Characterization of the Genomic Structure, Promoter Region, and a Tetranucleotide Repeat Polymorphism of the Human Neurotensin Receptor Gene*

(Received for publication, March 25, 1996, and in revised form, September 17, 1996)

Fei Le‡, Karen Groshan, Xiang P. Zeng‡, and Elliott Richelson§
From the Laboratory of Neuropsychopharmacology, Mayo Clinic, Jacksonville, Florida 32224

In the present study, we have cloned the human neurotensin receptor (NTR) gene, determined its structure, demonstrated that its promoter is functional in transfection experiments, and identified the start site of transcription and a tetranucleotide repeat polymorphism that locates at less than 3 kilobase pairs from the gene. The gene contains three introns, all in the coding regions. Several differences in genomic clones and previously characterized cDNA sequences are reconciled. The 5' regulatory region, which is rich in presumptive transcription factors, can drive luciferase expression in transfected CHO-K1 cells. Stepwise 5' deletions identify a positive modulator between −782 and −1309 and a negative modulator between −1309 and −1563. Southern blot analyses demonstrate a single copy gene for the NTR. The tetranucleotide repeat polymorphism is highly informative with at least 23 alleles and might serve as a very useful marker for genetic study of the relationship between the NTR and neuropsychiatric disorders.

Neurotensin (NT), an endogenous tridecapeptide, has a wide spectrum of biological activity in the central and peripheral nervous systems, including hypotension, hyperglycemia, hypothermia, antinociception, and regulation of intestinal motility and secretion (1). It is well established that the multiple functions of neurotensin are mediated through the neurotensin receptor (NTR), which is broadly distributed in central and peripheral tissues (2). Molecular biologists have recently cloned the rat (3) and human (4, 5) NTR. The human NTR cDNA encodes a 418-amino acid protein and rat a 424-amino acid protein. It has seven transmembrane spanning regions and a high degree of homology with other receptors that couple to G-proteins (guanine nucleotide-binding proteins). Therefore, it belongs to the large superfamly of receptors coupled to G-proteins. However, no information on the genomic structure of the NTR, the regulation of this gene, or genetic markers for this gene was available.

Growing evidence supports the hypothesis that NT and its receptor are pathophysiologically altered in some neuropsychiatric disorders such as Parkinson’s disease and schizophrenia. For example, a significant decrease was observed in the concentration of immunoreactive NT in the hippocampus of Parkinson’s disease patients but not other regions including substantia nigra (6). When 125I-NT was used as a radioligand to localize the NTR in postmortem brain tissues from Parkinson’s patients, significant NTR decreases were found in the substantia nigra (7). Most recently, Wolf et al. (8), in an autoradiographic study, found a 40% reduction in neurotensin receptors in the entorhinal cortex of schizophrenics compared with controls. Yet the mechanisms underlying these alterations remain to be elucidated.

Our laboratory previously reported that mRNA for the rat NTR was elevated throughout the substantia nigra/ventral tegmental area after chronic treatment with haloperidol, a typical neuroleptic (9). Whether the increase in NTR mRNA is mediated via regulation by neuroleptics of the NTR gene is unknown because information about the genomic structure of the NTR gene, especially the promoter regulatory region, had not been available. Therefore, analysis of the molecular regulation of the NTR may advance our understanding of how these receptors are modified by disease and by pharmacological manipulation. To understand these regulatory mechanisms requires knowledge of the genomic structure, including the regulatory region of this gene.

EXPERIMENTAL PROCEDURES

Materials—Materials were obtained from the following sources: human genomic libraries from Stratagene (La Jolla, CA); Colony/Plaque Screen from DuPont NEN; Hybrid N+ nylon membranes and multiprime DNA labeling system from Amersham Corp.; restriction endonucleases, T4 polynucleotide kinase, T4 DNA ligase, and luciferase assay kit from Promega (Madison, WI); pCR3lacZ plasmid from Invitrogen (San Diego, CA); DNA sequencing kit from United States Biochemical Corp.; GeneClean kit from Bio 101 (La Jolla, CA); G-25 spin columns from 5 Prime → 3 Prime Inc. (Boulder, CO); radiolabeled dCTP (3000 Ci/mmol; 1 Ci = 37 GBq) and α-32P-dATP (1563), [γ-32P]ATP (~3000 Ci/mmol; 1 Ci = 37 GBq) and α-32P-dATP (~1300 Ci/mmol) from DuPont NEN; Taq polymerase from Fisher; Total RNA isolation kit from Ambion (Austin, TX); Superscript™ reverse transcriptase and RNase H from Life Technologies, Inc.; HT-29 (human adenocarcinoma cell line) and CHO-K1 cells from American Type Culture Collection (Rockville, MD); Dulbecco’s modified Eagle’s medium from Life Technologies, Inc.; Fetal Clone II from Hyclone (human adenocarcinoma cell line) and CHO-K1 cells from American Type Culture Collection (Rockville, MD); Dulbecco’s modified Eagle’s medium from Life Technologies, Inc.; Fetal Clone II from Hyclone Laboratories, Inc. (Logan, UT); human cerebral cortex from Brain Bank of Mayo Clinic, Jacksonville, FL; the blood samples from subjects in a study conducted by Dr. Neill R. Graff-Radford (Mayo Clinic, Jacksonville, FL); and two CEPH families from Bios Laboratories (New Haven, CT). All reagents were of molecular biology grade.

Isolation and Characterization of Genomic Clones—Approximately 1 × 106 colony forming units of a human placenta genomic library in the cosmid SuperCos 1 vector (Stratagene) were screened at high stringency with the 32P-labeled full-length human NTR cDNA probe using

---

*This research was supported by the Mayo Foundation for Medical Education and Research and United States Public Health Service Grants MH27692. 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.

‡Present address: Hanson Centre for Cancer Research, Frome Rd., Adelaide, SA 5000, Australia.
§To whom correspondence should be addressed: Neuropsychopharmacology Research, Mayo Clinic, Jacksonville, 4500 San Pablo Rd., Jacksonville, FL 32224. Tel: 904-953-2439; Fax: 904-953-2482; E-mail: riche@mayo.edu.

1 The abbreviations used are: NT, neurotensin; NTR, neurotensin receptor; PCR, polymerase chain reaction; AA, amino acid; bp, base pair(s); kb, kilobase pair(s).
the previously reported methods (10). Filters were hybridized in 50% formamide, 5 × SSC, 5 × Denhardt’s, 0.1 mg/ml sheared and denatured salmon sperm DNA, 0.5% SDS, and 50 mM sodium phosphate, pH 7.0, and washed in 2 × SSC (1 × SSC = 0.15 M NaCl and 0.015 M sodium citrate) at room temperature for 10 min each, then in 0.1% SSC, 0.1% SDS at 65°C for 20 min. Filters were then washed in 0.1% SDS, 0.1× SSC for 5 min, then hybridized with the probe, and purified to homogeneity. Restriction enzyme digestion products were hybridized with the NTR cDNA probe and with oligonucleotide probes from the extreme end of 5’- and 3’-noncoding regions and coding regions of each of the NTR cDNAs. The hybridizing genomic fragments were subcloned into pBluescript for further restriction enzyme analysis and DNA sequencing. Oligonucleotides were synthesized by Ranson Hill Bioscience, Inc. (Ramona, CA).

Polymerase Chain Reaction—PCR was used to determine the intron size with the following three sets of oligonucleotides as primers: 1) 5'-TGGACGTGGAACACCGACATCCTA-3' (nucleotide 161–183 in Fig. 1A) and 5'-GGGATATGGAGGACGATAGGTTG-3' (745 to 720); 2) 5'-CAATGCGGACCTGGTATGAGC-3' (783–803) and 5'-GAGTCTCACTGCTGACGAG-3' (1006 to 897); 3) 5'-GGTGCATCGCCCTTGTTGTTGCT-3' (922–942) and 5'-TGGCACAGAAGGTTGGTACGCT-3' (1108 to 1089). For intron 1, PCR was performed according to the previous methods (11) with a slight modification. Briefly, genomic DNA (100 ng) was amplified in a 50-μl reaction buffer containing 1x Tris-HCl, pH 8.3, 2 mM MgCl₂, 5 pmol of synthetic oligonucleotides (100 ng) of each, and 0.005% SDS at room temperature for 10 min each, then in 0.1% SDS, 0.1× SSC at 65°C for 20 min. The PCR program consisted of 35 cycles as follows: 94°C for 1 min, 55°C for 1 min, and 72°C for 2 min. For introns 2 and 3, PCRs were carried out using eLONGase Amplification System provided by Life Technologies, Inc. Cosmid DNA (100 ng), which contained the genomic NTR clone, was amplified in a 50-μl reaction buffer that contained 30 mM Tris-Cl, pH 9.1, 16 mM (NH₄)₂SO₄, 2 mM MgSO₄, 100 ng of synthetic oligonucleotides, 200 μM each dNTP mix, and 2 μl of eLONGase Enzyme Mix, which contained a Taq/Pyrroccocus species GE-D DNA polymerase mixture. The PCR program included pre-amplification denaturation for 1 min at 94°C followed by 25 thermal cycles; 94°C for 40 s, 55°C for 40 s, and 70°C for 8 min. The PCR products were fractionated on 1% agarose gel and stained with ethidium bromide. To confirm whether the PCR products were specific ones, PCR products were transferred to Hybond-N membrane with 0.4M NaOH and then hybridized with the internal oligonucleotide probes. The intron sizes were determined by comparison of the PCR products to the DNA markers.

Cell Culture—HT29 or CHO-K1 cells were cultured at 37°C under 10% CO₂ in minimum essential medium supplemented with 5% Fetal Cline II and 1% minimum Eagle's medium nonessential amino acids. HT29 cells were cultured for RNA preparation and transient expression assays and CHO-K1 cells only for transient expression assays.

DNA Preparation—Total RNA preparation from human cerebral cortex or from HT29 human adenocarcinoma cell line, and extended using phenol/chloroform methods as described in a previous publication (12). For preparation of total RNA from HT29 cells, an RNA isolation kit was used. Approximately 5 × 10⁶ to 5 × 10⁷ cells were used for RNA isolation. The purity and integrity of RNA were analyzed by measuring the absorbance ratio at A₂₆₀/A₂₈₀ as well as by electrophoresis on a 1% agarose gel.

Primer Extension—Primer extensions were carried out using an oligonucleotide that corresponded to –283 to –309 base pairs (bp) upstream of the first methionine codon of the NTR (5'-TGGACGTGGAACACCGACATCCTA-3'), under conditions described in the instruction manual along with a kit for first strand cDNA synthesis (Life Technologies, Inc.). Briefly, the oligonucleotides were end-labeled with [γ-³²P]ATP, hybridized to 30 μg of total RNA from the human cerebral cortex or from HT29 human adenocarcinoma cell line, and extended using 200 units of SuperScript II reverse transcriptase in 20 μl of reaction solution. Thirty μg of yeast tRNA was used as a negative control. The extension reaction was incubated at 42°C for 60 min and then terminated at 70°C for 15 min followed by addition of 1 μl of RNaše H to digest mRNA at 37°C for 30 min. The primer-extended products were separated on an 8% native, 6% polyacrylamide gel and transferred to a nitrocellulose membrane. Autoradiography, Southern blotting, and single-stranded DNA from bacteriophage M13mp18 was used as a DNA marker.

Plasmid Constructions—To construct nested deletions in the promoter of the NTR gene, four DNA fragments, whose sizes were 303, 770, and 1297, and 1873 bp, were generated by PCR using sense primers located at –315 to –296, –782 to –762, –1309 to –1289, and –1885 to –1866, and an antisense primer at –10 to –28. The promoter region of the NTR gene (Fig. 2). The first three fragments were subcloned into the pGL3 luciferase reporter vector (pGL3-Basic), yielding clone pLUC (–315), pLUC (–782), and pLUC (–1309). The last fragment (1873 bp) was cut with HindIII. A large piece of DNA (∼1563 to –10) was purified and subcloned into the pGL3-Basic, yielding a clone pLUC (–1563). The clones were then sequenced to determine correct orientation, pLUC-Basic, which lacked a promoter and an enhancer, was used for base-line luciferase activity.

Transient Cell Transfections and Luciferase Assays—Transient transfections of CHO-K1 and HT29 cells were performed by the calcium phosphate procedure as described by Chen and Okayama (13). Briefly, 2 μg of pronouncing sequence amplified DNA, along with 0.1 μg of pCR31A2 plasmid DNA to normalize for transfection efficiency, were co-precipitated using CaPO₄. Cells were harvested 72 h after transfection in lysis buffer and assayed for luciferase expression according to the manufacturer’s instructions (Promega, Madison, WI). β-Galactosidase was assayed using a chemiluminescent detection procedure according to the manufacturer’s instructions (Tropix, Bedford, MA), except that all reagents were scaled down to one-half.

Genomic Southern Analysis—10 μg of genomic DNA bought from Promega (Madison, WI) was digested overnight with the restriction endonucleases described in the legend to Fig. 8. The digested samples were subjected to electrophoresis on a 1% agarose gel and then transferred to Hybond-N membrane using the alkaline (0.4 M NaOH) transfer method (14). The DNA samples were then hybridized using the 3²P-labeled NTR cDNA. The hybridization and wash conditions were the same as those described in the section on “Isolation and Characterization of Genomic Clones.”

Identification of Microsatellite Repeats—To determine whether any microsatellite repeat sequences were present in the sequence of or flanking the NTR gene, the cosmid containing the NTR gene were digested with Sac3AI and screened for the presence of di-, tri-, or tetranucleotide repeats by hybridization with 3²P-labeled dinucleotide repeat (AC)₁₅ or trinucleotide repeat (GGC)₁₅, (GTT)₁₅, (TA)₁₅, (GAA)₁₅, (GAT)₁₅, (GAT)₁₅, (GAC)₁₅, (GAC)₁₅, (GGA)₁₅ oligonucleotide probes or tetranucleotide repeat (AGAT)₅, (AAAT)₅, (CTT)₅, (CCT)₅, (CCG)₅, (GAC)₅, and (ACG)₅ oligonucleotide probes. The Southern and restriction fragments hybridized to (CCTT)₅ and (CCTT)₅ oligonucleotide probes were subcloned into pBluescript vector and sequenced.

Radioactive PCR Assays for Polymorphism Analysis—The genomic DNAs, which were prepared from the blood samples of 105 unrelated individuals according to a previously published method (11), were amplified by PCR using two primers flanking the microsatellite (5'-CCTCTCACAGGAACAGATGATGACTTCTCACAGGAACAGATG-3'). PCR amplification was carried out in a 25-μl reaction buffer that contained 100 ng of genomic DNA, 10 mM Tris-HCl, pH 8.3, 50 mM KCl, 1.5 mM MgCl₂, 4% dimethyl sulfoxide, 100 ng each unlabeled primer, 100 μCi of [³²P]dCTP, and 1 unit of AmpliTaq polymerase. Samples were heated to 95°C for 5 min to denature the DNA. The PCR program consisted of 35 cycles as follows: 94°C for 1 min, 55°C for 40 s, and 72°C for 40 s. The amplified PCR products were migrated on a 6% denaturing polyacrylamide gel for 5 h at 70 watts. The gel was dried and autoradiographed overnight. Sizes of the alleles were determined by comparison to M13mp18 DNA sequencing ladders.

RESULTS

Isolation and Characterization of the Human NTR Gene—We screened approximately 1 × 10⁷ recombinants from a human placenta genomic library using the 3²P-labeled human NTR cDNA probe and isolated ten positive colonies. These colonies were digested with endonuclease restriction enzyme EcoRI and then the products were fractionated on 1% agarose gel, followed by hybridization with the NTR cDNA probe and with oligonucleotide probes from the extreme end of 5’- and 3’-noncoding regions of the NTR cDNA. All ten positive clones contained the NTR sequence, among which we found that two clones, clones 1 and 5, were identical and contained exon 1, including the 5’-flanking region and part of the coding region. The other eight clones all contained exons 2, 3, and 4 including the remainder of the coding region, 3'-untranslated region, and the 3'-flanking region. The hybridizing genomic fragments from clones 1 and 5 were subcloned into pBluescript for further analysis and sequencing.
Characterization of the Human Neurotensin Receptor Gene

Genomic Structure of the Human NTR Gene—The NTR gene spanned more than 10 kb. Except for the intron sequence, the entire sequence of the NTR gene, including 5'- and 3'-flanking regions, coding region, and exon-intron junctions, was determined (Fig. 1A). Comparison of the sequence of the NTR gene from the genomic clones with that of the cDNA showed that the NTR gene consisted of four exons and three introns. All three introns were located within the coding region (Fig. 1B). The sequences of exon-intron boundaries and the placement of each intron interrupting the coding sequence are shown in Fig. 1. All of the exon-intron junction site sequences followed the consensus “AG-GT” rule. The sizes of the introns determined by PCR were approximately 0.3, 4.1 and 1.8 kb, respectively.

Comparison of the NTR genomic sequence with the sequence of the published NTR cDNA (5) revealed some minor differences (Fig. 1A). Three nucleotide differences were observed in the coding region, two conservative substitutions of G for A and G for A at the third base of codon 63 (nucleotide 189) and at the third base of codon 92 (nucleotide 276). A non-conservative substitution of G for A was found at the first base of codon 200 (nucleotide 598), which gave rise to a change from Thr to Ala at residue 200. Additionally, two nucleotide differences were found in the 5'-untranslated region with the substitution of A for G at nucleotide −309 and A for T at −472.

Analysis of the 5'-Flanking Sequence of the Human NTR Gene—We sequenced the 5'-flanking region 1887 bp 5' of the first translation codon, ATG (Fig. 2). A comparison of the 5'-flanking sequence with established consensus sequences in a transcription factors data base revealed that the 5'-flanking region lacked a typical TATA box or CAAT box but contained several potential Sp1 binding sites. It also revealed the presence of the putative binding sites for several other transcription factors, which included AP2 (14), GATA motifs (15), E47 (16), myeloid zinc finger gene 1 (17), P300 (18), histone promoter-associated nuclear factor-A (19), histone H4 transcription factor-2 (20), late SV40 transcription factor (21), and cAMP response element binding protein (22). In addition, at -1765 there was an octanucleotide sequence of CCATGTGC at -1765, which is an inverted sequence of the MyoD-E2A binding site (23), and at -1770, 1715, -1672, and -1039 four hexanucleotide sequences of TCCTCAG, which is an inverted sequence of the “acute-phase reactant regulatory element” (24) (Fig. 2).

To determine the transcription initiation site, we performed primer extension experiments using a primer (−282 to −309), which was located 281 bp 5' of the first codon, ATG. The results showed that an extended product was found at −438 with total RNA from both human cerebral cortex and HT29 cells used as templates. No signal was observed in the control groups, where the same amount of yeast RNA was used as a negative control (Fig. 3).

Analysis of the 3'-Flanking Sequence of the Human NTR Gene—There was a potential polyadenylation signal 30 bp 5' of the polyadenylation site of the cDNA, which was followed by four GT clusters (Fig. 1A). These two features have been reported to represent conserved areas for transcription termination and 3'-processing (25).

Functional Analysis of the NTR Promoter—To determine the promoter region that controls the human NTR gene expression, several 5' deletion mutants of the 1.887-kb promoter were generated. The deleted constructs, fused 5' to the luciferase gene of a reporter plasmid, were transfected into CHO-K1 cells for the transient expression analysis. Results of the luciferase assays in CHO-K1 cells (Fig. 4) indicated that the highest promoter activity was produced with the clone pLUC(−1309), containing a 1300-bp region (−10 to −1309) of the NTR promoter. Deletion of the promoter region 5' to −782 bp significantly decreased the promoter activity. When the deletion reached −315 bp, the promoter activity was completely abolished. Interestingly, the longest luciferase construct pLUC(−1563), which contained a 1554-bp region (−10 to −1563) of the NTR promoter, had less promoter activity than did pLUC(−1309), indicating that silencer elements existed in the region between −1309 and −1563 (Fig. 4).

Identification of Microsatellite Repeat Sequences in the NTR Gene—To determine whether any microsatellite repeat sequences were present in the NTR gene or in the sequences flanking the NTR gene, we digested the cosmids containing the NTR gene with Sau3AI and screened for the presence of microsatellite repeats by hybridization with 17P-labeled di-, tri- or tetranucleotide repeat oligonucleotide probes. Hybridization to the (CCTT)_n and (CTTT)_n oligonucleotide probes was observed in a fragment from clones 3 and 4. This microsatellite repeat, which was located at less than 3 kb from the poly(A) site (data not shown), was subcloned into pBluescript vector and then sequenced. Sequence analysis revealed an almost exclusive CT region (−360 bp), in which there existed two tetranucleotide microsatellite repeats. The first microsatellite had 9 perfect “CTTT” repeats and the second had 17 perfect “CTTT” repeats (Fig. 5).

Analysis of the Tetranucleotide Repeat Polymorphism in the NTR Gene—Using PCR with two primers flanking the tetranucleotide repeat to amplify the genomic DNA from 105 unrelated individuals, we found that this repeat was very polymorphic (Fig. 6). Further study demonstrated that 23 alleles were found in 210 chromosomes from 105 unrelated individuals who were 97 Caucasian, 2 Black, 4 Asian, and 2 others. Many alleles had very low frequencies, among which 11 were below 2%, 8 were between 2 and 8%, and only 4 were over 10% (range 11.90–12.86%) (Table I). The estimated heterozygosity was 0.914, and polymorphism information content value was 0.906. This tetranucleotide repeat polymorphism was also found to be stable and inherited in a Mendelian fashion in two three-generation pedigrees (Fig. 7).

Southern Blot Hybridization Analyses of the Human NTR Gene—To determine the number of genes encoding the human NTR, we digested human genomic DNA with EcoRI, HindIII, KpnI, or XbaI restriction enzymes, whose restriction sites were not present in the sequence of the NTR cDNA, and then probed with the NTR cDNA probe. A single band was detected in KpnI and XbaI digests, and two bands were detected in EcoRI and HindIII digests (Fig. 8). Since the NTR genomic clones, digested with EcoRI and HindIII, also showed two bands (data not shown), this indicated that the restriction sites for EcoRI and HindIII were in introns. Therefore, the Southern analysis revealed that only a single gene encoded the receptor.

DISCUSSION

In the present study we have clarified the organization, genomic structure, and promoter function of the human NTR gene. We have also identified a microsatellite polymorphism that locates at less than 3 kb from the poly(A) site. The present study demonstrates that the human NTR gene contained three introns, all of which were in the coding region (Fig. 1). The placement of these introns (Fig. 1B) is interesting, since we have shown that the intracellular loop 3 is involved with coupling to release of inositol phosphates (26) and that the extracellular loop 3 is likely involved with binding of agonists (27, 28). Among the molecularly cloned G-protein-coupled receptors, some, such as the α,β and αβ-adrenergic receptors (29, 30, 31), the D1-dopamine receptor (32), muscarinic receptors (33), the platelet-activating factor receptor (34), and IL-8 receptor A (35) have no introns in their coding regions. Others have an intron or multiple introns in the coding region, such as A2a.
Characterization of the Human Neurotensin Receptor Gene

Fig. 1. A, the nucleotide and deduced amino acid sequence of the human NTR gene. The deduced amino acid sequence is shown below the nucleotide sequence. The exon sequences are shown in uppercase letters. The intron and the flanking sequences are in lowercase letters. The size...
adenosine receptor (10), dopamine D_2, D_3, and D_4 receptors (36, 37, 38, 39), the 5-HT_2 receptor (40), the substance K receptor, the substance P receptor (41), the endothelin-A receptor (42), the neuromedin K receptor (43), opsins (44), and the luteinizing hormone receptor (45). Therefore, the human NTR gene elucidated in the present study belongs to the family of G-protein-coupled receptor genes that contain intronic sequences within their coding regions.

Comparison of the NTR genomic sequence with the sequence of the NTR cDNA from HT29 cells (5) revealed several nucleotide differences between the two sequences. Two were found in the 5'-noncoding region and three in the coding region (Fig. 1A). Compared with the coding region of the NTR cDNA published by our group (4), the genomic sequence showed two nucleotide differences. Table II summarizes the differences occurring in the coding region. Both the present results and the results from the previous work of our group (4) indicated that the third base of codon 63 (nucleotide 189) and the first base of codon 200 (nucleotide 598) were G rather than A, compared with the sequence reported by Vita et al. (5). By carefully resequencing the NTR cDNA clones from both the HT29 cell line and the substantia nigra (in the case of AA200 two more human DNA samples were sequenced), Watson et al. (4) suggested each intron is shown in parentheses. Nucleotides of the transcript are numbered sequentially from the first nucleotide of the translation start codon. The different nucleotides and amino acids (at AA200), compared with the sequence of the NTR cDNA reported by Vita et al. (5), are bold. The polyadenylation signal is bold and underlined. Four GT clusters are underlined.

**Fig. 2. Nucleotide sequence of the promoter region of the human NTR gene.** The transcription start site is indicated by an asterisk. The potential binding sites for transcription factors, Sp1, AP2, late SV40 transcription factor, GATA motifs, myeloid zinc finger gene 1, F900, E47, CAMP response element binding protein, histone promoter-associated nuclear factor-A, and histone H4 transcription factor-2 are bold and underlined. An inverted sequence of MyoD-E2A-binding site is shown in italics and underlined at position -1765, and four inverted sequences of acute-phase reactant regulatory element are underlined at position -2177, -1715, -1672, and -1039.

**Fig. 4. Deletion constructs of the NTR gene promoter and summary of luciferase activity.** Schematic representation of the 5'-stepwise deletion constructs used to test the functional activity of the NTR promoter in transient transfection assays and relative luciferase expression of each construct are shown. Numbers on the left of each construct indicate the 5' end of the promoter fragment relative to the first translation codon (+1). Each construct is fused to the luciferase reporter gene at position -10 relative to the first translation codon in exon 1. The regulatory motifs in NTR promoter are schematically represented at the top by various symbols, which are identified. Arrow represents the transcription start site. CHO-K1 cells were cotransfected with the various NTR promoter luciferase constructs and SV40-β-galactosidase plasmid. Luciferase activity was assayed as described in the text and is displayed as fold induction compared with pLUC-Basic. Values represent the mean ± S.D. of five to seven different experiments. *, p < 0.05, compared with pLUC(-1563) and pLUC(-782) constructs with a Student's t test.

with the sequence reported by Vita et al. (5). By carefully resequencing the NTR cDNA clones from both the HT29 cell line and the substantia nigra (in the case of AA200 two more human DNA samples were sequenced), Watson et al. (4) sug-
gested that the base changes at AA63 and AA200 were probably sequencing errors in the NTR sequence reported by Vita et al. (5). This suggestion was confirmed by the present study. Furthermore, we note that when the first base at AA200 is G and not A, not only is there an amino acid change but also there is a BglII restriction site introduced.

In order to be certain that the first base at AA200 was G rather than A in our genomic DNA, we further examined this DNA by a combination of PCR technique and restriction enzyme analysis. A 606-bp fragment including nucleotide 598 was

Fig. 6. Autoradiogram of denaturing acrylamide gel showing typing of 10 unrelated individuals for the tetranucleotide repeats in the NTR gene. The genotype of each individual is shown at the top of each lane. Methodology for radioactive PCR amplification of the genomic DNA from the unrelated individuals and for polymorphism analysis of PCR products was described under “Experimental Procedures.”

Table I

| Allele | Size (bp) | Frequency |
|--------|----------|-----------|
| A1     | 343      | 0.005     |
| A2     | 363      | 0.005     |
| A3     | 371      | 0.009     |
| A4     | 375      | 0.033     |
| A5     | 379      | 0.029     |
| A6     | 383      | 0.009     |
| A7     | 387      | 0.038     |
| A8     | 389      | 0.076     |
| A9     | 385      | 0.071     |
| A10    | 399      | 0.129     |
| A11    | 403      | 0.124     |
| A12    | 407      | 0.119     |
| A13    | 411      | 0.071     |
| A14    | 415      | 0.057     |
| A15    | 419      | 0.024     |
| A16    | 423      | 0.124     |
| A17    | 427      | 0.009     |
| A18    | 431      | 0.019     |
| A19    | 439      | 0.019     |
| A20    | 443      | 0.005     |
| A21    | 451      | 0.005     |
| A22    | 459      | 0.005     |
| A23    | 547      | 0.014     |

Fig. 7. Mendelian inheritance of the tetranucleotide repeat polymorphism in the NTR gene in two pedigrees from CEPH family 884 and 1331. Females are indicated by circles and males by squares. Numbers indicate the NTR gene alleles (Table I).

Table II

Differences in codon sequence between human NTR cDNAs and genomic DNA

| Source of sample | Codon sequence |
|-----------------|----------------|
|                 | AA63 | AA200 |
| Genomic DNA clone | AAG  | AAA  |
| Substantia nigra  | AAG  | AAA  |
| cDNA 1a          | AAG  | AAG  |
| cDNA 2a          | AAG  | AAG  |
| Human DNA Sample 1-2a | —    | —    |
| Human DNA Sample 3-17a | —    | —    |
| HT29 cDNA        | AAG  | AAG  |
| (Watson et al. (9)) | AAG  | AAG  |
| HT29 cDNA (Vita et al. (5)) | AAA  | AAA  |

a Data from Watson-Bolden et al. (9).

b — data not available.
amplified by PCR from the genomic DNA of 15 unrelated individuals. The resulting PCR products were digested with BglI. The results demonstrated that all samples tested had a BglI cutting site, which indicates all 15 samples had the base G at nucleotide 598 (Table II).

With the nucleotide difference at the first base of codon 194 (nucleotide 580), where C or T was observed in two published cDNA sequences, it was suggested that this is a polymorphic site (4). Further study in our group found that this polymorphic site has no effect on ligand binding, although it causes an amino acid change, Leu or Phe (1). The genomic clone sequenced in this study had a C base in this position or Leu for the amino acid. Another nucleotide difference was found at AA92 (nucleotide 276), where the third base was A in the genomic DNA, whereas it was G in the cDNA. However, this was only a conservative change. Additionally, two nucleotide differences were found in the 5’-untranslated region with the substitution of A for G at nucleotide −309 and A for T at −472. Whether these nucleotide differences are polymorphic sites remains to be determined.

In order to study the promoter function for the NTR gene, we sequenced the promoter region 1515 bp 5’ of the cDNA sequence thus far available (Fig. 2). We found that the 5’-flanking region was GC-rich. The overall GC content was greater than 68%. Eighty percent GC content was found in the region encompassing the transcription initiation site (±130 bp), which was located at 438 bp 5’ of the methionine initiation codon, as determined by the primer extension experiments (Fig. 3). This phenomenon seems to fit with the observation that the promoter of a gene is usually surrounded by GC-rich sequences. The 5’ regulatory region of the NTR gene is high in GC content and lacks a typical TATA or CAAT box but contains multiple potential Sp1 binding sites, among which two were found to be located within 40 bp 5’ of the transcription start site (Fig. 2). Functional study of the promoter region revealed that the two Sp1 transcription factors were very important for the activation of transcription. In addition, a large number of binding sites for other putative transcription factors are present in the promoter region of the NTR gene, suggesting that the regulation of this gene involves a complex array of regulatory factors.

We further investigated the promoter function of the human NTR gene using a transient expression system in CHO-K1 cells to define the areas of the 5’-flanking region that play a vital role in transcriptional regulation. Maximal luciferase activity was found in the plasmid construct pLUC(−1309) containing the 1300-bp promoter region. Construct pLUC(−315) bearing 306 bp of the promoter, not including the transcription start site, had no promoter activity. The other two constructs, pLUC(−782) and pLUC(−1563) having 773 and 1554 bp of the promoter, respectively, were found to have less luciferase activity, compared with the construct pLUC(−1309). This study, on the one hand, supported the location for the transcription start site determined by primer extension experiments. On the other hand, it suggested that the region between −315 and −782 plays an essential role in transcriptional initiation and activation, that the region from −782 to −1309 possesses positive regulatory elements, which enhance gene expression, and that the region between −1390 and −1563 contains negative regulatory elements.

The NTR gene is expressed in a tissue-specific manner. For example, results from our group demonstrated that NTR mRNA exists at high levels in substantia nigra pars compacta and the nucleus paranigralis but at background levels in the nucleus ruber, the colliculus inferior, the nucleus caudatus, the putamen, and the nucleus accumbens of both human and rat brain (46). To determine whether the portion of the promoter we cloned contains the sequence information for tissue-specific/cell-type-specific function, we tried to analyze the expression of the luciferase reporter in HT29 cells that are known to express the NTR. This would enable us to compare the promoter activity in HT29 cells with that in CHO-K1 cells that do not express the NTR naturally. However, repeated attempts to detect promoter activity in HT29 cells were not successful. It may be that the HT29 cell line is not suitable for the transient transfection assays in this case.

Recently, microsatellite polymorphisms have received more and more attention in mapping defective genes for genetic diseases, since they are highly informative, abundant, and uniformly distributed genetic markers. With the microsatellite polymorphisms used as genetic markers, the locations of genes for cystic fibrosis, Duchenne muscular dystrophy, and Huntington’s disease have been found.

In the present study, we demonstrated a tetranucleotide repeat polymorphism present in the 3’-flanking region of the NTR gene (less than 3 kb from the poly(A) site). This microsatellite polymorphism has at least 23 alleles as determined in 210 chromosomes from 105 unrelated individuals, which we examined. Many alleles were of very low frequencies (Table I). This polymorphism also had a high heterozygosity (0.914) and polymorphism information content value (0.906) with a Mendelian inheritance. Therefore, this is a highly informative polymorphism and will be very useful as a genetic marker.

The accumulated evidence has shown that there are striking interactions between NT or the NTR and central dopaminergic systems (47), suggesting that NT and the NTR play an important role in the etiology of some neurological and psychiatric disorders. These include Parkinson’s disease (46) and schizophrenia (48, 49). Especially for schizophrenia, studies have shown that genetic factors play an important role in the etiology of this disease (50), at least for some subgroups of schizophrenia. Therefore, the search for markers of vulnerability has been a recent focus in schizophrenia research.

In light of the studies suggesting an association between NT and the NTR with schizophrenia (48, 49, 50), it is reasonable to hypothesize that the gene for NT or the NTR receptor might be a candidate gene for schizophrenia. Recently, the NTR gene was mapped to chromosome 20q13 (51). It is noteworthy that some deletions involving the long arm of chromosome 20 have also been reported. These deletions seem to be associated with myeloid disorders, mental retardation, and severe malformation of the limbs (52). Therefore, it would be very interesting to look at the association genetically between the NTR and schizophrenia. However, until the present study, no genetic markers were available for NT or the NTR gene. With the availability of a highly informative microsatellite polymorphism in the NTR gene as demonstrated here, the genetic study of the relationship between the NTR gene and schizophrenia or other neuropsychiatric disorders has become feasible.

Acknowledgment—We thank Terrance Souder for his assistance in some techniques.

REFERENCES
1. Cusack, B., McCormick, D. J., Pang, Y.-P., Souder, T., Garcia, R., Faux, A., and Richelson, E. (1995) J. Biol. Chem. 270, 18359–18366
2. Quirion, R., Rowe, W. B., Lapchak, P. A., Araujo, D. M., and Beaudet, A. (1992) Ann. N. Y. Acad. Sci. 668, 109–119
3. Tanaka, K., Mauo, M., and Nakashii, S. (1990) Neuron 4, 847–854
4. Watson, M., Isackson, P., Makker, M., Yamada M. S., Yamada, M., Cusack, B., and Richelson, E. (1993) Mayo Clin. Proc. 68, 1043–1048
5. Vita, N., Laurent, P., Lefort, S., Chalon, P., Dumont, X., Kaghad, M., Guly, D., Lefur, G., Ferrara, P., and Caput, D. (1993) FEBS Lett. 317, 139–142
6. Nemeroff, C. B., and Bissette, G. (1985) Prog. Clin. Biol. Res. 192, 333–341
7. Chino, G., Probst, A., and Palacios, J. M. (1990) Neuroscience 39, 351–360
8. Wolf, S. S., Hyde, T. M., Saunders, R. C., Herman, M. M., Weinerberger, D. R., and Kleinman, J. E. (1990) J. Neuroat. Transm. 102, 55–65
9. Watson-Bolen, C., Watson, M. A., Murray, K. D., Isackson, P., and Richelson,
Characterization of the Human Neurotensin Receptor Gene

E. (1993) J. Neurochem. 61, 1141–1143
10. Le, F., Townsend-Nicholson, A., Baker, E., Sutherland, G. R., and Schofield, P. R. (1996) Biochem. Biophys. Res. Commun. 223, 461–467
11. Le, F., Mitchell, P., Vivero, C., Waters, B., Donald, J., Selfie, L., Shine, J., and Schofield, P. (1994) J. Affective Disord. 32, 187–195
12. Le, F., Wilce, P. A., Hume, D. A., and Shanley, B. C. (1992) J. Neurochem. 59, 1309–1315
13. Chen, C., and Okayama, H. (1987) Mol. Cell. Biol. 7, 2745–2752
14. Mitchell, P. J., Timmons, P. M., Hebert, J. M., Rigby, P. W. J., and Tjian, R. (1990) Genes Dev. 5, 105–119
15. Merika, M., and Orkin, S. H. (1993) Mol. Cell. Biol. 13, 3999–4010
16. Sun, X.-H., and Baltimore, D. (1991) Cell 64, 459–470
17. Moreia, J. F., Hromas, R., and Rauscher, F. J., III, (1994) Mol. Cell. Biol. 14, 1786–1796
18. Rikitake, Y., and Moran, E. (1992) Mol. Cell. Biol. 12, 2826–2836
19. van Wijnen, A. J., Stein, J. L., and Stein, G. S. (1987) J. Biol. Chem. 262, 6958–6962
20. Dailey, L. D., Hanly, S. M., Goeder, R. G., and Heintq, N. (1986) J. Biol. Chem. 261, 232–252
21. Murre, C., McCaw, P. S., Vaessin, H., Caudy, M., Jan, L. Y., Jan, Y. N., and Baltimore, D. (1989) Cell 58, 537–544
22. Inoue, A., Yanagisawa, M., Takuwa, Y., Mitsui, Y., Kobayashi, M., and Ishii, S. (1989) Proc. Natl. Acad. Sci. U. S. A. 86, 7241–7245
23. Kim, C. H., Heath, C., Bertuch, A., and Hansen, U. (1987) Proc. Natl. Acad. Sci. U. S. A. 84, 6025–6029
24. Murre, C., McCaw, P. S., Vaessin, H., Caudy, M., Jan, L. Y., Jan, Y. N., Cabrera, C. V., Buskin, J. N., Hauschka, S. D., Lassar, A. B., Weintraub, H., and Baltimore, D. (1989) Cell 56, 537–544
25. Inoue, A., Yanagisawa, M., Takawa, Y., Mitsui, Y., Kobayashi, M., and Masaki, T. (1989) J. Biol. Chem. 264, 14954–14959
26. Birnstiel, M., Busslinger, M., and Strub, K. (1985) Cell 41, 349–359
27. Yamada, M., Yamada, M., Watson, M. A., and Richelson, E. (1994) Mol. Pharmacol. 46, 470–476
28. Cusack, B., Groshan, K., McCormick, D. J., Pang, Y.-P., Perry, R., Phung, C.-T., Souder, T., and Richelson, E. (1996) J. Biol. Chem. 271, 15054–15059
29. Pang, Y.-P., Cusack, B., Groshan, K., and Richelson, E. (1996) J. Biol. Chem. 271, 15600–15608
30. Emorine, L. J., Marullo, S., Delavier-Klutcho, C., Kaveri, S. V., Durieu-Trottier, C., and Struberg, A. D. (1987) Proc. Natl. Acad. Sci. U. S. A. 84, 6995–6999
31. Koblika, B. K., Frielle, T., Dohlan, H. G., Bolanowski, M. A., Dixon, R. A. F., Keller, P., Caron, M. G., and Lefkowitz, R. J. (1987a) J. Biol. Chem. 262, 7321–7327
32. Koblika, B. K., Matsui, H., Koblika, T. S., Yang-Feng, T. L., Francke, U., Caron, M. G., Lefkowitz, R. J., and Gegen, J. W. (1987b) Science 238, 650–656
33. Sunahara, R. K., Niznik, H. B., Weiner, D. M., Stormann, T. M., Brann, M. R., Kennedy, J. L., Gerlunser, J. E., Rozmahel, R., Yang, Y., Israel, Y., Seeman, P., and O'Dowd, B. F. (1990) Nature 347, 80–83
34. Bonner, T. I., Buckley, N. J., Young, A. C., and Brann, M. R. (1987) Science 237, 527–532
35. Seyfried, C. E., Schweickart, V. L., Godiska, R., and Gray, P. W. (1992) Genomics 13, 832–834
36. Sprenger, H., Looy, A. R., Meyer, R. G., Johnston, J. A., and Kelvin, D. J. (1994) J. Immunol. 153, 2524–2532
37. Gandelman, K., Harmon, S., Todd, R. D., and O’Malley, K. L. (1991) J. Neurochem. 56, 1024–1029
38. Grandy, D., Marchionni, M., Makam, H., Stoiko, R., Alfano, M., Froshingham, L., Fisher, J., Burke-Howie, K., Bunzow, J., Server, A., and Civelli, O. (1988) Proc. Natl. Acad. Sci. U. S. A. 85, 9762–9766
39. Sokoloff, P., Bhat, P. S., Martres, M., Bouthenet, M., and Schwartz, J. (1990) Nature 347, 146–151
40. van Tol, H., Bunzow, J., Guan, H., Sunahara, R., Seeman, P., Niznik, H., and Civelli, O. (1991) Nature 350, 610–614
41. Chen, K., Yang, W., Grimsby, J., and Shih, H. C. (1992) Mol. Brain Res. 14, 20–26
42. Hershey, A. D., Dykema, P. E., and Krause, J. E. (1991) J. Biol. Chem. 266, 4366–4374
43. Hosoda, K., Nakao, K., Tamura, N., Arai, H., Ogawa, Y., Sugai, S., Nakahishi, S., and Imura, H. (1992) J. Biol. Chem. 267, 18797–18804
44. Takahashi, K., Tanaka, A., Han, M., and Nakahishi, S. (1992) Eur. J. Biochem. 204, 1025–1033
45. Nathans, J., and Hope, G. (1984) Proc. Natl. Acad. Sci. U. S. A. 81, 4851–4855
46. Tsai-Morris, C. H., Baracho, E., Wang, W., Xie, X.-Z., and Dufau, M. L. (1991) J. Biol. Chem. 266, 11355–11359
47. Yamada, M., Yamada, M., and Richelson, E. (1995) Neuroscience 64, 495–417
48. Deutch, A. Y., and Zahn, D. S. (1992) Annu. N. Y. Acad. Sci. 668, 252–252
49. Nemeroff, C. B., Bissette, G., Widerlov, E., Beckmann, H. H., Gerner, R., Manberg, P. J., Lindstrom, L., Prange, A. J., Jr., and Gattaz, W. (1989) J. Neuropsychiatry Clin. Neurosci. 1, 16–20
50. Garver, D. L., Bissette, G., Yao, J. K., and Nemeroff, C. B. (1991) Am. J. Psychiatry 148, 484–488
51. Tienari, P., Wynne, L. C., Moring, J., Lahti, I., Naarala, M., Sorri, A., Wahlberg, K. E., Saerante, O., Seitamaa, M., and Kaleva, M. (1994) Br. J. Psychiatry 165, (suppl.) 20–26
52. Laurent, P., Clerc, M. G., Forgez, P. M., Dumont, X., Ferrara, P., Caput, D., and Rostene, W. (1994) Genome 5, 303–306
53. Shahbat, B. F., Ben-Sasson, E., Ariel, E., and Grinblat, J. (1983) J. Med. Genet. 20, 171–173