

promoter activity by lactogenic hormones. Their study is focussed on the DNA regions located between -400 and the *tsp*, which include the highly conserved sequences we have observed. They suggested that the sequence 5'-ANTTCTTGGA located between nt -99 and -90 might represent an essential part of a specific mammary gland factor-recognition site. In the rabbit *Aslca* gene, a similar sequence AATTCTTAGAA was found at exactly the same position (-99, -90). To assess the importance of these regions in rabbit *Aslca* gene expression and to determine other regulatory sequences, the *Aslca* gene is currently being tested using cell transfection and transgenic mice.

Note

After submission of this manuscript a paper has been published describing the bovine α -S1 casein-encoding gene (Koczan et al., 1991).

ACKNOWLEDGEMENTS

We thank Dr. A. Langlois for performing genomic blots, and Dr. I. Dusanter-Fourt, Dr. P. Gaye and Dr. C. Cotinot for helpful suggestions on primer extension analysis.

REFERENCES

- Alexander, L.J., Stewart, A.F., Mackinlay, A.G., Kapelinskaya, T.V., Tkach, T.M. and Gorodetskiy, S.I.: Isolation and characterization of the bovine kappa casein gen. Eur. J. Biochem. 178 (1988) 395-401.
- Bonsing, J., Ring, J.M., Stewart, A.F. and Mackinlay, A.G.: Complete nucleotide sequence of the bovine β -casein gene. Aust. J. Biol. Sci. 41 (1988) 527-537.
- Cheng, J.F., Printz, R., Callaghan, T., Shuey, D. and Hardison, R.C.: The rabbit C family of short, interspersed repeats. Nucleotide sequence determination and transcriptional analysis. J. Mol. Biol. 176 (1984) 1-20.
- Devinoy, E., Hubert, C., Jolivet, G., Thépot, D., Clergue, N., Desaleux, M., Dion, M., Servely, J.L. and Houdebine, L.M.: Recent data on the structure of rabbit milk protein genes and on the mechanism of the hormonal control of their expression. Reprod. Nutr. Develop. 28 (1988) 1145-1164.
- Devinoy, E., Shaerer, E., Jolivet, G., Fontaine, L.M., Krachenbühl, J.P. and Houdebine, L.M.: Sequence of the rabbit α S1-casein cDNA. Nucleic Acids Res. 16 (1988) 11813.
- Gorodetskiy, S.I., Tkach, T.M. and Kapelinskaya, T.V.: Isolation and characterization of the *Bos taurus* β -casein gene. Gene 66 (1988) 87-96.
- Hall, L., Emery, D.C., Davies, M.S., Parker, D. and Craig, R.K.: Organization and sequence of the human α -lactalbumin gene. Biochem. J. 242 (1987) 735-742.
- Jahn, G., Djiane, J. and Houdebine, L.M.: Inhibition of casein synthesis by progestagens 'in vitro': modulation in relation to the intensity of response to prolactin. J. Steroid Biochem. 32 (1985) 373-379.
- Jones, W.K., Yu-Lee, L.Y., Clift, S.M., Brown, T.L. and Rosen, J.M.: The rat casein multigene family. Fine structure and evolution of the β -casein gene. J. Biol. Chem. 260 (1985) 7042-7050.
- Koczan, D., Hobom, G. and Seyfert, H.M.: Genomic organization of the alpha-S1 casein gene. Nucleic Acids Res. 19 (1991) 5591-5596.
- Limonta, J., Castro, F.O., Basulto, R., Aguilar, A., Martinez, R., Lieouart, R., Herrera, L. and de la Fuente, J.: Bovine α S1-casein gene promoter directs the expression of active human tissue-type plasminogen activator in mouse milk. Adv. Gene Technol. 1 (1991) 106.
- Maniatis, T., Hardison, R.C., Lacy, E., Lauer, J., O'Connell, C., Quon, D., Sim, G.K. and Efstratiadis, A.: The isolation of structural genes from libraries of eucaryotic DNA. Cell 15 (1978) 687-701.
- Maniatis, T., Fritsch, E.F. and Sambrook, J.: Molecular Cloning. A Laboratory Manual. Cold Spring Harbor Laboratory, Cold Spring Harbor, NY, 1982, pp. 320-323.
- Meade, H., Gates, L., Lacy, E. and Lonberg, N.: Bovine α lactalbumin gene sequences direct high level expression of active human urokinase in mouse milk. Bio-Technology 8 (1990) 443-446.
- Mercier, J.C., Grosclaude, F. and Martin, P.: Polymorphism, biosynthesis and evolution of caseins. Medicine Sciences 7 (1991) R1-R8.
- Sanger, F., Nicklen, S. and Coulson, A.R.: DNA sequencing with chain-terminating inhibitors. Proc. Natl. Acad. Sci. USA 74 (1977) 5463-5467.
- Schmitt-Ney, M., Doppler, W., Ball, R.K. and Groner, B.: β -Casein gene promoter activity is regulated by the hormone-mediated relief of transcriptional repression and a mammary-gland-specific nuclear factor. Mol. Cell. Biol. 11 (1991) 3745-3755.
- Thépot, D., Devinoy, E., Fontaine, M.L. and Houdebine, L.M.: Structure of the gene encoding rabbit β -casein. Gene 97 (1991) 301-306.
- Yoshimura, M. and Oka, T.: Isolation and structural analysis of the mouse β -casein gene. Gene 78 (1989) 267-275.
- Yu-Lee, L.Y., Richter-Mann, L., Couch, C.H., Stewart, A.F., Mackinlay, A.G. and Rosen, J.M.: Evolution of the casein multigene family: conserved sequences in the 5' flanking and exon regions. Nucleic Acids Res. 14 (1986) 1883-1902.

Gene. 113 (1992) 263-268
© 1992 Elsevier Science Publishers B.V. All rights reserved. 0378-1119/92/\$05.00

GENE 06403

Expression and genomic structure of the genes encoding FdI, the major allergen from the domestic cat

(Polymerase chain reaction; allergy; *Felis domesticus*; alternate splicing; differential tissue-specific expression)

Irwin J. Griffith, Sandra Craig, Joanne Pollock, Xu-Bo Yu, Jay P. Morgenstern and Bruce L. Rogers

ImmuLogic Pharmaceutical Corporation, Cambridge, MA 02139 (USA)

Received by S.T. Case: 8 November 1991
Revised/Accepted: 29 December 1991/8 January 1992
Received at publishers: 31 January 1992

SUMMARY

The genes encoding chain 1 (Ch1) and chain 2 (Ch2) of the major allergen of the domestic cat, *Felis domesticus* I, have been analyzed by genomic cloning and by polymerase chain reaction (PCR). Ch1 has two potential leader sequences, designated A and B. Analysis of a genomic clone encoding Ch1 demonstrated that one structural gene contains sequences corresponding to both leaders, which utilize different Met start codons. PCR analysis showed that genes encoding Ch1 and Ch2 are co-expressed in both the salivary glands and the skin, and that leader sequence A of Ch1 is utilized preferentially in both tissues. Ch2 was shown to have two dominant forms that are differentially expressed in the aforementioned tissues. The long form (Ch2L), composed of 92 amino acids (aa), is preferentially expressed in the salivary glands, while the short form (Ch2S), composed of 90 aa, is preferentially expressed in the skin. There is minor sequence polymorphism in both forms of Ch2. A genomic clone for Ch2 only contained sequences for Ch2S, suggesting that Ch2L is encoded by an exon not contained within this genomic clone.

INTRODUCTION

Clinical studies have shown that FdI, the major allergen from domestic cats, plays an important role as an airborne allergen in allergic diseases (Ohman et al., 1984; Sundin

et al., 1986) and has been implicated as a major asthma-inducing agent (Van Metre et al., 1986). FdI has been biochemically and allergenically characterized (Leitermann and Ohman, 1984; Lowenstein et al., 1985; Ohman et al., 1974) and has been shown to be a heterodimeric 39-kDa protein when affinity-purified in its native form (Chapman et al., 1988; Duffort et al., 1991; Van Milligen et al., 1991). Early studies detected FdI in cat pelts and saliva with the former attributed to deposition of the allergen during grooming (Brown et al., 1984; Ohman et al., 1974). Recent experiments suggest that FdI may also be produced by sebaceous glands in the pelt (Charpin et al., 1991).

Biochemical purification and the generation of monoclonal antibodies directed against FdI led to improved purification protocols (Chapman et al., 1988; De Groot et al., 1988; Duffort et al., 1988) and determination of the N-terminal aa sequence (Chapman et al., 1988; Duffort et al., 1991). Characterization of affinity-purified FdI demonstrated that this glycosylated protein is a heterodimer with 5-6-kDa and 16-kDa subunits (on SDS-polyacryl-

Correspondence to: Dr. I.J. Griffith, ImmuLogic Pharmaceutical Corporation, One Kendall Square, Cambridge, MA 02139 (USA). Tel. (617)494-0060; Fax (617)537-8686.

Abbreviations: aa, amino acid(s); bp, base pair(s); Ch1, chain 1 of FdI; Ch1, gene (DNA) encoding Ch1; Ch2, chain 2 of FdI; Ch2, gene (DNA) encoding Ch2; Ch2L, long form of Ch2; Ch2L, gene (DNA) encoding Ch2L; Ch2Lv, variant form of Ch2L; Ch2Lv, gene (DNA) encoding Ch2Lv; Ch2S, short form of Ch2; Ch2S, gene (DNA) encoding Ch2S; Ch2St, short truncated form of Ch2S; Ch2St, gene (DNA) encoding Ch2St; Ch2Sv, variant form of Ch2S; Ch2Sv, gene (DNA) encoding Ch2Sv; F, *Felis*; FdI, major allergen of *F. domesticus*; FdI, gene (DNA) encoding FdI; kb, kilobase(s) or 1000 bp; nt, nucleotide(s); oligo, oligodeoxynucleotide; PCR, polymerase chain reaction; Ug, uteroglobin; UT, untranslated.

Met start codon (M) is indicated for each cDNA clone. Leader sequences (A, B) are represented by striped, coding sequences by blackened and noncoding sequences by open boxes.



Fig. 3. Presence of *Ch1* mRNA leader sequences in salivary glands and skin. Total and poly(A)⁺ RNA were isolated from the pooled salivary glands of five cats and skin tissues from two individual cats using standard techniques (Aviv and Leder, 1972; Chirgwin et al., 1979), as described (Morgenstern et al., 1991). First-strand cDNA (1°) was synthesized (cDNA Synthesis System Plus Kit, Amersham, Arlington Heights, IL) from 1 µg of total or poly(A)⁺ RNA using oligo(dT) as a primer (Gubler and Hoffman, 1983) and amplified 1°, using the GeneAmp kit (US Biochemicals) as described (Rafnar et al., 1991; Morgenstern et al., 1991). Secondary PCR (2°) was performed on 10% (10 µl) of the primary PCR material using nested primers, as specified in section b. The presence of *Ch1* mRNA with leader A (lanes 2,5,8) and leader B (lanes 3,4,6,7,9,10) was analyzed in salivary gland (lanes 2-4) and skin (lanes 5-10) RNA samples. Skin 1 (lanes 5-7) was derived from one of the cats used in the salivary gland pool. Skin 2 (lanes 8-10) was derived from a sixth cat. Lane 1 contains 2.5 µg of *Hae*III-digested ϕ X174 DNA. Primary amplification of leader A with primers FI-6 and FI-8 and leader B with primers FI-7 and FI-8 gives predicted DNA bands of 321 and 324 bp, respectively. Secondary amplification of leader B with primers FI-7 and FI-5 gives a predicted DNA band of 227 bp. Samples were resolved on a 3% agarose gel (FMC BioProducts, Rockland, ME).

tissue samples (Fig. 3, lanes 4, 7 and 10). Five PCR-derived clones (four from salivary gland RNA and one from skin RNA) were sequenced and found to be identical to the published *Ch1* sequence, except for a single nt change in the codon for aa 29 (AAA→AAT) of the skin-derived clone that results in the substitution of Asn for Lys (Lys²⁹→Asn) (data not shown). While direct aa sequencing of affinity purified FdI (Morgenstern et al., 1991) and dideoxy sequencing of PCR amplified *Ch1* product or other *Ch1* cDNA clones (data not shown) show only Lys²⁹, the independent detection of Asn at this residue in genomic and mRNA-derived clones confirms that this difference is due to a true polymorphism. The frequency of this and other polymorphic residues in the outbred cat population is not known.

(c) Genomic structure and sequence of the Ch2-encoding gene

Hybridization screening of the cat genomic library with the *Ch2* probe F2P (Morgenstern et al., 1991) resulted in the isolation of clone CTGF2P #5, which contained a 6.8-kb *Eco*RI insert. A 2.3-kb DNA fragment was generated by PCR amplification of this genomic insert using *Ch2*-specific 5'-primer FII-8 and 3' *Ch2*-specific primer FII-7 (Table I). The complete nt sequence of the 2.3-kb fragment

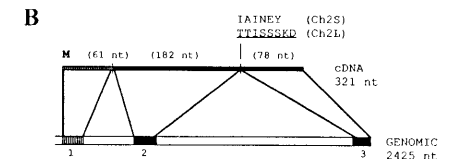
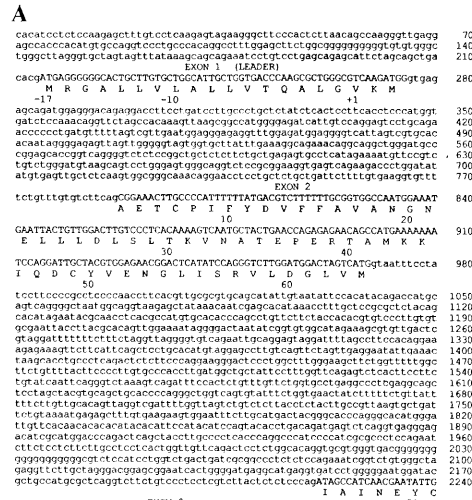


Fig. 4. Nucleotide sequence of the 5'-flanking regions and exons 1 to 3 of the *Ch2* gene and deduced aa sequence. (A) The nt and deduced aa sequences of *Ch2* from clone CTGF2P #5 (EMBL accession No. X62478). The EMBL library was screened with the *Ch2* probe F2P (Morgenstern et al., 1991). A 2.3-kb cDNA was amplified by PCR from genomic clone CTGF2P #5 using the oligo primers specified in section c and sequenced as described in the legend to Fig. 1. Exon and intron sequences are presented as in Fig. 1A. Sequence of the first 206 nt was determined by direct sequencing of the 6.8-kb genomic clone. (B) Schematic representation of the genomic *Ch2* exon/intron structure. Coding and noncoding sequences are presented as in Fig. 1B. The aa sequences differentiating Ch2S and Ch2L are indicated above the figure.

and deduced aa of *Ch2* are presented in Fig. 4A. Organization of *Ch2* is presented schematically in Fig. 4B.

The putative leader sequence and the first 3 aa of the mature protein are encoded by exon 1 (61 nt; 20 aa). The bulk of the mature protein is encoded by exons 2 and 3 (aa 4-64 and 65-90, respectively; numbering of aa is based upon the *Ch2* sequence published in Morgenstern et al., 1991). The first 18 nt of exon 3 encode the residues, IAINNEY (aa 65-70), rather than the published sequence, TTISSSKD (Fig. 4A and B), suggesting that *Ch2* has two forms (see section d, below). There were no other differ-

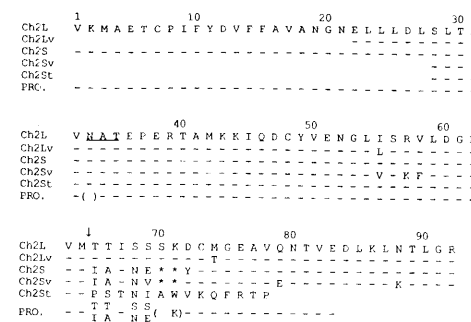


Fig. 5. Deduced aa polymorphisms of *Ch2*. Sequences are shown in standard one-letter code relative to *Ch2L*. Only partial aa sequences were obtained for *Ch2Lv*, *Ch2Sv* and *Ch2St* clones. PRO indicates *Ch2* sequence obtained by direct aa sequencing (Morgenstern et al., 1991). Asterisks indicate gaps inserted to maximize alignment; dashes indicate aa identity; super- and subscripts denote polymorphic aa sequence; parentheses indicate an unknown or probable sequence. The single N-linked glycosylation site of *Ch2* is underlined. The downward arrow indicates the junction of exons 2 and 3 in *Ch2S*.

ences detected between the genomic exons and published *Ch2* cDNA sequence.

(d) PCR analysis of *Ch2* mRNA

Primers FII-8 or FII-4 (5' *Ch2*-specific) and FII-9 (3' *Ch2*-specific) were used to amplify full-length or partial *Ch2* cDNA clones for nt sequence analysis. Examination of 52 clones derived from salivary gland RNA (24 clones) or skin RNA (28 clones) showed that *Ch2* has two dominant forms, a 92-aa *Ch2L* and a 90-aa *Ch2S* that differ from each other by 6 contiguous aa and a 2-aa insertion/deletion (Fig. 5; Table II). *Ch2L* matches the previously published *Ch2* aa sequence (Morgenstern et al., 1991), while *Ch2S* matches

TABLE II

Clones encoding *Ch2* variants isolated from different tissues^a

	Salivary gland	Skin	
		Skin 1	Skin 2
Long forms:			
<i>Ch2L</i>	23	0	0
<i>Ch2Lv</i>	0	0	3
Short forms:			
<i>Ch2S</i>	1	11	9
<i>Ch2Sv</i>	0	3	0
<i>Ch2St</i>	0	2	0

^a Fifty-two *Ch2* cDNA clones were isolated from salivary gland- or skin-derived RNA. The clones were classified as 'long' or 'short' based on direct nt and deduced aa sequence analysis.

that encoded by genomic clone CTGF2P #5 (Fig. 4A). The deduced aa differences between *Ch2L* and *Ch2S* resolve the sequence polymorphism observed in the original aa sequence analysis of *Ch2* (Morgenstern et al., 1991).

Ch2L and *Ch2S* may arise from differential splicing of an mRNA precursor. However, examination of the 1.25-kb intron between exons 2 and 3 of the genomic *Ch2* clone failed to disclose any mini-exon sequence encoding the aa unique to *Ch2L*. Further, there is no appropriate splice donor/acceptor site in exon 3 near the common sequence of *Ch2L* and *Ch2S*. This suggests that the 3' end of *Ch2L* is encoded by a completely separate exon, exon 4, downstream from exon 3 that has not yet been identified. PCR analysis of genomic clone CTGF2P #5 indicates that the purported fourth *Ch2* exon is not contained within this 6.8-kb *Eco*RI fragment (data not shown). Alternatively, *Ch2L* and *Ch2S* may represent different alleles of the *Ch2* gene.

Sequence analysis of the PCR-derived clones also demonstrated that minor sequence polymorphism is present in both the *Ch2L* and *Ch2S* genes. A short variant (*Ch2Sv*) found in the skin RNA of cat 1 was found to have 7 aa differences from *Ch2S* (Ile⁵⁵→Val, Arg⁵⁷→Lys, Val⁵⁸→Phe, Glu⁶⁹→Val, Tyr⁷²→Asp, Gln⁷⁹→Glu, and Asn⁸⁸→Lys; Fig. 5). In addition, there was a short truncated (*Ch2St*) form detected in the skin of cat 1 that results from use of a cryptic splice acceptor site in exon 3 and the subsequent deletion of 4 nt (Fig. 4A). The resulting reading frame shift would lead to a protein 78 aa long that lacks a potentially important Cys found in *Ch2S* and *Ch2L* and their other variants (Fig. 5). A long variant (*Ch2Lv*) encoded in skin RNA of the second cat was found to have two differences from *Ch2L* (Ile⁵⁵→Leu and Met⁷⁴→Thr; Fig. 5). All variants were detected in clones derived from at least two independent PCR amplifications, suggesting that they represent polymorphic forms of *Ch2L* and *Ch2S*, rather than PCR-generated artifacts.

The analysis of the 52 sequenced *Ch2* clones is summarized in Table II. The data suggest that the *Ch2L* gene is preferentially expressed in the salivary glands (23 of 24 clones, 96%) and *Ch2S* and its variants are preferentially expressed in the skin (25 of 28 clones, 89%; Table II). Further, the small number of variant cDNA clones identified (eight of 52 clones, 15%) suggests that *Ch2L* and *Ch2S* are the dominant proteins and that *Ch2Lv*, *Ch2Sv* and *Ch2St* are minor protein species. The dominance of *Ch2L* and *Ch2S* is supported by direct aa sequence analysis of FdI collected from cat pelt washes or carpets, which only showed sequence corresponding to *Ch2L* and *Ch2S* (Morgenstern et al., 1991). The differential expression of the two dominant forms of *Ch2* suggests that they may have different and, as of yet unknown, biological functions. The mechanism(s) regulating this differential expression are also

unknown. The relative importance of the two forms of Ch2 in the induction of allergy to FdI is under investigation.

(c) Conclusions

(1) A genomic clone containing the gene encoding Ch1 has been isolated and sequenced. This single structural gene contains sequences corresponding to the two previously described potential leader sequences, A and B (Morgenstern et al., 1991). Although both leader sequences are expressed in the tissues examined, *Ch1* mRNA with leader A is found more frequently than mRNA with leader B. Further, the genomic organization and deduced aa sequence of Ch1 with leader A are homologous with rabbit U_g.

(2) Sequence analysis of multiple *Ch2* cDNA clones revealed that there are two distinct forms. Ch2L matches the previously published Ch2 sequence (Morgenstern et al., 1991), while Ch2S matches the sequence encoded in a genomic clone described here. Interestingly, *Ch2L* and *Ch2S* are preferentially expressed in salivary gland and skin tissues, respectively.

ACKNOWLEDGEMENTS

The authors wish to thank Dr. Julian Bond for his kind gift of a skin-derived RNA sample; Dr. Rob Bronson for his assistance in collecting tissue samples; Dr. Jim Carpenter for helpful discussions and demonstrations; Dr. Mei-chang Kuo for helpful discussions; Drs. Roland Bue-low and Dean Falb for their critical evaluation of the manuscript; and Ms. Joanne Trio for her help in the manuscript preparation.

REFERENCES

- Aviv, H. and Leder, P.: Purification of biologically active globin messenger RNA by chromatography on oligo-thymidic acid-cellulose. *Proc. Natl. Acad. Sci. USA* 69 (1972) 1408–1412.
- Bailly, A., Atger, M., Atger, P., Cerbon, M.-A., Alizon, M., Hai, M.T.V., Logeat, F. and Milgrom, E.: The rabbit uteroglobin gene. *J. Biol. Chem.* 258 (1983) 10384–10389.
- Benton, W.D. and Davies, R.W.: Screening lgt recombinant clones by hybridization to single plaques in situ. *Science* 196 (1977) 180–182.
- Brown, P.R., Leitermann, K. and Ohman Jr., J.L.: Distribution of cat allergen 1 in cat tissues and fluids. *Int. Arch. Allergy Appl. Immunol.* 74 (1984) 67–70.
- Chapman, M.D., Aalberse, R.C., Brown, M.J. and Platts-Mills, T.A.: Monoclonal antibodies to the major feline allergen *Fel d 1*, II. Single step affinity purification of *Fel d 1*, N-terminal sequence analysis, and development of a sensitive two-site immunoassay to assess *Fel d 1* exposure. *J. Immunol.* 140 (1988) 812–818.
- Charpin, C., Mata, P., Charpin, D., Lavaut, M.N., Allasia, C. and Ver-vloet, D.: *Fel d 1* allergen distribution in cat fur and skin. *J. Allergy Clin. Immunol.* 88 (1991) 77–82.
- Chirgwin, J.M., Przybyla, A.E., MacDonald, R.J. and Rutter, W.J.: Isolation of biologically active ribonucleic acid from sources enriched in ribonuclease. *Biochemistry* 18 (1979) 5294–5299.
- De Groot, H., Van Swieten, P., Van Leeuwen, J., Lind, P. and Aalberse, R.C.: Monoclonal antibodies to the major feline allergen *Fel d 1*, I. Serologic and biologic activity of affinity-purified *Fel d 1* and of *Fel d 1*-depleted extract. *J. Allergy Clin. Immunol.* 82 (1988) 772–786.
- Duffort, O., Carreira, J. and Lombardero, M.: Monoclonal antibodies against *Fel d 1* and other clinically relevant cat allergens. *Immunol. Lett.* 17 (1988) 71–77.
- Duffort, O.A., Carreira, J., Nitti, G., Polo, F. and Lombardero, M.: Studies on the biochemical structure of the major cat allergen *Felis domesticus* I. *Mol. Immunol.* 28 (1991) 301–309.
- Frischauf, A.-M., Lehrach, H., Poustka, A. and Murray, N.: Lambda replacement vectors carrying polylinker sequences. *J. Mol. Biol.* 170 (1983) 827–842.
- Gubler, U. and Hoffman, B.J.: A simple and very efficient method for generating cDNA libraries. *Gene* 25 (1983) 263–269.
- Leitermann, K. and Ohman Jr., J.L.: Cat allergen I: biochemical, antigenic, and allergenic properties. *J. Allergy Clin. Immunol.* 74 (1984) 147–153.
- Lowenstein, H., Lind, P. and Weeke, B.: Identification and clinical significance of allergenic molecules of cat origin. *Allergy* 40 (1985) 430–441.
- Morgenstern, J.P., Griffith, L.J., Brauer, A.W., Rogers, B.L., Bond, J.F., Chapman, M.D. and Kuo, M.-c.: Determination of the amino acid sequence of *Fel d 1*, the major allergen of the domestic cat: protein sequence analysis and cDNA cloning. *Proc. Natl. Acad. Sci. USA* 88 (1991) 9690–9694.
- Mount, S.M.: A catalogue of splice junction sequences. *Nucleic Acids Res.* 10 (1982) 459–472.
- Ohman Jr., J.L., Lowell, F.C. and Bloch, K.J.: Allergens of mammalian origin. III. Properties of a major feline allergen. *J. Immunol.* 113 (1974) 1668–1677.
- Ohman Jr., J.L., Findlay, S.R. and Leitermann, K.M.: Immunotherapy in cat-induced asthma. Double-blind trial with evaluation of in vivo and in vitro responses. *J. Allergy Clin. Immunol.* 74 (1984) 230–239.
- Rafnar, T., Griffith, L.J., Kuo, M.-c., Bond, J.F., Rogers, B.L. and Klap-per, D.G.: Cloning of *Amb a 1* (antigen E), the major allergen family of short ragweed pollen. *J. Biol. Chem.* 266 (1991) 1229–1236.
- Saiki, R.K., Gelfand, D.H., Stoffen, S., Scharf, S.H., Higuchi, R., Horn, G.T., Mullis, K.B. and Erlich, H.A.: Primer-directed enzymatic amplification of DNA with a thermostable DNA polymerase. *Science* 239 (1988) 487–491.
- Sambrook, J., Fritsch, E.F. and Maniatis, T.: *Molecular Cloning. A Laboratory Manual*, 2nd ed. Cold Spring Harbor Laboratory Press, Cold Spring Harbor, NY, 1989.
- Sanger, F., Nicklen, S. and Coulson, A.R.: DNA sequencing with chain-terminating inhibitors. *Proc. Natl. Acad. Sci. USA* 74 (1977) 5463–5460.
- Sundin, B., Lilja, G., Graff-Lonnevig, V., Hedlin, G., Heilborn, H., Norr-lind, K., Pegelow, K.-O. and Lowenstein, H.: Immunotherapy with partially purified and standardized animal dander extracts, I. Clinical results from a double-blind study on patients with animal dander asthma. *J. Allergy Clin. Immunol.* 77 (1986) 478–487.
- Suske, G., Wenz, M., Cato, A.C.B. and Beato, M.: The uteroglobin gene region: hormonal regulation, repetitive elements and complete nucleotide sequence of the gene. *Nucleic Acids Res.* 8 (1983) 2257–2271.
- Van Metre, T.E., Marsh, D.G., Adkinson Jr., N.F., Fish, J.E., Kagey-Sabotka, A., Norman, P.S., Radden Jr., E.B. and Rosenberg, G.L.: Dose of cat (*Felis domesticus*) allergen 1 (*Fel d 1*) that induces asthma. *J. Allergy Clin. Immunol.* 78 (1986) 62–75.
- Van Milligen, F.J., Aalberse, R.C. and Van Swieten, P.: Variability of *Fel d 1* structure in commercial cat allergen extracts as revealed by monoclonal antibodies to denatured *Fel d 1*. *J. Allergy Clin. Immunol.* 87 (1991) 327.
- Gene*, 113 (1992) 269–274.
© 1992 Elsevier Science Publishers B.V. All rights reserved. 0378-1119/92/\$05.00

GENE 06425

Cloning and characterization of the human gene encoding p11: structural similarity to other members of the S-100 gene family

(Recombinant DNA; annexins; exon/intron structure; CpG island; Ca²⁺ ions; EF-hand)

Thomas Harder, Eckhard Kube and Volker Gerke

Department of Biochemistry, Max-Planck-Institute for Biophysical Chemistry, D-3400 Göttingen (Germany) Tel. (49-551) 201486

Received by H. van Ormondt: 18 September 1991

Revised/Accepted: 9 December/16 December 1991

Received at publishers: 7 February 1992

SUMMARY

The human gene (*CLP11*) encoding p11, the cellular ligand of the tyrosine kinase substrate, annexin II (AnxII), has been isolated from a human genomic library. Restriction mapping and sequencing reveals that *CLP11* covers a stretch of approx. 11 kb in the human genome. The structure of *CLP11* resembles that of the other genes encoding S-100 proteins which have been characterized so far: the transcribed region is divided by two introns, one in the 5'-nontranslated portion and the second in the protein-coding region. Interestingly, in *CLP11*, as well as all other *S-100* genes, the second intron separates the codons for two corresponding amino acids, which reside in the sequence connecting the two helix-loop-helix (EF-hand) motifs. The 5'-nontranscribed region, which most likely represents the *CLP11* promoter, is characterized by a high G+C content and probably is part of a CpG-island. Several putative binding sites for transcription factors can be identified in the 5'-nontranscribed region of *CLP11*. Among them, the β DRE element, which was first described in the β -globin promoter, is most notable, since it is also present in the promoter of *ANXII*. It could be responsible for the simultaneous induction of *CLP11* and *ANXII* expression during certain cell differentiation processes, e.g., the nerve growth factor-induced differentiation of the pheochromocytoma cell line, PC12.

INTRODUCTION

The response to elevation of cytoplasmic Ca²⁺ levels following extra- or intracellular stimuli is mediated by pro-

teins which are capable of binding Ca²⁺ ions. A particular class of these proteins is characterized by the so-called EF-hand, a helix-loop-helix motif involved in coordinating the Ca²⁺ ion (for review see Kretsinger et al., 1991). Within the EF-hand superfamily a distinct set of proteins is grouped in the so-called S-100 protein family, whose members share a high degree of sequence similarity with S-100a and S-100b, Ca²⁺-binding proteins originally isolated from cerebral fluid (Moore, 1965; Isobe et al., 1977). All S-100 proteins are characterized by two consecutive EF-hands (one in the N-terminal and one in the C-terminal half of the molecule), which are connected via a stretch of aa of a less defined secondary structure (for review on S-100 proteins see Hilt and Kligman, 1991).

Protein p11 is a member of the S-100 family but has several unique features. It has suffered crucial deletions and aa substitutions in both EF-hand motifs, which are thought to render both Ca²⁺-binding sites inactive (for review see

Correspondence to: Dr. V. Gerke, Department of Biochemistry, Max-Planck-Institute for Biophysical Chemistry, D-3400 Göttingen (Germany) Tel. (49-551) 201477; Fax (49-551) 201578.

Abbreviations: aa, amino acid(s); Anx, annexin; ANX, human gene encoding Anx; β DRE, β -globin direct repeat element; bp, base pair(s); cDNA, DNA complementary to mRNA; *CLP11*, human gene encoding p11; 50 × Denhardt's, 5% Ficoll/5% polyvinylpyrrolidone/5% bovine serum albumin; EF-hand, see INTRODUCTION; HTF, *HpaII* tiny fragments; kb, kilobase(s) or 1000 bp; NGF, nerve growth factor; nt, nucleotide(s); p11, cellular ligand of annexin II; SDS, sodium dodecyl sulfate; src, oncogene encoded by Rous sarcoma virus; SSC, 0.15 M NaCl/0.015 M Na₂ citrate pH 7.6; SSPE, 0.18 M NaCl/0.01 M Na₂HPO₄ pH 7.4/1 mM EDTA; TBE, 0.09 M Tris-borate pH 8.0/2 mM EDTA; tsp, transcription start point(s).