FEBS Letters 450 (1999) 126–130 FEBS 21935

A novel 45 kDa secretory protein from rat olfactory epithelium: primary structure and localisation

Maria I. Merkulova[a], Svetlana G. Andreeva[a], Tatyana M. Shuvaeva[a], Sergey V. Novoselov[b], Igor V. Peshenko[b], Marina F. Bystrova[b], Vladimir I. Novoselov[b], Evgeny E. Fesenko[b], Valery M. Lipkin[a,*]

[a] *Shemyakin and Ovchinnikov Institute of Bioorganic Chemistry, Russian Academy of Sciences, Miklukho-Maklaya St. 16/10, 117871 GSP, Moscow V-437, Russia*
[b] *Institute of Cell Biophysics, Russian Academy of Sciences, Pushchino 142292, Russia*

Received 15 March 1999

Abstract cDNA clones encoding the 45 kDa protein were isolated from a rat olfactory epithelium cDNA library and their inserts were sequenced. The reconstructed protein sequence comprises 400 amino acids with a calculated molecular mass of 46 026 Da. A homology was revealed between the amino acid sequence of the 45 kDa protein and the proteins involved in the transfer of hydrophobic ligands. Using in situ hybridisation, the 45 kDa protein mRNA expression was detected in the layer of supportive cells of olfactory epithelium, apical region of trachea, surface layer of the ciliated bronchial epithelium in lung and in skin epidermis.

Key words: 45 kDa secretory protein; cDNA cloning; Olfactory epithelium; Transfer protein

1. Introduction

The olfactory mucus contains a variety of water-soluble proteins, some of which are present at high levels. Among these proteins, the functional role for odour-binding proteins has been established [1]. Other proteins remain poorly investigated. We have previously isolated a major water-soluble protein from rat olfactory epithelium with a relative molecular mass of 45 kDa based on SDS-gel electrophoresis [2,3]. Apart from its recognition by an antiserum directed against the synthetic peptide corresponding to the α-$_{common}$ sequence of the known G-proteins [4], we have not found any other evidence indicating that the 45 kDa protein belongs to G-protein family [2,3]. Using immunohistochemistry methods, this protein has been localised mainly to olfactory mucus and apical regions of the olfactory epithelium cells. In this study, we performed cDNA cloning and molecular characterisation of the 45 kDa protein and revealed its homologues in databases.

2. Materials and methods

2.1. Isolation and partial sequencing of the 45 kDa protein

The 45 kDa protein was isolated as described [2]. 5 nmol of the purified 45 kDa protein was denatured in 6 M guanidinium-HCl and reduced with high purity 2-mercaptoethanol. Re-distilled 4-vinylpyridine was then added. The modified 45 kDa protein was cleaved with lysyl endopeptidase and/or protease from *Bacillus intermedius* [5]. The resultant peptides were separated by reverse-phase HPLC and sequenced by automatic Edman degradation using an Applied Biosystem 470A gas-phase sequenator.

2.2. Isolation of the 45 kDa protein cDNA

A rat olfactory epithelium cDNA library in bacteriophage λNM 1149 was a kind gift of Prof. H. Breer (Iniversitat Hohenheim, Stuttgart, Germany). cDNA was cloned into the *Eco*RI site of vector λNM 1149. The λ phage DNA was transfected into *Escherichia coli* C 600 Hfl. The recombinant bacteriophages were identified by in situ plaque hybridisation [6] as modified in [7]. The phage DNA was isolated from individual clones using the 'Wizard Lambda Preps DNA Purification System' kit (Promega, USA) in accordance with the manufacturer's recommendations. cDNA was sequenced on both strands by the Sanger method [8] using a 'Sequence version 2,0' kit (USB, USA) following the protocol of the manufacturer. PCR was performed using a *Tag* DNA polymerase kit (Perkin Elmer-Cetus, USA). PCR products were cloned in vector pGEM using the 'pGEM-T Easy Vector Systems' kit (Promega, USA). PCR primers were as follows: I, 5′-GAGGGIAA$^{T}/_{C}$TGGAA$^{A}/_{G}$GA; II, 5′-GATCAGGTIAA$^{A}/_{G}$AC$^{G}/_{A}/_{T}/_{C}$CA$^{A}/_{G}$TA; III, 5′-TTCGTATTGIGT$^{T}/_{C}$TT$^{G}/_{A}/_{T}/_{C}$AC$^{T}/_{C}$TG; IV, 5′-GTATTGTTGCATICC$^{T}/_{C}$TC$^{A}/_{G}$TC.

Other manipulations with DNA were carried out by the conventional methods [8].

2.3. In situ hybridisation

The PCR fragment A was cloned into the pGEM-T Easy Vector for in vitro transcription. This construction was linearised by *Sal*I for preparing a high specific activity antisense RNA probe or *Nco*I for preparing a sense RNA probe by means of a RNA labelling kit (Amersham) and [^{33}P]UTP (>1000 Ci/mmol) according to the supplier's conditions. Unincorporated nucleotides were removed by ethanol precipitation. The RNA probe was resuspended in hybridisation buffer containing 50% formamide, 10% dextran sulphate, 1×Denhardt's solution (0.02% Ficoll, 0.02% polyvinylpyrrolidone, 10 mg/ml bovine serum albumin), 0.3 M NaCl, 20 mM Tris-HCl, pH 7.4, 5 mM EDTA, 10 mM sodium phosphate, pH 8.0 and 50 μg/ml of baker's yeast tRNA.

Rats were perfused with 100 ml of buffered saline (20 mM phosphate buffer, pH 7.2, 140 mM NaCl) followed by 80–100 ml of 4% paraformaldehyde in phosphate buffer. After dissection, tissues from various organs were dehydrated with ethanol and immersed in Histoplast-S (Serva). Sections, 5–6 μm thick, were cut on microtome and mounted on slides. Sections were then processed through deparaffinisation and rehydration and were subsequently treated with 0.2 M HCl for 20 min, 2×SSC for 10 min, 1×SSC for 30 min at 50°C. Sections were fixed in 4% formaldehyde for 20 min and then acetylated with acetic anhydride (25 mM) and triethanolamine (10 mM) solution for 10 min at room temperature. Control sections were additionally treated with RNase (20 μg/ml in 10 mM Tris-HCl, 0.5 NaCl, pH 8,0). For in situ hybridisation, ^{33}P-labelled antisense and sense probes in hybridisation buffer were applied on sections that were then covered with coverslips and sealed with rubber cement. After 12 h hybridisation at 50°C, sections were washed three times with 4×SSC for 5 min, treated with RNase (20 μg/ml) for 20 min at 34°C and washed subsequently with 2×SSC for 30 min, 0.1×SSC for 45 min at 45°C. Finally, sections were dehydrated in ethanol series. Slides were im-

*Corresponding author. Fax: (7) (095) 310 7007.
E-mail: v.m.lipkin@ibch.siobc.ras.ru

PII: S0014-5793(99)00470-6

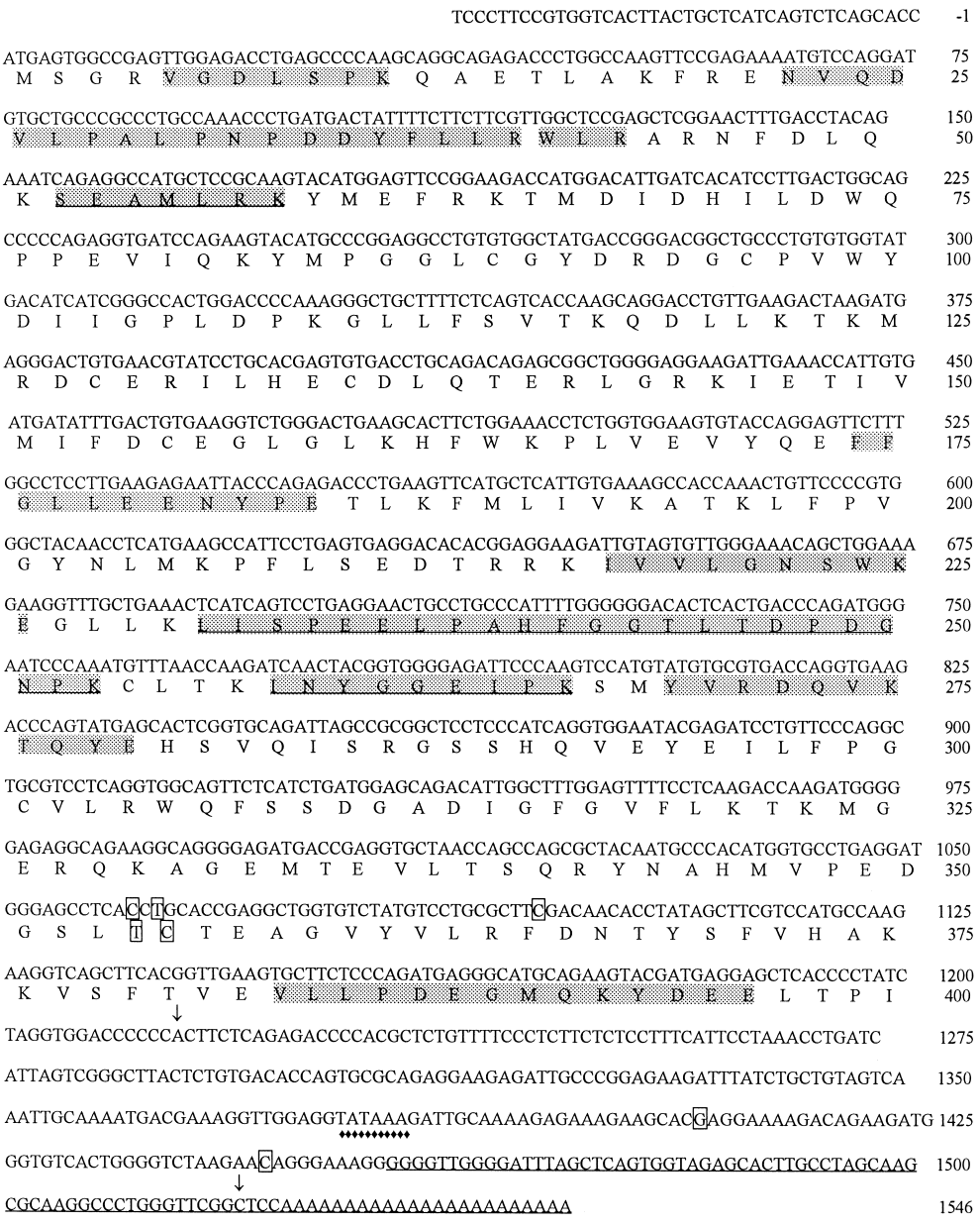

Fig. 1. The nucleotide sequence of cDNA of the 45 kDa protein from rat olfactory epithelium and the deduced amino acid sequence. Peptides obtained by digestion with *B. intermedius* are shaped and those digested with lysyl endopeptidase are shaped and underlined. The polyadenylation signal is underlined by rhombuses. The ID element located in the 3′-untranslated region is underlined. The mismatches to PCR fragment C and cDNA of clone λr22.1 are boxed. Arrows indicate the positions of the insertions found in cDNA of clone λr22.1. The nucleotide sequence presented here is registered in the EMBL and GenBank databases under the accession number AJ132352.

mersed in a liquid photographic emulsion and exposed for 21–28 days at 4°C. Slides were then developed and examined under a microscope.

2.4. Sequence analysis

DNA and protein databases were searched at the NCBI, using the BLAST network service. The alignment was generated using Multalin version 5.3.3 software [9].

3. Results and discussion

A total of 10 distinct peptides, three of which were recovered from the lysyl endopeptidase and seven from *B. intermedius* protease digestions of the 45 kDa protein, were determined by a combination of mass spectrometry and gas-phase Edman degradation, providing the partial sequence information throughout the protein (Fig. 1). The N-terminus of this protein appeared to be blocked, as repeated attempts to sequence the N-terminus failed so far. The nature of this block remains unknown. Degenerate oligonucleotides were designed based on the sequenced peptides and used as PCR primers to isolate the 45 kDa protein cDNA fragment from a rat olfactory epithelium cDNA library. Three PCR fragments (Fig. 2) were amplified and sequenced. The identity of these PCR fragments to the 45 kDa protein was confirmed by comparing their deduced amino acid sequences with peptide sequences. The rat olfactory epithelium cDNA library of 1×10^5 clones was screened using PCR fragment C as a probe and five positive cDNA clones (λr2.2, λr4.1, λr17.1, λr22.1, λr35.2) were isolated and sequenced (Fig. 2). The sequences

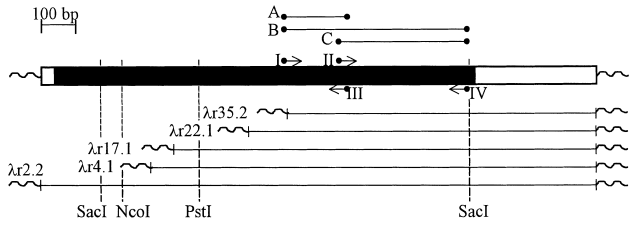

Fig. 2. Location of the isolated cDNA clones (λr2.2, λr4.1, λr17.1, λr22.1, λr35.2) and PCR fragments (A, B, C) on the restriction map of the 45 kDa protein cDNA. A solid box denotes an encoding sequence and open boxes represent 5′- and 3′-untranslated regions of cDNA. Arrows indicate the orientations of the oligonucleotide primers used for the PCR amplification of cDNA fragments.

of cDNA inserts of clones λr2.2, λr4.1, λr17.1, λr35.2 and PCR fragments A and B were identical and differed from that of clone λr22.1 in the encoding (the silent nucleotide replacement of C_{1095} for T) and 3′-untranslated regions (C insertion between C_{1214} and A_{1215}, the replacement of G_{1407} for A, C_{1447} for T, the insertion of the fragment TCCTCAG between G_{1519} and C_{1520}). We also found single base substitutions in the sequence of PCR fragment C which would result in amino acid substitutions at positions 354 (Thr(ACC) to Ile(ATC)) and 355 (Cys(TGC) to Arg(CGC)). The revealed mismatches may be attributed to genetic polymorphism. As shown in Fig. 1, the sequence of the 45 kDa protein cDNA contains an open reading frame, which begins with the initiation codon ATG at nucleotide positions 1–3 and is terminated by a TAG codon at positions 1201–1203. The potential polyadenylation signal is found at positions 1378–1383. We also revealed the presence of an ID element in the 3′-untranslated region of the 45 kDa protein cDNA sequence. ID elements are a major family of short interspersed repetitive elements found in the rodent genome (from a few 100 copies per haploid genome in guinea pigs to 130 000 copies per haploid genome in rats) [10]. It has been suggested that the ID element can act as an enhancer of RNA polymerase II transcription, although its precise function is not yet clear [11].

We deduced a primary structure of 400 amino acids from the cDNA sequence and confirmed that it contained all sequenced peptides (Fig. 1). The relative molecular mass of the predicted protein is 46 026 Da. The protein is characterised by a high content of negatively-charged residues, its calculated isoelectric point being 5.732.

Database searches revealed that the deduced protein sequence is homologous to some known and hypothetical proteins (Fig. 3). The 45 kDa protein shows the highest homology to human hypothetical protein AC004997 identified in PAC clone DJ130H16 (76.5% identity) and the retinal-binding

```
45 kD       1  MSGRVGDL---SPKQAETLAKFRENVQD----------VLPALPNPDDYFLLRWLRARNFDLQKSEAMLRKYMEFR-KTMDIDHI
AC004997    1  MSGRVGDL---SPRQKEALAKFRENVQD----------VLPALPNPDDYFLLRWLRARSFDLQKSEAMLRKHVEFR-KQKDIDNI
RALBP                                                                    MSLR-EQMGADTL
SEC14      22  LPGTPGNL---DSAQEKALAELRKLLED----------A-GFIERLDDSTLLRFLRARKFDVQLAKEMFENCEKWR-KDYGTDTI
α-TTP      10  VGKQLNEQPDHSPLVQPGLAELRRRAQEEG---------VPETPQPLTDAFLLRFLRARDFDLDLAWRLMKNYYKWRAECPELSAD
CRALBP     45  LQKAKDELNEKEETREEAVRELQELVQAQAASGEELALAVAERVQARDSAFLLRFIRARKFDVGRAYELLKGYVNFRLQYPELFDS

45 kD      72  L-D--WQPPE-VIQKYMPGGLCGYDRDGCPVWYDIIGPLDPKGLLFSVTKQDLLKTKMRDCERILH-ECDLQTERLGRKIETIVMI
AC004997   72  I-S--WQPPE-VIQQYLSGGMCGYDLDGCPVWYDIIGPLDAKGLLFSASKQDLLRTKMRECELLLQ-ECAHQTTKLGRKVETITII
RALBP      13  IAE--YTPPD-VIQKFMTGGDVGHDKDGSVLRIEPWGYLDMKGIMYSCKKSDLEKSKLLQCEKHLK-DLEAQSEKVGKPCTGLTVV
SEC14      92  LQDFHYDEKP-LIAKFYFQYYHKTDKDGRPVYFEELGAVNLHEMNKVTSEERMLKNLVWEYESVVQYRLPACSRAAGHLVETSCTI
α-TTP      87  L-H--PRSILGLLKAGYHGVLRSRDPTGSRVLIYRISYWDPK----VFTAYDVFRVSLITSELIVQ-E--VETQRNGVKA-----I
CRALBP    131  L-S--MEALRCTIEAGYFGVLSSRDKYGRVVMLFNIENWHCE----EVTFDEILQAYCFILEKLLE-N--EETQINGFCI-----V

45 kD      153 FDCEGLGLKHFWKPLVEVYQEFFGLLEENYPETLKFMLIVKATKLFPVGYNLMKPFLSEDTRRKIVVLGNSWKEGLLKLISPEELP
AC004997   153 YDCEGLGLKHLWKPAVEAYGEFLCMFEENYPETLKRLFVVKAPKLFPVAYNLIKPFLSEDTRKKIMVLGANWKEVLLKHISPDQVP
RALBP       95 FDMENVGSKHMWKPGLDMYLYLVQVLEDNYPEMMKRLFVINAPTLFPVLYKLVKPLLSEDMKNKIFVLGGDYKDTLLEYIDAEELP
SEC14      177 MDLKGISISSAYS-VMSYVREASYISQNYYPERMGKFYIINAPFGFSTAFRLFKPFLDPVTVSKIFILGSSYQKELLKQIPAENLP
α-TTP      158 FDLEGWQISHAFQITPSVAKKIAAVVTDSFPLKVRGIHLINEPVIFHAVFSMIKPFLTEKIKGRIHLHGNNYKSSLLQHF-FDILP
CRALBP     202 ENFKGFTMQQAAGLRPSDLKKMVDMLQDSFPARFKAIHFIHQPWYFTTTYNVVKPFLKNKLLQRVFVHGDDL-DGFFQEIDENILP

αcommon peptide                          GAGESGKSTIVK-QMK

45 kD      239 AHFGGTLTDPDGNPKCLTKINYGGEIPKSMYVRDQVKTQYEHSVQISRGSSHQVEYEILFPGCVLRWQFSSDGADIGFGVFLKTKM
AC004997   239 VEYGGTMTDPDGNPKCKSKINYGGDIPRKYYVRDQVKQQYEHSVQISRGSSHQVEYEILFPGCVLRWQFMSDGADVGFGIFLKTKM
RALBP      181 AYLGGTKSE--GDEKCSELICHGGEVPKEFYLENTDDFETMETITVGSGDKIYYEYEIENENTYIKWEYKTEEHDIGFGLF-----
SEC14      262 VKFGGKSEV--DESKGGLYLSDIGPWRDPKYIGPEGEAPEAFSMK
α-TTP      243 LEYGGNESSMEDI-CQEWT-NF-IMKSEDYLSSISETIQ
CRALBP     287 ADFGGTLPKYDG--KVVAEQLF-GPRAEVENTAL

45 kD      325 GERQKAGEMTEVLTSQRYNAHMVPEDGSLTCTEAGVYVLRFDNTYSFVHAKKVSFTVEVLLPDEGMQKYDEELTPI
AC004997   325 GERQRAGEMTEVLPNQRYNSHLVPEDGTLTCSDPGIYVLRFDNTYSFIHAKKVNFTVEVLLPDKASEEKMKQLGAGTPK
RALBP      260 --RKNGDEWEEVVPIERTDCSIMTLDGSHKCKDPGTYALCFDNSFSMMTSKNVRYTAEVMDPEVDSEDINKMIKDSTWDELSAQIS
```

Fig. 3. Alignment of the amino acid sequence of the 45 kDa protein and other related sequences: AC004997 (human, GenBank accession number AC004997); RALBP (retinal-binding protein, Japanese flying squid, Swiss-Prot number P49193); SEC14 (SEC14 cytosolic factor, yeast, Swiss-Prot number P24280); α-TTP (α-tocopherol transfer protein, rat, Swiss-Prot number P41034) and CRALBP (cellular retinaldehyde-binding protein, mouse GenBank number AF084638). Numbers on the left refer to the first amino acid of the line. Shaded letters denote amino acid residues identical to the 45 kDa protein. Well-conserved residues are in bold. Gaps (−) were introduced to optimise alignments.

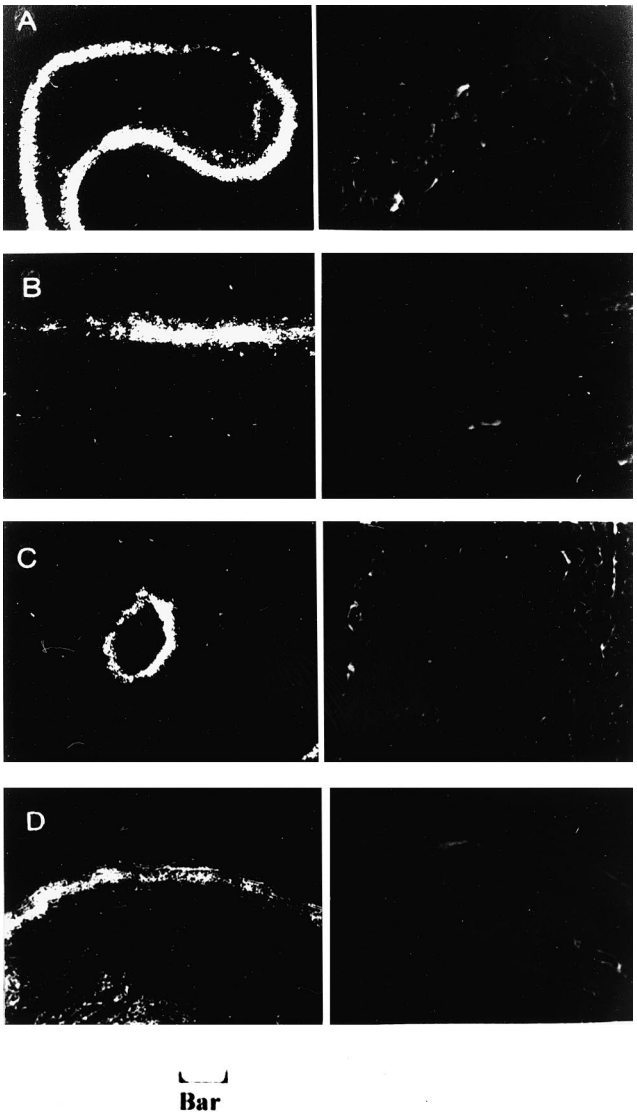

Fig. 4. Expression of the 45 kDa protein mRNA in various rat tissues revealed by in situ hybridisation with a ^{33}P-labelled antisense RNA probe. Dark field views of the sections of olfactory epithelium (A), trachea (B), lung (C) and skin (D). Specificity was determined by using a sense probe control to adjacent tissue sections (micrographs in right panel). Scale bar, 100 μm.

protein (RALBP) of the Japanese flying squid (38% identity between residues 60–387 of the 45 kDa protein and residues 1–327 of RALBP) [12]. The 45 kDa protein also shares a less significant although distinct homology with the yeast SEC14 cytosolic factor (also known as the phosphatidylcholine/phosphatidylinositol transfer protein) [13], the mammalian retinaldehyde-binding (CRALBP) and α-tocopherol transfer (α-TTP) proteins [14-16] and a non-catalytic domain of the MEG2 protein tyrosine phosphatase [17]. Most of these proteins have been shown to be capable of in vitro binding to hydrophobic ligands, suggesting their role in the transfer of the hydrophobic molecules [12–16]. The 45 kDa protein may have the similar function, as its homology to the afore-mentioned proteins extends across the entire primary structure, including the putative ligand-binding domain [13].

We have previously demonstrated that an antiserum directed against the synthetic peptide corresponding to the P-region of the α-subunit of G-protein (α-$_{common}$ sequence) was reactive with the 45 kDa protein [2,3]. This region, common not only in the G-protein α-subunit but also in some other GTP-binding proteins [18], has been suggested to be directly involved in the GTP hydrolytic process. The protein sequence analysis of the 45 kDa protein also revealed a region with a similarity to the α-$_{common}$ sequence (Fig. 3), which can account for both the ability of the 45 kDa protein to react with antiserum against the α-$_{common}$ peptide and to its GTP-binding. Nevertheless, these properties seem unlikely to be essential for the 45 kDa protein's functional role, as all its closest homologues lack the α-$_{common}$ sequence in their primary structure.

We examined the expression of the 45 kDa protein mRNA in various rat tissues by in situ hybridisation. Tissue sections were hybridised with ^{33}P-labelled antisense (Fig. 4, left panel) and sense (Fig. 4, right panel) RNA probes. In olfactory epithelium, the strong hybridisation signal was detected in the layer of supportive cells (Fig. 4A), consistent with the previous immunohistochemical data on the 45 kDa localisation in this tissue [3,19]. Expression of the 45 kDa protein mRNA was confined to the apical region of trachea (Fig. 4B) and to the surface layer of the ciliated bronchial epithelium in lung (Fig. 4C). In skin, the 45 kDa protein mRNA expression was detected in epidermis (Fig. 4D). Other rat tissues tested (liver, kidney and brain) showed a weak diffuse distribution of the hybridisation signal (data not shown). No signal was detected with the sense control probe. Thus, the 45 kDa protein mRNA is expressed mainly in tissues characterised by a high level of cell proliferation.

Acknowledgements: The authors thank Prof. H. Breer (Iniversitat Hohenheim, Stuttgart, Germany) for generously providing a rat olfactory epithelium cDNA library, N.S. Bystrova for the oligonucleotide synthesis. This study was supported by an International Grant from INTAS-RFBR (N 95-1111) and Grants from RFBR (N 96-04-48161, 99-04-48170).

References

[1] Pelosi, P. (1996) J. Neurobiol. 30, 3–19.
[2] Novoselov, V.I., Peshenko, I.V., Evdokimov, V.A., Nikolaev, Ju.V., Matveeva, E.V. and Fesenko, E.E. (1994) FEBS Lett. 353, 286–288.
[3] Novoselov, V.I., Peshenko, I.V., Evdokimov, V.A., Nikolaev, Ju.V., Matveeva, E.V. and Fesenko, E.E. (1996) Chem. Senses 21, 181–188.
[4] Mumby, S.M., Kahn, R.A., Manning, D.R. and Gilman, A.G. (1986) Proc. Natl. Acad. Sci. USA 83, 265–269.
[5] Leshchinskaya, I.B., Shakirov, E.V., Itskovitch, E.L., Balaban, N.P., Mardanova, A.M., Sharipova, M.R., Viryasov, M.B., Rudenskaya, G.N. and Stepanov, V.M. (1997) FEBS Lett. 404, 241–244.
[6] Woo, S.L.C. (1979) Methods Enzymol. 68, 389–395.
[7] Henikoff, S. (1984) Gene 28, 351–359.
[8] Maniatis, T., Fritsch, E.F. and Sambrook, J. (1982) Molecular Cloning: A Laboratory Manual, Cold Spring Harbor Laboratory, USA.
[9] Corpet, F. (1988) Nucleic Acids Res. 16, 10881–10890.
[10] Kim, J., Martignetti, J.A., Shen, M.R., Brosius, J. and Deininger, R. (1994) Proc. Natl. Acad. Sci. USA 91, 3607–3611.
[11] McKinnon, R.D., Shinnick, T.M. and Sutcliffe, J.G. (1986) Proc. Natl. Acad. Sci. USA 83, 3751–3755.
[12] Ozaki, K., Terakita, A., Ozaki, M., Hara, R. and Hara, T. (1994) J. Biol. Chem. 269, 3838–3845.
[13] Sha, B., Phillips, S.E., Bankaitis, V.A. and Luo, M. (1998) Nature 391, 506–510.

[14] Crabb, J.W., Goldflam, S., Harris, S.E. and Saari, J.C. (1988) J. Biol. Chem. 263, 18688–18692.
[15] Kennedy, B.N., Huang, J., Saari, J.C. and Crabb, J.W. (1998) Mol. Vis. 4, 14–21.
[16] Sato, Y., Arai, H., Miyata, A., Tokita, S., Yamamoto, K., Tanabe, T. and Inoue, K. (1993) J. Biol. Chem. 268, 17705–17710.
[17] Gu, M., Warshawsky, I. and Majerus, P.W. (1992) Proc. Natl. Acad. Sci. USA 89, 2980–2984.
[18] Kaziro, Y., Itoh, H., Kozasa, T., Nakafuku, M. and Satoh, T. (1991) Annu. Rev. Biochem. 60, 349–400.
[19] Novoselov, V.I., Peshenko, I.V., Evdokimov, V.A., Nikolaev, J,V., Popov, V.I., Matveeva, E.V. and Fesenko, E.E. (1995) Sens. Syst. 9, 85–95.