

# A rat brain cDNA encoding the neurotransmitter transporter with an unusual structure

Qing-Rong Liu, Sreekala Mandiyan, Beatriz López-Corcuera, Hannah Nelson and Nathan Nelson

*Roche Institute of Molecular Biology, Roche Research Center, Nutley, NJ 07110, USA*

Received 17 November 1992

A rat cDNA clone encoding the novel membrane protein of the neurotransmitter transporters family was cloned and sequenced. The cDNA was identified as a transcript of the gene NTT4 of which a partial genomic clone was previously sequenced. Alignment of the amino acid sequence of NTT4 with other members of the neurotransmitter transporter family revealed a marked deviation from the conserved structure of all other members of the family. The largest extracellular loop with a potential glycosylation site was identified between membrane segments 7 and 8. The protein retains the common glycosylated loop between transmembrane helices 3 and 4 in all members of the family. The transcript of NTT4 was found exclusively in the central nervous system and is more abundant in the cerebellum and the cerebral cortex.

Neurotransmitter transporter; Central nervous system; Molecular cloning

## 1. INTRODUCTION

Cloning of cDNA encoding the transporter of  $\gamma$ -aminobutyric acid (GABA) [1,2] and subsequent cloning of a cDNA encoding the noradrenaline transporter [3] suggested the existence of a proliferated family of neurotransmitter transporters. Recently cDNAs encoding more than ten different neurotransmitter transporters have been cloned and sequenced [1–16]. The family of these genes could be divided into three subfamilies including the GABA and taurine transporters [12,15], the amino acid (glycine and proline) transporters [8–10], and the catecholamine transporters [3–7]. The general structure of all these gene products is very similar. They contain twelve potential transmembrane helices and an extended external loop with 3–4 glycosylation sites between membrane segments three and four [1–16]. The calculated molecular weights of the transporters is about 70 kDa and both their C- and N-terminal peripheral peptides contain about 40 amino acids and may be located on the cytoplasmic side of the membrane. In GABA and catecholamine transporter subfamilies the amino acid sequence of each member is 60–80% identical to the other members within a subfamily and about 40% identical to members between the two subfamilies [16]. Amino acid transporters such as the glycine transporter and proline transporter share about 40–45% homology with all members of the neurotransmitter transporter superfamily. Sequence homology among the members of the neurotransmitter transporters family gave clear indication that they evolved from a common

ancestral gene. Moreover, partial genomic cloning of several neurotransmitter transporters revealed that in all of them the first intron in the reading frame is located in an identical position [16]. In this communication we report on the cloning and sequencing of cDNA of NTT4 which is a transcript of one of the reported genomic clones [16]. Extensive trials of expressing NTT4R cRNA in *Xenopus* oocytes failed to identify the specific substrate for the uptake. The structure of this gene product deviates from the common structure of the other transporters and a new mechanism of uptake may be involved by this transporter protein.

## 2. EXPERIMENTAL PROCEDURES

### 2.1. Screening of the rat brain cDNA library and sequencing

A rat brain  $\lambda$ ZAP cDNA library (Stratagene) was screened at high stringency with the  $^{32}$ P-labeled exon fragment of the genomic clone NTT4 [16] containing initiator methionine. The pBluescript plasmid was excised from the positive phages and analyzed by dot blot as previously described [15,17]. The cDNA was sequenced by the dideoxy chain termination method following serial deletions by exonuclease III [18,19]. Sequence verification was performed with synthetic oligonucleotide primers.

### 2.2. Northern blot analysis

Mouse brain was dissected into various parts and RNA was prepared from the various brain parts as well as from kidney and liver as previously described [15]. About 40  $\mu$ g of total RNA were applied into an RNA denaturing agarose gel, and following electrophoresis and transfer onto a nylon filter was probed by the  $^{32}$ P-labeled exon fragment of the genomic clone NTT4. The filters were washed and processed as previously described [15].

### 2.3. Expression in *Xenopus* oocytes

Six to 12 oocytes were assayed as previously described [15] for uptake of the following radioactive substances: adenosine,  $\beta$ -alanine, taurine, histamine, choline,  $\gamma$ -aminobutyric acid, glutamate, alanine,

Correspondence address: N. Nelson, Roche Institute of Molecular Biology, Roche Research Center, Nutley, NJ 07110, USA.

arginine, aspartate, glycine, histidine, isoleucine, leucine, lysine, phenylalanine, cysteine, methionine, glutamine, tryptophan, proline, threonine and valine.

1 GAATTCGGCAGGAGCGCTGGCGCTCGGCCAGCTCCGTTCCGCTCCGCGAGGCGGAGCGG 60  
 61 GACCGGCGCGGGGACAGCGGCACTCGGCTGGGGCTTCGCGAGCACGCGCTCCCTCCGG 120  
 121 GATCTTACCGGCGGCTCCCTAGAGCCAGAGCGGCGGCGAGAGCGGCGGCTGGGTTGGG 180  
 181 ACTGGCGGCAACAGGGAAGCGACGAGCGAGTTGATTCGCTCGCGCGGCGGCGGAGCT 240  
 241 CCGGGTCCCGCAAGCCCGGCTGGGGCTGTTGGCGCTGAGCTCAATGGCCCTCTGGAA 300  
 301 GGACGCTGCACCCGCGACCGTGGGGACACGGCGCTCGGCGGGAATCCAGGGGCGGA 360  
 361 GTCCGCGGCGCGGCCACGCGGCGAGCGCTTTGGATGAGAGAAAGCTGACAGGAGCTGA 420  
 421 GCCCACCTTCTCTGTGTGGGGAATAGGCGTACGAGCCCGAGTGGCACCCATCGCG 480  
 481 AAGAACAGCAAGGTGACCCAGCGTGAACACAGCAATGAGCATGTCAGTGCCTGGCT 540  
 541 GACCTGCTGGGCTCGAGGAGCTGTGGACTATAAGCAGAGTGTACTGACGCTGGCTGGT 600  
 601 GAGACAGTGGCAAGCAGAGGTGGCGGAGGAGGCTGGACGAGAGGACCGGCGGAGCC 660  
 661 TGGGCGKQKVAAEEEDAEEDRPA 720  
 721 TGGACAGCAAGCTGAGTACATCTGGCCCGAGTCCGCTTTCTGTGGGCTTGGCAAC 780  
 781 WNSKQLQYILAQIGFSSVGLGN 840  
 841 ATCTGGAGGTTCCTTACCTGTGCGAGAAAATGGAGGAGGTGCTTACCTGGTGGCCCTAC 840  
 841 IWRFPYLYLCQKNGGGAAYLV 900  
 901 CTGGTGTCTGATCATTATGGCATCCCGCTCTTCTTCTGGAGCTGGCTGTGGGCGAG 960  
 961 LVLLLIIGIFLFFLELAVG 1020  
 1021 AGGATCCGCGTGGCAGCATTTGGGCTTGGCAGTATGTGCCCCCGCTGGGGGCGATT 1080  
 1081 RLRSLGCVWHYVCFRLG 1140  
 1141 GCGTTTCTGCGTATTTCTGCTTCTTGGGCTGTACTACACGCTGATCATTGGA 1200  
 1201 GPFSSCTIVCLVFGYVNIIG 1260  
 1261 TGGAGCTCTTCTACTTCTCAAGTCTTCCAGTACCCACTGCGCTGGAGTGAATGCTCT 1320  
 1321 WSVFYFYFKSFQYPLPWSECP 1380  
 1381 GTCATCAGGAATGGGACTGTGGCAGTGGTGGAGCGTGGTGGAGAGGCTCAGCTACT 1440  
 1441 VIRNGTVAVVEPECEKSSAT 1500  
 1501 ACCTACTTCTGGTATCGAGGCGCTGGAGCTTCCAACTCCATCTCAGAGAGTGGGGC 1560  
 1561 TYFWYREALDISNSISESGG 1620  
 1621 CTCAGTGGAAAGTACAGCTTGGCTTCTGGGCTGGAGCATGGGCGATGGCTGTA 1680  
 1681 LNWKHTVCLVAWSIVGM 1740  
 1741 GTCAAGGCGTCACTCTCGAAAGTAAATGATTACGCTCCCTCTCCCTATGTG 1800  
 1801 VKGIQSSSGKVMYFSSLFYV 1860  
 1861 GTGTGGCTCTCTGGTCCGGGAGTCTGCTTCGAGGGGCGATTGATGGCATCCTG 1920  
 1921 VLAACFLVLRGLLLLRGAVDGIL 1980  
 1981 CACATGTTCACTCTAAGCTGGACAAGTGTGGAGCCCGAGGTGGCGGGAGGACGCG 2040  
 2041 HMFPTPKKLDKMLDPQVWREAA 2100  
 2101 ACGCAGGTCTTCTGGCCCTGGGGCTGGGCTTTGGAGGTGTATCGCTTTCCAGCTAC 2160  
 2161 TQVVFALGLGFGGVIAFSSY 2220  
 2221 AACAAACAGGACAACTGCGGCTTCGAGCTGCGCTTGGTGTCTTCACTCACTTCTC 2280  
 2281 NQDWNCHFDAAALVSPIN 2340  
 2341 ACCTCCGTGTGGCCACCTCGTGGTGTTCAGTGTGGGTTCAAGCTAACATCATG 2400  
 2401 TSVLAATLVVFAVLGFKANIM 2460  
 2461 AATGAGAAGTGTGGTTCGAGAAATGCTGAGAAATCTAGGGTACCTCACTCAATGTC 2520  
 2521 NEKCVVENAEKILGYLNSNV 2580  
 2581 TTGAGCGGCGGCTCATCCCGCCAGCTCACTTCTCACTGACCAAGGACTAC 2640  
 2641 LSRDLIPPHVNFSLHLLTKDY 2700  
 2701 TCAGAGTGTACAACTGTCATGACTGTTAAGGAGAAGCAGTCTCAGCGCTGGGGCTG 2760  
 2761 SMYNIHMTYKQKQFSAALGL 2820  
 2821 GATCTTGGCTCTGGAGTACGCTGGAGCACTGTCAGGGCGCAGGCTGGGCTTC 2880  
 2881 DDCLEDEKFSVGL 2940  
 2941 ATCCGCTTCACTGAGGCGATGACATTTCCCGGCTCTCGTCTTGGTGGTCTGTC 3000  
 3001 IAFTEAMTHFPASPFWSMVF 3060  
 3061 TTCTGATGCTCACTCACTGGGCTGGGCGAGTATGATGGGAGCTGCGAGGATCACC 3120  
 3121 FLMLINLGLGSMIGTMAGIT 3180  
 3181 ACACCTATCATCGACCTTCAAGGTGCCAAGGAGATGTCACAGTGGGCTGCTGTGTC 3240  
 3241 TPIIDTFKVPKEMFTVGGCCV 3300  
 3301 TTGCACTTCTGGGGCTATTTGCTTCCAGCGCTCGGAACTACTTTGTCACCATG 3360  
 3361 FIFFLFVGRSGN 3420  
 3421 TTGATGACTTTCGGCCACCTGCGGCTGCGGCTGCGGCTTCTGAGAACATCGCT 3480  
 3481 FDDYSSATLPLTVIIVILENIA 3540  
 3541 GTGGCTGGATTGAGCAACAAAGATTATGAGGAGTGCAGAGATGCTGGGGCTC 3600  
 3601 VAMIIYGTKKFMQELTEMLGF 3660  
 3661 CGACGCTACCGATTCTATTCTACATGTTGAGTGTGCTTCCAGTGTGCTGAGTGTG 3720  
 3721 RPYRFFYFYMFKFVSPCLMAV 2760  
 2761 CTCACACAGCGCATCATCGAGCTGGGGGTGTCACCCCGAGCTACAGTGGCTGGATT 2820  
 2821 LTTASISIIQLGVSPPGYSANI 2880  
 2881 AAGGAGGAGCGCGCGGCTACTTGTACTTCCCAATGGGCGATGGCAGTCTGATC 2940  
 2941 KEAAEAERYLFPFNWMAALIL 3000  
 3001 ACCCTGATGCTGGGCAACCTGCTATCTCTGGTGTTCATCTCGAGGACTTCCAC 3060  
 3061 LLAIVFLLRLRH 3120  
 3121 CTGCTCTGATGGTTCACACCTCTCGGTGCTCTCAAGAGGGCGGCTGATGATGA 3180  
 3181 LLSDGSNLTLSVSYKKCRM 3240  
 3241 GACATCTCAACCTGGAGGAGAGATGAGACAGCTTATCTCTCAGCAAGGTGCCAGT 3300  
 3301 DISNLEENDETRFILSKVPS 3360  
 3361 GAGGCACTCCCGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 3420  
 3421 EAPSPMPPTHRSYSLGPGSTSP 3480  
 3481 TTGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 3540  
 3541 LSSSHFPNGRYSGYLLAST 3600  
 3601 CCTGATCAGAGCTGAGCAGCTCCAGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 3660  
 3661 PSEEL 3720  
 3721 CCATTGGGCGGGCTGTCTCTTCTCTCTGTTGGGCGGCGGCGGCGGCGGCGGCGG 2760  
 2761 GATTGAGGAGTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 2820  
 2821 CAGGGGAGCTGGGAGCAGTTCAGTCCCTATTCCAGGCGGCGGCGGCGGCGGCGGCG 2880  
 2881 GAACCTTGGCAAGTTCAGGAGCTGTAACCAACAATCCAGTATGGGGGAGCTTCAAG 2940  
 2941 TGGGCACTTTAGTGTGACCTCTCTCTCTGCTGTAACCAATCACTTGTGTCTTGG 3000  
 3001 GTTTCAGGAGGAGTCCATTGTGCGCACTGGTCAATGTGGTGGTGGAGAGGCTTCAGG 3060  
 3061 CTGAGGCTCAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 3120  
 3121 GTCTAACGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 3180  
 3181 TCGGCTTTGCTTGGCAGGAGGATATGGGATCTCTCATCTGCTGCTATGAGGAAAG 3240  
 3241 CAGAGCTCTCAGGCTTAAGCATGAACGCTGGCATTGAAAGAGGAGGAGGAGGAGG 3300  
 3301 TGGAGTCTTGAATGTTCTCTGTACCCCAACAGGCGCAAGAGGAGGAGGAGGAGGAG 3360  
 3361 CTGAGCGAGAGCATAGTATGGGCTCACTTAACTAGGGTTAGCCATTTTCAATCTTT 3420  
 3421 GCTATCCAACTGGAGCGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 3480  
 3481 ACAGTCAAGATGAGGTTCCAGAGCAAAATTTGAGGTTCCAGGCAATCATGATGTT 3540  
 3541 TTGGAGCACAAGAGAAATCTTCCGCACTTCACTGGAGCTGATTTCTCTTATGTT 3600  
 3601 GAGCTTCTGCTGAGATGAGGAGTTCGAGGATTCAGAGGATGCGCCAGGAGGAGG 3660  
 3661 GTTCTAAGGAGAGAGGTAACCTACTGCGGCTTGGGCTGCGCACTCACTCTGTCGG 3720  
 3721 AATTC 3725

Fig. 1. Nucleotide and deduced amino acid sequence of cDNA encoding NTT4R. The cDNA was cloned from the rat brain library and both strands were sequenced using the Exonuclease III digestion method [19] and oligonucleotide primers.

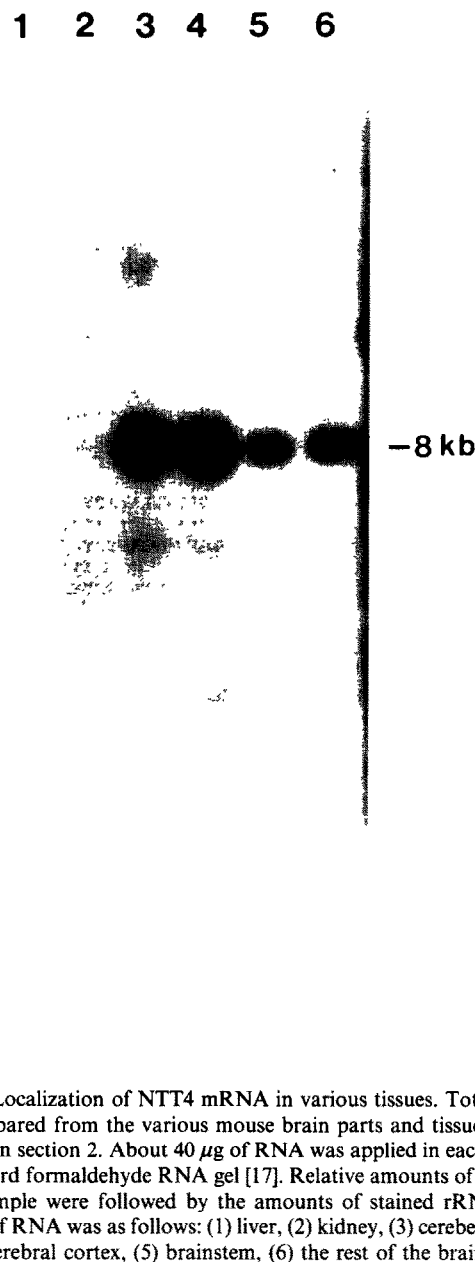


Fig. 2. Localization of NTT4 mRNA in various tissues. Total RNA was prepared from the various mouse brain parts and tissues as described in section 2. About 40  $\mu$ g of RNA was applied in each well in a standard formaldehyde RNA gel [17]. Relative amounts of RNA in each sample were followed by the amounts of stained rRNA. The source of RNA was as follows: (1) liver, (2) kidney, (3) cerebellum, (4) cerebral cortex, (5) brainstem, (6) the rest of the brain.

### 3. RESULTS AND DISCUSSION

Sequence analysis revealed that four of the rat brain cDNA clones were highly homologous to the mouse NTT4 genomic clone [16]. We concluded that these cDNAs resulted from an identical gene of the rat genome and named the rat cDNA NTT4R. Fig. 1 shows the nucleotide and deduced amino acid sequence of the NTT4R clones. This sequence contains the entire reading frame but is missing parts of the 3' and 5' flanking regions. As shown in Fig. 2, the transcript size of NTT4 is about 8 kb. The size of the upstream flanking sequence of one of the NTT4R cDNA clones is estimated to be about 2 kb. This unusually long 5' sequence may present difficulties for the expression of this



Fig. 3. Alignment of the amino acid sequences of NTT4R, the serotonin and taurine transporters. Deduced amino acid sequences of NTT4R (Fig. 1), the serotonin and taurine transporters [6,12] were aligned using a DNASTAR program where NTT4R was aligned in two parts artificially separated in the middle of the loop between membrane segments 7 and 8 (see Fig. 4). Identical amino acids are indicated by (\*).

cDNA. The *EcoRI* site starting the sequence depicted in Fig. 1 is present in the NTT4 transcript and does not result from a cloning artifact. Fig. 2 shows that the NTT4 RNA is quite abundant in the brain and is not present in the kidney or liver. The RNA is much more abundant in the cerebellum and cerebral cortex than in the brainstem or the rest of the brain. This suggests a neuronal function for NTT4. NTT4R encodes 727 amino acids which is about 100 residues larger than average size (about 600 amino acids) of other members of this family. The calculated molecular weight of the transporter (81 kDa) is greater than all known neurotransmitter transporters by about 10 kDa. The difference in molecular weight was demonstrated by comparison of in vitro translation of the cRNAs of NTT4R and

the taurine transporter in rabbit reticulocyte lysate (not shown). The alignment of NTT4R, the serotonin and taurine transporters [6,12] is shown in Fig. 3. It is apparent that the extra sequences in NTT4 are present in the N- and C-termini and in the loop between membrane segments 7 and 8 (Fig. 4). The loop between membrane segments 3 and 4 contains 70–90 amino acids in all transporters sequenced so far. It also contains 3–4 potential glycosylation sites [1–15]. In NTT4R this loop was shortened to about 60 amino acids and contains only a single glycosylation site. The loop between membrane segments 7 and 8 enlarged from 28 amino acids in the other transporters to 90 amino acids in NTT4R, and this loop of NTT4R exclusively contains a potential glycosylation site. Fig. 4 depicts a schematic proposal

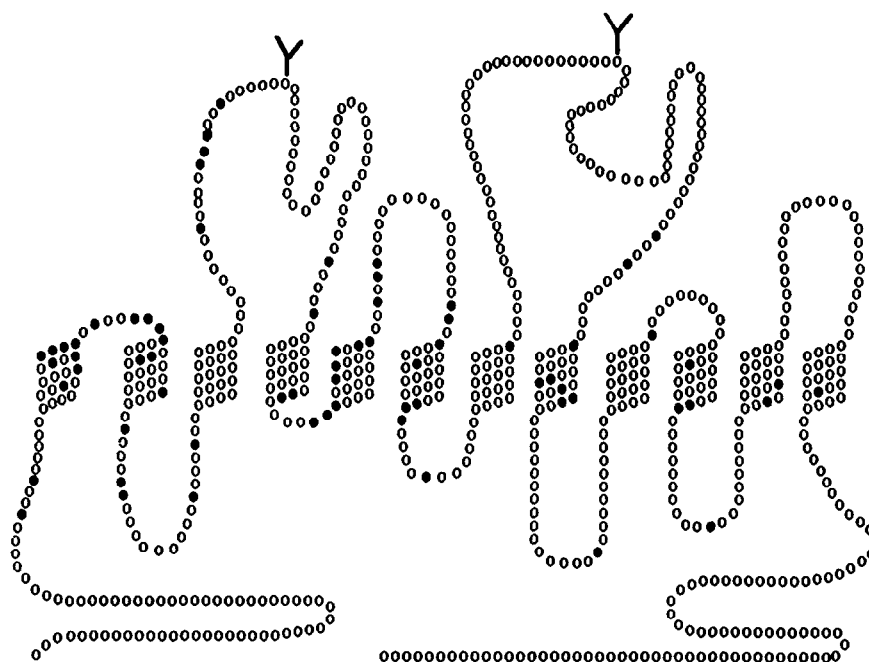


Fig. 4. Schematic presentation of the membrane structure of NTT4. Full symbols represent conserved amino acids that are present at the same position in all neurotransmitter transporters [1–15]. Y indicates a potential glycosylation site.

for the structure of NTT4. The full symbols indicate the position of amino acids that are conserved in all neurotransmitter transporters sequenced so far. NTT4R protein is the largest and least conserved member of the superfamily. The homology falls between 35–38% with other members of the family. The highest homology is with the choline transporter (37.9% identity) and the lowest homology is with the glycine transporters (35% identity). It has the sequence characteristics of amino acid transporters. The six transmembrane helices at the COOH-terminal half are less conserved for NTT4R than those at the amino-terminal half. Only 70 out of its 727 amino acids were strictly conserved among all the transporters. Their position is scattered all over the protein. The concentration of conserved amino acids in membrane segments 1 and 2 suggest a function common to all transporters. In addition, a notable conservation is present in the external loops between membrane segments 3 and 4 as well as segments 5 and 6. The protein folding may bring together these amino acid stretches, and together with the two conserved amino acids in the external loop near membrane segment 8, may form a vital recognition site on the external surface of the membrane.

An extensive attempt to identify the transported substances failed to reveal its identity. *Xenopus* oocytes injected with synthetic NTT4R mRNA show no uptake of individual and combination of the various amino acids, catecholamines and several other substances. The extended 5' untranslated sequence may be a factor affecting the expression. NTT4R might transport an unknown neurotransmitter or require other facilitative

ions besides  $\text{Na}^+$ ,  $\text{Cl}^-$  and  $\text{K}^+$ . Alternatively, cofactors or hetero-oligomers may be involved in NTT4R functional expression. A more interesting possibility is that this protein is expressed in intracellular organelles depending on  $\text{H}^+$  gradient and therefore evaded detection of its transport activity in *Xenopus* oocytes. Recently three laboratories reported on the cloning of an organelar transporter (SV2) [20–22]. Even though SV2 should be less specific than NTT4, no uptake could be detected in cells expressing this protein. The localization of NTT4 in the neuronal parts of the central nervous system and an unusual protein structure suggest an important role in neurotransmission.

## REFERENCES

- [1] Gaustella, J., Nelson, N., Nelson, H., Czyzyk, L., Keynan, S., Miedel, M.C., Davidson, N., Lester, H.A. and Kanner, B.I. (1990) *Science* 249, 1303–1306.
- [2] Nelson, H., Mandiyan, S. and Nelson, N. (1990) *FEBS Lett.* 269, 181–184.
- [3] Pacholczyk, T., Blakely, R.D. and Amara, S.G. (1991) *Nature* 350, 350–354.
- [4] Shimada, S., Kitayama, S., Lin, C.-L., Patel, A., Nanthakumar, E., Gregor, P., Kuhar, M. and Uhl, G. (1991) *Science* 254, 576–578.
- [5] Kilty, J.E., Lorang, D. and Amara, S.G. (1991) *Science* 254, 578–579.
- [6] Hoffman, B.J., Mezey, E. and Brownstein, M.J. (1991) *Science* 254, 579–580.
- [7] Blakely, R.D., Berson, H.E., Freneau Jr., R.T., Caron, M.G., Peek, M.M., Prince, H.K. and Bradley, C.C. (1991) *Nature* 354, 66–69.
- [8] Liu, Q.-R., Nelson, H., Mandiyan, S., López-Corcuera, B. and Nelson, N. (1992) *FEBS Lett.* 305, 110–114.

- [9] Smith, K.E., Borden, L.A., Hartig, P.R., Branchek, T. and Weinschank, R.L. (1992) *Neuron* 8, 927-935.
- [10] Fremeau Jr., R.T., Caron, M.G. and Blakely, R.D. (1992) *Neuron* 8, 915-926.
- [11] López-Corcuera, B., Liu, Q.-R., Mandiyan, S., Nelson, H. and Nelson, N. (1992) *J. Biol. Chem.* 267, 17491-17493.
- [12] Liu, Q.-R., López-Corcuera, B., Nelson, H., Mandiyan, S. and Nelson, N. (1992) *Proc. Natl. Acad. Sci. USA* (in press).
- [13] Mayser, M., Schloss, P. and Betz, H. (1992) *FEBS Lett.* 305, 31-36.
- [14] Mayser, W., Betz, H. and Schloss, P. (1991) *FEBS Lett.* 295, 203-206.
- [15] Liu, Q.-R., López-Corcuera, B., Mandiyan, S., Nelson, H. and Nelson, N. (1993) *J. Biol. Chem.* (in press).
- [16] Liu, Q.-R., Mandiyan, S., Nelson, H. and Nelson, N. (1992) *Proc. Natl. Acad. Sci. USA* 89, 6639-6643.
- [17] Sambrook, J., Fritsch, E.F. and Maniatis, T. (1989) *Molecular Cloning: A Laboratory Manual*, 2nd edn., Cold Spring Harbor Laboratory, Cold Spring Harbor, NY.
- [18] Sanger, F., Nicklen, S. and Coulson, A.R. (1977) *Proc. Natl. Acad. Sci. USA* 74, 5463-5467.
- [19] Henikoff, S. (1984) *Gene* 28, 351-359.
- [20] Bajjalieh, S.M., Peterson, K., Shinghal, R. and Scheller, R.H. (1992) *Science* 257, 1271-1273.
- [21] Feany, M.B., Lee, S., Edwards, R.H. and Buckley, K.M. (1992) *Cell* 70, 861-867.
- [22] Gingrich, J.A., Andersen, P.H., Tiberi, M., El Mestikawy, S., Jorgensen, P.N., Fremeau, R.T. and Caron, M.G. (1992) *FEBS Lett.* 312, 115-122.