Volume 13 Number 5 1985 Nucleic Acids Research

Nucleotide sequence of a chicken δ-crystallin gene

Mutsuhito Ohno, Hiroshi Sakamoto, Kunio Yasuda, T.S.Okada and Yoshiro Shimura*

Department of Biophysics, Faculty of Science, Kyoto University, Kyoto 606, Japan

Received 14 January 1984; Revised and Accepted 1 February 1985

ABSTRACT

We have determined the complete nucleotide sequence of one of the two non-allelic δ-crystallin genes in the chicken, arbitrarily designated δ-gene 1, using a genomic clone (λgδ106) containing the entire gene sequence. By comparison of the genomic sequence and the δ-crystallin cDNA sequence previously determined, we have identified exon sequences in the genomic sequence. Thus, the presence of 17 exons and 16 introns in the gene has been clarified. The δ-crystallin polypeptide deduced from the exon sequences consists of 465 amino acids which is larger, by 19 amino acid residues, than the polypeptide deduced from the cDNA sequence previously reported. Re-examination of the cDNA sequence using the same cDNA clone previously used shows that the present exon sequences are correct and the molecular weight of the deduced δ-crystallin polypeptide is 50,615 daltons instead of the previously reported value of 48,447 daltons. In addition, some structural features of the δ-crystallin gene including putative expression signals are discussed.

INTRODUCTION

Crystallins are a group of soluble proteins that represent the major constituents of the vertebrate ocular lens (1,2). Each of four immunologically distinct crystallin families (α, β, γ and δ) is known to have a specific spatial and temporal expression in the developing tissues (3-5). δ-Crystallin, found in avians and reptilians, is the first major soluble protein that appears in the developing chicken lens (5,6), and has been a useful index for studying differential gene expression in cell differentiation (7-9). Chicken δ-crystallin is a tetrameric protein of about 200,000 daltons, which is composed of at least two similar polypeptides of approximate molecular weights of 48,000-50,000 daltons (10).

Recently, considerable efforts have been directed toward the study of chicken δ-crystallin genes (11). Many chicken δ-crystallin cDNA clones have been constructed and analyzed (12-15), and the nucleotide sequences of full-length cDNAs have been determined (16,17). A number of DNA fragments containing δ-crystallin gene sequences have also been cloned and characterized (15, 18-20). The presence of two non-allelic δ-crystallin genes has been

thus demonstrated. The two δ-crystallin genes are oriented on the same chromosome with similar 5'-3' polarity and separated from each other by approximately 4 kilobases (20,21). Of the two chicken δ-crystallin genes, the gene which is located directionally upstream has been arbitrarily designated δ-gene 1 (or δ1 according to the designation by Hawkins et al. (20)), while the other has been designated δ-gene 2 (or δ2). Both genes are highly interrupted with initial estimates of 13-15 intervenining sequences (15,18, 19). However, the estimate of intron number has varied with an upward trend and has recently been reported to be 16 (20). Presumably, these variable results are mainly due to inherent limitations of the electron microscopic analyses employed in the previous studies. Thus, the precise organization and structure of δ-crystallin genes still remain to be clarified.

In this paper we report the nucleotide sequences of δ-gene 1 and its flanking regions. The organization and structural features of the chicken δ-crystallin gene have been thus unraveled.

MATERIALS AND METHODS

DNA Sources. As the source of δ-crystallin gene (δ-gene 1), a genomic DNA clone (λCδ106) isolated previously (15) was used. This clone contains an insert of approximately 15 kb, which includes the entire δ-crystallin gene sequences (15).

DNA Sequencing. The DNA sequences of the δ-crystallin gene and its flanking regions were determined by the dideoxynucleotide chain termination method (22) using restriction fragments cloned into M13mp8, M13mp9, M13mp18, or M13mp19 (23,24) as template. Some DNA fragments were sequenced by the chemical degradation method (25).

Materials. DNA primers and other reagents for DNA sequencing were purchased from Takara Shuzo, Inc. Restriction enzymes and other enzymes were from Takara Shuzo, Inc. and Nippon Gene, Inc. Radioactive materials were from Amersham.

RESULTS

Nucleotide Sequences of δ-Gene 1 and Its Flanking Regions. We have previously shown that one of our genomic DNA clones, λCδ106, contains the entire gene sequences complementary to δ-crystallin mRNA except for the 3' terminal poly(A) tail (15). Subsequently, the δ-crystallin gene in the clone has been shown to represent δ-gene 1, the directionally 5'-side gene of the two non-allelic δ-crystallin genes (21). The insert of λCδ106 consists of four EcoRI fragments of 7.6 kb, 4.0 kb, 2.6 kb, and 0.8 kb, and the gene

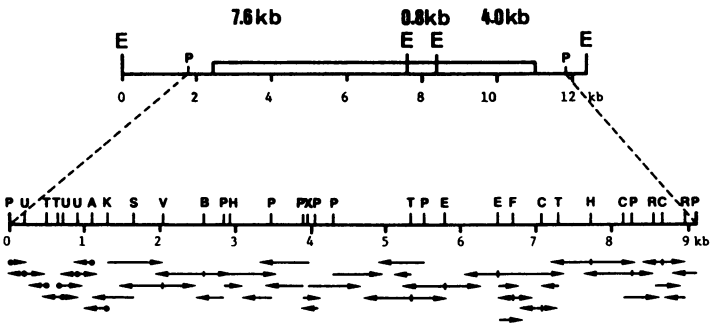
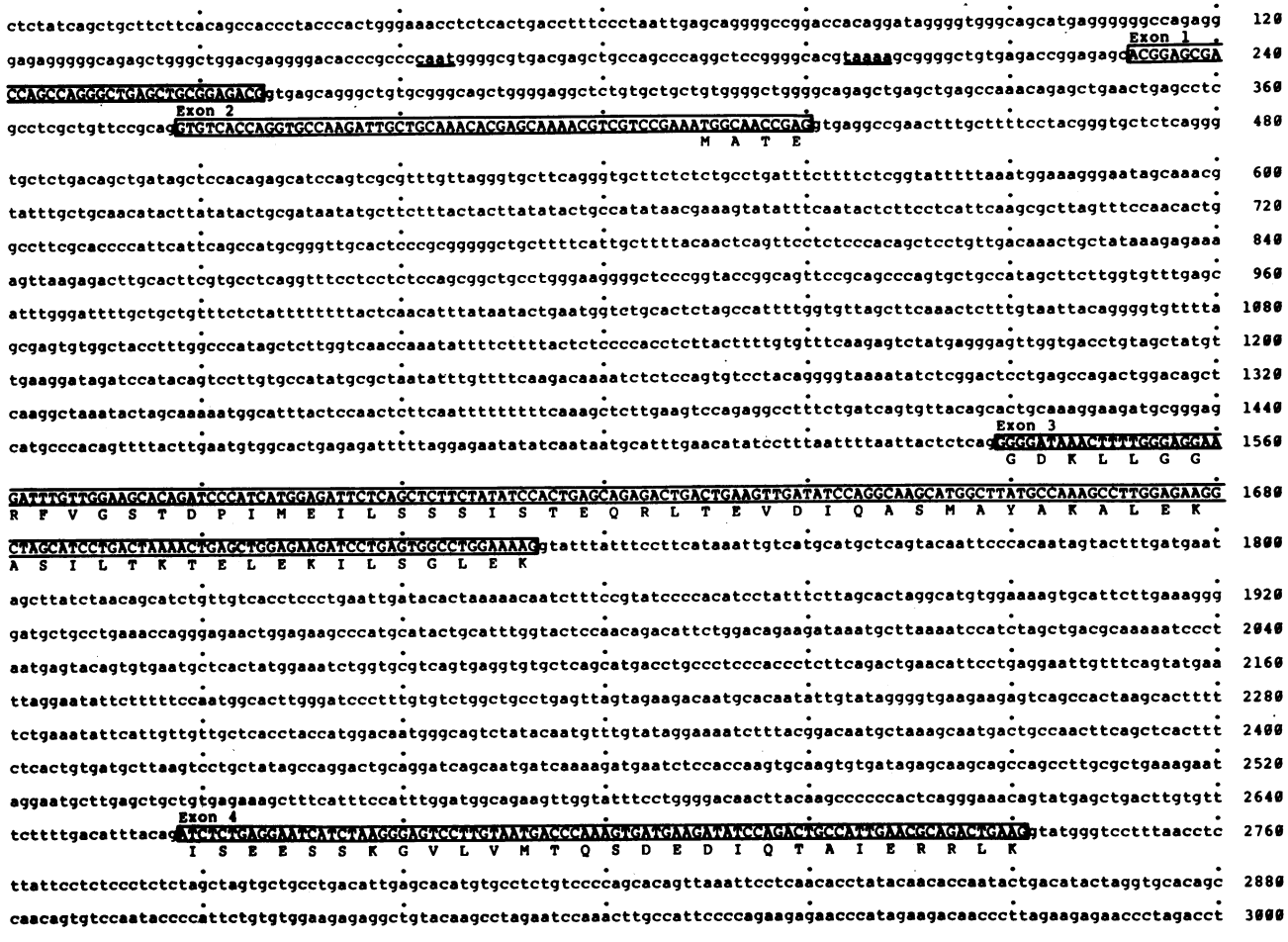


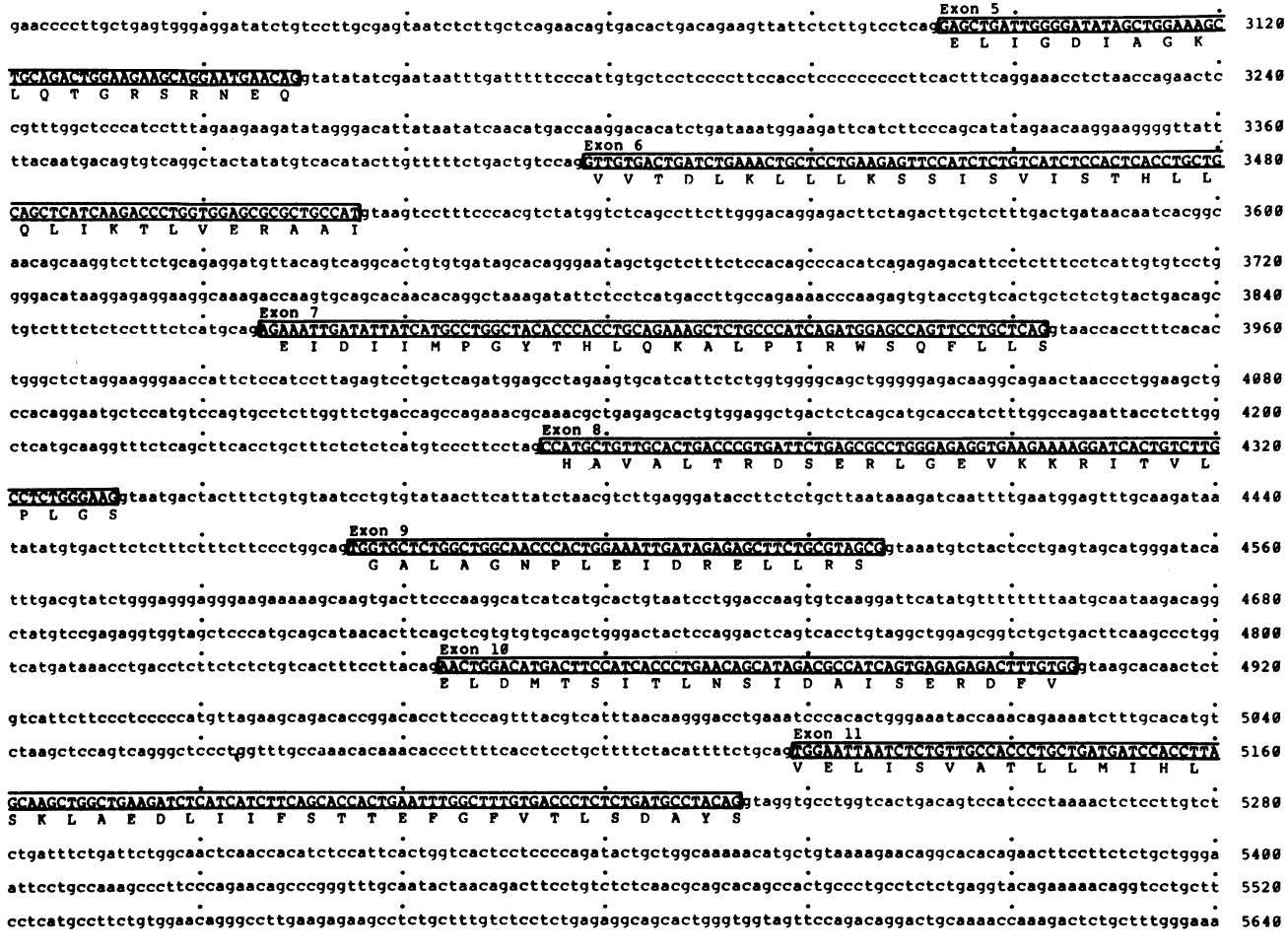
Figure 1. Sequencing strategy for δ-gene 1 and its flanking regions. In the upper part of the figure, the open bar represents the δ-gene 1 sequence. The following restriction endonuclease cleavage sites are shown: E, EcoRI; P, PstI; K, KpnI; H, HindIII; B, BamHI; V, EcoRV; X, XbaI; T, HpaII; A, HaeIII; U, PvuII; S, Sau3AI; C, HincII; F, FokI; R, RsaI. The sequencing strategy is shown below the restriction map. The region between the PstI site at 0 and the KpnI site at 1.3 kb in the lower part of the figure was sequenced by the chemical degradation method. Small dots show the 5' termini of restriction fragments labeled with ^{32}P. The remaining portions were sequenced by the dideoxy method. In both cases, arrows show the directions of sequencing. The lines behind the arrowheads indicate portions that actually yielded satisfactory sequences.

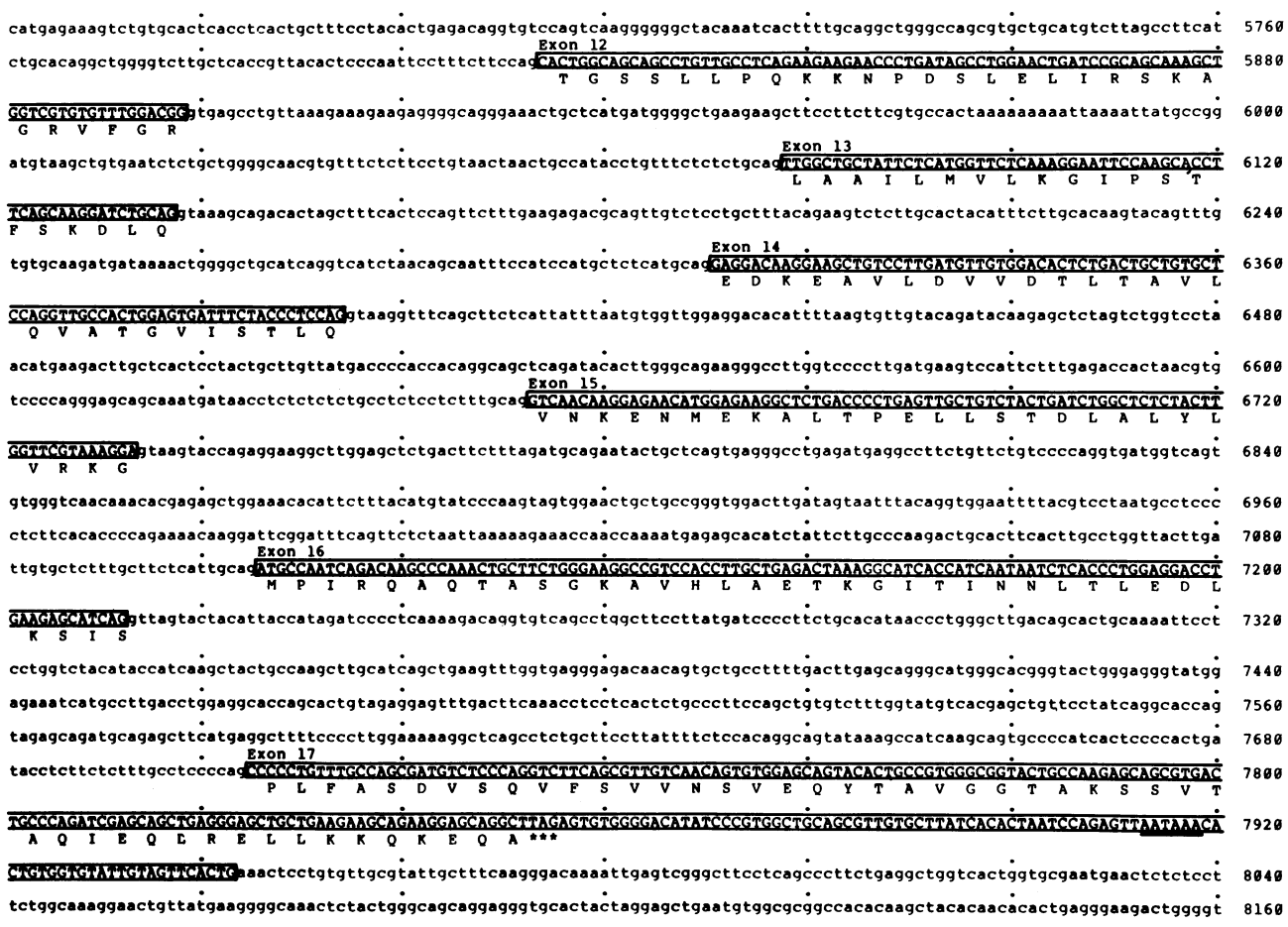
sequences reside in the (5')7.6 kb - 0.8 kb - 4.0 kb(3') fragments (15).

Using the three EcoRI fragments derived from the genomic clone, we determined the nucleotide sequences of the δ-gene 1 and its flanking regions. Figure 1 shows the strategy for sequencing the regions in the fragments. The majority of the sequences were determined by the dideoxynucleotide chain termination method and only a few portions of the sequences were determined by the chemical degradation method. It was possible to generate fragments and strands that would overlap most of the restriction sites within the region between the PstI site at position 1.7 kb and the PstI site at position 11.8 kb in the figure, except for the two EcoRI sites. As one of the EcoRI sites is present within an exon (Exon 13), the sequences flanking the EcoRI site were checked with the δ-crystallin cDNA sequence which had been determined previously (16). The other EcoRI site was found to be within an intron (Intron K) and was not overlapped in the analysis. The nucleotide sequence thus determined is shown in Fig. 2.

Structure of δ-Gene 1. Comparison of the genomic DNA sequence with the δ-crystallin cDNA sequence allowed us to identify the exon sequences within the genomic sequence. Thus, we have found that cDNA is split into 17 pieces as indicated in Fig. 2. The exact chain lengths of 17 exons and 16 introns


Exon 1
Exon 2
M A T E
Exon 3
G D K L L G G
R F V G S T D P I M E I L S S S I S T E Q R L T E V D I Q A S M A Y A K A L E K
A S I L T K T E L E K I L S G L E K
Exon 4
I S E E S S K G V L V M T Q S D E D I Q T A I E R R L K


Exon 5
E L I G D I A G K
L Q T G R S R N E Q
Exon 6
V V T D L K L L L K S S I S V I S T H L L
Q L I K T L V E R A A I
Exon 7
E I D I I M P G Y T H L Q K A L P I R W S Q F L L S
Exon 8
H A V A L T R D S E R L G E V K K R I T V L
P L G S
Exon 9
G A L A G N P L E I D R E L L R S
Exon 10
E L D M T S I T L N S I D A I S E R D F V
Exon 11
V E L I S V A T L L M I H L
S K L A E D L I I F S T T E F G F V T L S D A Y S


Exon 12
Exon 13
Exon 14
Exon 15
Exon 16
Exon 17
T G S S L L P Q K K N P D S L E L I R S K A
G R V F G R
L A A I L M V L K G I P S T
F S K D L Q
E D K E A V L D V V D T L T A V L
Q V A T G V I S T L Q
V N K E N M E K A L T P E L L S T D L A L Y L
V R K G
M P I R Q A Q T A S G K A V H L A E T K G I T I N N L T L E D L
K S I S
P L F A S D V S Q V F S V V N S V E Q Y T A V G G T A K S S V T
A Q I E Q L R E L L K K Q K E Q A ***

are listed in Table 1. As noted in the table, the lengths of introns are more variable than those of exons; the smallest (Intron A) and largest (Intron B) introns are 111 bp and 1,098 bp, respectively, in length. Although the lengths of exons are within a range of 35 bp and 239 bp, most exons are 60-90 bp in length. The organization of the δ-gene 1 is schematically shown in Fig. 3. It is worth noting that the G-C contents of exons are generally higher than those of introns (Table 1). The G-C contents of Exon 1 (68.6%) which represents the 5' non-coding region, and of Intron A (66.7%) which adjoins Exon 1 are much higher than those of other exons and introns.

The exon sequences determined in the present study differ from the δ-crystallin cDNA sequence reported previously (16) by about a dozen of nucleotides. Most of the differences are base transitions. The transitions at the third letters of codons (amino acid residues 40, 77, 135, 251, 293, 410, and 441) result in no change in amino acids, whereas those at the first and/or second letters of codons (amino acid residues 118, 129, 274, 298, 345, and 348) lead to amino acid changes. The most significant difference in the present exon sequences is the presence of an A residue at position 86 from the 5' end of Exon 17, which was absent in the previous cDNA sequence. This leads to shift of the reading frame of the cDNA sequence from the position. Consequently, the stop codon (UGA) of the previous cDNA sequence is no longer in frame and, instead, 19 additional amino acids are encoded before appearance of a new stop codon (UAG) 149-151 bp downstream from the 5' end of Exon 17. Thus, the total number of amino acid residues of δ-crystallin is 465 excluding the first methionine residue instead of 446, the value we reported previously (16). The molecular weight of the deduced δ-crystallin polypeptide must be 50,615 daltons excluding the first methionine instead of the previously reported value of 48,447 daltons. The insertion of the A residue is not due to a polymorphism within the δ-crystallin locus but rather to an experimental error in the cDNA sequencing, since upon re-examination of the cDNA sequence in pBδ11, the clone that we used in the previous sequence analysis, the A residue was also identified in the sequence (data not shown).

There are introns that split exon sequences at the boundaries of two codons (Introns B, C, D, E, L, M, N, and O), whereas some introns split exons

Figure 2. Nucleotide sequences of δ-gene 1 and its flanking regions. The sequence (sense strand) is presented from the 5' to 3' direction and is numbered arbitrarily. The exon sequences are enclosed in boxes. Amino acid residues encoded are indicated by single letter desigantions. The intron sequences are presented in small letters. The CAAT and Goldberg-Hogness boxes and the consensus polyadenylation signal are underlined.

Table 1. Size and G-C content of exons and introns of δ-gene 1

Exon			Intron		
	Size (bp)	% G-C		Size (bp)	% G-C
1	35	68.6			
			A	111	66.7
2	63	54.0			
			B	1098	43.6
3	195	47.2			
			C	924	42.8
4	84	44.1			
			D	351	47.6
5	57	49.1			
			E	268	41.0
6	98	54.1			
			F	350	48.6
7	78	50.0			
			G	310	53.2
8	78	55.1			
			H	143	35.0
9	53	54.7			
			I	316	47.8
10	63	49.2			
			J	212	46.7
11	115	47.8			
			K	580	51.7
12	85	56.5			
			L	179	45.3
13	60	46.7			
			M	173	43.9
14	84	53.6			
			N	258	47.3
15	81	49.4			
			O	372	45.7
16	107	51.4			
			P	492	51.0
17	239	54.4			

inside of particular codons (Introns F, G, H, I, J, K, and P). In the latter case, four out of the seven codons that are split internally are those for serine (AGY), in which splitting occurs between the second and third letters. For all the introns in the δ-crystallin gene, the sequences of the splice points follow the GU-AG rule, and the junction sequences of the donor and acceptor sites are in good agreement with the consensus sequences (26-28) as shown in Table 2.

In all the introns of the δ-crystallin gene, we have found sequences (CTPuAC) that are homologous to the yeast TACTAAC box (29) and also to the PyXPyTPuAPy consensus sequence of globin introns (30) near the 3' splice

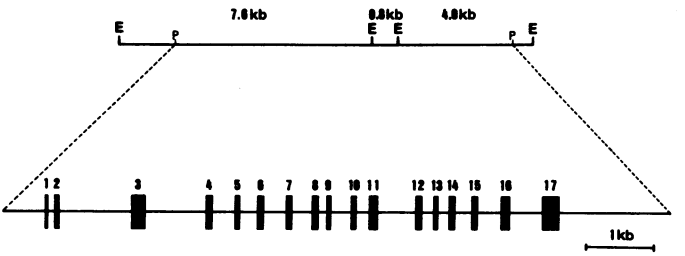


Figure 3. Schematic representation of δ-gene 1. In the upper part of the figure, the EcoRI 7.6kb-0.8kb-4.0kb fragments are shown. The following restriction sites are shown: E, EcoRI; P, PstI. In the lower part of the figure, the thick solid bars represent exons and the thin lines attached to them show intervening and flanking sequences.

sites (Table 3). Also noted in intron sequences is the presence of short stretches of sequences (6-22 bp) that are complementary to the 3' terminal regions of immediately preceding exons. These intron sequences are present between the CTPuAC sequences described above and the 3' termini of intron

Table 2. Nucleotide sequences of splice sites of δ-gene 1

Donor Site	Acceptor Site
(Exon 1)---GACG ↓gtgagc---(Intron A)---ttccgcag↓ GTG---(Exon 2)	
(Exon 2)---CGAG gtgagg---(Intron B)---actctcag GGG---(Exon 3)	
(Exon 3)---AAAG gtattt---(Intron C)---atttacag ATC---(Exon 4)	
(Exon 4)---GAAG gtatgg---(Intron D)---gtcctcag GAG---(Exon 5)	
(Exon 5)---ACAG gtatat---(Intron E)---ctgtccag GTT---(Exon 6)	
(Exon 6)---CCAT gtaagt---(Intron F)---tcatgcag AGA---(Exon 7)	
(Exon 7)---TCAG gtaacc---(Intron G)---cttcctag CCA---(Exon 8)	
(Exon 8)---GAAG gtaatg---(Intron H)---cctggcag TGG---(Exon 9)	
(Exon 9)---AGCG gtaaat---(Intron I)---ccttacag AAC--(Exon 10)	
(Exon 10)--GTGG gtaagc---(Intron J)---ttctgcag TGG--(Exon 11)	
(Exon 11)--ACAG gtaggt---(Intron K)---tcttccag CAC--(Exon 12)	
(Exon 12)--ACGG gtgagc---(Intron L)---ctctgcag TTG--(Exon 13)	
(Exon 13)--ACAG gtaaag---(Intron M)---tcatgcag GAG--(Exon 14)	
(Exon 14)--CCAG gtaagg---(Intron N)---ctttgcag GTC--(Exon 15)	
(Exon 15)--AGGA gtaagt---(Intron O)---cattgcag ATG--(Exon 16)	
(Exon 16)--TCAG gttagt---(Intron P)---ctccccag CCC--(Exon 17)	

Capital letters and small letters represent exon and intron sequences, respectively.

Table 3. Characteristic pentanucleotide sequences and sequences complementary to the 3' terminal regions of immediately preceding exons in the introns of δ-gene 1

Intron A	-----(CTGAG)---13bp----------TTCCGCAG/	Exon 2
Intron B	-----(ATAAT)---16bp----------TTtaATTT-13bp/	Exon 3
Intron C	-----(CTGAC)---23bp----------CaTTTaCAG/	Exon 4
Intron D	-----(CTGAC)---11bp----------TCTTGTC--5bp/	Exon 5
Intron E	-----(GTCAC)------ATACTTGTTTTTCTGaCTgTCCAG/	Exon 6
Intron F	-----(CTGAC)---21bp----------CATGCAG/	Exon 7
Intron G	-----(CTGAC)---44bp----GCAAGgTTTCTCAG-34bp/	Exon 8
Intron H	-----(GTGAC)---11bp----------TCTTCCCtGGCAG/	Exon 9
Intron I	-----(CTGAC)----5bp--CTCTCTGTCAcTTTCC--6bp/	Exon 10
Intron J	-----(CTAAG)---11bp--GTTTGCCaaaCACAAA-39bp/	Exon 11
Intron K	-----(TTCAT)---34bp----------CCcAATTC-11bp/	Exon 12
Intron L	-----(CTAAC)----6bp--------ACCtGTTTCT--8bp/	Exon 13
Intron M	-----(CTAAC)---20bp----------TCaTGCAG/	Exon 14
Intron N	-----(CTAAC)----6bp--CCagGGAGcaGcAAAT-33bp/	Exon 15
Intron O	-----(TTGAT)----6bp----------CTTTGC-12bp/	Exon 16
Intron P	-----(CTGAT)----2bp-----CTCTTC-8bp-TCCCCAG/	Exon 17

The pentanucleotide sequences are shown in parentheses. In the sequences at or near the 3' splice sites, capital letters represent sequences complementary to the exon regions.

sequences (Table 3).

Putative Expression Signals of δ-Gene 1. In the genomic DNA sequence shown in Fig. 2, a pentanucleotide sequence TAAAA and a tetranucleotide sequence CAAT are present 24-28 bp and 67-70 bp, respectively, upstream from the 5' end of the first exon. These sequences are consistent with the TATA and CAAT boxes, respectively. As will be described elsewhere, when a 3-kb fragment derived from the EcoRI 7.6-kb fragment by digestion with KpnI was transcribed in vitro with a HeLa cell extract and the transcription initiation site of the gene was mapped by the S1 nuclease protection method, the site was indistinguishable from the 5' terminus of the first exon (Exon 1). In addition, deletions of the TAAAA sequence of the gene described above completely abolished the synthesis of the run-off transcript in the in vitro transcription system (data not shown). It is highly likely that the two sequences consistent with the TATA and CAAT boxes represent signals for transcription initiation of the δ-crystallin gene.

In the 3' non-coding region, there is consensus poly(A) addition signal, AATAAA, 209-214 bp downstream from the 5' end of Exon 17. In the cDNA

sequence, the poly(A) tail begins 26 bp downstream of the sequence (16). In the genomic DNA sequence, there are three As 26-28 bp downstream of the signal sequence. It is not known, however, whether the AAA sequence in the genomic DNA are encoded in δ-crystallin mRNA as a part of the poly(A) tail or the sequence is removed from the transcript of the gene by processing before the addition of poly(A) tail.

DISCUSSION

We have determined the nucleotide sequences of a chicken δ-crystallin gene (δ-gene 1) and its flanking regions. The gene represents one of the two non-allelic δ-crystallin genes in the chicken and is expressed actively in the developing lens. The earlier electron microscopic analyses showed that the gene is interrupted with 13-15 intervening sequences (15,18,19). More recently, the presence of 16 intervening sequences in the gene has been reported on the basis of comparative R-loop analyses of the two chicken δ-crystallin genes (20). The present study has unequivocally shown that the δ-crystallin gene contains 16 introns and 17 exons. A close examination of our sequence data and of the electron microscopic data of Hawkins *et al.* (20) indicates that the exon-intron organization reported by the latter group is in fairly good agreement with ours, except for those in the terminal regions. According to Hawkins *et al.* (20), their first exon (Exon 1) is 95 bp long that is followed by an intron of 977 bp (Intron A). Apparently, their Exon 1 corresponds to our Exon 2. It is highly likely that our Exon 1 was missed in the electron microscopic analysis, since the exon is only 35 bp long and hard to visualize by the electron microscopic method. It is also possible that such a short terminal exon may not form stable RNA-DNA duplex due to branch migration. Exons 16 (70 bp) and 17 (70 bp) and Intron P (70 bp) which interrupts the two exons of Hawkins *et al.* (20) seem to correspond to our Exon 17 of 239 bp in length, as judged by the sizes, although the possiblity that the difference is due to a genetic polymorphism within the δ-crystallin genes analyzed by the two groups may not be ruled out. In any case, the results of the R-loop analysis of such a short exon-intron sequences must be interpreted with extreme caution.

The first intron (Intron A) is present within the 5'non-coding sequence. The remaining 15 introns interrupt the coding regions of the gene. In the approximate 5' half of the gene, introns (B, C, D, and E) interrupt the coding sequences at the boundaries of codons, so that codons are not interrupted internally. In the central portion of the gene, introns (F, G, H, I, J, and K) interrupt the coding sequences within single codons. The remaining

downstream introns (L, M, N, and O) except for Intron P interrupt the coding sequences between codons again. These interruption patterns of the δ-crystallin introns are of particular interest in view of the suggested notion that present structures of the chicken δ-crystallin genes have resulted from repeated duplicaion of an ancestral exon encoding a protein segment of stable structure (31).

The splice junctions of all of the introns in the δ-crystallin gene have the consensus sequences for the donor and acceptor sites of eukaryotic pre-mRNA (26-28). In all cases, the 5' and 3' termini of introns are GT and AG, respectively. All the introns contain sequences (CTPuAC) homologous to the yeast TACTAAC box sequence which is essential for splicing (29) or to the consensus sequence (PyXPyTPuAPy) in globin introns (30) near the 3' splice sites. The globin consensus has been proposed to represent the branch point sequences for splicing reaction (30). It remains to be seen whether the homologous sequences in the δ-crystallin gene play a role in pre-mRNA processing. In this connection, the intron sequences complementary to the 3' terminal regions of immediately preceding exons are of particular interest, since base pairings between the intron and exon regions may position two neighbouring exons to be ligated in close proximity. These features of the δ-crystallin gene will be detailed elsewhere.

The exon sequences determined in the present study are largely consistent with the δ-crystallin cDNA sequence previously reported (16,17), although there are some differences between the two sequences. Most of the differences are base transitions, which may reflect, at least in part, a polymorphism within the chicken δ-crystallin locus. The A residue located 86 bp downstream from the 5' end of Exon 17 is absent in the cDNA sequences that were reported previously by us (16) and also by Nickerson and Piatigorsky (17). When we re-examined the cDNA sequence of pBδ11, a cDNA clone previously employed in cDNA sequencing, the A residue was identified. Thus, at least in our case, the absence of the A residue in the cDNA sequence must be due to an experimental error. The insertion of the A residue in the cDNA sequences leads to disappearance of the stop codon previously assigned in the sequences (16). Instead, 19 additional codons have emerged before the new stop codon, UAG. A tryptic peptide Glu-Leu-Leu-Lys-(Lys) that is present in the additional sequence of 19 amino acids is consistent with the analytical results of de Jong *et al*. (32). Thus, we conclude that the number of amino acid residues of δ-crystallin is 465 excluding the first methionine and the molecular weight of the deduced δ-crystallin polypeptide is 50,615 daltons. The major chicken δ-crystallin polypeptide has been reported to have a molecular weight

of approximately 48,000 daltons, as judged by SDS-polyacrylamide gel electrophoresis (33). It is likely that the δ-crystallin polypeptide whose amino acid sequence has been deduced in the present study corresponds to the major δ-crystallin polypeptide reported previously, although the difference in the molecular weights remains to be clarified.

It has been reported that the two non-allelic δ-crystallin genes are arranged in a similar and characteristic pattern of 17 exons/16 introns (20). We have sequenced approximately one half of the δ-gene 2. Comparison of the two δ-crystallin gene sequnces shows that the sequences of corresponding exons are highly homologous, while those of corresponding introns differ considerably from each other. The sequence of δ-gene 2 will be detailed elsewhere. It is anticipated that further analyses of the two δ-crystallin genes will provide much insights into the mode of their expression.

ACKNOWLEDGEMENTS

We thank Dr. H. Ozeki for valuable discussions. This work was supported in part by research grants from the Ministry of Education, Science, and Culture of Japan to Y. S., K. Y. and T. S. O.

*To whom correspondence should be addressed

REFERENCES

1. Harding, J. J. and Dilley, K. J. (1976) Exp. Eye Res. 22, 1-73.
2. Bloemendal, H. (1977) Science 197, 127-138.
3. Clayton, R. M. (1974) in The Eye, Davson, H. and Graham, L. T. Eds., Vol. 5, pp. 399-494, Academic Press, London.
4. Papaconstantinou, J. (1967) Science 156, 338-346.
5. Zwaan, J. and Ikeda, A. (1968) Exp. Eye Res. 7, 301-311.
6. Rabey, M. (1962) Exp. Eye Res. 1, 310-316.
7. Katoh, A. and Yoshida, K. (1973) Exp. Eye Res. 15, 353-360.
8. Okada, T. S. (1980) in Current Topics in Developmental Biology, Moscona, A. A., Monroy, A., and Hunt, H. K. Eds., Vol. 16, pp. 349-380, Academic Press, New York.
9. Piatigorsky, J. (1981) Differentiation 19, 134-153.
10. Piatigorsky, J., Zelenka, P. and Simpson, R. T. (1974) Exp. Eye Res. 18, 435-446.
11. Piatigorsky, J. (1984) Mol. Cell Biochem. 59, 33-56.
12. Bhat, S. P. and Piatigorsky, J. (1979) Proc. Natl. Acad. Sci. USA. 76, 3299-3303.
13. Nickerson, J. M. and Piatigorsky, J. (1982) FEBS Lett. 144, 289-292.
14. Bower, D. J., Errington, L. H., Wainwright, N. R., Sime, C. Morris, S. and Clayton, R. M. (1981) Biochem. J. 201, 339-344.
15. Yasuda, K., Kondoh, H., Okada, T. S., Nakajima, N. and Shimura, Y. (1982) Nucleic Acids Res. 10, 2879-2891.
16. Yasuda, K., Nakajima, N., Isobe, T., Okada, T. S. and Shimura, Y. (1984) EMBO J. 3, 1397-1402.

17. Nickerson, J. M. and Piatigorsky, J. (1983) Proc. Natl. Acad. Sci. USA. 81, 2611-2615.
18. Jones, R. E., Bhat, S. P., Sullivan, M. A. and Piatigorsky, J. (1980) Proc. Natl. Acad. Sci. USA. 77, 5879-5883.
19. Bhat, S. P., Jones, R. E., Sullivan, M. A. and Piatigorsky, J. (1980) Nature (London) 284, 234-238.
20. Hawkins, J. W., Nickerson, J. M., Sullivan, M. A. and Piatigorsky, J. (1984) J. Biol. Chem. 259, 9821-9825.
21. Yasuda, K. and Okada, T. S. (1983) Develop. Growth and Differ. 25, 422.
22. Sanger, F., Nicklen, S. and Coulson, A. R. (1977) Proc. Natl. Acad. Sci. USA. 74, 5463-5467.
23. Messing, J. and Vieira, J. (1982) Gene 19, 269-276.
24. Morrander, J., Kempe, T. and Messing, J. (1983) Gene 26, 101-106.
25. Maxam, A. M. and Gilbert, W. (1980) Methods in Enzymol. 65, 499-560.
26. Breathnach, R. and Chambon, P. (1981) Ann. Rev. Biochem. 50, 349-383.
27. Sharp. P. A. (1981) Cell 23, 643-646.
28. Mount, S. M. (1982) Nucleic Acids Res. 10, 459-472.
29. Langford, C. J. and Gallwitz, D. (1983) Cell 33, 519-527.
30. Ruskin, B., Krainer, A. R., Maniatis, T. and Green, M. R. (1984) Cell 38, 317-331.
31. Moormann, R. J. M., den Dunnen, J. T., Bloemendal, H. and Schoenmakers, J. G. G. (1983) J. Mol. Biol. 171, 353-368.
32. de Jong, W. W., Stapel, S. O. and Zweers, A. (1981) Comp. Biochem. Physiol. 69B, 593-598.
33. Reszelbach, R., Shinohara, T. and Piatigorsky, J. (1977) Exp. Eye. Res. 25, 583-593.