A New Member of the Glutamine-rich Protein Gene Family Is Characterized by the Absence of Internal Repeats and the Androgen Control of Its Expression in the Submandibular Gland of Rats*

(Received for publication, November 6, 1989)

Isabelle Rosinski-Chupin‡ and François Rougeon‡§
From the U2Unité de Génétique et Biochimie du Développement, Laboratoire associé Centre National de la Recherche Scientifique 361, Institut Pasteur, 75724 Paris Cedex 15, France

A cDNA, corresponding to a rat submandibular mRNA which is accumulated at a 20-fold higher level in males than females, has been isolated. The predicted protein, SMR2, has a calculated molecular mass of 15.4 kDa and is rich in glutamine/glutamic acid, proline, and asparagine/aspartic acid, a characteristic of the so-called salivary glutamine-rich proteins (GRPs) of the submandibular gland of rats. Nucleotide sequence comparisons indeed revealed strong similarities between the sequences of the SMR2 mRNA and that of GRPs, except in the region encoding the carboxy-terminal part of the proteins. In particular, the SMR2 mRNA contains the 5'-untranslated region and the signal peptide region shared by both groups of GRPs and proline-rich proteins (PRPs). A major difference is that, in SMR2, the peptide motif which is repeated four or five times in GRPs, is only found once. The SMR2 gene is about 3.0 kilobases in length and contains 4 exons. The second intron, which does not exist in characterized GRP genes, splits the "transition" region which separates the repetitive sequences from the signal peptide. This structure is reminiscent of that found in most PRP genes, strengthening the hypothesis that GRP and PRP genes have the same ancestral origin.

Two different functions are frequently assigned to the submandibular gland (SMG) of rodents. One is an exocrine function consisting in the constitution of the salivary fluids. The other is an endocrine function which leads to the release of certain growth factors and hormones.

Exocrine secretions mainly involve the acinar cells of the SMG. Among the secreted proteins are families of tissue specific proteins characterized by highly repetitive contiguous peptide sequences. According to their predominant amino acids, these proteins have been classified into proline-rich proteins (PRPs) or glutamine-rich proteins (GRPs). Their role in the salivary fluids has not been elucidated. On account of their high affinity for calcium phosphate (1, 2), it has been proposed that they could be involved in the protective proteinaceous structure of teeth surfaces. In addition, a role in the detoxification of certain substances, such as tannins, has been postulated for PRPs (3, 4).

The cDNAs corresponding to several mouse (5), rat (5, 6), and human (7) PRPs and rat GRPs (2, 8) have been cloned, and their sequences have been determined. In addition, the structure of some PRP and GRP genes (8-12) has been established. The peptide sequence of GRPs and PRPs can similarly be divided into four regions: a signal peptide, a "transition" region (which separates the repetitive region from the signal peptide), a repetitive region, and a carboxyl-terminal region. The organization of GRP and PRP genes is very similar and, in particular, the sequence of the first exon (corresponding to 5'-untranslated region and signal peptide) is highly conserved among these genes (8, 9, 11, 12). This suggests that GRP and PRP genes may derive from a common ancestor.

PRPs are encoded by a multigenic family mapped on chromosome 12 in man (13). In mouse, the PRP genes were firstly assigned to chromosome 8 (14) on the basis of results with mouse × hamster somatic cell hybrids but new linkage data indicate that they are on chromosome 6 (15). Evolutionary models for this gene family include a series of internal duplications of a 42-bp unit (9). Diversity would have been generated by recruitment or deletion of three bases from the ancestral unit during the duplication events, leading to a final length which varies between 42 and 69 bp. Finally, gene conversion would have homogenized the divergence between the internal repeats.

GRP genes differ from PRP genes, in particular, by the length (69 bp) and the sequence of the repeats. They are also part of a multigenic family, more than 10 GRP genes have been detected by Southern blot analysis in rats (8). The sequences of the two characterized GRP mRNAs are identical except for the number of repetitive motifs and the carboxyl-terminal part of the proteins, probably due to recent gene conversion events (8).

We are interested in the androgen regulation of genes expressed in the SMG of rodents. A number of growth factors, hormones, and other proteins with biologically defined properties are synthesized in large amounts in the SMG of rodents under androgen control (16, 17). It is the case, for instance, for renin, epidermal growth factor, and nerve growth factor in the SMG of mice. The role of these peptides in the saliva is unclear. Curiously, the pattern of proteins expressed at a higher level in the SMG of males than females appears to be...
species-specific. In an attempt to characterize some of the peptides whose expression is regulated by androgens in the SMG of rats, we have compared the patterns of in vitro translation products directed by the SMG mRNAs prepared from males and females (17). We have shown that several polypeptides are translated in higher amounts from male than female SMG mRNAs. One of them, SMR1, was shown to have the structure of a hormonal precursor, which could potentially give rise to a thyrotropin releasing hormone-like peptide after processing (18).

Here we report the characterization of another SMG mRNA, accumulated under androgen control. This mRNA encodes a protein, SMR2, which is related to the GRPs but contains only once the repetitive unit present in GRPs. The structure of the corresponding gene has been studied and is very similar to that of GRP except for the presence of an additional intron in the transition region coding sequence, reminiscent of that found in most PRP genes.

EXPERIMENTAL PROCEDURES AND RESULTS

DISCUSSION

In this paper, we report the characterization of a gene which is expressed under androgen control in the SMG of rats. The product of this gene, SMR2, belongs to the family of salivary glutamine-rich proteins. The nucleotideic sequence of the SMR2 mRNA is about 75% homologous to that of GRP mRNAs (except in the regions encoding the carboxy-terminal part of the proteins). A similar intron-exon structure is found in GRP and SMR2 genes.

Like the GRPs, SMR2 contains a relatively high proportion of glutamine + glutamic acid (20%), proline (12%), and asparagine + aspartic acid (13%). In addition, SMR2 and GRPs share a certain number of structural properties. They display an overall negative charge and a similar distribution of charges along the sequence with an excess of negative charges in the central part of the proteins and an excess of positive charges in the carboxyl-terminal region. Analysis of SMRP primary structure by the method of Hopp and Woods (32) (data not shown and Ref. 2) reveals that they are hydrophilic, except in the amino-terminal (signal peptide) and in the carboxy-terminal regions which are more hydrophobic. Curiously, both GRPs and SMR2 have the same aberrant behavior on NaDodSO4-PAGE. Their predicted secondary structure is reminiscent of the structure of a hormonal precursor, which could represent a species-specific difference in PRP gene structure.

Curiously, in the region where intron II occurs in SMR2 gene, GRP mRNAs contain a 15-bp sequence which does not align with the SMR2 mRNA sequence (see dot matrix of homology on Fig. 4). In an attempt to optimize these alignments, we found that the sequence present in GRP mRNA and absent in SMR2 mRNA is highly homologous to the 3′ end of the intron. As shown on Fig. 4, one possibility is that this sequence was originally a part of the exonic sequence. We propose that, after insertion of intron II into a PRP-like gene, this sequence has been released inside intron II by the use of a new more 3′ splicing site, leading to the present structure of the SMR2 gene. Since the sequences of PRP genes have too much diverged in this region, they cannot have the structure of a hormonal precursor, which could represent a species-specific difference in PRP gene structure.
be used to verify the validity of such a model. One should therefore also consider the possibility that this region of the ancestral gene had the same structure as in the SMR2 gene and that the GRP structure was created by the imprecise excision of the intron II.

Exons I of SMR2 and GRP genes, are 92% homologous and are also highly homologous to that of PRP genes (Fig. 8). The potential significance of this surprising conservation of exon I in PRPs has been discussed by Ann et al. (9). This sequence could be critical for the synthesis and/or secretion of these proteins.

It has been proposed that the transition region and the carboxyl-terminal region of the PRPs could have emerged, like the repeats, during the internal duplication events (9). A surprising finding is that SMRP mRNA is about 75% homologous in terms of nucleotidic sequences. Heinrich and Habener (8) have therefore proposed that the two genes and that the GRP structure was created by the imprecise excision of the intron II.

References

1. Bennick, A. (1982) Mol. Cell. Biochem. 45, 83-99
2. Mirels, L., Bedi, G. S., Dickinson, D. P., Gross, K. W., and Tabak, L. A. (1987) J. Biol. Chem. 262, 7289-7297
3. Mehansho, H., Hagerman, A., Clements, S., Butler, L., Rogler, J., and Carlson, D. M. (1983) Proc. Natl. Acad. Sci. U. S. A. 80, 3948-3952
4. Mehansho, H., Ann, D. K., Butler, L. G., Rogler, J., and Carlson, D. M. (1987) J. Biol. Chem. 262, 12344-12350
5. Clements, S., Mehansho, H., and Carlson, D. M. (1985) J. Biol. Chem. 260, 13471-13477
6. Ziemer, M. A., Swain, W. F., Rutter, W. J., Clements, S., Ann, D. K., and Carlson, D. M. (1984) J. Biol. Chem. 259, 10475-10480
7. Maeda, N., Kim, H. S., Ann, E. A., and Smithies, O. (1985) J. Biol. Chem. 260, 11123-11130
8. Heinrich, G., and Habener, J. F. (1987) J. Biol. Chem. 262, 5202-5207
9. Ann, D. K., and Carlson, D. M. (1985) J. Biol. Chem. 260, 15683-15687
10. Kim, H. S., and Maeda, N. (1986) J. Biol. Chem. 261, 6712-6718
11. Ann, D. K., Smith, M. K., and Carlson, D. M. (1988) J. Biol. Chem. 263, 10887-10893
12. Ann, D. K., Gadbois, D., and Carlson, D. M. (1987) J. Biol. Chem. 262, 3958-3963
13. Ann, E., Goodman, P. A., and Lalley, P. A. (1985) Am. J. Hum. Genet. 37, 418-424
14. Ann, P. A., Carlson, D. M., Clements, S., Lalley, P. A., and Vanin, E. (1984) Science 226, 967-969
15. Lush, I. A., and Holland, G. (1988) Genet. Res. 52, 207-212
16. Banks, T. (1980) J. Histochem. Cytochem. 28, 836-850
17. Greski, E. W. (1980) J. Histochem. Cytochem. 28, 860-870
18. Rosinski-Chupin, I., Tronik, D., and Rougeon, F. (1988) Proc. Natl. Acad. Sci. U. S. A. 85, 8653-8657
19. Tronik, D., Dreyfus, M., Babinet, C., and Rougeon, F. (1987) EMBO J. 6, 983-987
20. Pelham, H. R. B., and Jackson, R. J. (1976) Eur. J. Biochem. 67, 247-256
21. Paterson, B. M., Roberts, B. E., and Kuff, E. L. (1977) Proc. Natl. Acad. Sci. U. S. A. 74, 4370-4374
22. Sanger, F., Nicklen, S., and Coulson, A. R. (1977) Proc. Natl. Acad. Sci. U. S. A. 74, 5463-5467
23. Claverie, J. M., and Bricault, L. (1986) Proteins Struct. Func. Genet. 1, 60-65
24. Lipman, D. J., and Pearson, W. R. (1985) Science 227, 1435-1440
25. Grossberger, D. (1987) Nucleic Acids Res. 15, 6937
26. Maniatis, T., Fritsch, E. F., and Sambrook, J. (1982) in Molecular Cloning: A Laboratory Manual, pp. 77-83, Cold Spring Harbor Laboratory, Cold Spring Harbor, NY
27. Thomas, P. (1980) Proc. Natl. Acad. Sci. U. S. A. 77, 5201-5205
28. Proudfoot, N. J., and Brownlee, G. G. (1976) Nature 263, 211-214
29. Mount, S. M. (1982) Nucleic Acids Res. 10, 459-472
30. Brentnach, R., and Chabom, P. (1981) Annu. Rev. Biochem. 50, 349-363
31. McLaughlan, J., Gaffney, D., Whitson, J. L., and Clements, J. B. (1985) Nucleic Acids Res. 13, 1347-1368
32. Hopp, T. P., and Woods, K. R. (1981) Proc. Natl. Acad. Sci. U. S. A. 78, 9234-9239
33. Troxler, R. F., Belford, H. S., and Oppenheim, F. G. (1983) Biochem. Biophys. Res. Commun. 111, 239-246
34. Pachmunay, H., and Timpl, R. (1971) Annu. Rev. Biochem. 41, 510-516
35. Benedum, U., Baeverle, P. A., Konecki, D. S., Frank, R., Powell, J., Mallet, J., and Huttner, W. B. (1986) EMBO J. 5, 1495-1502
36. Brawerman, G. (1987) Cell 48, 5-6

Continued on next page.
Supplemental Material in:
A new member of the glutamine-rich protein family is characterized by the absence of internal repeats and the undercontrol of its expression in the submandibular gland of rats
Isabelle Ronchi-Chapron and François Roingeon

EXPERIMENTAL PROCEDURES

Animals and hormonal treatments. Ten week-old male and female Wistar rats were purchased from Iffa Credo (Lyon, France). Androgens were withdrawn by castration and, where indicated, 35 mg of testosterone (Sorbo) was injected subcutaneously 10 days later. Where indicated the same dose of testosterone was administered to female rats. Rats were killed one week after androgen treatment.

RNA extraction and in vitro translation. RNA was prepared from rat tissues as described (19). In vitro translation of RNAs was performed with the mRNA-dependent reticulocyte lysate translation system (20). The products were analyzed by denaturing NaDodSO4-polyacrylamide gel electrophoresis.

Cloning and characterization of the SMR2 cDNA. The submandibular gland cDNA library was constructed and screened as described in (18). The recombinant clones were identified by DNA-mRNA hybridizations in cell-free translation experiments (21). A cDNA insert was subcloned in M13 mp9 vector and sequenced by the dideoxy-fragment chain-termination method (22). The sequence was compared homologies between SMR2 and the promoter of the P53/CRP library (22) or between SMR2 mRNA and the nucleoic sequences of the EMRL data bank was done with the FASTP and the FASTN program according to Lipman and Pearson (23).

Construction and screening of rat genomic library. DNA of high molecular weight was prepared from rat spleen, partially digested with Sau3AI and loaded on a 5-25% NaCt gradient in buffer 10 mM Tris-HCl, 3 mM EDTA, pH 8. After 4 h 30 of centrifugation in a SW 41 (Beckman) rotor at 37 000 rpm, the gradients were collected and fractions were analyzed by electrophoresis in a 0.5% agarose gel.

Approximately 106 recombinant phages were screened by hybridization with a 32P-labeled cDNA insert after transfered on nitrocellulose filter. Minipreparations of DNA from positive clones were sequenced according to Crossbergcr (25) and analyzed after Southern blotting. Androgens were withdrawn by castration and, where indicated, 35 mg of desoxycorticosterone (Sovrandyl Retard, Hoechst-Roussel) was injected s'cubane 10 days later. Where indicated the same dose of testosterone was administered to female rats. Rats were killed one week after androgen treatment.

RESULTS

Isolation of cDNA complementary to a mRNA accumulated in the submandibular gland of rats under androgen control. To isolate mRNAs accumulated under androgen control in the submandibular gland of rats, cDNA library screening was performed and sequenced as described (18). The positive clones were characterized by DNA-mRNA hybridizations in cell-free translation experiments (19). As shown on figure 1, one of the cDNA inserts abolished the in vivo synthesis of a polyprotein with an apparent Mr 25 kDa on NaDodSO4 PAGE. SMR2.

Northern blot analysis (figure 2 A) using the full length cDNA insert as a probe revealed that it corresponds to a short mRNA of about 600 nucleotides, accumulated in higher amounts in the SMG of male rats than in female rats. In addition, there was some cross-hybridization with a mRNA of 950 nucleotides present at the same level in males and females. This cross-hybridization can been almost completely abolished by use of a shorter probe corresponding to the 3' end of the cDNA insert, indicating that only the 600 nucleotides mRNA corresponds to SMR2. Since SMR2 and CRP mRNAs have highly homologous 5' ends (see below), the 950 nucleotides mRNA probably corresponds to one of the CRP mRNAs.

In order to verify whether SMR2 mRNA accumulation in male rats was due to a regulation by androgens, castration experiments and androgen treatment of castrated males and of females were performed as described in the experimental procedures. As shown in figure 2 B, SMR2 mRNA accumulates to about 20-fold higher levels in male than in female rats. The amount of SMR2 mRNA in males is reduced about 10-fold 20 days after castration. Transcriptional treatment of these castrated males or of females induces SMR2 mRNA accumulation to a level similar to that of males.

The tissue distribution of the full-length SMR2 cDNA was investigated by Northern blot analysis. Among the different rat organs studied: SMG, prostate, seminal vesicles, kidney, liver, lung, SMR2 mRNA was detected only in the SMG (not shown).

Fig. 1. Hybrid arrest in vitro translation analysis of SMR2 cDNA clones. Five micrograms of total RNA prepared from SMG of male rats (lane A, C, D), female rats (lane B) or testosterone treated male rats (lane B) was translated in a reticulocyte cell free system either directly (lanes A and D) or after libylation with 1000 units of purified SMR2 cDNA insert 30 (lane C) or 600 nanograms of pUC9 DNA (lane D). The in vitro translation products were electrophoresed in a NaDodSO4 12.5% polyacrylamide gel and autoradiographed.

Northern blot analysis (figure 2 A) using the full length cDNA insert as a probe revealed that it corresponds to a short mRNA of about 600 nucleotides, accumulated in higher amounts in the SMG of male rats than in female rats. In addition, there was some cross-hybridization with a mRNA of 950 nucleotides present at the same level in males and females. This cross-hybridization can been almost completely abolished by use of a shorter probe corresponding to the 3' end of the cDNA insert, indicating that only the 600 nucleotides mRNA corresponds to SMR2. Since SMR2 and CRP mRNAs have highly homologous 5' ends (see below), the 950 nucleotides mRNA probably corresponds to one of the CRP mRNAs.

In order to verify whether SMR2 mRNA accumulation in male rats was due to a regulation by androgens, castration experiments and androgen treatment of castrated males and of females were performed as described in the experimental procedures. As shown in figure 2 B, SMR2 mRNA accumulates to about 20-fold higher levels in male than in female rats. The amount of SMR2 mRNA in males is reduced about 10-fold 20 days after castration. Transcriptional treatment of these castrated males or of females induces SMR2 mRNA accumulation to a level similar to that of males.

The tissue distribution of the full-length SMR2 cDNA was investigated by Northern blot analysis. Among the different rat organs studied: SMG, prostate, seminal vesicles, kidney, liver, lung, SMR2 mRNA was detected only in the SMG (not shown).
Evolutionary Aspects of the Glutamine-rich Protein Family

Computer search for sequence homologies in the data banks revealed a significant degree of homology between SMR2 and mRNAs encoding the glutamine-rich proteins (GRPs) of rats. These proteins are characterized by the presence of contiguous identical peptides, rich in glutamine/aspartic acid. Figure 4A shows a dot matrix of homology between SMR2 and CRP2, the longer GRP mRNA sequence reported by Henrich and Habener (8). Parallel diagonals are indicative of the presence of the internal repeat in the sequences of GRPs.

By analogy with GRPs, the sequence of SMR2 mRNA can be divided into five regions (Figures 3 and 4). The first 70 nucleotides, corresponding to the 5' untranslated region and a highly hydrophobic segment characteristic of signal peptides for secretion are strongly conserved (92% homology) between SMR2 and CRP2. The next segment, 141 bp long in CRP2 and 134 bp long in SMR2 defines the "transition" region which displays approximately 75% nucleotide sequence conservation between both mRNAs. The different length of the two sequences in this region is due to a 15 bp gap in the SMR2 sequence. The third segment corresponds to the region of contiguous repeats in GRPs. Each repeated motif is 69 bp long and the number of repeats is variable in the two characterized GRP mRNAs. In SMR2 mRNA, the motif is only found once (Figure 4B) and is 91% homologous to that of GRPs in nucleotide sequence, but only 71% homologous with respect to amino-acid sequence. Regions corresponding to the carboxy-terminal part of the proteins are highly divergent both in terms of nucleotide and peptide sequences. This region is also the most divergent between both characterized GRPs, GRP2 and CRP2 (8). Surprisingly, a high degree of similarity is again found in the 3' untranslated regions of SMR2 and GRP sequences.

![Fig. 4.](image_url)

**Fig. 4.** (A) Homology matrix comparison of GRP (CRP2) and SMR2 mRNA sequences. Similarity of comparison is the ratio of required matches to the length of comparison. Different regions defined in each mRNA sequence are indicated: 1: 3' untranslated region and signal peptide; 2: transition region; 3: contiguous repeat region; 4: carboxy-terminal region; 5: 3' untranslated region.

![Fig. 5.](image_url)

**Fig. 5.** (A) Nucleotide sequence of SMR2 gene. The exons are underlined by wavy lines. Splice sites are indicated by triangles. Exon transcription initiation sites are marked by an asterisk. Polyadenylation signal and YGTGTTYY resembling sequence are underlined by thin bars. The translation initiation codon is boxed. (B) Restriction map of the 4.5 kb integrated region and restriction enzymes are: B: BglII, BgII, BgIII; K: KpnI; E: EcoRI, EcoRV, EcoRV, PstI; H: HindIII.

![Fig. 8.](image_url)

**Fig. 8.** Nucleotide homologies at the 5' end of different PRP, CRP2, and SMR2 genes. Nucleotide sequences of different PRP cDNAs from human (cP1, cP2), mouse (cP3, cP4, cP5). cDNA and rat (pRP2, pRP3, pRP4) were aligned in homologous sequences. The ATG initiation codon is underlined and the initiation codons are aligned with the A designated as nucleotide 1.
A new member of the glutamine-rich protein gene family is characterized by the absence of internal repeats and the androgen control of its expression in the submandibular gland of rats.

I Rosinski-Chupin and F Rougeon

J. Biol. Chem. 1990, 265:10709-10713.

Access the most updated version of this article at http://www.jbc.org/content/265/18/10709

Alerts:
- When this article is cited
- When a correction for this article is posted

Click here to choose from all of JBC's e-mail alerts

This article cites 0 references, 0 of which can be accessed free at http://www.jbc.org/content/265/18/10709.full.html#ref-list-1