlaAlaPheMetAspAsnProPheGluPheAsnProGluAspProIleProAspThrAsn

```

FIG. 2. (A) Nucleotide and deduced amino acid sequences of *TRK-T3* cDNA. Nucleotides derived from a genomic clone containing the 5' end of the *TFG* gene (see the text) are lowercased. The proposed ATG initiation codon is underlined. In frame TAG termination codons are indicated by asterisks. The *NTRK1* transmembrane (TM) and TK domains are indicated. The putative coiled-coil domain and the codon created by the rearrangement are boxed. The nucleotide differences between *TRK-T3* and *NTRK1* are shown above the sequence. Nucleotide (far right) and amino acid positions are numbered on the right. (B) Nucleotide and amino acid sequences of the *TRK-T3* alternatively spliced exon (lightface). The oligonucleotide (oligo 12) used for detection of the alternative exon is indicated.

cDNA (Fig. 2A). A 295-bp fragment, representative of the shortest mRNA, was detected by ethidium bromide staining (data not shown). However, by Southern blot hybridization of the PCR products with an oligonucleotide whose sequence is present only in the longest *TRK-T3* mRNA (Fig. 2B), a 373-bp fragment was visualized (data not shown), thus confirming the existence of the two mRNAs. In the longest mRNA the translation proceeds in frame with that of *NTRK1* (Fig. 2B), thus suggesting that it could represent a mature mRNA produced by alternative splicing. In extracts from NF874 cells, the TRK-T3 oncoprotein is detectable, by Western blot analysis, as

a single band of approximately 68 kDa. We cannot assess the presence of a protein derived from the longest mRNA given that, under our experimental conditions, it would comigrate with the smallest one; in addition, it would be less abundant than the shortest one, considering the relative amount of the corresponding mRNA.

TRK-T3 contains a potential coiled-coil domain within the TFG region. Computer-assisted analysis of the TFG amino acid sequence contained in TRK-T3 did not show homology with proteins in the EMBL data bank. Furthermore, we investigated the presence of coiled-coil domains in the TFG protein,

FIG. 3. Western blot analysis of TRK-T3 oncoprotein. (A) Cell extracts from HeLa cells transiently transfected with *TRK-T3* antisense (lane 1) and sense (lane 2) cDNAs, NIH 3T3 cells transfected by *TRK-T3* cDNA (lanes 3), and the NF874 primary transformant (lane 4) were processed for Western blot analysis as described in Materials and Methods. The molecular weight of the TRK-T3 oncoprotein (p68) is indicated. IP, immunoprecipitation; α TRK, anti-TRK. (B) Constitutive phosphorylation of TRK-T3 oncoprotein. Cell extracts from HeLa cells transiently transfected with sense and antisense *TRK-T3* cDNAs, treated (+) with sodium orthovanadate (Na₃VO₄) or untreated (-), were analyzed by Western blot as described in Materials and Methods. α pTyr, antiphosphotyrosine.

since such domains, conferring dimerization capability, have been recently proposed to play a crucial role in the activation of *TK* oncogenes (35). The TFG amino acid sequence was analyzed by an automated method based on the statistical evaluation of sequence similarity to a coiled-coil sequence profile (20). By this method, TFG was predicted to contain a coiled-coil domain of approximately four heptads from Leu-97 to Glu-124 (Fig. 5). A search for the most likely *cap* residues using the frequencies of Richardson and Richardson (33) indicated that the coiled-coil helix probably extended for just over three heptads (Gln-101 to Leu-123), thus being shorter than the typical leucine zippers characterized in GCN4 (8) and Max (9). Like leucine zippers, TFG contains only Leu residues

in position d of the heptad repeat, but, unlike them, it also contains only hydrophobic residues in position a, a feature that increases its predicted stability of association and may explain its comparatively short length. In typical leucine zippers, at least one position a towards the middle of the zipper contains a hydrophilic residue (generally, Asn) that increases the specificity of dimer formation at the expense of stability (13). In TFG, the specificity of dimerization may derive from the alternation of Val and Leu residues in position a of the core.

TRK-T3 protein forms complexes in vivo. Ligand-independent activation of *RTK* oncogenes is believed to be mediated by dimerization. The presence of coiled-coil domains within several *RTK* activating sequences supports this model (35). To

FIG. 4. Cellular localization of TRK-T3 oncoprotein. NIH 3T3 cells expressing *TRK-T3* (A and C), *TRK* (B), or the wild type (D) were subjected to immunofluorescence as described in Materials and Methods. Cells were treated with the NTRK1 antiserum (A, B, and D) or with the secondary antibody only (C).

FIG. 5. Predicted coiled-coil propensity of TFG in a 21-residue score. The inset shows the sequence of the putative coiled coil with its corresponding heptad frame; the core three heptads most likely to form the coiled coil are boldfaced.

examine whether dimerization is important for *TRK-T3* activation, we investigated if dimers of *TRK-T3* oncoprotein form in vivo. Proteins isolated from the NF874 cell line were separated by sucrose density gradient fractionation as reported in Materials and Methods. Cells were lysed in the presence of iodoacetamide to prevent postlysis aggregation. Cell lysates were fractionated on 5-to-20% sucrose gradient in the absence of SDS. Fractions were collected from the gradient and analyzed for the presence of *TRK-T3* oncoprotein by Western blot. The *TRK-T3* product was found distributed within the first three fractions (Fig. 6). In a control gradient run in parallel, a 220-kDa molecular mass marker was found distributed with a peak on the third fraction (Fig. 6). These results indicate that *TRK-T3* forms, in vivo, complexes containing three or four molecules. Such arrangement can be promoted by the coiled-coil motif; indeed, it is known that coiled coils can form dimers, trimers, and tetramers (13). Moreover, a trimeric or tetrameric *TRK-T3* complex is still consistent with the proposed mechanism of *RTK* activation by autotransphosphorylation of the TK domains.

TFG gene expression. As part of studies aimed at characterizing the *TFG* gene we have investigated its expression pattern. Northern blot analysis of mRNA from adult and fetal human tissues using a *TFG* cDNA probe (from nucleotide 251 to 598; Fig. 2A) was performed. As shown in Fig. 7 the probe detects a transcript of about 2.4 kb in all tissues tested. In addition, *TFG* expression has been detected in human thyroid cells (data not shown). This result shows that the *TFG* gene is ubiquitously expressed, a common feature shared with other *NTRK1*-activating genes and genes activating other receptor tyrosine kinases.

The *TFG* gene is localized on human chromosome 3. A panel of human-rodent somatic cell hybrids was analyzed by PCR amplification using two combinations of primers specific for *TFG* genomic sequences (primers 7 and 8 and primers 9 and 5; see Materials and Methods) amplifying fragments of 402 and 170 bp, respectively, from human DNA. The presence of the amplified fragments was concordant with the presence of chromosome 3 and was highly discordant with the presence of other

human chromosomes, including chromosome 1 (Table 1). In particular, the amplified fragments were absent in DNA from hybrid p7A2, containing only the p arm of chromosome 3 as a consequence of a t(3;17) translocation (see Table 1, footnote e), thus indicating that 3p does not contain the *TFG* locus.

Molecular cloning and sequence analysis of the genomic rearrangement. In order to unveil the molecular mechanism producing the *TRK-T3* oncogene, we have cloned the 14- and 18-kb *EcoRI* *NTRK1* rearranged fragments detected by Southern blot analysis in DNA from tumor T75 (Fig. 1). The 18-kb *EcoRI* fragment, containing the oncogenic rearrangement, was isolated by screening a genomic library from the NF874 transformant cell line with the pLM6 probe, containing the full-length *NTRK1* cDNA (22). A positive clone, phage λ 4C, was subjected to further characterization. The 12-kb *EcoRI*-*Bam*HI subfragment containing the rearrangement (Fig. 6A) was isolated from phage λ 4C and inserted into the pGEM4 plasmid vector, to produce plasmid pGT3R3'.

The 14-kb *EcoRI* rearranged fragment containing the 5' portion of *NTRK1* (Fig. 1) was cloned from a genomic library of the T75 tumor DNA. Plaques positive for the pLM6 probe that could contain either the germ line or the rearranged *NTRK1* fragments, were subsequently hybridized with the pDM56 and pDM17 probes, containing the 5' and the 3' portions of *NTRK1* cDNA, respectively (10). Phage 15A was positive for probe pDM56 and negative for pDM17, indicating that it contains the rearranged 5' portion of *NTRK1*. The 14-kb *EcoRI* rearranged fragment was excised from phage 15A and inserted into the pGM4 vector, to produce plasmid pGT3R5'.

A restriction analysis of the genomic rearranged fragments is shown in Fig. 8A. Comparison of the resulting pattern with that of the corresponding germ line *NTRK1* region allowed us to localize the *TRK-T3* breakpoint within the 1.4-kb *Xba*I *Hind*III fragment. Breakpoint sites of all the other *TRK* and *TRK-T* oncogenes have been detected within the upstream 1.5-kb *Hind*III-*Sma*I *NTRK1* fragment (10). The sequence analysis of the genomic rearranged fragments compared with *NTRK1* cDNA (22) allowed us to localize the breakpoint within an exon of *NTRK1* (Fig. 8B) (10). As shown in Fig. 6B, the rearrangement joins the last 68 nucleotides of *NTRK1* exon 1 (10) to the sequence of the *TFG* gene. The 9 nucleotides of *TFG* preceding the junction with *NTRK1* belong to an exon, as deduced from comparison between genomic and cDNA *TRK-T3* sequences (Fig. 2B and 8B). Thus, the rearrangement creates a 78-nucleotide chimeric exon that can be alternatively spliced, as reported above.

In order to investigate whether the *TRK-T3* oncogene was generated by a reciprocal recombination involving two loci, we tried to ascertain whether the foreign sequences linked to *NTRK1* in the two rearranged fragments belong to the same locus. If this is the case, then oligonucleotides derived from the rearranged sequences flanking the breakpoint should be able to amplify a specific fragment from human DNA. By using

FIG. 6. Detection of *TRK-T3* complexes in vivo. NF874 cells, expressing the *TRK-T3* oncogene, were lysed in the presence of iodoacetamide; proteins were separated on a 5-to-20% sucrose gradient. Fractions were collected and *TRK-T3* protein was detected by Western blot as described in Materials and Methods. The numbers above the lanes correspond to the fraction numbers from the heaviest region to the lightest region of the gradient. Molecular size standards were included in a control gradient run in parallel. The migration positions of myosin (220 kDa), phosphorylase *b* (97 kDa), bovine serum albumin (68 kDa), and ovalbumin (46 kDa) are indicated.

FIG. 7. Northern blot analysis of adult (A) and fetal (B) human mRNAs for the expression of the *TFG* gene. A multitissue Northern (MTN) blot (Clontech) was hybridized to a *TFG*-specific probe (from nucleotide 264 to 598 [Fig. 2]).

primers 5 and 6 (see Materials and Methods) (Fig. 8A), a fragment of approximately 660 bp was amplified from HeLa DNA (data not shown) and shown to contain the novel sequences of both *NTRK1* rearranged fragments, thus proving that they originate from the same locus. Comparison of the nucleotide sequence around the breakpoint in the rearranged fragments and in the germ line counterparts (Fig. 8B) revealed that the rearrangement did not produce gross alterations, except for a deletion of 11 nucleotides within the *TFG* sequence and a duplication of two C's belonging to *NTRK1*. To try to elucidate the molecular mechanism generating the *TRK-T3* oncogene, we performed a computer analysis of the sequences surrounding the rearrangement. In particular, we looked for sequences known to be relevant for recombination. The 2.9-kb *NTRK1* sequence surrounding the *TRK* breakpoint does not

present Alu repetitive sequences (7). The sequence around the *TRK-T3* breakpoint is GC rich (61% in 1,100 bp). A Pu-Py tract, d(GT)_n (3), of about 35 nucleotides, is present in an intronic region, 1,800 nucleotides downstream from the *TRK-T3* breakpoint (data not shown). Several repeats of the guanine-rich element GGAGGG (17) have been reported for the intron containing most *TRK* oncogene breakpoints (7), downstream from *TRK-T3*. Moreover, the *TRK-T3* breakpoint occurs within a region containing palindromes and direct and inverted repeats. In particular, two sequences matching (7 of 8 nucleotides) the variable tandem repeat (VTR) consensus (19), located 21 nucleotides upstream and 6 nucleotides downstream from the breakpoint, respectively, have been noted. The presence of short repeats has also been detected. With respect to the *TFG* gene, a sequence of 1,689 nucleotides

TABLE 1. Presence of the human *TFG* gene in 12 human-rodent somatic hybrid cell lines

Human chromo- some	Presence or absence in ^a :													No. of discordant hybrids/total no. of hybrids
	CTP34B4 ^b (+)	CTP41A2 ^b (+)	DT1.2.4 ^c (+)	FG10 ^b (-)	3W4CLS ^b (+)	MOG34A4 ^b (-)	DUR4.3 ^c (+)	SIR74W ^b (+)	SIF4A31 ^b (-)	FASC13C112 ^b (-)	TWIN19-D12 ^b (+)	P7A2 ^c (-)		
1	+	-	-	-	-	+	-	+	-	— ^d	+	-	5/11	
2	+	+	-	+	-	-	-	+	-	-	-	-	8/12	
3	+	-	+	-	-	+	+	Tr	+	-	+	— ^e	0/11	
4	-	-	-	-	-	+	-	+	+	-	+	Tr	4/11	
5	+	-	-	+	-	+	+	-	+	-	-	-	3/12	
6	+	+	-	-	-	+	-	-	+	-	+	-	3/12	
7	+	+	-	-	+	+	-	-	-	-	-	-	6/12	
8	+	-	-	+	-	+	-	-	-	-	+	-	4/12	
9	-	-	-	-	-	-	-	-	-	-	-	-	6/12	
10	-	-	— ^d	+	+	+	+	-	-	-	-	-	5/11	
11	-	-	+	-	+	+	+	-	-	-	+	-	3/12	
12	+	-	-	-	+	+	+	+	-	-	+	-	4/12	
13	-	-	-	-	-	+	+	+	-	-	-	-	5/12	
14	+	+	-	-	+	+	+	+	+	-	+	-	5/12	
15	-	-	-	+	+	+	+	-	-	-	-	-	8/12	
16	+	+	-	-	-	+	-	-	+	-	+	-	3/12	
17	+	-	-	-	+	-	+	+	-	-	+	-	4/11	
18	+	-	+	+	-	+	+	Tr	-	-	+	-	2/11	
19	-	-	-	-	-	+	-	-	-	-	-	-	5/12	
20	-	-	+	-	-	-	+	-	-	-	+	-	3/12	
21	-	-	+	+	+	+	+	+	-	-	-	-	6/12	
22	-	-	-	-	-	-	+	Tr	-	-	-	-	4/10	
X	+	+	+	+	+	+	+	+	+	+	-	-	6/11	

^a Results of *TFG* PCR detection are indicated in parentheses. Tr, trace (presence in <10% of cells).
^b See the text.
^c Reference 39a.
^d —, chromosome fragments present.
^e This hybrid contains a t(3;17) 3pter-3p11:17p11-17pter.

A

B

NTRK1 GTCAACAACGGCAACTACACGCTGCTGGCTGCCAACCCCTTCGGCCAGGCCCTCCGCTCCATCATGGCTGCCTTCATGGACAACCCCTTTCGAGTTCAACCCCGAGGACCCCATCCCTGgtgcgagggc
 NTRK1/TFG GTCAACAACGGCAACTACACGCTGCTGGCTGCCAACCCCTTCGGCCAGGCC-----AAAAGAGAATTGTGTGAAATTTGATGAGGTCAGACTTTTAAATTTGATTAAATGTAGTATATTCCA
 TFG atagttaaatgtgtacagtccattttgtatttttagagACTTTCATGGAATTACTCACAAGAGAATTGTGTGAAATTTGATGAGGTCAGACTTTTAAATTTGATTAAATGTAGTATATTCCA
 TFG/NTRK1 atagttaaatgtgtacagtccattttgtatttttagagACTTTCATG-CCTCGCCTCCATCATGGCTGCCTTCATGGACAACCCCTTTCGAGTTCAACCCCGAGGACCCCATCCCTGgtgcgagggc
 NTRK1 GTCAACAACGGCAACTACACGCTGCTGGCTGCCAACCCCTTCGGCCAGGCCCTCCGCTCCATCATGGCTGCCTTCATGGACAACCCCTTTCGAGTTCAACCCCGAGGACCCCATCCCTGgtgcgagggc

FIG. 8. (A) Molecular cloning of NTRK1 rearranged fragments from tumor T75. *TFG/NTRK1* and *NTRK1/TFG* rearranged fragments were isolated from the NF874 transformant and T75 tumor DNA, respectively. Portions belonging to the *NTRK1* gene are shown as lines; those from the *TFG* gene are represented by open boxes. a and b, positions of primers 5 and 6 used in PCR amplification to isolate the *TFG* germ line fragment spanning the breakpoint (see the text). Ac, *AccI*; B, *Bam*HI; E, *Eco*RI; H, *Hind*III; S, *Sma*I; Sc, *Sca*I; X, *Xba*I. (B) Nucleotide sequence analysis of the breakpoint regions and their alignments to corresponding germ line regions. Exon sequences are uppercased; intron sequences are lowercased. Repeats (lines and dashed lines), palindromes (tildes), sequences matching the VTR consensus (triangles), and the stretch of 8 nucleotides (stars) present in both rearranging genes are indicated.

around the breakpoint has been determined (EMBL data library accession number X85959), and it showed a high AT content (65%). Two guanine-rich elements (17) are present, 254 and 313 nucleotides downstream from the breakpoint, respectively. As shown in Fig. 8B, the break occurs in a region containing several repeats and palindromes and, in particular, involves a sequence (TTCATGGA) that is also present in the *NTRK1* gene, 23 nucleotides downstream from the breakpoint. No other homologies between the *TFG* and *NTRK1* sequences around the breakpoint have been detected.

DISCUSSION

We have characterized a new rearrangement of the *NTRK1* gene that occurs in a human papillary thyroid tumor and produces the oncogene *TRK-T3*. A cDNA clone was isolated from a primary transformant and was shown to induce transformation of NIH 3T3 cells. Sequence analysis indicated that *TRK-T3* contains 1,412 nucleotides of the *NTRK1* proto-oncogene preceded by 598 nucleotides that do not match any sequence in the data bank. Thus, the activating portion of *TRK-T3* is provided by a novel gene that we have named *TFG*.

TFG represents the third gene capable of *NTRK1* activation, the other two being the tropomyosin gene (*TPM3*) (21, 39) and *TPR* (12). Here we have reported evidence that the *TFG* gene

is ubiquitously expressed, being detectable in all the tissues analyzed including normal thyroid cells. Computer analysis of the *TRK-T3* amino acid sequence contributed by the *TFG* gene indicated the presence of a leucine repeat motif that is predicted to form a coiled-coil structure, involved in protein-protein interaction. Both the ubiquitous expression and the capability to fold into a coiled-coil structure are features shared by all the *TRK*-activating genes as well as genes activating other *TK* proto-oncogenes (34). It has been recently reported that the *TPR* portion present in the *MET* oncogene contains a coiled-coil motif; this region is essential for efficient transforming activity and has been shown to mediate the constitutive dimerization and activation of the *MET* kinase in a ligand-independent fashion (35). Similarly, all the other *TRK*-derived oncogenes could be activated by constitutive dimerization, judging from the presence of coiled-coil motifs within the activating portions. Sedimentation gradient experiments indicated that *TRK-T3* protein forms multimeric complexes in vivo, most likely trimers or tetramers. Both aggregate forms are consistent with the proposed mechanism of activation, requiring the interaction of at least two *TK* domains.

Cellular localization may also play a role in the transforming activity of *TRK*-derived oncogenes. The products of the *TRK* oncogenes do not span the cell membrane, because of the loss of the signal sequence and, in some cases, also of the trans-

membrane domain from the normal receptor. We have shown that TRK-T3 protein is localized in the cytoplasm, as is the TRK protein (25). In addition, the chimeric oncogenes could acquire a novel localization domain from the fusion partner. It is worth mentioning that TPM3 is a cytoskeletal protein, while TPR has been recently identified as a component of the nuclear pore complex and localized on its cytoplasmic surface (2). Therefore, their putative novel cellular localizations could allow the TKs to interact with unusual substrates, perhaps modifying their functional properties.

Several lines of evidence provided by biological and biochemical studies suggest that the *TRK* oncogenes behave as the normal NTRK1 receptor stimulated by NGF. The *TRK-T1* oncogene induces neuronal differentiation of PC12 cells (11), as does the *NTRK1* receptor upon NGF stimulation. In a preliminary study of the signal transduction triggered by the TRK oncoproteins we have demonstrated that the latter are able to interact with shc and grb2 adaptor molecules. In addition, in cells containing TRK oncoproteins as well as in cells expressing the *NTRK1* proto-oncogene and treated with NGF, the different shc forms are phosphorylated on tyrosine (6). However, the significance of such interactions and the modalities of the subsequent steps of the signal transduction, i.e., whether the TRK oncoproteins, located in the cytoplasm, are still capable of interacting with the membrane-anchored Ras-dependent signalling pathway, remain to be elucidated.

We have localized the *TFG* gene on chromosome 3, most likely on 3q; thus, *TRK-T3* provides the first example of a *TRK* oncogene created by a rearrangement between genes located on different chromosomes. Inversion on chromosome 1, involving either the tropomyosin gene or the *TPR* gene, is the predominant mechanism of *TRK* oncogene activation (7, 9a). Interestingly, chromosomal inversion is the most frequent type of chromosome rearrangement in thyroid tumors. In fact, the *RET* thyroid oncogenes are created mostly by chromosome 10 inversions, involving either the *H4* (30) or the *ELE1* (23) gene, except for two cases of rearrangement with the *RI α* gene on chromosome 17 (4). No cytogenetic data are available for tumor T75; thus, the precise nature of the chromosome rearrangement producing the *TRK-T3* oncogene remains unknown. However, the features of the *NTRK1* rearranged fragments of the T75 tumor, i.e., the molecular evidence of the existence of a fusion fragment reciprocal (5'*NTRK1*-3'*TFG*) to that forming the oncogenic sequence (5'*TFG*-3'*NTRK1*) with the loss of only 11 nucleotides of germ line sequences, would suggest a reciprocal and balanced chromosome translocation t(1q;3) as the most likely rearrangement mechanism.

The *TRK-T3* breakpoint was found to occur within exons of both the *NTRK1* gene and the *TFG* gene. The rearrangement produces a 78-bp chimeric exon that can be alternatively spliced, producing two different mRNAs. We cannot prove the presence of a protein encoded by the longest TRK-T3 mRNA; however, the additional amino acids would not affect the oncogenic capability.

We have recently reported that all the *TRK* oncogenic rearrangements fall within a 2.9-kb region of the *NTRK1* gene; the one producing *TRK-T3* occurs upstream from all the others (10). The analysis of the sequence flanking the TRK-T3 breakpoint in both the *TFG* gene and the *NTRK1* gene has revealed the presence of several sequences involved in DNA rearrangement. Guanine-rich elements (17), direct and inverted repeats (14), and palindromes are present in both rearranging genes; the VTR consensus (19) and Pu-Py tract d(GT)_n (3, 41) have been detected within *NTRK1* sequences. The lack of homology between the two rearranging genes, except for a stretch of 8 nucleotides, suggests that the *TFG/NTRK1* fusion could be

produced by nonhomologous recombination. This mechanism, in fact, is known to require little or no homology and has been proposed as capable of sealing complex DNA breaks caused by ionizing radiation or carcinogens (28).

In conclusion, by studying a new activated form of the *TRK* oncogene, we have discovered a novel gene, *TFG*. The understanding of the physiological role of *TFG* will require extensive molecular and biochemical studies. However, it is tempting to speculate that by analogy with TPM3 and TPR, involved in cytoskeletal assembly and nuclear transport, respectively, the ubiquitously expressed TFG protein also could accomplish an essential cellular function. Oncogenic chromosomal rearrangements are thus providing an unexpected and new tool for unveiling novel cellular components that could be involved in fundamental cell processes.

ACKNOWLEDGMENTS

A. Greco and C. Mariani contributed equally to this work.

We thank Anna Grassi and Giovanna Raineri for secretarial work and Mario Azzini and Maria Teresa Radice for technical assistance.

This work was partially supported by the Associazione and the Fondazione Italiana per la Ricerca sul Cancro, the Special Project of the CNR ACRO, and the Ministero della Sanità, Rome, Italy.

REFERENCES

- Barbacid, M., F. Lamballe, D. Pulido, and R. Klein. 1991. The *trk* family of tyrosine protein kinase receptors. *Biochim. Biophys. Acta* **1072**:115-127.
- Bird, D. A., D. J. Sweet, N. Pante', K. N. Konstantinov, T. Guan, A. C. S. Shapire, P. J. Mitchell, C. S. Cooper, U. Aebi, and L. Gerace. 1994. Tpr, a large coiled coil protein whose amino terminus is involved in activation of oncogenic kinases, is localized to the cytoplasmic surface of the nuclear pore complex. *J. Cell Biol.* **127**:1515-1526.
- Boehm, T., L. Mengle-Gaw, U. R. Kees, N. Spurr, I. Lavenir, A. Forster, and T. H. Rabbitts. 1989. Alternating purine-pyrimidine tracts may promote chromosomal translocations seen in a variety of human lymphoid tumors. *EMBO J.* **8**:2621-2631.
- Bongarzone, I., N. Monzini, M. G. Borrello, C. Carcano, G. Ferraresi, E. Arighi, P. Mondellini, G. Della Porta, and M. A. Pierotti. 1993. Molecular characterization of a thyroid tumor-specific transforming sequence formed by the fusion of *ret* tyrosine kinase and the regulatory subunit RI α of cyclic AMP-dependent protein kinase A. *Mol. Cell. Biol.* **13**:358-366.
- Bongarzone, I., M. A. Pierotti, N. Monzini, P. Mondellini, G. Manenti, R. Donghi, S. Pilotti, M. Grieco, M. Santoro, A. Fusco, G. Vecchio, and G. Della Porta. 1989. High frequency of activation of tyrosine kinase oncogenes in human papillary thyroid carcinoma. *Oncogene* **4**:1457-1462.
- Borrello, M. G., G. Pelicci, E. Arighi, L. De Filippis, A. Greco, I. Bongarzone, M. G. Rizzetti, P. G. Pelicci, and M. A. Pierotti. 1994. The oncogenic versions of the Ret and Trk tyrosine kinases bind Shc and Grb2 adaptor proteins. *Oncogene* **9**:1661-1668.
- Butti, M. G., I. Bongarzone, G. Ferraresi, P. Mondellini, M. G. Borrello, and M. A. Pierotti. 1995. A sequence analysis of the genomic regions involved in the rearrangements between TPM3 and NTRK1 genes producing TRK oncogenes in papillary thyroid carcinomas. *Genomics* **28**:15-24.
- Ellenberger, T. E., C. J. Brandl, K. Struhl, and S. C. Harrison. 1992. The GCN4 basic region leucine zipper binds DNA as a dimer of uninterrupted α helices: crystal structure of the protein-DNA complex. *Cell* **71**:1223-1237.
- Ferré-Amaré, A. R., G. C. Prendergast, E. B. Ziff, and S. K. Burley. 1993. Recognition by Max of its cognate DNA through a dimeric b/HLH/Z domain. *Nature (London)* **363**:38-45.
- Greco, A., et al. Unpublished data.
- Greco, A., C. Mariani, C. Miranda, S. Pagliardini, and M. A. Pierotti. 1993. Characterization of the NTRK1 genomic region involved in chromosomal rearrangements generating TRK oncogenes. *Genomics* **18**:397-400.
- Greco, A., R. Orlandi, C. Mariani, C. Miranda, M. G. Borrello, A. Cattaneo, S. Pagliardini, and M. A. Pierotti. 1993. Expression of TRK-T1 oncogene induces differentiation of PC12 cells. *Cell Growth Differ.* **4**:539-546.
- Greco, A., M. A. Pierotti, I. Bongarzone, S. Pagliardini, C. Lanzi, and G. Della Porta. 1992. TRK-T1 is a novel oncogene formed by the fusion of TPR and TRK genes in human papillary thyroid carcinomas. *Oncogene* **7**:237-242.
- Harbury, P. B., T. Zhang, P. S. Kim, and T. Alber. 1993. A switch between two-, three-, and four-stranded coiled coils in GCN4 leucine zipper mutants. *Science* **262**:1401-1407.
- Jego, N., G. Thomas, and R. Hamelin. 1993. Short direct repeats flanking deletions, and duplicating insertions in p53 gene in human cancers. *Oncogene* **8**:209-213.
- Kaplan, D. R., B. L. Hempstead, D. Martin-Zanca, M. V. Chao, and L. F.

- Parada.** 1991. The *trk* proto-oncogene product: a signal transducing receptor for nerve growth factor. *Science* **252**:554–558.
16. **Klein, R., S. Jing, V. Nanduri, E. O'Rourke, and M. Barbacid.** 1991. The *trk* proto-oncogene encodes a receptor for nerve growth factor. *Cell* **65**:189–197.
 17. **Knapp, S., G. Delle Karth, J. Friedl, B. Purtscher, G. Mitterbauer, P. H. Dömer, and U. Jaeger.** 1994. Guanine-rich (GGNNGG) elements at chromosomal breakpoints interact with a loop-forming, single-stranded DNA-binding protein. *Oncogene* **9**:1501–1505.
 18. **Kozak, M.** 1987. An analysis of 5'-noncoding sequences from 699 vertebrate messenger RNAs. *Nucleic Acids Res.* **15**:8125–8132.
 19. **Kroczynska, A. M., R. A. Rudders, and T. Krontiris.** 1990. The human minisatellite consensus at breakpoints of oncogene translocations. *Nucleic Acids Res.* **18**:1121–1127.
 20. **Lupas, A., M. Van Dyke, and J. Stock.** 1991. Predicting coiled coils from protein sequences. *Science* **252**:1162–1164.
 21. **Martin-Zanca, D., S. H. Hughes, and M. Barbacid.** 1986. A human oncogene formed by the fusion of truncated tropomyosin and protein tyrosine kinase sequences. *Nature (London)* **319**:743–748.
 22. **Martin-Zanca, D., R. Oskam, G. Mitra, T. Copeland, and M. Barbacid.** 1989. Molecular and biochemical characterization of the human *trk* proto-oncogene. *Mol. Cell. Biol.* **9**:24–33.
 23. **Minoletti, F., M. G. Butti, S. Coronelli, M. Miozzo, G. Sozzi, S. Pilotti, A. Tunncliffe, M. A. Pierotti, and I. Bongarzone.** 1994. The two genes generating *RET/PTC3* are localized in chromosomal band 10q11.2. *Genes Chrom. Cancer* **11**:51–57.
 24. **Miozzo, M., M. A. Pierotti, G. Sozzi, P. Radice, I. Bongarzone, N. K. Spurr, and G. Della Porta.** 1990. Human TRK proto-oncogene maps to chromosome 1q32-q41. *Oncogene* **5**:1411–1414.
 25. **Mitra, G., D. Martin-Zanca, and M. Barbacid.** 1987. Identification and biochemical characterization of p70TRK, product of the human TRK oncogene. *Proc. Natl. Acad. Sci. USA* **84**:6707–6711.
 26. **Morris, C. M., Q. L. Hao, N. Heisterkamp, P. H. Fitzgerald, and J. Groffen.** 1991. Localization of the TRK proto-oncogene to human chromosome bands 1q23-1q24. *Oncogene* **6**:1093–1095.
 27. **Park, M., M. Dean, C. S. Cooper, M. Schmidt, S. J. O'Brien, D. G. Blair, and G. F. Vande Woude.** 1986. Mechanism of *met* oncogene activation. *Cell* **45**:895–904.
 28. **Pfeiffer, P., S. Thode, J. Hancke, and W. Vielmetter.** 1994. Mechanisms of overlap formation in nonhomologous DNA end joining. *Mol. Cell. Biol.* **14**:888–895.
 29. **Pierotti, M. A., E. Arighi, I. Bongarzone, M. G. Borrello, M. G. Butti, A. Greco, C. Mariani, M. Miozzo, C. Miranda, and G. Sozzi.** 1993. *RET/ptc* and TRK oncogenes in papillary thyroid carcinoma, p. 87–98. *In* L. M. G. Heilmeyer (ed.), Tyrosine phosphorylation/dephosphorylation and downstream signalling. Springer-Verlag, Heidelberg, Germany.
 30. **Pierotti, M. A., M. Santoro, R. B. Jenkins, G. Sozzi, I. Bongarzone, M. Grieco, N. Monzini, M. Miozzo, M. A. Herrmann, A. Fusco, I. D. Hay, G. Della Porta, and G. Vecchio.** 1992. Characterization of an inversion on the long arm of chromosome 10 juxtaposing D10S170 and *ret* and creating the oncogenic sequence *ret/ptc*. *Proc. Natl. Acad. Sci. USA* **89**:1616–1620.
 31. **Rabbitts, T. H.** 1994. Chromosomal translocations in human cancer. *Nature (London)* **372**:143–149.
 32. **Radice, P., G. Sozzi, M. Miozzo, V. De Benedetti, T. Ciarani, I. Bongarzone, N. K. Spurr, M. A. Pierotti, and G. Della Porta.** 1991. The human tropomyosin gene involved in the generation of the *trk* oncogene maps to chromosome 1q31. *Oncogene* **6**:2145–2148.
 33. **Richardson, J. S., and D. C. Richardson.** 1988. Amino acid preferences for specific locations at the ends of α helices. *Science* **240**:1648–1652.
 34. **Rodriguez, G. A., and M. Park.** 1993. Oncogenic activation of tyrosine kinases. *Curr. Opin. Genet. Dev.* **13**:6711–6722.
 35. **Rodriguez, G. A., and M. Park.** 1993. Dimerization mediated through a leucine zipper activates the oncogenic potential of the *met* receptor tyrosine kinase. *Mol. Cell. Biol.* **13**:6711–6722.
 36. **Sambrook, J., E. F. Fritsch, and T. Maniatis.** 1989. Molecular cloning: a laboratory manual. Cold Spring Harbor Laboratory Press, Cold Spring Harbor, N.Y.
 37. **Sanger, F., S. Nicklen, and A. R. Coulson.** 1977. DNA sequencing with chain-terminating inhibitors. *Proc. Natl. Acad. Sci. USA* **74**:5463–5467.
 38. **Santoro, M., R. Rosati, M. Grieco, M. T. Berlingieri, G. L. Colucci-D'Amato, V. de Franciscis, and A. Fusco.** 1990. The *ret* proto-oncogene is consistently expressed in human pheochromocytomas and thyroid medullary carcinomas. *Oncogene* **5**:1595–1598.
 39. **Sozzi, G., I. Bongarzone, M. Miozzo, C. T. Ciarani, P. Mondellini, C. Calderone, S. Pilotti, M. A. Pierotti, and G. Della Porta.** 1992. Cytogenetic and molecular genetic characterization of papillary thyroid carcinomas. *Genes Chromosomes Cancer* **5**:212–218.
 - 39a. **Spurr, N. K.** Personal communication.
 40. **Szentirmay, Z., Y. Ishizaka, H. Ohgaki, T. Tahira, M. Nagao, and H. Esumi.** 1990. Demonstration by in situ hybridization of *ret* proto-oncogene mRNA in developing placenta during mid-term of rat gestation. *Oncogene* **5**:701–705.
 41. **Wahls, W. P., L. J. Wallace, and P. D. Moore.** 1990. The Z-DNA motif d(TG)₃₀ promotes reception of information during gene conversion events while stimulating homologous recombination in human cells in culture. *Mol. Cell. Biol.* **10**:785–793.