Chromosomal organization and expression analysis of two distinct genes encoding glutamine/glutamic acid-rich proteins

Kelly G. TEN HAGEN*†, Thomas M. BERES*, Josiane SZPIRER‡, Claude SZPIRER‡, and Lawrence A. TABAK*†‡§

Departments of * Dental Research and † Biochemistry, School of Medicine and Dentistry, University of Rochester, 601 Elmwood Avenue, Box 611, Rochester, NY 14642, U.S.A. and ‡ Laboratoire de Génétique, Université Libre de Bruxelles, Brussels, Belgium

GRP-Ca and GRP-Cb are two genes that encode glutamine/glutamic acid-rich proteins of the rat. These genes are very similar in structure and sequence, differing only within an approx. 90 bp segment of exon 3. We have used distinct oligonucleotide probes to unambiguously distinguish GRP-Ca and GRP-Cb gene expression. The two genes are expressed to relatively equivalent levels only in the submandibular gland. Chronic daily exposure to the β -adrenergic agonist, isoprenaline, resulted in a statistically significant decrease in GRP-Ca expression, with no effect on GRP-Cb, in contrast with previous reports. Furthermore

it was determined by PCR analysis of both submandibular-gland cDNA and genomic DNA that the GRP-Cb gene shows inter-animal variability in the number of 69 bp tandem repeats found within exon 3; GRP-Cb genes were shown to contain four, five or six repeats. GRP-Ca shows no such variability, containing only five tandem repeats in all animals examined. The two genes were localized to within 450 kb of one another at q43-q44 of rat chromosome 4 using somatic cell hybrid analysis, pulsed-field gel analysis and fluorescent *in situ* hybridization.

INTRODUCTION

Several families of proteins secreted by salivary exocrine glands are encoded by genes that contain tandem-repeat sequences. Expression of these tandem repeats often results in a repetitive amino acid sequence, which is important for the function of the protein. In the case of salivary mucins, these tandem-repeating amino acid sequences serve as the principal site for attachment of *O*-glycans [1]. The multiple tandem repeats observed for proline-rich proteins (PRPs) are thought to enhance PRP-mediated tannin neutralization [2]. Similarly, glutamine/glutamic acid-rich proteins (GRPs) may bind to tooth mineral via the highly charged tandem-repeat region of the protein [3]. Once bound to the tooth, GRPs appear to mediate the adhesion of specific micro-organisms to the tooth surface [4].

A subset of the GRPs secreted from the rat submandibular gland are encoded by two genes, GRP-Ca and GRP-Cb. Genomic clones encoding GRP-Ca [5] and cDNA clones encoding GRP-Ca and GRP-Cb [3,6,7] have been isolated and sequenced. These two genes were found to be very similar in sequence: they contain identical leader regions, transition regions, Gln/Glu-rich 69 bp repeat regions and 3' untranslated regions. Within the coding regions, they differ only by an approx. 90 bp segment of exon 3. In this report, we used distinct oligonucleotide probes within this 90 bp region of exon 3 to unambiguously distinguish GRP-Ca and GRP-Cb expression as well as to identify their positions within the rat genome. We found that GRP-Ca and GRP-Cb are expressed to relatively equivalent levels and both are restricted in expression to the rat submandibular gland. However, these two similarly expressed genes display dramatically different patterns of expression in the presence of the β -adrenergic agonist, isoprenaline. We also found that GRP-Cb has variability in the number of the tandem repeats found within exon 3. Through somatic cell hybrid analysis and fluorescent *in situ* hybridization (FISH), both genes were localized to the long arm of rat chromosome 4 (q43-q44). Pulsed-field gel analysis suggests that

the two genes may lie within 450 kb of one another on chromosome 4.

MATERIALS AND METHODS

Isoprenaline treatment

Male Wistar rats weighing 175 g were obtained from Charles River Laboratories. All animals were fed *ad libitum*. After 7 days of acclimation, rats were given daily intraperitoneal injections of either saline (140 mM NaCl) or ($\pm$)-isoprenaline hydrochloride (0.033 g/kg prepared in 140 mM NaCl immediately before use) for up to 48 h or 10 days as described previously [7]. After treatment, submandibular glands were removed and snap-frozen in liquid nitrogen for subsequent RNA isolation. RNA from five animals was used for each treatment. Dot-blots were hybridized sequentially with a GRP-Ca-specific oligonucleotide (*Ca90:30*), a GRP-Cb-specific oligonucleotide (*Cb95:30*) and an oligonucleotide that detects the 18 S ribosomal subunit (*18Santi*). GRP-Ca- and GRP-Cb-specific hybridizations were normalized to 18 S hybridization to control for sample-loading variations. The normalized densitometric values for the five individuals from each treatment were then averaged to give the values shown. *P* values stated in the text were calculated using a two-sided *t* test.

RNA isolation, Northern-blot and slot-blot analysis

RNA was isolated from rat tissues as described by Sambrook et al. [8]. It was electrophoresed on formaldehyde-agarose gels and transferred to either nitrocellulose (Schleicher and Schuell) or Hybond-N (Amersham) membranes according to the manufacturer's instructions. The agarose concentrations of each gel are listed in each accompanying Figure legend. The following oligonucleotides were end-labelled with [γ -³²P]ATP using T4 polynucleotide kinase (NEB) according to standard procedures

Abbreviations used: GRP, glutamine/glutamic acid-rich protein; FISH, fluorescence *in situ* hybridization; PRP, proline-rich protein; G3PDH, glyceraldehyde-3-phosphate dehydrogenase.

§ To whom correspondence should be addressed.

[8]: the GRP-Ca-specific oligonucleotide *Ca90:30* (AAGGAG-GGGCTTCTTTGGACGACCGCTGTC) (nt position 668–697) of GRP-Ca cDNA [7], the GRP-Cb-specific oligonucleotide *Cb95:30* (GCCTGAATACTACTACTGGTCTTCCTCTGT) (nt position 711–740) of GRP-Cb cDNA [7], the antisense 18 S ribosomal subunit oligonucleotide *18Santi* (TATTGGAGCTG-GAATTACCGCGGCTGCTGG) and *GRP3'exon3* (GAGGT-GACTCTACATTTCCTTCCTCAGGTT) (nt position 546–575) [7], which detects GRP-Ca and GRP-Cb equally. A fragment of the glyceraldehyde-3-phosphate dehydrogenase (G3PDH) gene was amplified using the G3PDH Control Amplifier Sets (Clontech) according to the manufacturer's instructions and labelled with [α - 32 P]dCTP using the Nick Translation System (Promega). Dot-blots were performed using a Minifold® Filtration Manifold (Schleicher and Schuell) according to the manufacturer's instructions. All hybridizations were performed in $5 \times$ SSPE/50% formamide at 42 °C with final washes for oligonucleotide probes in $2 \times$ SSC/0.1% SDS at 55 °C and final washes for all other probes in $1 \times$ SSC/0.1% SDS at 65 °C (where $1 \times$ SSPE is 0.15 M NaCl/10 mM sodium phosphate (pH 7.0)/1 mM EDTA and $1 \times$ SSC is 0.15 M NaCl/0.015 M sodium citrate). Densitometry was performed on a Molecular Dynamics PhosphorImager.

PCR analysis

Rat submandibular gland cDNA was prepared using Clontech's 1st-STRAND® cDNA Synthesis Kit according to the manufacturer's instructions; 1 μ g of rat submandibular-gland total RNA was used in 20 μ l of reaction mixture. The 20 μ l was then diluted to 100 μ l and 2.5 μ l of this diluted sample was then used in each PCR. Rat genomic DNA was prepared as described previously [9]; 100 ng of genomic DNA was used in each PCR. PCR was carried out using the GRP-Ca- and GRP-Cb-specific oligonucleotides described above as well as *GRP181:25* (TGCAGATGCAGAGAATGTTCAAGAG) (nt position 181–205) [7], a sense strand primer that lies 22 bp 5' of the beginning of the repeat region of exon 3 in both GRP-Ca and GRP-Cb, and *GRP3'int2* (GTGCTCAGATCAGGGAAAG-AAACAA) (nt position 1838–1862 of GRP-Ca genomic clone) [5], a sense strand primer specific for the 3' end of intron 2 in both GRP-Ca and GRP-Cb. For each PCR, the reaction mixture contained 0.125 μ l of AmpliTaq (Perkin-Elmer), 0.2 mM dNTPs, 1.0 μ M each oligonucleotide and 2.0 mM MgCl₂. PCR conditions of the *GRP3'int2* oligonucleotide consisted of 40 cycles of 94 °C for 45 s, 57 °C for 45 s and 72 °C for 2 min. Conditions for reacting with the *GRP181:25* oligonucleotide consisted of 30 cycles at 94 °C for 45 s, 67 °C for 45 s and 72 °C for 2 min. A hot start of 94 °C for 3 min followed by 75 °C for 4 min was used for every PCR. PCR products were run on 2.0% MetaPhor® agarose (FMC)/TBE gels at 4 °C at a rate of approx. 4V/cm.

Somatic cell hybrid analysis and FISH

The panel of mouse $\times$ rat cell hybrids segregating rat chromosomes has been described previously and used to map numerous rat genes [10,11]. FISH was carried out as described previously with chromosomes being identified by DAPI banding [11,12]. The probe used was a 2.1 kb *HhaI*–*HincII* genomic fragment from plasmid H7.5 [5], which detects both GRP-Ca and GRP-Cb equally. H7.5 is a plasmid that contains the entire coding region of GRP-Ca in pBluescript® II KS ($\pm$) (Stratagene).

Pulsed-field gel electrophoresis

Pulsed-field gel electrophoresis was performed as described previously [13]. Briefly, the DNA from rat submandibular-gland

cells (A5 cell line) suspended in agarose blocks was digested overnight at 37 °C with either *SacII* (*KspI*) or *NotI* (Boehringer-Mannheim) after being left at 4 °C for 20 min. Gels consisted of 1% agarose in $0.5 \times$ TBE and were run at 200 V for 15 h with a 60 s switching time and then for 9 h with a 90 s switching time on the CHEF-DR II apparatus (Bio-Rad) at 14 °C. DNA in the gels was acid-nicked in 0.5 M HCl for 30 min and transferred to nitrocellulose membranes after standard denaturation and neutralization procedures [8]. All hybridizations were carried out in $5 \times$ SSPE/50% formamide at 42 °C with final washes in $1 \times$ SSC/0.1% SDS at 65 °C. The 2.1 kb *HhaI*–*HincII* fragment from the plasmid H7.5 was labelled with 32 P using the Nick Translation System (Promega); this fragment contains approx. 200 bp of intron 2, all of exon 3 and intron 3 and 40 bp of exon 4 of the GRP-Ca gene and detects both GRP-Ca and GRP-Cb. The GRP-Cb gene was specifically detected using a PCR labelling technique described previously [14]. Briefly, the plasmid, pCb-96, was digested with *NsiI* to release an approx. 200 bp fragment containing the region of exon 3 unique to GRP-Cb. This fragment was then annealed with the antisense primer, *762-lower* (TCT-AAAATCTTCTGGCGA) (nt position 761–778 of GRP-Cb cDNA) [7] to synthesize a single-stranded probe with [α - 32 P]dCTP through 30 cycles of PCR under previously described conditions [14]. The specificity of this probe for the GRP-Cb gene was verified by Southern-blot analysis. Conventional Southern blots were prepared by digesting 5 μ g of rat genomic DNA with either *EcoRI* or *HindIII*, electrophoresing the DNA through a 0.7% agarose/TAE gel at 1.5 V/cm and transferring to Hybond-N membranes.

RESULTS

Expression of GRP-Ca and GRP-Cb

In previous reports, expression analysis of GRP-Ca and GRP-Cb were performed simultaneously using a probe that detects both gene transcripts as a single band on Northern blots. To examine the expression patterns of each gene individually, we used a region of exon 3 displaying sequence divergence between GRP-Ca and GRP-Cb to synthesize gene-specific oligonucleotides. These oligonucleotides were end-labelled with [γ - 32 P]ATP and used to probe Northern blots containing RNA from the following tissues of male Wistar rats: submandibular gland, sublingual gland, parotid gland, liver and kidney. When the GRP-Ca-specific oligonucleotide (*Ca90:30*) was used as a probe, a single transcript of approx 1000 nt was seen only in RNA from the submandibular gland (Figure 1A). No signal from any other tissue was visible even after prolonged exposure, including small intestine and stomach (results not shown). A similar tissue-specific expression pattern was seen after stripping and reprobing this Northern blot with the GRP-Cb-specific oligonucleotide (*Cb95:30*) (Figure 1B), except that two distinct transcripts of approx. 1100 and 900 nt were visible in the submandibular-gland RNA lane. Submandibular-gland-specific expression of both genes was also seen in female rats (results not shown). Northern blots were stripped and reprobed with a fragment of the rat G3PDH gene to control for the integrity of RNA in each lane (Figure 1C).

We wished to determine whether the multiple GRP-Cb transcripts seen in the previous Northern blot were consistently present in different animals or whether they represent some form of variability within this gene in a population. RNA was isolated from the submandibular glands of individual male Wistar rats, fractionated by gel electrophoresis and Northern blotted. Each lane represents the RNA isolated from the submandibular gland of an individual rat. On probing with the GRP-Cb-specific

Figure 1 Northern-blot analysis of GRP-Ca and GRP-Cb expression in rat tissues

Total RNA was extracted from parotid gland (P), sublingual gland (SL), submandibular gland (SM), liver (L) and kidney (K) of male Wistar rats. After electrophoresis on a formaldehyde–1.2% agarose gel and transfer to Hybond-N membranes (Amersham), the RNA was hybridized with a GRP-Ca-specific oligonucleotide (*Ca90:30*) (A), a GRP-Cb-specific oligonucleotide (*Cb95:30*) (B) and finally, the G3PDH gene (C) as a control for RNA integrity. Each lane contains 15 μ g of total RNA. Size markers are indicated on the right of each gel.

Figure 2 Northern-blot analysis of GRP-Cb and GRP-Ca transcript sizes in seven different animals

Submandibular-gland total RNA was extracted from seven individual male Wistar rats, electrophoresed on a formaldehyde–1.2% agarose gel, blotted and hybridized with *Cb95:30* (A) or *Ca90:30* (B). Each lane contains 15 μ g of total RNA from a single animal. Size markers are indicated on the right of each gel.

Figure 3 Comparison of Northern-blot analysis of expression levels of GRP-Ca and GRP-Cb

Total RNA from submandibular glands of three individual male Wistar rats (lanes 1, 2 and 3) was electrophoresed on a formaldehyde–1.6% agarose gel, blotted and hybridized with an oligonucleotide probe that detects both GRP-Ca and GRP-Cb (*GRP3' exon 3*). Each lane contains 15 µg of total RNA. Gene-specific transcripts and size markers are indicated on the right.

oligonucleotide, different-sized bands were seen for RNA derived from each individual rat, indicating that the GRP-Cb gene exhibits some form of variability within the population (Figure 2A). The same Northern blot stripped and reprobed with the GRP-Ca-specific oligonucleotide (Figure 2B) revealed a single GRP-Ca band of the same size across all lanes. In total, submandibular-gland RNA from 22 individual rats showed transcript size variability for the GRP-Cb gene but no sign of variability for the GRP-Ca gene.

We took advantage of the transcript size difference between GRP-Ca and GRP-Cb in certain rats to examine the expression levels of each gene relative to one another. Submandibular-gland RNA from three individual male Wistar rats, which were previously shown to have a GRP-Ca-specific transcript of 1000 nt and a GRP-Cb-specific transcript of 1100 nt, was electrophoresed on a formaldehyde–agarose gel and Northern blotted. This Northern blot was then hybridized in greater than fivefold molar excess with antisense oligonucleotide *GRP3' exon 3*, which detects both GRP-Ca and GRP-Cb. As seen in Figure 3, the expression levels of the two genes are relatively similar, with GRP-Cb being anywhere from 5 to 25% less than GRP-Ca in the individual animals examined.

Isoprenaline is a β -adrenergic agonist that has been shown to affect the expression of many genes that are responsive to cAMP. Previously in our laboratory it was suggested that GRP-Ca expression was decreased on exposure to isoprenaline, whereas GRP-Cb expression was increased. We repeated these experiments with the GRP-Ca- and GRP-Cb-specific oligonucleotides. Male Wistar rats were given daily intraperitoneal injections of

Table 1 GRP-Ca and GRP-Cb expression after isoprenaline treatment

Male Wistar rats were given daily injections of either saline (S) or isoprenaline (I) as described in the text. Submandibular-gland RNA was harvested after 48 h (48 h) or 10 days (10 d) of treatment and dot-blotted. Each value represents the mean $\pm$ S.D. of normalized densitometric values from five animals.

	Expression			
	48h-S	48h-I	10d-S	10d-I
GRP-Ca	2519 $\pm$ 1298	1146 $\pm$ 192	2081 $\pm$ 578	449 $\pm$ 236
GRP-Cb	655 $\pm$ 390	407 $\pm$ 69	613 $\pm$ 142	571 $\pm$ 232

A.

B.

C.

Figure 4 PCR analysis of GRP-Cb transcript size variation

(A) Schematic diagram of PCR strategy for each gene. 1, *GRP3' int2* primer; 2, *GRP181:25* primer; Ca, *Ca90:30* primer; Cb, *Cb95:30* primer; B, *Bsu36I*. [□], Region of tandem repeats; [▨], GRP-Ca-specific area; [▩], GRP-Cb-specific area. Only the 3' end of intron 2 is shown. (B) Agarose gel of PCR products from rat submandibular-gland cDNA amplifications. Lanes 1 and 10 are the 100 bp ladder (Gibco–BRL). Lanes 2–5, cDNA from rat 426; lanes 6–9, cDNA from rat 104. PCR mixtures and digests are as follows: 2 + Ca (lanes 2 and 6), 2 + Cb (lanes 3 and 7), 2 + Ca *Bsu36I* (lanes 4 and 8), 2 + Cb *Bsu36I* (lanes 5 and 9). (C) Agarose gel of PCR products from rat genomic DNA amplifications. Lanes 1, 6 and 11 are the 100 bp ladder. Lanes 2–5, genomic DNA from rat 426; lanes 7 and 8, cDNA from rat 426; lanes 9 and 10, no DNA. PCR mixtures and digests are as follows: 1 + Ca (lanes 2, 7 and 9), 1 + Cb (lanes 3, 8 and 10), 1 + Ca *Bsu36I* (lane 4), 1 + Cb *Bsu36I* (lane 5). Size markers are indicated to the left of each gel.

either saline or isoprenaline as described previously [7] for up to 10 days. Five animals were used for each time point and each type of treatment. Submandibular glands were removed and RNA was extracted after either 48 h or 10 days of treatment.

Table 2 Rat chromosome constitution of the rat × mouse hybrids and segregation of GRP-Ca and GRP-Cb

In the first column, + or — indicates the presence or absence of the rat GRP-Ca and GRP-Cb genes. The two genes, defined by *Hind*III restriction fragments of 8.0 and 6.5 kb respectively co-segregated together and with rat chromosome 4. In the remaining columns, + indicates that rat chromosomes are present in more than 55% of the metaphases, (+) indicates that rat chromosomes are present in 25–55% of the metaphases, (—) indicates that rat chromosomes are present in less than 25% of the metaphases and — indicates that rat chromosomes are absent. Independent hybrid clones are clones derived from distinct fusion events. In this Table, all clones are independent, except LB210-B and LB210-I. When a chromosome was present in less than 25% of the metaphases [(—)] the hybrid in question was not taken into account to establish the number of discordancies for that particular chromosome.

Hybrids	Presence of rat Ca/Cb genes	Presence of rat chromosomes																				
		X	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18	19	20
LB20	—	+	—	(+)	(+)	—	—	—	+	—	—	—	—	+	+	—	—	+	(+)	+	+	—
LB150-1	+	+	—	—	+	+	—	—	+	—	+	(+)	+	+	+	—	—	(+)	(+)	+	+	—
LB161	+	+	—	+	+	+	+	+	+	—	+	+	—	(+)	+	+	+	+	+	+	+	(+)
LB210-B	+	+	—	—	—	+	—	—	—	—	—	—	—	—	+	+	—	—	—	+	—	—
LB210-I	—	+	—	—	—	—	—	—	—	—	—	—	—	—	+	+	—	—	—	+	—	—
LB251	+	+	+	+	—	+	—	(+)	+	—	—	+	—	+	+	—	—	—	+	—	+	—
LB330TG3	+	—	—	—	—	+	—	—	—	—	—	—	—	+	—	—	—	—	—	—	—	—
LB510-6	+	+	—	+	+	+	—	—	+	—	—	—	—	+	+	+	+	+	+	+	—	—
LB600	+	+	+	+	+	+	+	(+)	+	—	(—)	+	+	+	+	+	+	+	—	+	+	—
LB630	+	+	(—)	—	+	+	(+)	+	+	—	+	—	+	+	+	(+)	+	+	—	+	+	(—)
LB780-6	—	+	—	+	—	—	—	—	—	—	—	+	+	—	+	—	—	—	—	+	—	—
LB810	+	+	—	+	+	+	—	+	+	+	—	+	+	+	+	+	+	+	+	—	+	(+)
LB860	+	+	—	+	+	+	—	—	+	—	+	—	+	+	+	—	+	+	+	+	—	(+)
LB1040TG1	+	—	—	—	—	+	—	+	—	—	—	—	+	(+)	—	—	+	+	—	+	—	—
<i>Independent discordant clones:</i>		5	8	7	5	<u>0</u>	8	5	4	10	6	7	6	2	5	6	4	4	6	6	6	7

Submandibular-gland RNA from each animal was then slot-blotted and probed sequentially with the GRP-Ca-specific oligonucleotide, the GRP-Cb-specific oligonucleotide and an antisense oligonucleotide specific for the 18 S ribosomal subunit. After densitometric scanning and normalization for RNA loading using the 18 S ribosomal signal, the values for all five animals at each time point and treatment were averaged and are shown in Table 1. As can be seen, a slight decrease in GRP-Ca expression occurred after 48 h of isoprenaline treatment relative to the saline control ($P = 0.048$). However, a much more pronounced decrease in GRP-Ca steady-state transcript levels was seen after 10 days of exposure to isoprenaline ($P = 0.0004$). When the same blot was stripped and reprobed with the GRP-Cb-specific oligonucleotide, no significant increase or decrease in GRP-Cb steady-state transcript levels was seen after 48 h ($P = 0.199$) or 10 days of exposure to isoprenaline ($P = 0.738$) relative to the saline controls (Table 1), in contrast with previous reports [7]. Therefore GRP-Ca steady-state transcript levels show a marked decrease after 10 days of exposure to isoprenaline whereas GRP-Cb steady-state transcript levels remain unchanged.

Variability in transcript size

We next wanted to investigate the nature of the transcript size variability of GRP-Cb, which we had detected on previous Northern blots. Given that the difference in size of the various transcripts was approximately the length of the repeating unit within exon 3 (~70 bp), we initially decided to focus on this portion of the coding region containing the tandem repeats to see if it was responsible for the size variability. First-strand cDNA was synthesized from rat 426, displaying GRP-Cb transcripts of ~900 nt and ~1100 nt (Figure 2A, lane 1) and a GRP-Ca transcript of ~1000 nt (Figure 2B, lane 1), and rat 104, displaying a single GRP-Cb band of ~1100 nt (Figure 2A, lane 2) and a GRP-Ca band of ~1000 bp (Figure 2B, lane 2). The strategy

and primers used in this study are shown in Figure 4(A). *GRP3'int2* is a sense strand primer specific for the 3' end of intron 2 in both GRP-Ca and GRP-Cb. *GRP181:25* is a sense strand primer that lies 22 bp 5' of the beginning of the repeat region of exon 3 in both GRP-Ca and GRP-Cb. *Ca90:30* is a GRP-Ca-specific antisense primer that lies 118 bp 3' of the end of the repeat region in exon 3 and *Cb95:30* is an antisense primer specific for GRP-Cb and lies 161 bp 3' of the end of the repeat region in exon 3. The first set of reactions was performed on the cDNA from rat 426 using primers *GRP181:25* and *Ca90:30* (to specifically amplify the repeat region of GRP-Ca) and primers *GRP181:25* and *Cb95:30* (to specifically amplify the repeat region of GRP-Cb) (Figure 4A). Half of each reaction mixture was then digested with *Bsu36I* to cut away the region 3' of the repeats so that the bands seen on the gel would represent the repeat region alone. The products of these reactions were electrophoresed on a 2% MetaPhor agarose/TBE gel (FMC) and are shown in Figure 4B (lanes 2–5). As seen in the uncut GRP-Ca-specific mixture (lane 2), a single band of 518 bp is present. In the uncut GRP-Cb-specific mixture (lane 3), two bands of 630 and 492 bp are seen. Once digested with *Bsu36I*, the GRP-Ca-specific mixture (lane 4) shows a band of 370 bp (representing a repeat region consisting of five repeating units) and a band of 148 bp (representing the region 3' to the repeats). The *Bsu36I*-digested GRP-Cb-specific mixture (lane 5) shows two bands of 439 and 301 bp (representing repeat regions consisting of six and four repeating units respectively) and a band of 191 bp (representing the region 3' to the repeats). Therefore these PCR experiments support the conclusion that the GRP-Cb transcript size variability seen in this rat is due to variable numbers of tandem repeats.

The next set of reactions was performed on the cDNA from rat 104, again using primers *GRP181:25* and *Ca90:30* (to specifically amplify GRP-Ca) and primers *GRP181:25* and *Cb95:30* (to specifically amplify GRP-Cb). These reaction mixtures were also

Figure 5 Regional localization of the rat GRP-Ca and GRP-Cb genes by FISH

Displayed is a portion of a metaphase showing the signal generated by the 2.1 kb *HhaI*–*HincII* probe. The diagram illustrates the position of the signals at the end of chromosome 4q.

digested with *Bsu36I*, electrophoresed and are seen in Figure 4(B) (lanes 6–9). Again, in the uncut GRP-Ca-specific mixture (lane 6), a single band of 518 bp is present, and, when digested with *Bsu36I*, the GRP-Ca-specific mixture (lane 8) shows a band of 370 bp (representing a repeat region consisting of five repeating units). In the uncut GRP-Cb-specific mixture (lane 7), one band of 630 bp is present. Once digested with *Bsu36I*, lane 9 shows a GRP-Cb-specific band of 439 bp (representing a repeat region consisting of six repeating units). Once again, these results are in agreement with the transcript sizes seen for this rat on previous Northern blots and support the notion that the transcript size variability seen for GRP-Cb is due to variable numbers of tandem repeats within this gene in different animals.

To demonstrate that this tandem-repeat variability is present

in the genome and is not due to a splicing variation, we also performed PCR analysis on the genomic DNA from rat 426, using, as the upstream primer, *GRP3'int2*, which lies at the 3' end of intron 2 in both GRP-Ca and GRP-Cb. As seen in Figure 4(C), the uncut GRP-Ca-specific reaction (lane 2) produces a band of 663 bp, whereas the *Bsu36I*-digested reaction mixture (lane 4) shows a band of 515 bp consistent with five repeating units within GRP-Ca. The lower band in the GRP-Ca digested mixture (lane 4) represents the region 3' of the repeats. The uncut GRP-Cb reaction mixture (lane 3) shows two bands, of 775 and 637 bp. When digested with *Bsu36I*, bands of 584 and 446 bp are present (representing six and four repeating units within the tandem repeat region respectively) and a band of 191 bp (representing the region 3' of the repeats). No bands are seen with either of these sets of primers when cDNA was used as a template (lanes 7 and 8). Negative controls for each reaction are shown in lanes 9 and 10. Therefore the variability seen in the number of tandem repeats within the GRP-Cb gene is present at the level of the genomic DNA and is transcribed to generate the transcript size variability seen on Northern blots.

Localization of GRP-Ca and GRP-Cb

The GRP-Ca and GRP-Cb genes were localized using mouse × rat cell hybrids which segregate rat chromosomes. The 2.1 kb *HhaI*–*HincII* probe from plasmid H7.5 [5] detected *HindIII* restriction fragments in rat DNA of 8.0 and 6.9 kb (see Figure 6B, Ca + Cb), corresponding to GRP-Ca and GRP-Cb respectively. Mouse DNA showed very faint signals on hybridization with this probe, allowing us to follow the rat fragments unambiguously in the hybrids. The two rat fragments co-segregated and the results were therefore combined (Table 2). Both GRP-Ca

Figure 6 Pulsed-field gel analysis of GRP-Ca and GRP-Cb

(A) High-molecular-mass rat genomic DNA was digested with either *SacII* (S) or *NotI* (N), electrophoresed on a CHEF apparatus (Bio-Rad), blotted and probed with either a GRP-Cb-specific probe (Cb) or a probe that detects both GRP-Ca and GRP-Cb (Ca + Cb). Each lane contains the DNA from approx. 10^6 cells. (B) Rat genomic DNA was digested with *EcoRI* (E) or *HindIII* (H), electrophoresed on conventional agarose gels, blotted and hybridized with the same probes as listed above. Each lane contains 5 μ g of genomic DNA. Size markers are indicated on the left of each gel.

and GRP-Cb co-segregated with rat chromosome 4; several discordant clones were counted for each of the other chromosomes (Table 2). Therefore GRP-Ca and GRP-Cb both reside on rat chromosome 4.

Regional chromosome localization of GRP-Ca and GRP-Cb was determined by FISH, again using the 2.1 kb *HhaI-HincII* probe, which detects both GRP-Ca and GRP-Cb. The probe generated bright double-chromatid signals on chromosome 4 only (chromosomes were identified and oriented by DAPI staining). The signals were subtelomeric, and unambiguously mapped at 4q43-q44 (Figure 5). The fact that the probe did not generate two distinct signals along each chromatid indicated that the two GRP genes do not map to cytogenetically distinguishable positions.

To attempt to characterize the distance between these two genes further, we obtained P1 clones for each gene (Genome Systems, Inc.) but were unable to find any clones that carried both genes on their ~80 kb inserts (results not shown). We then performed pulsed-field gel analysis of high-molecular-mass rat genomic DNA digested with rare cutting restriction enzymes to determine if the two genes could be localized to the same high-molecular-mass restriction fragments. Rat genomic DNA was digested with either *SacII* or *NotI*, electrophoresed in duplicate sets of lanes on a CHEF apparatus (Bio-Rad) and subsequently blotted to nitrocellulose membranes (Schleicher and Schuell). One set of lanes was probed with an asymmetric PCR probe specific for GRP-Cb, which was generated within the ~90 bp region of difference between the two genes. As seen in Figure 6(A) (Cb), *SacII* bands of 680 and 450 kb and a *NotI* band of 650 kb are detected by the GRP-Cb-specific probe. The two *SacII* fragments detected are probably the product of the two alleles for this locus. To verify the specificity of the GRP-Cb probe, a control Southern blot of rat genomic DNA cut with *EcoRI* and *HindIII* was also hybridized (Figure 6B, Cb). A GRP-Cb-specific *EcoRI* band of 6.1 kb and a *HindIII* band of 6.9 kb were the only fragments detected, indicating the gene specificity of this probe. On probing the other set of lanes with a probe that detects both GRP-Ca and GRP-Cb equally, *SacII* fragments of 680 and 450 kb and a *NotI* fragment of 650 kb were again detected (Figure 6A, Ca + Cb). To verify that both genes were being detected equally by this probe, a control Southern blot of rat genomic DNA was also hybridized. In the *HindIII* lane both the 6.9 kb GRP-Cb band and the 8.0 kb GRP-Ca band are seen at relatively equal intensities. GRP-Ca bands of 2.9 kb and 1.3 kb and a GRP-Cb band of 6.1 kb are seen in the *EcoRI*-digested lane as well. Therefore these genes may lie within the same 450 kb *SacII* fragment on chromosome 4.

DISCUSSION

Through the use of gene-specific oligonucleotides, we have unambiguously determined the expression pattern of the rat GRP-Ca and GRP-Cb genes. Both genes were found to be restricted in expression to the rat submandibular gland with no expression seen in liver, kidney, sublingual gland, parotid gland, small intestine or stomach. These results are in agreement with previous work using probes that detected both genes simultaneously in salivary-tissue RNA as well as immunohistochemical studies on salivary glands using antibodies specific for the protein products of these genes [3,6,15,16].

Although GRP-Ca and GRP-Cb display similarly restricted tissue-specific expression, their response to the β -adrenergic agonist isoprenaline is markedly different. Steady-state GRP-Ca transcript levels were drastically decreased after 10 days of chronic exposure, whereas steady-state levels of GRP-Cb tran-

scripts remained unchanged. Previous reports suggested that GRP-Cb expression increased after 10 days of exposure [7]. The reason for these contrasting results may be the animal-to-animal variation seen in the actual levels of both GRP-Ca and GRP-Cb, necessitating careful analysis to reach a statistically significant difference between the control and experimental groups. Potential cross-hybridization of GRP-Ca and GRP-Cb may have also contributed to the discrepancy.

During the course of these studies, we detected transcript size variability from animal to animal when probing specifically for GRP-Cb. Using primers that flank the tandem-repeat region of GRP-Ca and GRP-Cb based on previous sequencing, we were able to show that the transcript size variation is due to variable numbers of the repeating units within the tandem-repeat region of GRP-Cb; repeat numbers varied from animal to animal. No such variation was seen for GRP-Ca. This result is in part supported by the work of Heinrich and Habener [6], who sequenced a gene homologous to GRP-Cb that contained four tandem repeats, as opposed to the five tandem repeats found by Cooper et al. [7]. There are several examples of certain PRP genes that show little or no repeat length variation whereas others are highly variable [17]. Variable-number tandem-repeat polymorphisms have also been detected in a number of mucin genes [18]. The under-representation of repeat-length variants at certain loci could be due to selection against such variants or different rates of recombination within and around the repeat region of certain genes. We searched the existing sequences of both genes (2.7 kb of 5' region and all of the coding region) for the differential presence of Chi sequences which might suggest disparate propensities for recombination; the only Chi sequence found lies within exon 3 of both genes.

Although the actual function of these proteins remains unknown, it has been proposed that the highly charged Gln/Glu-rich repeat region may bind to the tooth surface, acting as a substrate for bacterial binding and colonization of the tooth. Increased numbers of repeats may therefore confer a stronger binding capacity and could potentially provide a selective advantage. However, this explanation for repeat number variation remains speculative until the actual role of these proteins is discovered.

Taking advantage of the transcript size variability between GRP-Ca and GRP-Cb, we directly compared their expression levels by Northern-blot analysis. We have found that GRP-Ca and GRP-Cb transcript levels are relatively similar to one another, with GRP-Cb being only 5–25% lower than GRP-Ca, varying with the individual animals examined.

Using somatic cell hybrid analysis and FISH, GRP-Ca and GRP-Cb were both localized to q43-44 of rat chromosome 4. We next turned to pulsed-field gel analysis to see if the two genes co-localized to the same high-molecular-mass restriction fragments. Both a GRP-Cb-specific probe and a probe that recognizes both genes detected 680 and 450 kb *SacII* fragments and a 650 kb *NotI* fragment. This result suggests that these two genes lie within the same genomic region encompassed by these restriction fragments. More specifically, they may lie within 450 kb of one another on the same *SacII* fragment of chromosome 4. This result is similar to the localization of many related gene families, including the human PRPs, where all family members lie within a 750 kb region on human chromosome 12 [19], the mouse PRPs, where members are clustered on chromosome 8 [20] and the VCS gene family, the members of which all lie within p21-22 of rat chromosome 14 [21].

Given the sequence similarity, in both coding and non-coding regions, between GRP-Ca and GRP-Cb and their relatively close proximity to one another in the genome, one probably arose as

a duplication of the other, as has been postulated for the members of other related gene families [19]. It has been suggested that genome diversity may be generated in part by gene duplication and subsequent divergence of sequences. There are many examples of gene duplications, where family members have evolved to display different developmental and tissue-specific regulation [22–24]. Such appears to be the case with GRP-Ca and GRP-Cb, given their proximity, sequence similarity and differential expression in the presence of isoprenaline. This scenario of gene duplication, recombination within tandem repeats and sequence divergence represents a potential method whereby gene products and their regulation may be subtly altered over time, resulting in genome expansion and expressional diversification.

This work was supported by grant DE 10490 to L. A. T., and also the FRSM (Belgium) and the Communauté Française de Belgique (ARC). C. S. is a research director of the FNRS (Belgium). We thank Marlene Balys, Michele Riviere and Pascale Vanvooren for excellent technical help.

REFERENCES

- 1 Tabak, L. A. (1995) *Annu. Rev. Physiol.* **57**, 547–564
- 2 Charlton, A. J., Baxter, N. J., Lilley, T. H., Haslam, E., McDonald, C. J. and Williamson, M. P. (1996) *FEBS Lett.* **382**, 289–292
- 3 Mirels, L., Bedi, G. S., Dickinson, D. P., Gross, K. W. and Tabak, L. A. (1987) *J. Biol. Chem.* **262**, 7289–7297
- 4 Kopec, L. and Bowen, W. H. (1995) *Caries Res.* **29**, 507–512
- 5 Cooper, L. F. and Tabak, L. A. (1991) *Gene* **106**, 261–266
- 6 Heinrich, G. and Habener, J. F. (1987) *J. Biol. Chem.* **262**, 5262–5270
- 7 Cooper, L. F., Elia, D. and Tabak, L. A. (1991) *J. Biol. Chem.* **266**, 3532–3539
- 8 Sambrook, J., Fritsch, E. F. and Maniatis, T. (1989) *Molecular Cloning: A Laboratory Manual*, Cold Spring Harbor University Press, Cold Spring Harbor, NY
- 9 Ten Hagen, K. G., Gilbert, D. M., Willard, H. F. and Cohen, S. N. (1990) *Mol. Cell. Biol.* **10**, 6348–6355
- 10 Szpirer, J., Levan, G., Thorn, M. and Szpirer, C. (1984) *Cytogenet. Cell Genet.* **38**, 142–149
- 11 Stephanova, E., Tissir, F., Dusetti, N., Iovanna, J., Szpirer, J. and Szpirer, C. (1996) *Cytogenet. Cell Genet.* **72**, 83–85
- 12 Pinkel, D., Landegent, J., Collins, C., Fuscoe, J., Segreaves, R., Lucas, J. and Gray, J. (1988) *Proc. Natl. Acad. Sci. U.S.A.* **85**, 9138–9141
- 13 Ten Hagen, K. G., Ravnán, J.-B. and Cohen, S. C. (1995) *J. Mol. Biol.* **254**, 119–129
- 14 Bednarczuk, T. A., Wiggins, R. C. and Konat, G. W. (1991) *Biotechniques* **10**, 478–479
- 15 Mirels, L., Kopec, L., Yagil, C., Dickinson, D. P., Dziejman, M. and Tabak, L. A. (1990) *Arch. Oral Biol.* **35**, 1–5
- 16 Moreira, J. E., Tabak, L. A., Bedi, G. S., Culp, D. J. and Hand, A. R. (1989) *J. Histochem. Cytochem.* **37**, 515–528
- 17 Lyons, K. M., Stein, J. H. and Smithies, O. (1988) *Genetics* **120**, 267–278
- 18 Van Klinken, B. J.-W., Dekker, J., Büller, H. A. and Einerhand, A. W. C. (1995) *Am. J. Physiol.* **269**, G613–G627
- 19 Kim, H., Smithies, O. and Maeda, N. (1990) *Genomics* **6**, 260–267
- 20 Ann, D. K., Smith, M. K. and Carlson, D. M. (1988) *J. Biol. Chem.* **263**, 10887–10893
- 21 Rosinski-Chupin, I., Kuramoto, T., Courty, Y., Rougeon, F. and Serikawa, T. (1995) *Mamm. Genome* **6**, 153–154
- 22 Layfield, R., Bannister, A. J., Pierce, E. J. and McDonald, C. J. (1992) *Eur. J. Biochem.* **204**, 591–597
- 23 Carlson, D. M., Zhou, J. and Wright, P. S. (1991) *Prog. Nucleic Acid Res. Mol. Biol.* **41**, 1–22
- 24 Dickinson, D. P., Abel, K., Near, J., Taylor, B. A. and Gross, K. W. (1989) *Biochem. Genet.* **27**, 613–637