Volume 314, number 3, 375–380 FEBS 11901 December 1992

The organisation of the mouse secretogranin II gene

Angelika Schimmel[a], Otto Bräunling[a], Ulrich Rüther[b,*], Wieland B. Huttner[a] and Hans-Hermann Gerdes[a]

[a]*Institute for Neurobiology, University of Heidelberg, Im Neuenheimer Feld 364, D-6900 Heidelberg, Germany* and [b]*European Molecular Biology Laboratory, Postfach 10.2209, D-6900 Heidelberg, Germany*

Received 27 October 1992

We have characterized the gene which encodes mouse secretogranin II (previously also referred to as chromogranin C), a tyrosine-sulfated secretory protein belonging to the granin (chromogranin/secretogranin) family which is found in secretory granules of most endocrine cells and neurons. The secretogranin II gene was found to contain 2 exons. In contrast to chromogranin A and chromogranin B, the two previously characterized granin genes, the entire secretogranin II protein is encoded by a single exon, exon 2, with exon 1 containing only a 5'-untranslated sequence. Consistent with previous data on the expression of secretogranin II, the putative promoter region was found to contain a cAMP-responsive element and a potential AP-1 binding site.

Granins; Secretogranin; Gene structure; Neurosecretory vesicle; Transcription factor; Evolution

1. INTRODUCTION

The granins (chromogranins/secretogranins) are a family of acidic secretory proteins found in secretory granules of a wide variety of endocrine cells and neurons (for review, see [1]). The selective neuroendocrine expression of these proteins has been documented, at both the protein and mRNA level, by numerous studies (for review, see [2]). However, little is known about the molecular mechanism underlying the regulation of transcription of the granin genes. A first step towards the analysis of this regulation is the cloning of these genes. Of the three classical members of the granin family, chromogranin A (CgA), chromogranin B (CgB, originally called secretogranin I [3]) and secretogranin II (SgII), only the genes for CgB [4] and CgA [5,6] have been cloned so far.

The elucidation of the primary structure of CgA [7,8], CgB [9] and SgII [10,11] has revealed that CgA and CgB are more closely related to each other than either protein is to SgII. The cloning of the genes for CgB [4] and CgA [5,6] showed that the structural homology between the CgA and CgB proteins is also reflected at the level of the gene in a very similar genomic organization. Given the structural differences between SgII and the two chromogranins, it is important to investigate the organization of the SgII gene and to compare it with the genes for CgA and CgB. We report here the cloning of the mouse SgII gene, and describe its promoter and exon–intron organization.

Correspondence address: H.-H. Gerdes, Institute for Neurobiology, Im Neuenheimer Feld 364, D-6900 Heidelberg, Germany.

**Present address:* Medizinische Hochschule Hannover, Institut für Molekularbiologie, Konstanty-Gutschow-Straße 8, D-3000 Hannover, Germany.

Abbreviations: Cg, chromogranin; Sg, secretogranin.

The sequences detailed in this paper have been submitted to the EMBL Data Library under the accession numbers X68837 (exon 1), and X68838 (exon 2).

2. MATERIALS AND METHODS

Standard molecular biology procedures [12] were used throughout.

2.1. *Cloning of the mouse SgII gene*

A genomic library of F_1(C57BL/6xSJL) mouse DNA was constructed using the lambda DASH vector (Stratagene). This library was screened with a full-length rat SgII cDNA, which extended ≈ 40 nucleotides beyond the 5'-end and ≈ 440 nucleotides beyond the 3'-end of the published cDNA sequence [10]. This cDNA had been released from its cloning vector by *Hin*dIII and *Eco*RI and, due to the presence of an internal *Eco*RI site, was used in the form of two DNA fragments of ≈ 0.6 kb (5') and ≈ 1.8 kb (3') size. Alternatively, the cDNA was released with *Xho*I, which yields a 5' ≈ 1.7 kb and a 3' ≈ 0.7 kb fragment. The insert of the genomic SgII clone was released from the lambda vector by either *Eco*RI or *Xba*I. The *Eco*RI- or the *Xba*I-released insert was subjected to restriction mapping using various enzymes. *Eco*RI, *Xba*I, and *Eco*RI–*Xba*I fragments were subcloned into Bluescript (Stratagene) or pSP73 (Promega) vectors. Southern analysis after digestion of the insert with *Eco*RI, *Xba*I or other enzymes was performed using, as probes, either the two fragments of the full-length rat SgII cDNA, the 5' ≈ 0.6 kb *Hin*dIII–*Eco*RI fragment of the rat SgII cDNA, the ≈ 0.7 kb *Xho*I fragment derived from the very 3'-end of the rat SgII cDNA, or selected *Eco*RI–*Xba*I fragments of the mouse SgII genomic DNA.

2.2. *Sequence analysis*

For sequencing, subcloned DNA fragments were subjected to deletion cloning when necessary. Double-stranded DNA sequencing was performed by the dideoxy-chain-termination method, using oligonu-

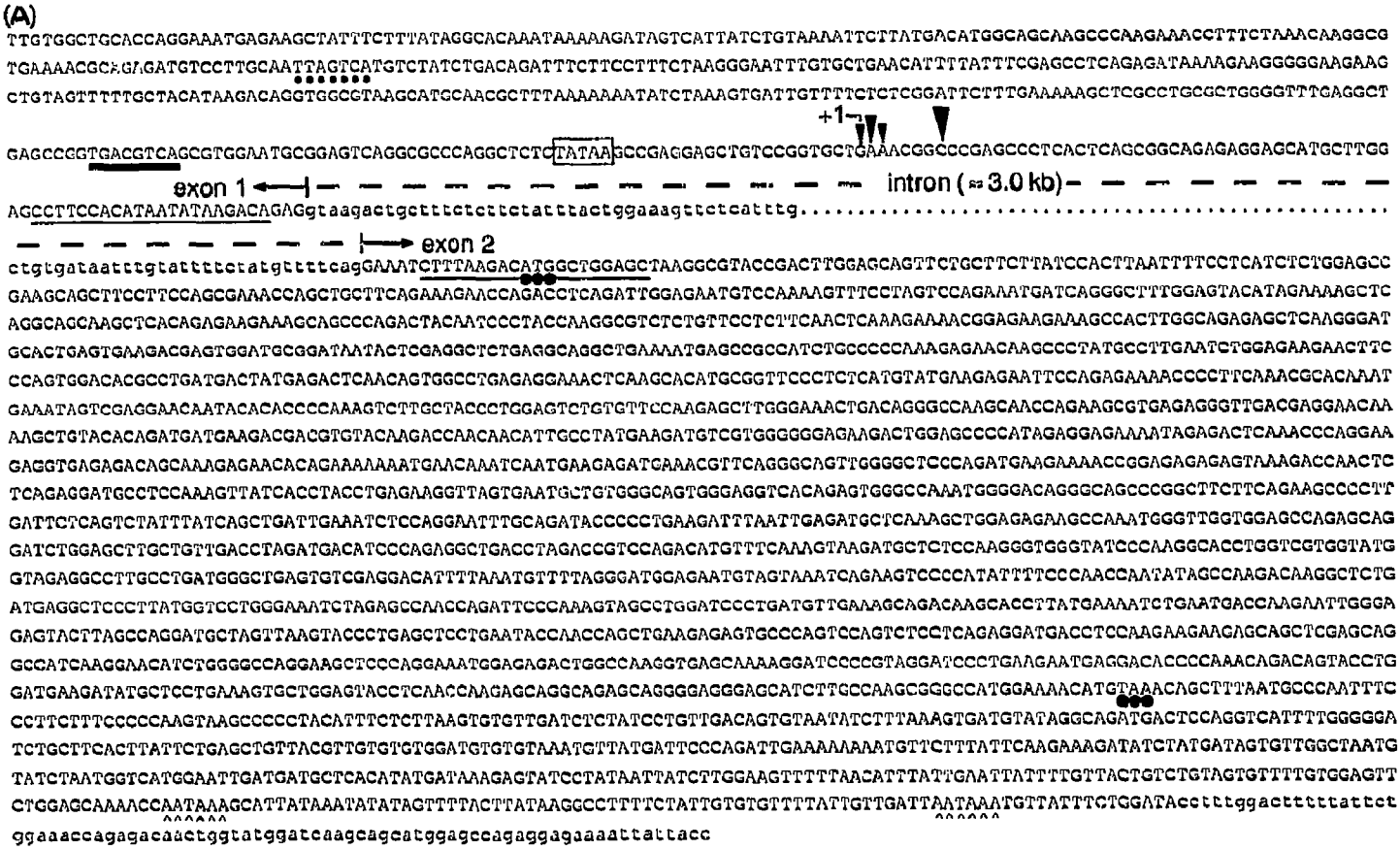

Fig. 1. The mouse SgII gene. (A) The DNA sequence shown includes: (i) the putative promotor region with a potential AP-1 binding site (dotted line), a cyclic AMP-responsive element (thick bar) and a TATA box; (ii) exon 1; the 5′-ends of the various primer extension products are indicated by the arrowheads (sizes reflect their abundance, see bottom panel); the thin bar indicates the position of the oligonucleotide P1 primer used for primer extension analysis; (iii) intron 1 (dashed line, lower case letters); (iv) exon 2 with the start and stop codons of translation (dots) and two polyadenylation signals (open triangles); the most 3′ capital 'A' corresponds to the beginning of the poly-A track of the rat mRNA; the thin bar indicates the position of the oligonucleotide P2 primer used for primer extension analysis. (B) Mapping of the SgII mRNA start sites by primer extension analysis using the P1 oligonucleotide. Lane 1: mouse liver RNA. Lane 2: AtT-20 RNA. Lanes GATC: an unrelated sequencing reaction used to calibrate single nucleotide steps. Arrowheads correspond to those in (A).

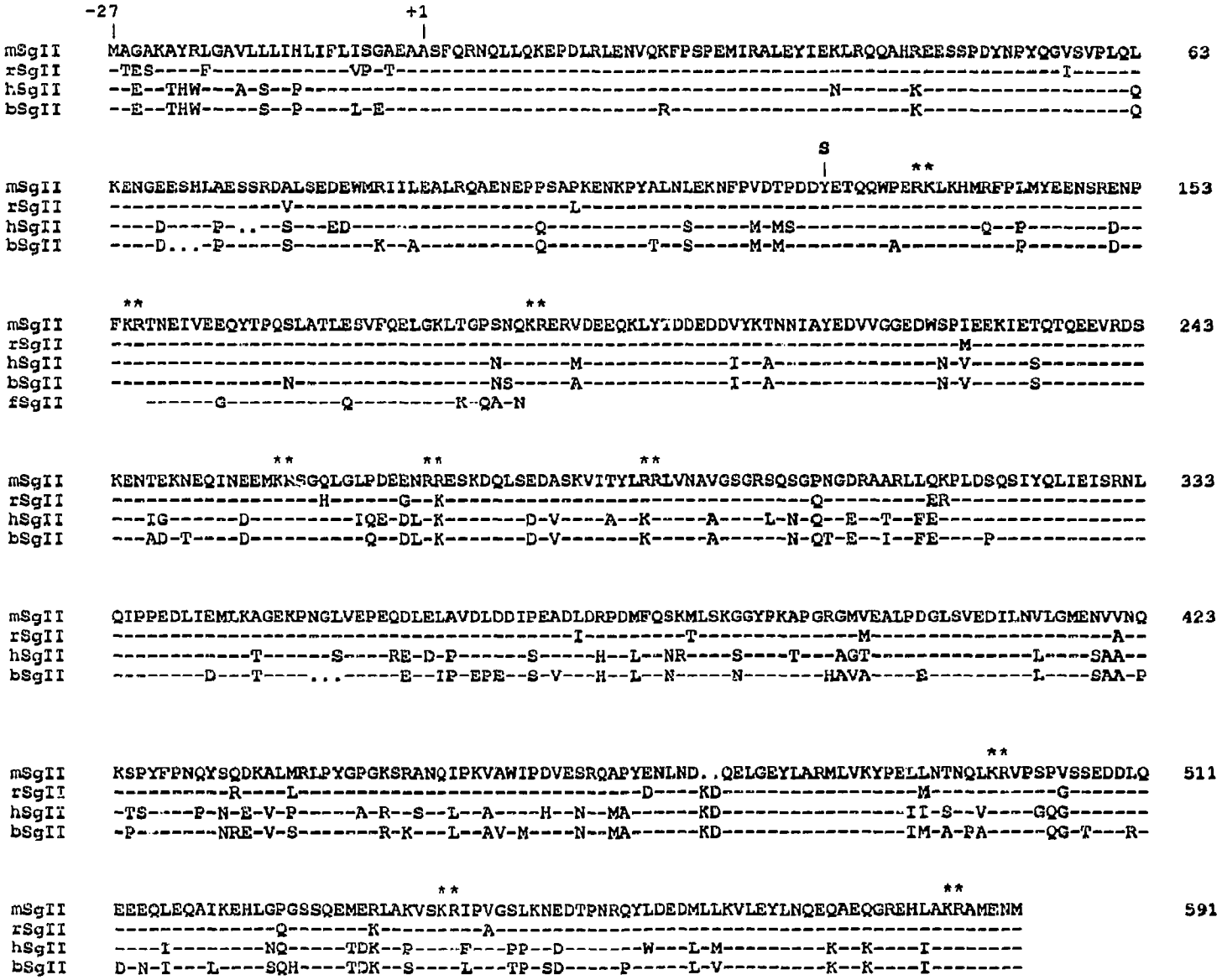

Fig. 2. Comparison of the amino acid sequences of SgII from mouse (mSgII), rat (rSgII) [10], human (hSgII) [11] and bovine (bSgII) [14], and of the SgII-derived frog peptide (fSgII) [17]. Gaps introduced for aligning the sequences are indicated by dots. Amino acid residues in rat, human, bovine and frog SgII that are identical to the corresponding residues in mouse SgII are shown as dashes. The number −27 indicates the beginning of the signal peptide, and +1 indicates the putative amino-terminal residue after signal peptide cleavage. The numbers on the right refer to this amino terminus. Double asterisks: dibasic sites; Bold S: putative tyrosine sulfation site.

cleotide primers corresponding either to vector sequences or to mouse genomic SgII sequences.

2.3. *Primer extension*

The 5′-end of the mouse SgII mRNA was determined by primer extension as described [13], using AMV reverse transcriptase at 42°C. A 21-mer oligonucleotide with a sequence complementary to part of exon 1 (referred to as P1 primer, see Fig. 1A) and a 20-mer oligonucleotide complementary to the 5′-end of exon 2 (referred to as P2 primer, see Fig. 1A) of the mouse SgII gene were used to prime the reaction on the RNA. Thirty μg of denatured RNA from the mouse neuroendocrine cell line AtT-20 or the rat neuroendocrine cell line PC12 and, as a negative control, 30 μg of denatured RNA from mouse liver were used per reaction.

3. RESULTS AND DISCUSSION

3.1. *Cloning and sequencing of the mouse SgII gene*

Screening of 1.8 million plaques of a mouse genomic lambda library with the full-length rat SgII cDNA [10] as probe led to the identification of one positive clone. Southern analysis of this clone and, for comparison, of mouse DNA, both digested with either *Eco*RI or *Xba*I and probed with the full-length rat SgII cDNA, showed the same pattern (data not shown). This suggested that the isolated clone contained the complete mouse SgII gene.

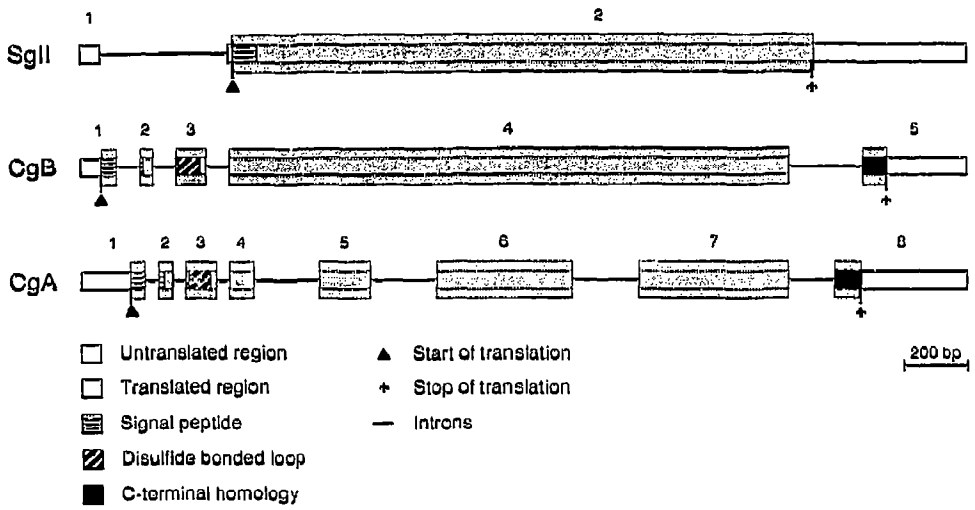

Fig. 3. Diagram comparing the structural organisation of the genes encoding mouse SgII, mouse CgB [4] and mouse CgA [6]. Exons are shown as numbered boxes and are drawn to scale. The introns (thin lines) are not drawn to scale.

*Eco*RI digestion of the genomic SgII lambda clone generated 4 fragments of ≈ 1.6 kb (referred to as E1), ≈ 2.4 kb (referred to as E2), ≈ 2.9 kb (referred to as E3) and ≈ 6.3 kb (referred to as E4). *Xba*I digestion generated 3 fragments of ≈ 2.7 kb (referred to as X1), ≈ 4.5 kb (referred to as X2) and ≈ 6.0 kb (referred to as X3). These fragments were subjected to Southern analysis using, as probes, either the full-length rat SgII cDNA, the 5′ ≈ 0.6 kb *Hin*dIII–*Eco*RI fragment of the rat SgII cDNA, the ≈ 0.7 kb *Xho*I fragment derived from the very 3′-end of the rat SgII cDNA, or selected *Eco*RI–*Xba*I fragments of the mouse SgII genomic DNA. This analysis, together with the results of the sequencing of the subcloned fragments, revealed that the 5′-to-3′ arrangement of the *Eco*RI fragments that were recognized by the various SgII probes was E4-E1-E2, with the non-hybridizing fragment E3 being located 3′ to E2.

The exon–intron structure of the mouse SgII gene was obtained by comparing its nucleotide sequence to that of the rat SgII cDNA [10]. This revealed the presence of 2 exons (Fig. 1A). Exon 1 (72 nucleotides; for the initiation site of transcription (+1), see below) contains most of the 5′-untranslated region but does not encode any translated sequence of SgII. Exon 2 (2,350 nucleotides long, if assumed to extend up to the A corresponding to the beginning of the poly-A track of the mRNA) contains the entire open reading frame of SgII and the 3′-untranslated region. The intron separating exons 1 and 2 is ≈ 3.0 kb long and contains the typical donor and acceptor splice sites (Fig. 1A).

3.2. *The mouse SgII promoter*

To determine the start site of transcription of the mouse SgII gene, primer extension was performed. By using the P1 primer, whose sequence is complementary to a portion of exon 1 (see Fig. 1A), and RNA from the mouse neuroendocrine cell line AtT-20, four primer extension products of different abundance were observed (Fig. 1B, lane 2, arrowheads). The nucleotide corresponding to the 5′-end of the longest extension product was defined as +1 (Fig. 1A). The same four primer extension products with a similar pattern of abundance were obtained when RNA from PC12 cells, a rat neuroendocrine cell line, was used as template for primer extension with the P1 oligonucleotide (data not shown). None of the four primer extension products were observed when RNA from mouse liver, a tissue lacking SgII mRNA [14], was used as template (Fig. 1B, lane 1). Primer extension using the P2 oligonucleotide, which is complementary to the 5′-end of exon 2, and AtT-20 RNA also yielded four products which had the predicted lengths (data not shown). It is unclear why the primer extension yielded four products rather than one. It is possible that transcription starts at multiple initiation sites. Alternatively, we cannot exclude that the reverse transcription reaction did not always extend to the very 5′-end of the mRNA.

A TATA-box is present approximately 30 nucleotides upstream of the putative initiation site(s) for transcription (Fig. 1A), as has been described for other promoters. We have analyzed ≈ 400 base pairs of the putative promoter region for the presence of transcription factor binding sites. The most striking result of this analysis is the existence of a typical cyclic AMP-responsive element (TGACGTCA) ≈ 30 nucleotides upstream of the TATA-box (thick bar in Fig. 1A). This may well explain, in molecular terms, the increase in SgII mRNA by agents which raise cAMP levels [14]. Phorbol esters have also been reported to increase SgII mRNA levels [14,15]. Although a 'classical' AP-1 site (TGAC/GTCA)

was not found in the ≈400 base pairs analyzed, a TGACTAA motif, which may also constitute an AP-1 binding site [16], is present ≈260 nucleotides upstream of the TATA-box in reverse orientation, on the complementary strand (dotted line in Fig. 1A). Furthermore, just 5′ to this site, we note the presence of the serum response element GATGTCC. Whether or not these elements actually function in the regulation of expression of the SgII gene requires further investigations.

3.3. *Species comparison of the secretogranin II protein sequence*

A comparison of the SgII protein sequence predicted from the mouse SgII gene exon 2 with those predicted from the rat [10], human [11] and bovine [14] SgII cDNAs is shown in Fig. 2. Calculation of the sequence similarity of SgII (without signal peptide) across the four species using the 'GAP' program of the UWGCG revealed a value of 98.5% similarity between mouse and rat, 90.5% similarity between mouse and human, and 89.2% similarity between mouse and bovine SgII. Thus, SgII is more conserved than the corresponding regions of CgB and CgA, i.e. exon 4 of CgB (mouse/human 73.2% similarity; mouse/bovine 66.7% similarity) and exons 4–7 of CgA (mouse/human 77.2% similarity; mouse/bovine 68.1% similarity). The conservation of SgII across species is particularly high between the N terminus and the third dibasic site of the protein (see Fig. 2): between mouse and human (or bovine) SgII, this portion shows 95% (94%) similarity, whereas the rest of the protein shows 88% (87%) similarity.

Mouse SgII contains 9 dibasic sites which are potential cleavage sites for endopeptidases to generate biologically active peptides (Fig. 2). The sequence position and dibasic nature of these sites is completely conserved between mouse, rat, human and bovine. An SgII-derived peptide previously isolated from frog brain [17] is generated by cleavage at the second and third dibasic site (Fig. 2). Another conserved feature is the putative tyrosine sulfation site at position 126, in line with the previous observations that SgII from various species is tyrosine-sulfated [3,18,19].

3.4. *Structural comparison of the SgII gene with the genes for CgA and CgB*

The observation that the entire SgII protein is encoded by one exon, exon 2, distinguishes the organisation of the SgII gene from that of the genes for the other two classical members of the granin family, CgB [4] and CgA [5,6], which are encoded by 5 and 8 exons, respectively (see Fig. 3). In the case of CgB, the elucidation of its genomic organisation revealed a striking correspondence between the exons and the predicted structural domains of the protein [4]. In particular, the disulfide-bonded loop near the N terminus, and the C-terminal region, where CgB is highly homologous to CgA [9], are each encoded by individual exons [4]. Also in the case of CgA, these two protein domains are encoded by individual exons [5,6].

In contrast, the elucidation of the primary structure of SgII [10,11] revealed that SgII lacks the disulfide-bonded loop found in CgA and CgB and that the C-terminal region of SgII is not related to the homologous C-terminal domain of CgA and CgB. Thus, the fact that the SgII protein is encoded by a single exon is fully consistent with it lacking the homologous N-terminal and C-terminal domains present in CgB and CgA. It has previously been noted [11] that SgII contains a weak homology to the C-terminal region of CgA and CgB at a position ≈120 amino acid residues upstream from its C terminus. The present data, showing that there is no separate exon corresponding to this weak homology, suggest that this homology may not be biologically significant.

Comparing the genomic structures of SgII and CgB, it appears that SgII as a whole (except for its signal peptide) corresponds to exon 4 of CgB. This is remarkable because exon 4 of CgB contains the structural features which are common to CgB and SgII, and which are general characteristics of the granin family (for reviews, see [1,2]). These include: (i) an abundance of acidic amino acids, (ii) calcium binding sites, and (iii) multiple potential dibasic cleavage sites. Given its simple genomic organisation, SgII may thus be considered as the prototype of a granin protein. In this context, it will be interesting to determine the appearance in evolution of SgII in comparison to that of CgB and CgA.

Acknowledgements: We thank Frank van der Hoeven for help in screening the lambda library, Christof Kaether for help with DNA subcloning, Rainer Frank for oligonucleotides, Alan Summerfield for excellent artwork and photography, and Francis Barr for his helpful comments on the manuscript. W.B.H and H.-H.G. were both recipients of grants from the Deutsche Forschungsgemeinschaft (SFB 317, C2 and C7).

REFERENCES

[1] Huttner, W.B., Gerdes, H.-H. and Rosa, P. (1991) Trends Biochem. Sci. 16, 27–30.
[2] Huttner, W.B., Gerdes, H.-H. and Rosa, P. (1991) in: Markers for Neural and Endocrine Cells (Gratzl, M. and Langley, K. Eds.) pp. 93–131, VCH, Weinheim, New York.
[3] Rosa, P., Hille, A., Lee, R.W.H., Zanini, A., De Camilli, P. and Huttner, W.B. (1985) J. Cell Biol. 101, 1999–2011.
[4] Pohl, T.M., Phillips, E., Song, K., Gerdes, H.-H., Huttner, W.B. and Rüther, U. (1990) FEBS Lett. 262, 219–224.
[5] Iacangelo, A.L., Grimes, M. and Eiden, L.E. (1991) Mol. Endo. 5, 1651–1660.
[6] Wu, H.-J., Rozansky, D.J., Parmer, R.J., Gill, B.M. and O'Connor, D.T. (1991) J. Biol. Chem. 266, 13130–13134.
[7] Benedum, U.M., Baeuerle, P.A., Konecki, D.S., Frank, R., Powell, J., Mallet, J. and Huttner, W.B. (1986) EMBO J. 5, 1495–1502.
[8] Iacangelo, A., Affolter, H.-U., Eiden, L.E., Herbert, E. and Grimes, M. (1986) Nature 323, 82–86.
[9] Benedum, U.M., Lamouroux, A., Konecki, D.S., Rosa, P., Hille, A., Baeuerle, P.A., Frank, R., Lottspeich, F., Mallet, J. and Huttner, W.B. (1987) EMBO J. 6, 1203–1211.

[10] Gerdes, H.-H., Phillips, E. and Huttner, W.B. (1988) Nucleic Acids Res. 16, 11811.
[11] Gerdes, H.-H., Rosa, P., Phillips, E., Baeuerle, P.A., Frank, R., Argos, P. and Huttner, W.B. (1989) J. Biol. Chem. 264, 12009–12015.
[12] Sambrook, M., Fritsch, E.F. and Maniatis, T. (1989) Molecular Cloning: a Laboratory Manual, Cold Spring Harbor Laboratory, Cold Spring Harbor, NY.
[13] D'Onofrio, C., Colantuoni, V. and Cortese, R. (1985) EMBO J. 4, 1981–1989.
[14] Fischer-Colbrie, R., Gutierrez, J., Hsu, C.-M., Iacangelo, A. and Eiden, L.E. (1990) J. Biol. Chem. 265, 9208–9213.
[15] Weiler, R., Meyerson, G., Fischer-Colbrie, R., Laslop, A., Pahlman, S., Floor, E. and Winkler, H. (1990) FEBS Lett. 265, 27–29.
[16] Lee, W., Mitchell, P. and Tijian, R. (1987) Cell 49, 741–752.
[17] Vaudry, H. and Conlon, J.M. (1991) FEBS Lett. 284, 31–33.
[18] Lee, R.W.H. and Huttner, W.B. (1983) J. Biol. Chem. 258, 11326–11334.
[19] Rosa, P., Fumagalli, G., Zanini, A. and Huttner, W.B. (1985) J. Cell Biol. 100, 928–937.