Volume 10 Number 11 1982 Nucleic Acids Research

Comparison of the nucleotide sequence of cloned human and guinea-pig pre-α-lactalbumin cDNA with that of chick pre-lysozyme cDNA suggests evolution from a common ancestral gene

Len Hall, Roger K.Craig*, Mark R.Edbrooke and Peter N.Campbell

Courtauld Institute of Biochemistry, The Middlesex Hospital Medical School, London W1P 7PN, UK

Received 6 April 1982; Revised and Accepted 13 May 1982

ABSTRACT

Nucleotide sequence analyses of essentially full-length copies of human and guinea-pig pre-α-lactalbumin cDNAs contained within recombinant plasmids, (i) confirm the presence of 19 amino acid hydrophobic amino terminal peptide extensions encoded within each mRNA; and (ii) provides evidence for the existence of a minor variant of guinea-pig α-lactalbumin mRNA encoding a protein with a 36 residue carboxyl-terminal extension. Comparison of the nucleotide sequence within the coding region of the human, and the predominant guinea-pig pre-α-lactalbumin mRNAs, with the analogous region of hen pre-lysozyme mRNA provides compelling evidence that all have evolved from a common ancestral gene.

INTRODUCTION

Gene duplication and subsequent sequence divergence forms the basis of two related but nevertheless distinct evolutionary processes (1). The first of these is typified by the globins and immunoglobulins, where several related proteins have evolved to perform a similar function. The second process involves divergence to form proteins with different functions; an example being provided by the peptide hormones prolactin, chorionic somatomammotropin and growth hormone. Recent sequence analysis of the mRNAs encoding these three proteins, together with an analysis of the distribution of silent and replacement substitutions, have suggested that their genes have arisen from a common ancestor (2,3). Furthermore, analysis of the sequence divergence at replacement sites has established an evolutionary clock for this gene family.

It has often been suggested that the structurally related, but functionally distinct proteins, α-lactalbumin and lysozyme, have also arisen from a common ancestral gene (4). Although superficially the functions of these two proteins seem to be related, in that they are involved in the catalysis of a β1 → 4 glycosidic linkage, closer examination reveals that they achieve this in quite distinct ways. α-Lactalbumin is a mammary

0305-1048/82/1011-3503$ 2.00/0

gland-specific protein which modifies the substrate specificity of galactosyl transferase thereby promoting the synthesis of lactose during lactation (5,6). While this is probably the primary function of α-lactalbumin, its high concentration in some milks may indicate a secondary, nutritional role, perhaps as a source of cysteine (which is absent from characterised guinea-pig caseins, see ref 7). Lysozyme, on the other hand, catalyses the hydrolysis of a β1 → 4 glycosidic linkage in the mucopolysaccharide component of some bacterial cell walls (8). Both mammalian and phage encoded lysozymes have been characterised but only the former exhibits amino acid homology with α-lactalbumin.

The complete amino acid sequence of a number of α-lactalbumins including human (9), guinea-pig (10), bovine (11), goat (12) and rabbit (13) are known, as well as several lysozymes (14-16). These sequences have frequently been compared with a view to understanding the relationship between structural differences and the evolutionary divergence of function within the α-lactalbumin/lysozyme family. However, such studies are limited in that one can only monitor mutational changes which lead to a different amino acid residue, and consequently give no indication of the extent of silent nucleotide substitutions at the level of the mRNA. A detailed comparative analysis of the corresponding mRNAs should therefore provide more information than the primary protein sequence alone. Towards this and other ends, we have recently constructed and partially characterised a number of recombinant plasmids containing human (17) and guinea-pig (18) α-lactalbumin cDNA sequences. Here we present the complete nucleotide sequence of the coding, 3'-noncoding, and part of the 5'-noncoding regions of human and guinea-pig α-lactalbumin cDNAs, contained within two of these recombinant plasmids, and compare them with each other, and also with the published nucleotide sequence of hen egg-white lysozyme (19). From the distribution of silent and replacement nucleotide substitutions we provide compelling evidence that the α-lactalbumin and lysozyme genes have evolved from a common ancestor.

MATERIALS AND METHODS

Materials

Restriction endonucleases BstNI and RsaI were purchased from BioLabs through CP Labs Ltd., Bishop's Stortford, Herts, U.K.; all other restriction enzymes, T4 DNA polymerase and T4 polynucleotide kinase were from Bethesda Research Laboratories, Cambridge, U.K. AMV reverse

transcriptase was provided by Dr. J.W. Beard, Life Sciences Inc., St. Petersburg, FL 33707, U.S.A.

Chemical DNA sequencing

Separation and elution of DNA restriction fragments from polyacrylamide gels, and exchange labelling of protruding 5'-ends with T4 polynucleotide kinase and [γ-^{32}P]ATP were exactly as described by Maxam and Gilbert (20). 3'-ends were labelled by incubating 1-10 µg of restricted DNA in 65 mM tris-HCl, pH 8, 6.5 mM $MgCl_2$, 5 mM dithiothreitol, 30 µM [α-^{32}P]dNTP (> 250 Ci/mmol) and all other unlabelled dNTPs (as appropriate) at 0.1 mM, with 5 units of T4 DNA polymerase at 11°C for 30 min. Flush ended, 5'-protruding and 3'-protruding DNA fragments could all be efficiently labelled in this way. Chemical sequencing of the end-labelled DNA fragments was exactly as described by Maxam and Gilbert (20).

Dideoxy sequencing by primer extension

Single-stranded DNA primers were prepared by electrophoresing and eluting suitable recombinant plasmid restriction fragments from polyacrylamide/urea sequencing gels (20). A three-fold excess of primer was hybridised (21) to total polyadenylated RNA from lactating mammary tissue and then used to synthesize cDNA in the presence of dideoxynucleotides (22).

RESULTS AND DISCUSSION

Sequence analysis of recombinant plasmids containing human and guinea-pig α-lactalbumin cDNAs

Using the sequencing strategies shown in Fig. 1, the complete nucleotide sequences of the human and guinea-pig α-lactalbumin cDNAs contained within the recombinant plasmids phB-35 and pgpK-9 were determined by the method of Maxam and Gilbert (Fig. 2). It can be seen that both contain part of the 3' poly(A) tail of the corresponding mRNA, the entire 3'-noncoding region, as well as the complete coding region. Furthermore, the guinea-pig sequence also includes 84 nucleotides of the 5'-noncoding region. Consequently phB-35 contains about 90% of the human α-lactalbumin mRNA sequence, and pgpK-9 contains virtually all of the guinea-pig sequence, assuming an overall length for pre-α-lactalbumin mRNA [including the poly(A) tail] of 800-850 nucleotides as indicated by Northern transfer (18; Davies, Hall & Craig, unpublished data), and a poly(A) tail length of about 50 nucleotides.

Comparison of the published amino acid sequences of human (9) and guinea-pig (10) α-lactalbumins isolated from milk, with those deduced from

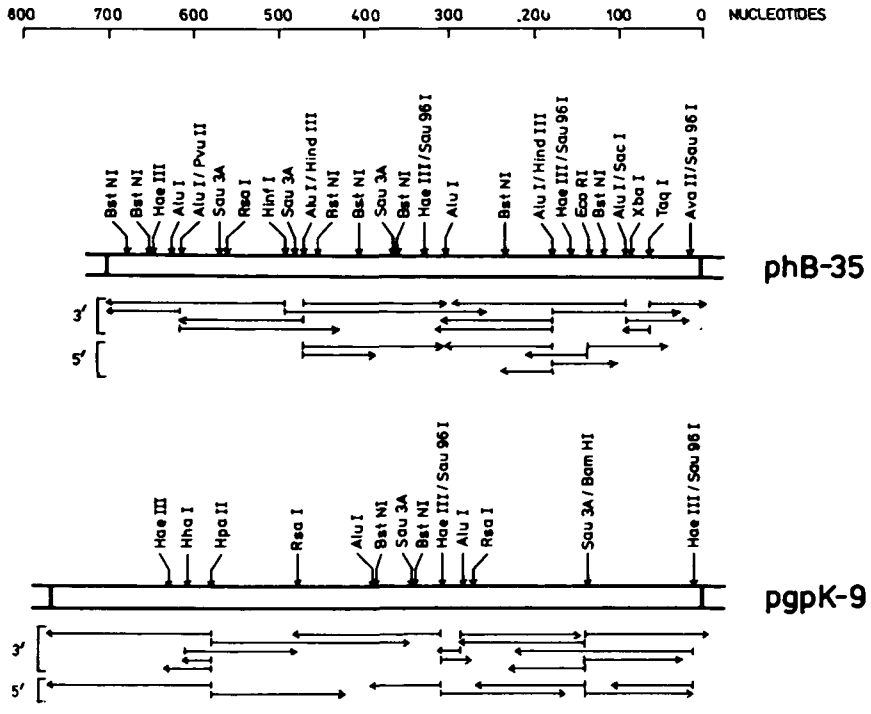


Figure 1. Restriction maps and sequencing strategies for the α-lactalbumin cDNA sequences contained within recombinant plasmids phB-35 (human α-lactalbumin) and pgpK-9 (guinea-pig α-lactalbumin). Arrows indicate the direction of sequencing of individual fragments; 3' and 5' indicate which end of the DNA was labelled (see Materials and Methods).

the nucleotide sequences (Fig. 2), confirm that both are synthesized as precursor proteins with characteristic amino-terminal extensions, a property common to the majority of secretory proteins (23,24), and previously indicated for human (25) and guinea-pig (26,27) α-lactalbumins on the basis of mRNA-directed cell-free protein synthesis. In general the amino-terminal presequences of various preproteins, while not exhibiting any significant sequence homology, do contain a common physical property in that they contain a region rich in nonpolar, hydrophobic amino acids (28), necessary for the interaction of the nascent polypeptide chain with putative receptors on the endoplasmic reticulum. This results ultimately in cotranslational processing of the signal sequence, and sequestration of the processed protein within the lumen of the endoplasmic reticulum. In this respect, the α-lactalbumins have typical presequences. Of the 19 amino

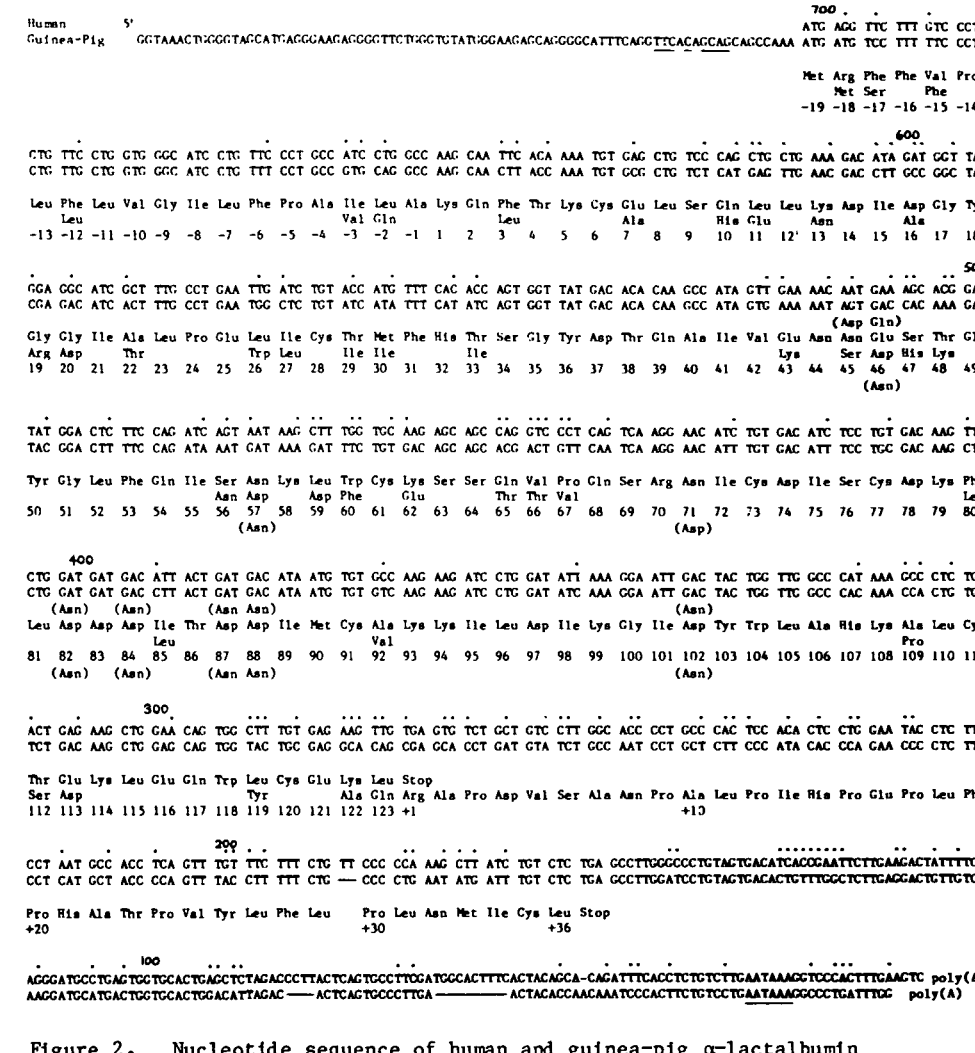


Figure 2. Nucleotide sequence of human and guinea-pig α-lactalbumin cDNAs contained within recombinant plasmids phB-35 and pgpK-9 respectively. The deduced amino acid sequence of human pre-α-lactalbumin is shown in full (residues -19 to -1 refer to the presequence, residues 1 to 123 represent the secreted protein). Residues which differ in guinea-pig pre-α-lactalbumin are shown directly below the human sequence. Amino acids shown in brackets are those incorrectly assigned by protein sequencing studies (10,11). Base substitutions are indicated by dots after introduction of gaps to maximize homology. Underlined sequences in the 5' and 3' noncoding regions indicate a possible ribosome binding site and the supposed polyadenylation enzyme recognition sequence (see text).

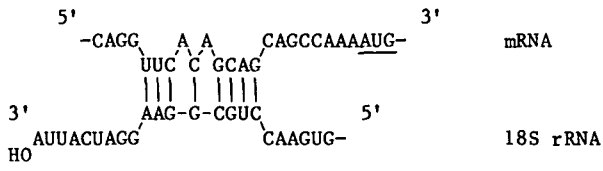
acid residues present in each signal peptide, 18 are hydrophobic (and one basic) in the human, whereas 16 are hydrophobic (and three neutral) in the guinea-pig preprotein. In particular, there is a leucine-rich region towards the amino-terminus of the presequence (residues -13 to -10), a feature found in many other signal sequences including those of other milk proteins (28,29). Furthermore the signal peptidase cleavage site (Ala-Lys) is also typical of secretory proteins with cleavage occurring after an amino acid with a small neutral side chain (usually Ala, Ser, Cys or Gly; see ref. 28).

Examination of the nucleotide-deduced amino acid sequence of the 123 residues comprising the mature human and guinea-pig α-lactalbumins (as found in milk), confirms the reported protein sequencing data (9,10) except for the commonly identified discrepancies between glutamic acid and glutamine or aspartic acid and asparagine residues (see Fig. 2), which can be difficult to distinguish from each other using protein sequencing techniques. Although the human nucleotide sequence has a termination codon (TGA) at the expected position immediately after residue 123, the guinea-pig sequence derived from the recombinant plasmid pgpK-9 does not. Instead there exists a CGA codon (Arg) at this position, the first in phase termination signal being 37 codons downstream, suggesting a 36 amino acid extension after the expected carboxyl-terminal of the major secreted protein. However, analysis of five other plasmids containing guinea-pig α-lactalbumin cDNA sequences derived from the same cDNA library and four other plasmids containing human α-lactalbumin sequences, all revealed a termination codon (TGA) in the expected position immediately following residue 123.

There are a number of possible explanations for this unexpected result. First, it is conceivable that the CGA codon in pgpK-9 arose by mistake due to misreading *in vivo* by RNA polymerase, *in vitro* by reverse transcriptase, or during plasmid replication. However, to the best of our knowledge this is the only discrepancy in our sequences. Apart from the easily accountable Glu/Gln and Asp/Asn errors, the nucleotide sequence confirms the protein sequence in all respects, and the nucleotide sequences contained within the other five recombinant plasmids containing guinea-pig α-lactalbumin cDNA were identical throughout the region examined, except for the termination codon. Interestingly, it has recently been reported that rat α-lactalbumin has a 17 residue carboxyl-terminal extension as revealed by cDNA sequence analysis (30), although in this instance all 14 α-

lactalbumin clones examined were identical. Consequently, although there are reasons for believing that the CGA codon in pgpK-9 is not an artifact, in vitro translation of α-lactalbumin-mRNA selected using immobilized pgpK-9 plasmid DNA (18), and subsequent gel electrophoresis of the products, revealed only a single protein band with the expected mobility of the 142 residue preprotein. This suggests that the synthesis of a larger form of α-lactalbumin containing a carboxyl-terminal extension, if it exists, is at a very low level, moreover, the biological significance of a proline-rich, hydrophobic extension has yet to be established. It is possible that α-lactalbumin may have an additional intracellular role, presumably related to glycosylation events, in which case the hydrophobic carboxyl-terminal extension may be required to anchor the functional part of the protein to a suitably hydrophobic membrane component within the secretory pathway. The existence of such a sequence would imply either multiple α-lactalbumin genes or alternative RNA splicing patterns. These possibilities have yet to be examined.

Examination of the noncoding regions of the human and guinea-pig α-lactalbumin cDNA sequences revealed a number of interesting structural features. First, in common with other eucaryotic mRNAs, both α-lactalbumin sequences possessed the supposed polyadenylation enzyme recognition sequence AAUAAA close to the site of poly(A) addition (31). Second, the guinea-pig sequence contained a region 9-18 nucleotides upstream from the initiating AUG codon (see Fig. 2) capable of base pairing with the highly conserved purine-rich region near the 3' end of 18S ribosomal RNA (32):



Although many of the sequenced eucaryotic mRNAs exhibit some degree of complementarity with this purine-rich region it is by no means universal. Furthermore, the distance between the postulated base-pairing and the initiating AUG is highly variable, much more so than with bacterial mRNAs (33). Consequently, although it seems unlikely that mRNA-18S rRNA base pairing is essential for translational activity, it is possible that it may be one of a series of contributory factors involved in the mRNA-ribosome

interaction, particularly in the case of mRNAs with relatively long 5' noncoding sequences (32). In this respect it may play a part in ensuring that initiation occurs at the correct AUG codon in guinea-pig α-lactalbumin mRNA, and not at either of the two additional AUG codons found upstream (see Fig. 2).

A further, somewhat unusual feature of the 5'-noncoding region of the guinea-pig α-lactalbumin cDNA sequence was the presence of six tracts of 3 or 4 G residues. Although the biological significance of these, if any, is unknown, it has proved to be a major factor in preventing the generation of full-length cDNA clones, since reverse transcription of mRNA is frequently terminated at these tracts of G residues, presumably due to stable secondary structure in this region, as indicated by primer extension experiments. Using a 51 nucleotide Hae III - Hpa II DNA fragment from the 5' end of pgpK-9 (see Fig. 1) as a primer for dideoxynucleotide sequencing of guinea-pig α-lactalbumin mRNA (see Materials and Methods), we have confirmed the 5' sequence of the cloned cDNA shown in Fig. 2 as far as the first tract of G residues (28 residues upstream from the initiating AUG codon). Beyond this point reverse transcription, and consequently isotope incorporation, was extremely low. Similar problems have been encountered during primer extension experiments with human α-lactalbumin mRNA (which also seems to possess tracts of G residues in similar positions in the 5'-noncoding region) though we have nevertheless been able to confirm the nucleotide sequence encoding the amino acids of the amino-terminal extension sequence together with an additional three A residues (as in the guinea-pig sequence) immediately preceding the initiating AUG codon.

Homology between human and guinea-pig α-lactalbumin cDNA sequences

Comparison of the nucleotide sequence within the common coding regions (residues -19 to 123) of human and guinea-pig pre-α-lactalbumins revealed greater nucleotide (80%) than amino acid (70%) sequence homology (Table 1, Fig. 2). Over half (75/142) of the codons were identical and a further 51 differed by only a single base. Of these, 24 (47%) represented silent substitutions, a silent mutation frequency far in excess of the value of 25% expected on a purely random basis (34). Similar high rates of silent mutations have often been reported between homologous nucleotide sequences from different organisms (35), and may reflect an evolutionary pressure to conserve certain structural elements. Indeed it can be seen from Fig. 2 that nucleotide and amino acid differences were not distributed randomly throughout the coding region, but instead there was a tendency for the

Table 1 Nucleotide and deduced amino acid sequence comparisons between human and guinea-pig pre-α-lactalbumin and hen pre-lysozyme cDNAs

	hu pre-αLA/ gp pre-αLA		hu pre-αLA/ hen pre-lysozyme		gp pre-αLA/ hen pre-lysozyme	
	No.	%	No.	%	No.	%
Amino acid sequence						
Identical amino acids	100/142	70	53/140	38	45/140	32
Nucleotide sequence						
Identical nucleotides - total	526/682	77	-	-	-	-
- coding region	343/429	80	223/423	53	204/423	48
- 3' noncoding region	183/253	72	-	-	-	-
Identical codons	75/142	53	31/141	22	24/141	17
Two identical nucleotides/codon	51/142	36	42/140	30	40/140	29
- silent substitutions	24/51	47	23/42	55	22/40	55
- replacement substitutions	27/51	53	19/42	45	18/40	45

Notes: cDNA sequences were aligned with gaps introduced to maximize homologies as shown in Figs. 2 and 3. Silent and replacement substitutions were calculated only for codons differing by a single base. The putative carboxyl-terminal extension of guinea-pig α-lactalbumin suggested by sequence analysis of pgpK-9 was not included as part of the coding region for homology comparisons.

centre of the protein (particularly residues 69 to 108) to be more highly conserved than the ends. If so, this presumably represents a structurally-imposed conservation since the residues considered to be essential for α-lactalbumin activity, namely Phe-31, His-32, Tyr-50, Trp-104 and Trp-118 (36,37), are mostly outside this central region. In contrast, examination of the signal sequence alone (residues -19 to -1), demonstrated that although the degree of homology was still high, all but one of the six single base-change codon differences led to an altered amino acid. However, substitutions were such that the overall hydrophobicity of the signal sequence was not affected. This supported the view that the hydrophobic nature, and not the precise amino acid sequence of the signal peptide, is important in the interaction of nascent secretory proteins with the membranes of the endoplasmic reticulum.

Analysis of the 3'-noncoding regions of the human and guinea-pig α-lactalbumin cDNA sequences after the introduction of gaps to maximize alignment (Fig. 2), also revealed extensive nucleotide homology (72%), although not as high as that of the coding regions (80%). This is consistent with the general hypothesis that nucleotide sequence divergence is slower in structurally important regions, but faster in areas where there are no RNA or protein sequence constraints (35,38).

Codon usage in human and guinea-pig α-lactalbumin mRNA

The codon usage in human and guinea-pig α-lactalbumin mRNA was nonrandom and followed the pattern found with other eucaryote mRNAs (39,40). In particular, the dinucleotide -CG-, which is very infrequent in eucaryotic DNA, appeared in only one codon in the human α-lactalbumin and three times in the guinea-pig α-lactalbumin. Also, in common with many other eucaryotic mRNAs, there was a marked preference for C or G in the third position of the codon (human α-LA: 61% G+C, guinea-pig α-LA: 59% G+C), despite a value of only 42% for the overall G+C content of total vertebrate DNA (41). The least common nucleotide in the third position was A, appearing in only 14% of the α-lactalbumin codons.

Homology between α-lactalbumin and chick lysozyme cDNA sequences

Comparison of the nucleotide sequence of the coding region of human pre-α-lactalbumin cDNA with that of chick pre-lysozyme cDNA (19) after the introduction of a number of gaps to maximise the alignment again revealed greater nucleotide (53%) than amino acid (38%) homology (Figure 3, Table 1). Furthermore, of the 42 codons which differed by only a single base, 55% represented silent substitutions. These observations strengthen

Figure 3. Comparison of the coding regions of human α-lactalbumin and hen lysozyme cDNAs. Base substitutions are indicated by dots after introducing gaps to maximize homology. The hen lysozyme cDNA sequence was obtained from ref 19.

the view that the α-lactalbumin and lysozyme genes have evolved from a common ancestor, and argue against two distinct ancestral genes giving rise to proteins of similar sequence. Comparison of the coding region of the guinea-pig pre-α-lactalbumin cDNA sequence (excluding the putative carboxyl-terminal extension) with that of chick pre-lysozyme cDNA produced similar results (Table 1).

In contrast, comparison of the 3'-noncoding region of either human or guinea-pig α-lactalbumin cDNA with that of chick lysozyme produced little, if any homology with the possible exception of the first 27 nucleotides after the termination codon. Furthermore, the lengths of the 3'-noncoding regions were markedly different; 276 nucleotides [excluding poly(A) tail] for human α-lactalbumin cDNA but only 118 nucleotides for chick lysozyme cDNA (19), suggesting that the 3'-noncoding region has diverged more rapidly than the coding region, in common with other published data (35,38).

Conclusion

We have shown that both the coding and 3'-noncoding regions of human and guinea-pig α-lactalbumin mRNAs exhibit a high degree of homology, and this extends to the coding region of chick lysozyme mRNA, observations consistent with evolution from a common ancestral gene. In addition, sequence analysis has provided evidence for the presence of a minor guinea-pig α-lactalbumin mRNA species encoding a protein with a carboxyl-terminal extension. Analysis at the genomic level will confirm whether or not multiple α-lactalbumin genes exist, and also to what extent, if any, the positions and sequences of the introns and hence protein domains are similar to those in the lysozyme genes (see ref. 19). The construction and sequencing of human and guinea-pig lysozyme cDNA clones and cloned α-lactalbumin and lysozyme cDNA from other species will extend these studies, and will enable predictions to be made as to the approximate date at which the divergence of α-lactalbumin and lysozyme genes occurred as a result of gene duplication.

ACKNOWLEDGEMENTS

We thank Miss Fleur Kreel and Dr. M.S. Davies for providing the primer extension data, and Dr. Janette Allison and Mr. D. Parker for performing preliminary plasmid restriction analysis. We also thank the Cancer Research Campaign for supporting this work, and Dr. J. Fox and Mr. R. Young for providing access to the ICRF computer facilities for analysis of DNA sequencing data.

*To whom correspondence should be addressed.

REFERENCES

1. Dayhoff, M.O. (1972) in Atlas of Protein Sequence and Structure, vol. 5, National Biomedical Research Foundation, Silver Spring, Maryland, USA.
2. Cooke, N.E., Coit, D., Shine, J., Baxter, J.D. & Martial, J.A. (1981) J. Biol. Chem. 256, 4007-4016.
3. Cooke, N.E., Coit, D., Weiner, R.I., Baxter, J.D. & Martial, J.A. (1980) J. Biol. Chem. 255, 6502-6510.
4. Brew, K., Vanaman, T.C. & Hill, R.L. (1967) J. Biol. Chem. 242, 3747-3749.
5. Brodbeck, U. & Ebner, K.E. (1966) J. Biol. Chem. 241, 762-764.
6. Brew, K., Vanaman, T.C. & Hill, R.L. (1968) Proc. Natl. Acad. Sci. USA 59, 491-497.
7. Craig, R.K., McIlreavy, D. & Hall, R.L. (1978) Biochem. J. 173, 633-641.
8. Phillips, D.C. (1967) Proc. Natl. Acad. Sci. USA 57, 484-495.

9. Findlay, J.B.C. & Brew, K. (1972) Eur. J. Biochem. 27, 65-86.
10. Brew, K. (1972) Eur. J. Biochem. 27, 341-353.
11. Brew, K., Castellino, F.J., Vanaman, T.C. & Hill, R.L. (1970) J. Biol. Chem. 245, 4570-4582.
12. Macgillivray, R.T.A., Brew, K. & Barnes, K. (1979) Arch. Biochem. Biophys. 197, 404-414.
13. Hopp, T.P. & Woods, K.R. (1979) Biochemistry 18, 5182-5191.
14. Canfield, R.E., Kammerman, S., Sobel, J.H. & Morgan, F.J. (1971) Nature New Biol. 232, 16-17.
15. Canfield, R.E. & Liu, A.K. (1965) J. Biol. Chem. 240, 1997-2002.
16. White, T.J., Mross, G.A, Osserman, E.F. & Wilson, A.C. (1977) Biochem. 16, 1430-1436.
17. Hall, L., Davies, M.S. & Craig, R.K. (1981) Nucl. Acids Res. 9, 65-84.
18. Craig, R.K., Hall, L., Parker, D. & Campbell, P.N. (1981) Biochem. J. 194, 989-998.
19. Jung, A., Sippel, A.E., Grez, M. & Schutz, G. (1980) Proc. Natl. Acad. Sci. USA 77, 5759-5763.
20. Maxam, A.M. & Gilbert, W. (1980) in Methods in Enzymology, Grossman, L. and Moldave, K. Eds., Vol. 65, pp. 499-560, Academic Press.
21. Agarwal, K.L., Brunstedt, J. & Noyes, B.E. (1981) J. Biol. Chem. 256, 1023-1028.
22. Hamlyn, P.H., Gait, M.J. & Milstein, C. (1981) Nucl. Acids Res. 9, 4485-4494.
23. Blobel, G. & Dobberstein, B. (1975) J. Cell Biol. 67, 835-851.
24. Blobel, G. & Dobberstein, B. (1975) J. Cell Biol. 67, 852-862.
25. Hall, L., Craig, R.K. & Campbell, P.N. (1979) Nature 277, 54-56.
26. Craig, R.K., Brown, P.A., Harrison, O.S., McIlreavy, D. & Campbell, P.N. (1976) Biochem. J. 160, 57-74.
27. Pascall, J.C., Boulton, A.P., Parker, D., Hall, L. & Craig, R.K. (1981) Biochem. J. 196, 567-574.
28. Steiner, D.F., Quinn, P.S., Chan, S.J., Marsh, J. & Tager, H.S. (1980) Ann. N.Y. Acad. Sci. 343, 1-16.
29. Mercier, J-C. & Gaye, P. (1980) Ann. N.Y. Acad. Sci. 343, 232-251.
30. Dandekar, A.M. & Qasba, P.K. (1981) Proc. Natl. Acad. Sci. USA 78, 4853-4857.
31. Proudfoot, N.J. & Brownlee, G.G. (1974) Nature 252, 359-362.
32. Hagenbuchle, O., Santer, M. & Steitz, J.A. (1978) Cell 13, 551-563.
33. Steitz, J.A. (1978) in Biological Regulation and Development, Goldberger R. Ed., Plenum Publ. Co., New York.
34. Jukes, T.H. & King, J.L. (1979) Nature 281, 605-606.
35. Miyata, T., Yasunaga, T. & Nishida, T. (1980) Proc. Natl. Acad. Sci. USA 77, 7328-7332.
36. Prieels, J-P., Bell, J.E., Schindler, M., Castellino, F.J. & Hill, R.L. (1979) Biochemistry 18, 1771-1776.
37. Richardson, R.H. & Brew, K. (1980) J. Biol. Chem. 255, 3377-3385.
38. Nishioka, N. & Leder, P. (1979) Cell 18, 875-882.
39. Grantham, R., Gautier, C., Gouy, M., Jacobzone, M. & Mercier, R. (1981) Nucl. Acids Res. 9, r43-r74.
40. Wain-Hobson, S., Nussinov, R., Brown, R.J. & Sussman, J.L. (1981) Gene 13, 355-364.
41. Sueoka, N. (1961) J. Mol. Biol. 3, 31-40.