Molecular cloning and characterisation of the gene encoding the murine D_4 dopamine receptor

C. Simone Fishburn, Shari Carmon, Sara Fuchs*

Department of Chemical Immunology, The Weizmann Institute of Science, Rehovot 76100, Israel

Received 12 January 1995; revised version received 7 February 1995

Abstract The murine D_4 dopamine receptor was isolated from a murine genomic DNA library. The receptor's entire coding region was contained within a 6 kb *Eco*RI genomic fragment, indicating that the murine D_4 receptor gene is significantly smaller than the corresponding D_2 and D_3 receptor genes, the coding regions of which each stretch over 30 kb. The murine D_4 receptor gene has three introns and four exons, in common with the rat and human D_4 receptor genes. RT-PCR on mRNA from different brain regions shows that the D_4 receptor mRNA is expressed in various areas of the brain, with some differences from the rat and human receptor homologues.

Key words: Murine D_4 receptor; Antipsychotic drug; Neuropathological disease; Gene organisation; Dopamine; G protein-linked receptor

1. Introduction

Dopamine receptors have been associated with a number of neuropathological disorders, and represent the primary therapeutic target in schizophrenia, Parkinson's disease and Huntington's chorea [1,2]. Receptors of the D2** subclass have shown a particular connection with these disorders, since many of the drugs used in their treatment act as specific agonists or antagonists at these receptors [3]. Molecular cloning studies have thus far ascribed three subtypes to this class, termed the D_2, D_3 and D_4 receptors, and these receptors are found to bind a range of antipsychotic drugs with differing orders of affinity (reviewed in [4,5]). While typical neuroleptics, such as haloperidol, bind all three subtypes with a dissociation constant in the nanomolar range, the atypical neuroleptic clozapine binds the D_4 subtype more potently than the D_2 and D_3 subtypes by one order of magnitude [6]. The main drawback of the use of typical neuroleptics in psychotherapy is that they frequently lead to the development of tardive dyskinesia, a hyperkinetic movement disorder similar to Parkinson's disease. Since treatment of schizophrenic patients with clozapine is generally not accompanied by these side effects, its selectivity and high affinity for the D_4 receptor have aroused much interest in this receptor subtype. Moreover, following the cloning of the human D_4 dopamine receptor, it was found to exist in several polymorphic forms in the human population [7], further fuelling interest in this receptor subtype.

Dopamine receptors all belong to the superfamily of G protein-linked receptors, and their sequences therefore display the predicted seven transmembrane topology. D2-type dopamine receptors, however, are unusual in this family, as they are encoded by genes containing introns within the coding region, in contrast to D1-type receptors and the majority of G protein-linked receptor genes which are encoded by intronless genes [8]. The presence of introns in the coding region has been shown, for the D_2 and D_3 receptor subtypes, to give rise to alternative splicing in the region of the putative third cytoplasmic loop, generating 'long' and 'short' functional receptor isoforms [9–12]. Previously, we performed a detailed analysis of the structural organisation of the murine D_3 dopamine receptor gene [13], and showed that, like the murine D_2 receptor gene [14], it contains six introns within its coding region, one more than the five introns reported for the homologous human D_3 receptor gene [15]. The D_4 receptor gene has been isolated from human and rat sources and found to differ from the D_2 and D_3 receptors in the number of introns and exons [6,16]. In order to characterise further the functional differences and evolutionary relationship between the various D2-type dopamine receptors, we have cloned and sequenced the murine D_4 dopamine receptor gene, and have examined its intron–exon organisation.

2. Materials and methods

2.1. Materials

All molecular biology reagents, unless stated otherwise, were purchased from Boehringer (Mannheim Germany). PCR on cDNA was performed using *Taq* polymerase (Promega). Oligonucleotides and PCR primers were prepared by the laboratory of oligonucleotide synthesis, and sequence reactions were performed on both strands of DNA by the DNA sequencing facility, both at the Weizmann Institute of Science. Nitrocellulose membranes were purchased from Schleicher and Schuell, and [γ-^{32}P]dATP (3,000 Ci/mmol) was purchased from Amersham.

2.2. Isolation of the murine D_4 dopamine receptor gene

A λFixII 129/sv mouse genomic library (Stratagene) was screened using 5′ labelled oligonucleotides: (1) 5′-CCACGGAAAGTGGCCCAGTAAAGC-3′ (rat D_4 739–761]); (2) 5′-TTCCTGATGTGTTGGACGCCTTTC-3′ (rat D_4 1091–1114); (3) 5′-CAACAGCCCACCGTCACCAGTAGCGCTGCTGTTCCCCAT-3′ (rat D_4 131–169); (4) 5′-AGCAGCGGAGACGAAGAGTCTTGCGGAAGACACTTC-3′ (rat D_4 1251–1285); based on the rat D_4 receptor sequence [16]. Probe labelling, hybridisation and washing conditions were performed according to standard procedures [17]. DNA from positive clones was isolated using Lambda Maxi-prep columns (Qiagen), and inserts were excised using the enzyme *Not*I. A 6 kb *Eco*RI fragment from a positive clone was subcloned into the vector pBluescript (Stratagene) for sequence analysis.

*Corresponding author. Fax: (972) (8) 344 141.

*To avoid confusion of nomenclature, the terms D1 and D2 shall be used to refer to the two classes of dopamine receptor originally defined by pharmacological and biochemical studies at the protein level. The cloned receptor subtypes, defined by their DNA sequence, shall be denoted by a subscript, e.g. D_2.

SSDI 0014-5793(95)00168-9

2.3. Analysis of D_4 receptor mRNA distribution

For analysis of the tissue distribution of the murine D_4 receptor, total RNA was prepared from different murine tissues and brain regions using the guanidium thiocyanate-CsCl method [18]. RT-PCR was performed using primers based on sequences from the receptor's putative third cytoplasmic loop: (5) 5′-GCCACTTTCCGCGGCCTGC-3&prime (mouse D_4 627–645); (6) 5′-GGACTCTCA-TTGCCTTGC-3′ (mouse D_4 929–945). PCR products were electrophoresed on agarose gels and Southern blotted as described previously [12].

3. Results

In order to clone the murine form of the D_4 dopamine receptor and examine its genomic organisation, we screened a mouse (129/sv) genomic library using oligonucleotide probes 1 and 2, based on the sequence of the third cytoplasmic loop of the rat D_4 receptor [16], since this is the region found to differ most between the respective dopamine receptors. Six clones were isolated, one of which also hybridised with oligonucleotide probes 3 and 4, derived from the amino- and carboxy-termini of the rat D_4 receptor, respectively, suggesting that the full coding region of the murine D_4 receptor was present. Restriction digestion and Southern blotting of this clone revealed a 6 kb *Eco*RI fragment the hybridisation profile of which suggested it contained the receptor's entire coding region. This fragment was subcloned into the vector pBluescript, and its identity as the murine form of the D_4 dopamine receptor was confirmed by sequence analysis.

The murine D_4 dopamine receptor is a protein of 387 amino acids, the sequence of which predicts the seven transmembrane topology expected for members of the G protein-linked recep-

```
1    ATG GGG AAC AGC AGC GCT ACT GAG GAC GGT GGG CTG TTG GCC GGG CGT GGG CCA GAA TCC 60
1    M  G  N  S  S  A  T  E  D  G  G  L  L  A  G  R  G  P  E  S  20
                *
61   CTG GGG ACT GGC GCC GGG CTT GGG GGC GCG GGC GCG GCG GCG CTG GTG GGG GGC GTG CTG 120
21   L  G  T  G  A  G  L  G  G  A  G  A  A  A  L  V  G  G  V  L  TM1
121  CTC ATC GGC TTG GTG TTG GCA GGG AAC TCG CTC GTG TGC GTG AGC GTG GCC TCC GAG CGC 180
41   L  I  G  L  V  L  A  G  N  S  L  V  C  V  S  V  A  S  E  R  60
181  ACG CTG CAG ACA CCC ACC AAC TAC TTC ATC GTG AGC CTG GCT GCT GCC GAC CTC CTC CTC 240
61   T  L  Q  T  P  T  N  Y  F  I  V  S  L  A  A  A  D  L  L  L  TM2
241  GCG GTG CTG GTG CTG CCT CTC TTT GTC TAC TCC GAG GTC CAG GGT GGC GTG TGG CTC CTG 300
81   A  V  L  V  L  P  L  F  V  Y  S  E  V  Q  G  G  V  W  L  L  100
301  AGC CCC CGT CTC TGT GAC ACG CTC ATG GCC ATG GAC GTC ATG CTG TGC ACC GCC TCC ATC 360
101  S  P  R  L  C  D  T  L  M  A  M  D  V  M  L  C  T  A  S  I  TM3
361  TTC AAC CTG TGC GCC ATC AGC GTG GAC AGG TTC GTG GCC GTG ACC GTG CCG CTG CGC TAC 420
121  F  N  L  C  A  I  S  V  D  R  F  V  A  V  T  V  P  L  R  Y  140
421  AAC CAG CAG GGT CAG TGC CAG CTG CTG CTC ATC GCC GCC ACG TGG CTG CTG TCC GCC GCG 480
141  N  Q  Q  G  Q  C  Q  L  L  L  I  A  A  T  W  L  L  S  A  A  TM4
481  GTG GCT TCG CCA GTG GTG TGT GGC CTC AAT GAT GTG CCC GGC CGC GAT CCG GCC GTG TGC 540
161  V  A  S  P  V  V  C  G  L  N  D  V  P  G  R  D  P  A  V  C  180
541  TGC CTG GAG AAC CGA GAC TAT GTG GTC TAC TCG TCC GTC TGC TCC TTC TTC CTG CCC TGT 600
181  C  L  E  N  R  D  Y  V  V  Y  S  S  V  C  S  F  F  L  P  C  TM5
600  CCG CTC ATG CTA CTG CTT TAC TGG GCC ACT TTC CGC GGC CTG CGG CGC TGG GAG GCA GCC 660
201  P  L  M  L  L  L  Y  W  A  T  F  R  G  L  R  R  W  E  A  A  220
661  CGG CAC ACC AAA CTT CAC AGC CGC GCG CCG CGC CGA CCC AGC GGC CCC GGC CCG CCG GTG 720
221  R  H  T  K  L  H  S  R  A  P  R  R  P  S  G  P  G  P  P  V  240
                                    ●
721  TCG GAC CCT ACT CAG GGT CCC TTC TTC CCA GAC TGC CCA CCT CCC TTA CCC AGC CTC CGG 780
241  S  D  P  T  Q  G  P  F  F  P  D  C  P  P  P  L  P  S  L  R  260
781  ACG AGC CCC AGC GAC TCC AGC AGG CCG GAG TCA GAG CTC TCT CAG AGA CCC TGC AGC CCC 840
261  T  S  P  S  D  S  S  R  P  E  S  E  L  S  Q  R  P  C  S  P  280
841  GGG TGT CTG CTC GCT GAT GCA GCT CTC CCG CAA CCT CCT GAG CCG TCT TCC CGC AGA AGG 900
281  G  C  L  L  A  D  A  A  L  P  Q  P  P  E  P  S  S  R  R  R  300
901  AGA GGC GCC AAG ATC ACG GGA AGG GAG CGC AAG GCA ATG AGA GTC CTG CCG GTG GTA GTC 960
301  R  G  A  K  I  T  G  R  E  R  K  A  M  R  V  L  P  V  V  V  TM6
961  GGG GCC TTC CTG GTG TGT TGG ACG CCT TTC TTC GTG GTG CAC ATC ACG CGG GCG CTG TGT 1020
321  G  A  F  L  V  C  W  T  P  F  F  V  V  H  I  T  R  A  L  C  340
1021 CCG GCT TGC TTC GTG TCT CCG CGC CTG GTC AGT GCC GTC ACC TGG CTA GGC TAT GTC AAC 1080
341  P  A  C  F  V  S  P  R  L  V  S  A  V  T  W  L  G  Y  V  N  TM7
1081 AGT GCC CTC AAC CCC ATC ATC TAC ACC ATC TTC AAC GCG GAG TTT CGA AGT GTC TTC CGC 1140
361  S  A  L  N  P  I  I  Y  T  I  F  N  A  E  F  R  S  V  F  R  380
1141 AAG ACT CTC CGT CTC CGC TGC TGA 1284
381  K  T  L  R  L  R  C  *  387
```

Fig. 1. Sequence of the coding region of the murine D_4 dopamine receptor. The deduced amino acid sequence is presented using the one-letter code. The seven putative transmembrane domains (TM1– TM7) are underlined. Positions of introns are shown by arrows. The potential N-linked glycosylation site is marked by an asterisk, and the putative site for phosphorylation by protein kinase A is marked by a filled circle. This sequence has been submitted to the Genbank and assigned the accession number U19880.

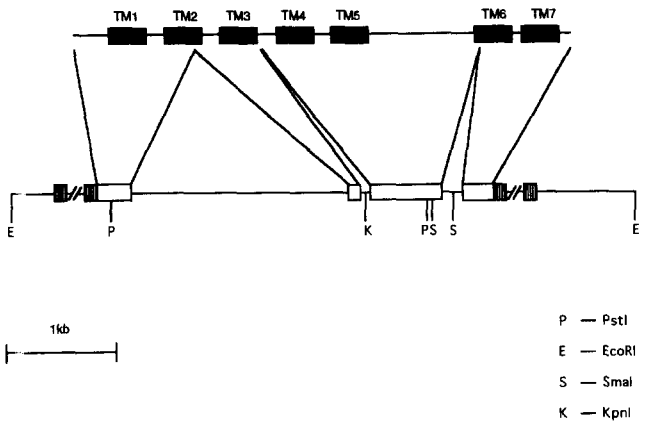

Fig. 2. Structure of the murine D_4 dopamine receptor gene. The top shows the proposed arrangement of the transmembrane segments (dark boxes) in the receptor. Receptor regions are connected to exons represented on the gene map below. Exons within the coding region are shown by white boxes, exons from the 5′ and 3′ non-translated regions are shown by hatched boxes. Introns are represented by a solid line. The gene map is drawn to scale as indicated. Restriction sites are indicated. E, *Eco*RI; K, *Kpn*I; P, *Pst*I; S, *Sma*I.

tor superfamily (Fig. 1). In common with the D_2 and D_3 receptors, the D_4 subtype displays a putative N-linked glycosylation site in the extracellular amino-terminus, and a consensus sequence for phosphorylation in the third cytoplasmic loop, both of which are conserved in the human and rat forms of the receptor. The murine D_4 receptor is highly homologous to the rat receptor [16], but has an extra two residues (VQ) at the end of the second transmembrane domain, which are also present in the human D_4 receptor [6]. This discrepancy is due to the fact that the sequence GTCCAG immediately following intron 1 was interpreted for the mouse and human receptors as the beginning of exon 2, representing the residues Val and Gln, whereas for the rat receptor they were seen as part of the intron. Since in all three species the sequence of this section was derived from a genomic rather than a cDNA clone, either interpretation could be correct, as the consensus sequence for an acceptor site (T/C_{11} NCAG) [19] is conserved in both cases. However, sequence comparison with the murine D_2 and D_3 dopamine receptors shows them to have an intron at exactly the same position as that determined here [13, 14], supporting the interpretation that the GTCCAG sequence represents the beginning of the second exon, and encodes the residues valine and glutamine.

We have used this genomic clone to examine the structural organisation of the murine D_4 receptor gene. As is shown in Fig. 2, this gene contains three introns and four exons in the coding region, in common with the rat and human D_4 receptor genes. Introns 1, 2 and 3 in the murine D_4 receptor gene are located at identical positions to introns 1, 2 and 6 in the murine D_2 and D_3 receptor genes, supporting a common evolutionary origin [13,14]. However, these introns are significantly smaller than their parallels in the other D2-like dopamine receptor genes, accounting for the comparatively small size of the D_4 receptor gene. As can be seen from Table 1, the GT:AG consensus for intron–exon boundaries is conserved for introns 1 and 2, at both donor and acceptor sites, but varies slightly for intron 3, as has been reported for the corresponding introns in the human D_4 receptor gene [6].

In order to examine the tissue distribution of the murine D_4 dopamine receptor, PCR was performed on different murine tissues and brain regions using primers (5 and 6) derived from the region of the third cytoplasmic loop. Southern analysis of the PCR products is shown in Fig. 3. The murine D_4 receptor mRNA is present in the olfactory bulb, hypothalamus, olfactory tubercle, brainstem and striatum, in common with the rat and monkey homologues [6,16]. Low levels of receptor mRNA were also detected in cortex (data not shown). In contrast with the rat and monkey, however, a relatively strong signal is obtained in the murine cerebellum. No signal could be detected in murine heart tissue (data not shown), unlike the rat heart in which relatively high levels of D_4 receptor mRNA have been reported [16].

4. Discussion

In this study we have cloned and sequenced the murine D_4 dopamine receptor gene, and have analysed its intron–exon organisation. The structural organisation of the murine D_4 receptor gene shows it to have three introns and four exons within the coding region, at analagous positions to those in the human and rat D_4 receptors [6,16]. The presence of the entire coding region within a 6 kb *Eco*RI fragment demonstrates that the murine D_4 receptor gene is significantly smaller than the murine D_2 and D_3 receptor genes, whose coding regions each extend over more than 30 kb [13,14]. This is in accordance with the human and rat D_4 receptor genes, which were found to be contained within 5 kb and 3.5 kb restriction fragments, respectively.

The murine D_2 and D_3 receptors show directly analogous gene organisations, and are closer to each other than to the murine D_4 receptor, which differs from them by the absence of

Table 1
Exon–intron boundaries of murine D_4 receptor gene

Intron no.	Donor		Acceptor		Position of intron[a]	Intron size	Size of corresponding human D_4 intron[b]	Size of corresponding murine D_3 intron[c]
1	TCCGAG	gtgagcctcg	tgcgtgaccag	GTCCAG	275	~2000 bp	~2000 bp	>12 kb
2	GGACAG	gtgggtaccc	tgctcctctag	GTTCGT	389	88 bp	103 bp	18 kb
3	TAGTCG	gtgggatcct	cccatccccag	GGGCCT	1022	174 bp	92 bp	1.3 kb

Upper case letters denote exon sequences, lower case letters denote intron sequences.
[a]Number refers to position in murine D_4 receptor cDNA sequence (Fig. 1).
[b]Van Tol et al. [6].
[c]Park et al. [13].

introns following the fourth transmembrane domain and in the putative third cytoplasmic loop [13,14]. Nevertheless, the structural organisation of the D_4 receptor gene remains closer to the D_2 and D_3 receptors than to any other receptor gene thus far cloned, supporting the idea that all three genes may have evolved from a common ancestor gene by successive gene duplication events. The presence or absence of introns in the coding region represents one of the fundamental differences between the D2-type and the D1-type receptor genes. While most G protein-linked receptors thus far studied are encoded by intronless genes, there is a growing number of receptors in this family which display introns in the coding region, suggesting that there may be subfamilies of receptors in this superfamily. Within the class of intron-containing G protein-linked receptors, the location of the introns with respect to the seven transmembrane domains is found to coincide among different neurotransmitter and hormone receptors [20]. It has been postulated that the intron-exon organisation of such G protein-linked receptors may signify an arrangement according to discrete functional domains, suggesting that gene conversion and duplication events may have taken place to generate structurally similar and functionally different receptors.

Sequence comparison of the murine D_4 subtype with the human and rat homologues shows that at the nucleotide level, the mouse D_4 receptor displays 79% and 93% homology with the human and rat receptors, respectively. The amino acid sequence of the murine D_4 receptor displays 80% and 95% homology with the human and rat D_4 receptors, respectively, the most conserved regions being observed in the transmembrane domains, consistent with the proposal that these domains form the ligand-binding site. The least conserved regions between the murine and human forms of the D_4 receptor are found in the amino-terminal domain (76% identity) and the putative third cytoplasmic loop (72% identity). The diversity between the murine and human D_4 receptors in the third cytoplasmic loop is significantly greater than that observed for the D_2 and D_3 receptors in this region, which show 93% and 83% sequence identity, respectively, between these species.

The human D_4 receptor has been found to exist in several polymorphic variations, which differ by the number of repeat units (composed of 16 amino acids), also located in the third cytoplasmic loop, and these polymorphic variants are found to display slightly different pharmacological properties [7]. If the human D_4 receptor's polymorphic variants underlie different D_4-related behavioural properties in the human population,

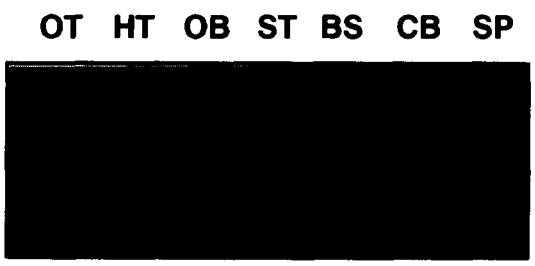

Fig. 3. Tissue distribution of the murine D_4 dopamine receptor. Southern analysis was performed on PCR products amplified using primers 5 and 6 from different mouse tissues, as described in section 2. OB, olfactory bulb; HT, hypothalamus; OT, olfactory tubercle; BS, brainstem; ST, striatum; CB, cerebellum; SP, spleen.

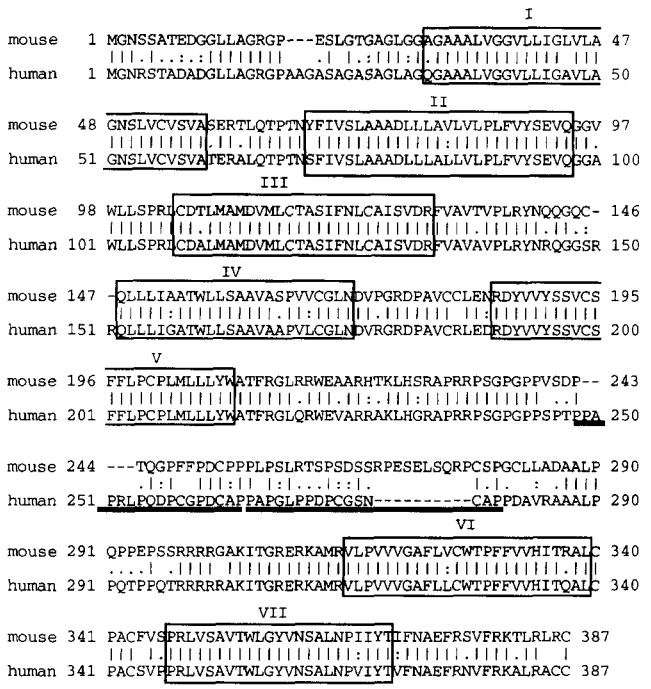

Fig. 4. Sequence alignment of the mouse and human D_4 dopamine receptors. Transmembrane domains are boxed and indicated in Roman numerals. The 16 amino acid repeat, present twice, in the human receptor is denoted by a solid black line.

this region of diversity may represent a functional domain which is responsible for these different properties. Fig. 4 shows a comparison of the murine receptor with the human D_4 receptor variant containing two repeat units ($D_{4.2}$), which is the closest form to the murine and rat receptors. It can be seen from this comparison that the area of greatest diversity between the murine and human D_4 receptors overlaps with the sequence representing the two 16 amino acid repeats in the human receptor (Fig. 4, bold underline). Given that complex behaviour and its molecular basis may be expected to differ between mice and humans, the weak degree of conservation among these residues between the species supports the notion that behavioural properties of the D_4 receptor may relate to this functional domain.

It has been suggested that the low levels of D_4 receptor mRNA may not directly reflect the level of receptor protein, and that higher levels of D_4 receptor protein are detected than would be predicted from the mRNA levels [21]. The availability of the murine D_4 receptor sequence enables us to develop anti-peptide antibodies specific for the murine D_4 dopamine receptor, which can be used in studies to determine the distribution and properties of the D_4 receptor protein in the mouse. Furthermore, the current movement towards analysis of a protein's role by gene targetting techniques generally uses the mouse as the species of choice [22]. Thus, it may be of particular benefit to have information regarding the characteristic features of the murine form of the D_4 dopamine receptor.

Acknowledgements: This work was supported by grants from the United States–Israel Binational Sciences Foundation, the Ernst and Anne Chain Research Programme and the Leo and Julia Forchheimer Center for Molecular Genetics at the Weizmann Institute of Science.

References

[1] Seeman, P., (1987) in: Dopamine Receptors, vol. 8 (I. Creese, C.M. Fraser, eds) pp. 233–245, Liss, New York.
[2] Seeman, P. (1987) Synapse 1, 133–52.
[3] Seeman, P. and Van Tol, H.H. M. (1994) Trends Pharmacol. Sci. 15, 264–270.
[4] Sibley, D.R. and Monsma, F.J. (1992) Trends Pharmacol. Sci. 13, 61–9.
[5] Civelli, O., Bunzow, J.R. and Grandy, D.K. (1993) Annu. Rev. Pharmacol. Toxicol., 32, 281–307.
[6] Van Tol, H.H., Bunzow, J.R., Guan, H.C., Sunahara, R.K., Seeman, P., Niznik, H.B. and Civelli, O. (1991) Nature 350, 610–4.
[7] Van Tol, H.H.M., Caren, M.W., Guan, H.-C., Ohara, K., Bunzow, J.R., Civelli, O., Kennedy, J., Seeman, P., Niznik, H.B. and Javanovic, V. (1992) Nature 358, 149–152.
[8] Dohlman, H.G., Caron, M.G. and Lefkowitz, R.J. (1987) Biochemistry, 26, 2657–64.
[9] Dal Toso, R., Sommer, B., Ewert, M., Herb, A., Pritchett, D.B., Bach, A., Shivers, B.D. and Seeburg, P.H. (1989) EMBO J. 8, 4025–4034.
[10] Monsma, F.J.J., McVittie, L.D., Gerfen, C.R., Mahan, L.C. and Sibley, D.R. (1989) Nature 342, 926–9.
[11] Chio, C.L., Hess, G.F., Graham, R.S. and Huff, R.M. (1990) Nature 343, 266–9.
[12] Fishburn, C.S., Belleli, D., David, C., Carmon, S. and Fuchs, S. (1993) J. Biol. Chem. 268, 5872–8.
[13] Park, B.-H., Fishburn, C.S., Carmon, S., Accili, D. and Fuchs, S. (1995) J. Neurochem. (in press).
[14] Mack, K.J., Todd, R.D. and O'Malley, K.L. (1991) J. Neurochem. 57, 795–801.
[15] Sokoloff, P., Giros, B., Martres, M.P., Bouthenet, M.L. and Schwartz, J.C. (1990) Nature 347, 146–51.
[16] O'Malley, K.L., Harmon, S., Tang, L. and Todd, R.D. (1992) New. Biol. 4, 137–46.
[17] Sambrook, J., Fritsch, E.F. and Maniatis, T. (1989) Molecular Cloning: A Laboratory Manual, Cold Spring Harbor Laboratory Laboratory Press, NY.
[18] Chirgwin, J.M., Przybyla, A.E., MacDonald, R.J. and Rutter, W.J. (1979) Biochemistry 18, 5294–5299.
[19] Padgett, R.A., Grabowski, P.J., Konarska, M.M., Seiler, S. and Sharp, P.A. (1986) Annu. Rev. Biochem. 55, 1119–1150.
[20] Stam, N.J., Vanderheyden, P., van Alebeek, C., Klomp, J., de Boer, T., van Delft, A.M.L. and Olijve, W. (1994) Eur. J. Pharmacol. 269, 339–348.
[21] Strange, P.G. (1994) Trends Pharmacol. Sci. 15, 317–319.
[22] Capecchi, M.R. (1989) Trends Genet. 5, 70–76