SHORT COMMUNICATION

Molecular Cloning and Chromosomal Localization of Human Genes Encoding Three Closely Related G Protein-Coupled Receptors

ZHAO-HUI SONG,* WILLIAM MODI,† AND TOM I. BONNER*,[1]

*Laboratory of Cell Biology, NIMH National Institutes of Health, Bethesda, Maryland 20892; and †Biological Carcinogenesis and Development Program, Program Resources Inc./Dyncorp, National Cancer Institute, Frederick Cancer Research and Development Center, Frederick, Maryland 21702-1201

Received January 23, 1995; accepted May 5, 1995

Cosmids containing human genes for three orphan G protein-coupled receptors, GPR12, GPR6, and GPR3, were isolated using their rat homologs as probes. Previous studies of the mouse and rat cDNAs have shown the receptors to be expressed primarily in brain but have failed to identify their ligands. The three receptor proteins of 334, 363, and 330 amino acids, respectively, are encoded by a single exon in each gene. Excluding the divergent sequences preceding the first transmembrane domain, they have ~60% amino acid identity with each other. Fluorescence *in situ* hybridization of GPR12, GPR6, and GPR3 localized these three genes to human chromosomal regions 13q12, 6q21, and 1p34.3–p36.1, respectively. © 1995 Academic Press, Inc.

A variety of extracellular signals are transmitted into cells through integral membrane receptors that couple to heterotrimeric G proteins (4). These G protein-coupled receptors therefore are crucial for the normal functions of many cell types, including neurons, endocrine cells, cardiac and smooth muscle cells, immune cells, and sensory cells for detection of light, taste, and smell. Recently, both activating and loss-of-function mutations in G protein-coupled receptors have been discovered to cause human diseases (2). Cloning of human genes encoding novel G protein-coupled receptors and mapping of these genes to human chromosomes may help to reveal the roles of these receptors in health and disease as well as provide potential targets for novel pharmacological therapies.

Recently, we and others have cloned rodent cDNAs for three closely related G protein-coupled receptors whose ligands have yet to be identified (3, 7, 8). Sequence alignment showed that these three receptors share ~60% amino acid identity and a number of common structural characteristics. Among receptors whose ligands are known, these receptors are mostly related to the cannabinoid and melanocortin receptors, although not so closely related as to suggest the identity of their ligands or their physiological function. Northern and *in situ* hybridization analyses demonstrated that they are expressed predominantly in the brain, each with a unique distribution pattern. These data suggest that they belong to a new subfamily of neurotransmitter/neuromodulator receptors with identical or closely related endogenous ligands. In this communication we report the cloning of human genes encoding these three receptors and the assignment of these genes to specific chromosomal regions by fluorescence *in situ* hybridization.

Partial rat cDNAs for the three closely related receptors GPR12, GPR6, and GRP3 were obtained through the application of a PCR cloning method based on conserved sequence motifs (8). These rat cDNA fragments were labeled with [^{32}P]dCTP by nick-translation and were used as probes at moderate stringency (hybridization, 3× SSPE, 1× Denhart's, 60°C; washing, 1× or 3× SSPE, 60°C) to isolate cosmids clones from a human genomic library (Stratagene 951202) by Southern blots of pools of cosmids followed by colony hybridization to positive pools of 2000 to 4000 cosmids (1). Hybridizing restriction fragments were subcloned into pUC18 and sequenced as previously described (8). Comparison of the rat cDNA sequences with the human genes indicates that, as with many other G protein-coupled receptors, there are no introns within the coding regions of these three receptor genes. However, all three genes have introns in their 5′ noncoding sequences.

For GPR12, the single coding exon is contained within a 1.3-kb *Hin*dIII–*Eco*RI fragment (GenBank Accession No. U18548) with the initiation codon located ~11 kb from the 5′ end of the cosmid. Comparison of the sequence with two different rat cDNAs (3, 8) and a mouse cDNA (7) suggests that there is a splice acceptor site located either 9 or 16 bases 5′ of the initiation codon (Fig. 1a). The site at 9 bases before the initiation codon

Sequence data from this article have been deposited with the EMBL/GenBank Data Libraries under Accession Nos. 618548–618550.

[1] To whom correspondence should be addressed. Telephone: (301) 496-8907. Fax: (301) 402-1748. E-mail: tibonner@helix.nih.gov.

GENOMICS **28**, 347–349 (1995)
0888-7543/95 $12.00
Copyright © 1995 by Academic Press, Inc.
All rights of reproduction in any form reserved.

a

```
                                            -16     -9
                                             ↓      ↓        MetAsnGluAspLeuLysValAsnLeuSerGly
human GPR12 gene  TAAACATTAAAAAGCCTGTCTTTTCGTTGAAGAGGACAGGGGTTAAAATGAATGAAGACCTGAAGGTCAATTTAAGCGGG
rat cDNA U12184   C.....................C.......................C.......C..................
rat cDNA X61496   AAGGAGCCCCCCCTCGGTCTGGGCGTGCC.......................C.......C..................
mouse cDNA D21061 AAGGAGCCACCCCTCGGCCAGGGCGTGCCA-------...............C.......C..................
```

b

```
GPR12                                                       MNEDLKVNLSGLPRDYLDAAAAENISAAVSSRVPAVEPE
GPR6                      MNASAASLNDSQVVVVAAEGAAAAATAAGGPDTGEWGPPAAAALGAGGGANGSLELSSQLSAGPPGL
GPR3                                                               MMWGAGSPLAWLSAGSGNVNVSSVGPAEGPTGPAA

                    I                            II                          III
        ------------------------------  -------------------------     -------------------------
GPR12   PELVVNPWDIVLCTSGTLISCENAIVVLIIFHNPSLRAPMFLLIGSLALADLLAGIGLITNFVFAYLLQSEATKLVTIGLIVASFSASVCSLLAITVDRY
GPR6    LLPAVNPWDVLLCVSGTVIAGENALVVALIASTPALRTPMFVLVGSLATADLLAGCGLILHFVFQYLVPSETVSLLTVGFLVASFAASVSSLLAITVDRY
GPR3    PLPSPKAWDVVLCISGTLVSCENALVVAIIVGTPAFRAPMFLLVGSLAVADLLAGLGLVLHFAAVFCIGSAEMSLVLVGVLAMAFTASIGSLLAITVDRY
               **  ** ***    *** **  *   *  * *** * **** ****** **    *        *   *   *    * **  **********

                          IV                                    V
                ------------------------              --------------------
GPR12   LSLYYALTYHSERTVTFTYVMLVMLWGTSICLGLLPVMGWNCLRDESTCSVVRPLTKNNAAILSVSFLFMFALMLQLYIQICKIVMRHAHQIALQHHFLA
GPR6    LSLYNALTYYSRRTLLGVHLLLAATWTVSLGLGLLPVLGWNCLAERAACSVVRPLARSHVALLSAAFFMVFGIMLHLYVRICQVVWRHAHQIALQQHCLA
GPR3    LSLYNALTYYSETTVTRTYVMLALVWGGALGLGLLPVLAWNCLDGLTTCGVVYPLSKNHLVVLAIAFFMVFGIMLQLYAQICRIVCRHAQQIALQRHLLP
        **** **** *  *       *   *     ******  ****      * ** **      *   *   *  ** **  **  * *** ***** * *

                    VI                        VII
           ------------------------    -----------------------
GPR12   TSHYVTTRKGVSTLAIILGTFAACWMPFTLYSLIADYTYPSIYTYATLLPATYNSIINPVIYAFRNQEIQKALCLICCGCIPSSLAQRARSPSDV
GPR6    PPHLAATRKGVGTLAVVLGTFGASWLPFAIYCVVGSHEDPAVYTYATLLPATYNSMINPIIYAFRNQEIQRALWLLLCGCFQSKVPFRSRSPSEV
GPR3    ASHYVATRKGIATLAVVLGAFAACWLPFTVYCLLGDAHSPPLYTYLTLLPATYNSMINPIIYAFRNQDVQKVLWAVCCCCSSSKIPFRSRSPSDV
          *   ****  ***  ** * * * ** *       * *** ********* *** *******  *  *     * * *    * **** *
```

FIG. 1. (**a**) A putative splice acceptor site in the 5′ noncoding region of GPR12. Arrows are the possible 3′ ends of intron. Dots indicate identity with GPR12. Dashes indicate gaps in mouse cDNA to align with rat cDNA X61496. (**b**) Comparison of the amino acid sequences of the three human orphan G protein-coupled receptors, GPR12, GPR6, and GPR3. Putative transmembrane regions are indicated by dashed lines. Amino acid sequences that are identical in three sequences are indicated by asterisks.

would end in the usual CAG but would be unusually purine rich, while the site 7 bases further 5′ would have the usual pyrimidine-rich sequence but end with an atypical AAG. One rat cDNA called 6–7 (GenBank Accession No. U12184) (8) appears not to be spliced at this point, while the other rat cDNA called R334 (GenBank Accession No. X61496) (3) and a mouse cDNA called GPCR01 (GenBank Accession No. D21061) (7) are spliced.

For GPR6, a 2.7-kb *Sst*I–*Eco*RI fragment from the middle of the cosmid was sequenced (GenBank Accession No. U18549) and by comparison with the rat cDNA called rCNL3 (GenBank Accession No. U12006) (8) contains a 5′ untranslated exon of >66 bp (70% identity to bases 66–130 of rCNL3), an intron of 628 bp, and a coding exon of 1.6 kb that extends to the poly(A) site.

For GPR3, a 3.3-kb segment extending to the 3′ end of the cosmid was sequenced (GenBank Accession No. U18550). A 2.51-kb segment of GPR3 is homologous to the whole 2.49-kb mouse cDNA called GPCR21 (GenBank Accession No. D21062) (7). This uninterrupted homology suggests that this gene is intronless. However, a recent report describes a cDNA with 27 bases of 5′ untranslated sequence that defines an intron located 6 bases 5′ of the initiation codon and a 5′ untranslated exon (5). Our sequence spans the gap between their reported exon sequences (GenBank Accession Nos. L32830 and L32831) and defines the intron as being 1056 bases. Thus, the mouse cDNA appears to be incompletely spliced, as appears to be the case with the rat GPR12 cDNA (GenBank Accession No. U12184) (3). The apparent poly(A) site of the mouse cDNA maps by sequence homology to a poly(A) signal in the human gene as well as in a cDNA isolated from a rat cortex library (unpublished). Since the mouse cDNA contains no obvious poly(A) signal, it is probable that the cDNA was primed by oligo(dT) binding to an A-rich sequence overlapping the poly(A) signal. The amino acid sequence of GPR3 is identical to the sequences described in two recent reports, which appeared during the preparation of this paper (5, 6).

As shown in Fig. 1b, human genes GPR12, GPR6, and GPR3 encode proteins of 334, 363, and 330 amino acids, respectively. Between each of these receptors, there is ~60% amino acid identity from transmembrane region 1 through the carboxyl-termini. As previously pointed out (8), this degree of sequence identity in rhodopsin-like G protein-coupled receptors suggests that they have the same, or at least closely related, endogenous ligands. In addition, GPR12, GPR6, and GPR3 have 92, 88, and 93% amino acid identity, respectively, with their rodent homologs. It is thus highly probable that the three genes have existed since well

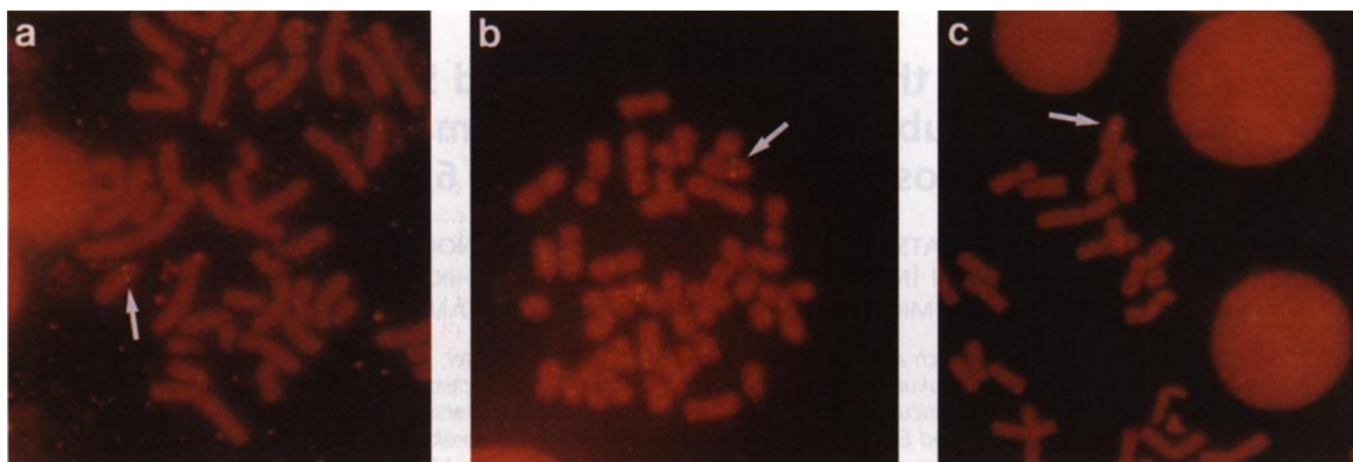

FIG. 2. Partial metaphase cells following fluorescence *in situ* hybridization with the (**a**) GPR12 cosmid, (**b**) GPR6 cosmid, and (**c**) GPR3 cosmid. Arrows indicate hybridization signals at 13q12, 6q21, and 1p34.3–p36.1, respectively. Chromosome identification was carried out using QFH banding by simultaneous Hoechst 33258 staining (not shown).

before the evolutionary divergence of rodents and primates. It is therefore not surprising that they reside on different chromosomes, as is the case with most closely related G protein-coupled receptor genes other than the olfactory receptors.

Fluorescence *in situ* hybridization was performed following the method of Tory *et al.* (9). Chromosomes were identified with QFH banding by simultaneously staining with Hoechst 33258 (Fig. 2). After hybridization with the GPR12 cosmid, a total of 57 metaphase cells were examined, with 22 of these cells exhibiting paired hybridization signals and an additional 11 cells showing one hybridization signal at 13q12. For the GPR6 cosmid probe 65 metaphase cells were examined, with 20 showing paired hybridization signals and 19 having one hybridization signal at 6q21. The GPR3 cosmid hybridization was examined in 90 metaphase cells with 18 showing paired signals and 9 showing one signal at 1p34.3–p36.1. No significant background was noted at any other chromosomal location with any clone. It should be noted that two recent reports of *in situ* hybridization by GPR3 localized it to 1p35–p36.1 and 1p34.3 (5, 6).

REFERENCES

1. Bonner, T. I., Buckley, N. J., Young, A. C., and Brann, M. R. (1987). Identification of a family of muscarinic acetylcholine receptor genes. *Science* **237:** 527–532.
2. Coughlin, S. R. (1994). Expanding horizons for receptors coupled to G proteins: Diversity and disease. *Curr. Opin. Cell Biol.* **6:** 191–197.
3. Eidne, K. A., Zabavnik, J., Peters, T., Yoshida, S., Anderson, L., and Taylor, P. L. (1991). Cloning, sequencing and tissue distribution of a candidate G protein-coupled receptor from rat pituitary gland. *FEBS Lett.* **292:** 243–248.
4. Gilman, A. G. (1987). G proteins: Transducers of receptor-generated signals. *Annu. Rev. Biochem.* **56:** 615–649.
5. Iismaa, T. P., Kiefer, J., Liu, M. L., Baker, E., Sutherland, G. R., and Shine, J. (1994). Isolation and chromosomal localization of a novel human G-protein-coupled receptor (GPR3) expressed predominantly in the central nervous system. *Genomics* **24:** 391–394.
6. Marchese, A., Docherty, J. M., Nguyen, T., Heiber, M., Cheng, R., Heng, H. H. Q., Tsui, L.-C., Shi, X., George, S. R., and O'Dowd, B. F. (1994). Cloning of human genes encoding novel G protein-coupled receptors. *Genomics* **23:** 609–618.
7. Saeki, Y., Ueno, S., Mizuno, R., Nishimura, T., Fujimura, H., Nagai, Y., and Yanagihara, T. (1993). Molecular cloning of a novel putative G protein-coupled receptor (GPCR21) which is expressed predominantly in mouse central nervous system. *FEBS Lett.* **336:** 317–322.
8. Song, Z. H., Young, W. S., III, Brownstein, M. J., and Bonner, T. I. (1994). Molecular cloning of a novel candidate G protein-coupled receptor from rat brain. *FEBS Lett.* **351:** 375–379.
9. Tory, K., Latif, F., Modi, W., Schmidt, L., Wei, M. H., Li, H., Cobler, P., Orcutt, M. L., Delisio, J., Geil, L., Zbar, B., and Lerman, M. I. (1992). A genetic linkage map of 96 loci on the short arm of human chromosome 3. *Genomics* **13:** 275–286.