Volume 11 Number 16 1983 Nucleic Acids Research

Nucleotide sequence of the constant region of a chicken μ heavy chain immunoglobulin mRNA

Auriel Dahan+*, Claude-Agnès Reynaud+ and Jean-Claude Weill+*

+Laboratoire d'Immunodifférenciation, Institut Jacques Monod, CNRS – Université Paris VII, Tour 43, 2 Place Jussieu, 75251 Paris, Cedex 05, and *Institut d'Immunobiologie, INSERM U20 and CNRS LA 143, Hôpital Broussais, 96, Rue Didot, 75674 Paris Cedex, France

Received 21 June 1983; Accepted 22 July 1983

ABSTRACT

We have recently reported the sequence of a chicken Ig λ light chain cDNA clone, isolated from a spleen partial cDNA library (1). In this paper, we describe the characterization of a cDNA clone coding for the chicken constant (C) region of the secreted μ chain. This is the first report on the nucleotide and amino acid sequence of a chicken Ig heavy chain constant region. It contains the 3' untranslated region of the μ mRNA up to the poly(A) tail, and, in comparison with the mouse Cμ sequence, displays the overall domain size and organization of a secreted μ chain, i.e. : a characteristic COOH-terminal region, a Cμ4, a Cμ3, a Cμ2, and part of a Cμ1 domain. The sequence homology between these two species ranges from 45 % for the Cμ4 to 18 % for the Cμ2. Thus, the Cμ sequence appears much less conserved between chicken and mouse than their respective λ light chain constant regions (1). These results, together with some distinctive features of the Cμ2 domain, may be of evolutionary relevance and will be further discussed.

INTRODUCTION

The IgM molecule is the first immunoglobulin (Ig) to appear in vertebrate evolution. It is also the first to appear during the ontogenic development of the B cell lineage.

There has been little information on the protein and gene structure of this Ig below the mammalian level. Among vertebrates, the chicken Ig may be of interest for two main reasons. When compared with the mammalian homologous sequences, they may give, like other lower species, some evolutionary insight into the different steps which led to the duplication and the diversification of the Ig gene family. When purified as cDNA probes, they may be used to study the early molecular events which control the rearrangement of Ig genes in the embryonic Bursa of Fabricius.

We have recently reported the characterization of a chicken λ light chain cDNA clone which was isolated from a spleen partial cDNA library. We report here the sequence of a cDNA clone coding for the constant

region of the chicken μ heavy chain (Cμ). This sequence shows a moderate homology with the mouse Cμ amino acid sequence ; however it presents the general characteristic domain organization observed for mammalian IgM constant regions.

MATERIALS AND METHODS

1) Enzymes and chemicals.

T_4 polynucleotide kinase and restriction endonucleases were purchased from Boehringer-Mannheim or Bethesda Research Laboratories, and DNA polymerase I from Boehringer-Mannheim.

$|\gamma\text{-}^{32}P|$ATP and $|\alpha\text{-}^{32}P|$ dCTP were obtained from the Radiochemical Centre (Amersham, England).

2) Construction and screening of a chicken spleen partial cDNA library.

Isolation of poly(A)-containing RNA from Brown Leghorn chicken stimulated spleen, construction and screening of recombinant cDNA plasmids, *in vitro* translation of mRNA fractions, immunoprecipitation and SDS-polyacrylamide gel electrophoresis were performed as previously described (1).

3) RNA blotting experiments.

Poly(A)-containing RNA was electrophoresed on 1 % agarose gel containing 5 mM methyl mercuric hydroxide (2) and transferred to diazobenzyloxymethylated-paper (3). Prehybridization and hybridization were as described (3), using a probe nick-translated at a specific activity of 10^8 cpm/μg with $|\alpha\text{-}^{32}P|$ dCTP. Washing conditions of Wahl et al. (4) were followed.

4) Plasmid DNA preparation.

Plasmid DNA was extracted by the alkaline procedure of Birnboim and Doly (5) and purified by hydroxylapatite chromatography according to Colman et al. (6). Restriction fragments fractionated by electrophoresis on polyacrylamide gels were prepared by electroelution.

5) DNA sequence analysis.

5'-labelled restriction fragments were purified from polyacrylamide-derived impurities by precipitation with 3 mM spermidine (7) before DNA sequence determination by the partial chemical degradation procedure of Maxam and Gilbert (8). Five base-specific reactions were used : G, A + G, C + T, C and A>C.

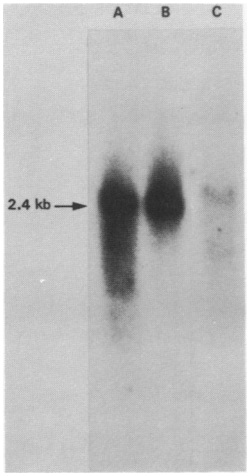

Figure 1. *Size determination of chicken μ heavy chain mRNA.*
Poly(A)-containing RNA was fractionated by electrophoresis on 1 % agarose gel containing 5 mM methyl mercuric hydroxide, transferred to diazobenzyloxymethylated-paper and hybridized with the nick-translated chicken Cμ cDNA clone.
(A) 10 μg of total chicken spleen poly(A)-containing RNA. (B) 4 μg of a chicken spleen mRNA fraction enriched for 19S mRNA sequences by two successive sucrose gradient centrifugations. (C) 24 μg of poly(A)-containing RNA from the TLT_1 chicken lymphoma cell line (9). The arrow shows the size of the major mRNA species of mouse spleen poly(A)-containing RNA hybridizing with a mouse Cμ probe.

RESULTS AND DISCUSSION

1) Isolation of a Cμ cDNA clone.

We previously reported the construction of a partial cDNA library from chicken spleen poly(A)-containing RNA (1). This library was screened for μ heavy chain mRNA sequences.

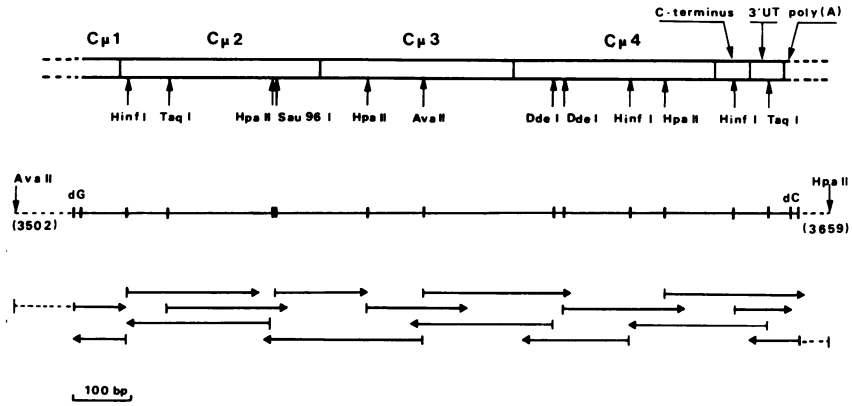

Figure 2. *Nucleotide sequencing strategy of a chicken Cμ cDNA clone.*
The chicken μ heavy chain mRNA is schematized in the upper part : first, second, third and fourth domains of the μ constant region (Cμ1, Cμ2, Cμ3, Cμ4), carboxy-terminal region (C-terminus), 3' untranslated region (3' UT) and poly(A) tract. Restriction sites used for 5'-end labelling of fragments are indicated, with Sutcliffe's coordinates (10) for those located in the pBR322 sequence. Arrows represent the extent of the DNA sequence determined.

```
                                    10                                      20                        Cu1   ||  Cu2 30
      Lys Gln Pro Phe Arg Cys Arg Ala Ala His Pro Arg Gly Asn Val Glu Val Ser Val Met Asn Pro Gly Pro Ile Pro Thr Pro Asn Gly Ile Pro Leu Phe Val Thr Met His
(Ch)  AG AAA CAG CCC TTC CGG TGT CGG GCG GCG CAT CCA CGC GGC AAC GTC GAG GTC AGC GTG ATG AAC CCA GGC CCG ATT CCC ACC CCG AAT GGC ATC CCC CTT TTC GTC ACC ATG CAC
(Mo)  CA G-T G-A TA- C-G GTA --C AAA ATC CAC T-C GG- G-- AAA --- AGA[   ]GAT C-- CAT GTG --C ATT --A GC- GT- G-A GA- -TG AA- CC- AAT G-A AAT --G TT- G-C[ ]
      Asp Glu Tyr Leu Val     Lys Ile His Tyr Gly Gly Lys     Arg[   ]Asp Leu His Val     Ile     Ala Val Ala Glu Met Asn Pro Asn Val Asn     Phe Val[ ]

      40                                      50                                      60                                      70
      Pro Pro Ser Arg Glu Asp Phe Glu Gly Pro[ ]Phe Arg Asn Ala Ser Ile Leu Cys Gln Thr Arg Gly Arg Arg Arg[ ]Pro Thr Glu Val Thr Trp Tyr Lys Asn Gly Ser Pro
(Ch)  CCC CCG TCC CGC GAG GAC TTC GAA GGC CCC[ ]TTC CGC AAC GCC TCC ATC CTC TGC CAG ACC CGC GGG CGC CGC CGT[ ]CCC ACC GAG GTC ACG TGG TAC AAA AAT GGC AGC CCC
(Mo)  --A --A CGG GAT -GC TT-[ ]TCT --- --T GCA CCA --- --G T-T AAA C-- A-- --- G-- G-- ACG AAC TT- ACT -CA AAA --G -T- ACA --A T-C --- CTA --G G-- --G -AG -T-
              Arg Asp Gly Phe[ ]Ser         Ala Pro     Lys Ser Lys Leu Ile     Glu Ala Thr Asn Phe Thr Pro Lys     Ile Thr     Ser     Leu     Asp     Lys Leu

                                          80                                      90                                      100
      Val[                        ]Ala Ala Ala Ala Thr Thr Ala Thr Thr Val Gly Pro Glu Val Val Ala Glu Ser Arg Ile Ser Val Thr Glu Ser Glu Trp Asp Thr Gly Ala
(Ch)  GTC[                        ]GCC GCC GCC GCC ACC ACC GCC ACC ACC GTC GGC CCC GAA GTG GTG GCC GAG AGC CGC ATC AGC GTC ACC GAA AGC GAA TGG GAC ACC GGG GCC
(Mo)  --G GAA TCT GGC TTC ACC ACA GAT C-G -TG A-- AT- GAG -A- AAA GGA T-- ACA CC- -AA ACC TAC AA- -T- ATA --- ACA C-T -C- A-- T-T --- -T- --C --- CTG -A- CT- AAT
          Glu Ser Gly Phe Thr Thr Asp Pro Val Thr Ile Glu Asn Lys Gly Ser Thr Pro Gln Thr Tyr Lys Val Ile     Thr Leu Thr Ile Ser     Ile Asp     Leu Asn Leu Asn

              110                                     120                         Cu2   ||  Cu3                               140
      Thr Phe Ser Cys Val Val Glu Gly Glu Met Arg Asn Thr Ser Lys Arg Met Glu Cys Gly Leu Glu Pro Val Val Gln Gln Asp Ile Ala Ile Arg Val Ile Thr Pro Ser Phe Val
(Ch)  ACC TTC AGC TGC GTC GTG GAG GGG GAG ATG AGG AAC ACC AGC AAG AGG ATG GAG TGC GGA TTA GAA CCC GTC GTG CAG CAG GAC ATC GCC ATC CGC GTC ATC ACG CCG TCC TTC GTG
(Mo)  GTG -A- -C- --- CGT --- --T CAC AG- GGT CTC -C- TT- TTG --- -AC G-- TCC -C- AC- -GT -CT G-- AGT CCC TCC ACA --- --- CTA -C- TT- AC- --- C-C --C --- --T -CC
      Val Tyr Thr     Arg     Asp His Arg Gly Leu Thr Phe Leu     Asn Val Ser Ser Thr Cys Ala Ala Ser Pro Ser Thr         Leu Thr Phe Thr     Pro             Ala

              150                                     160                                     170                                     180
      Asp Ile Phe Ile Ser Lys Ser Ala Thr Leu Thr Cys Arg Val Ser Asn Met Val Asn Ala Asp Gly Leu Glu Val Ser Trp Trp Lys Glu Lys Gly Gly Lys Leu Glu Thr Ala Leu
(Ch)  GAC ATC TTC ATC AGC AAA TCG GCC ACG CTG ACG TGC CGG GTG AGC AAC ATG GTG AAC GCC GAC GGC CTG GAG GTG TCG TGG TGG AAG GAG AAG GGG GGC AAA CTG GAG ACG GCG TTG
(Mo)  --- --- --- C-- --- --G --C --T -AC --- --C --T -T- --C TCA --- C-- -CA -C- TAT --A AC- --- A-T A-C --C --- GCT TCT C-A -GT --T -AA CC- --- --A --C AAA A-T
                  Leu             Asn             Leu             Leu Ala Thr Tyr Glu Thr     Asn Ile         Ala Ser Gln Ser     Glu Pro             Lys Ile

                      190                                     200                                     210                                     220
      Gly Lys Arg Val Leu Gln Ser Asn Gly Leu Tyr Thr Val Asp Gly Val Ala Thr Val Cys Ala Ser Glu Trp Asp Gly Gly Asp Gly Tyr Val Cys Lys Val Asn His Pro Asp Leu
(Ch)  GGG AAG AGG GTC CTG CAA AGC AAC GGC CTC TAC ACG GTG GAC GGG GTG GCC ACG GTG TGC GCC AGC GAA TGG GAC GGA GGG GAT GGC TAC GTG TGT AAG GTG AAC CAC CCC GAT CTG
(Mo)  AAA -TC -T- -AA AGC --T CC- --T --- AC- -T- -GT -CT A-G --T --- --T -GT --T --T -TG GAA --C --- A-T AAC A-- A-G -AA -TT --- --- -CT --- -CT --- AGG --- ---
      Lys Ile Met Glu Ser His Pro         Thr Phe Ser Ala Lys         Ser         Val Glu Asp     Asn Asn Arg Lys Glu Phe         Thr     Thr     Arg

                      230                         Cu3   ||  Cu4 240                                   250                                     260
      Leu Phe Pro Met Glu Glu Lys Met Arg Lys Thr Lys Ala Ser Asn Ala Arg Pro Pro Ser Val Tyr Val Phe Pro Pro Pro Thr Glu Gln Leu Asn Gly Asn Gln Arg Leu Ser Val
(Ch)  CTC TTC CCC ATG GAG GAG AAG ATG AGG AAG ACG AAA GCC AGC AAC GCC CGC CCC CCA TCC GTC TAC GTC TTC CCC CCC CCC ACG GAA CAA CTG AAC GGC AAC CAA CGG CTC AGC GTC
(Mo)  -CT -CA --A CA- A-- A-A TTC --C TCA --A C-C --T -AG GTG C-- AAA -AT --A --T G-T --G --- C-G C-G --A --A G-T CGT --G --- --- --- CTG -GG G-G TCA GC- -CA ---
      Pro Ser     Gln Lys Lys Phe Ile Ser     Pro Asn Glu Val His Lys His         Ala             Leu Leu         Ala Arg                 Leu Arg Glu Ser Ala Thr

                      270                                     280                                     290                                     300
      Thr Cys Met Ala Gln Gly Phe Asn Pro Pro His Leu Phe Val Arg Trp Met Arg Asn Gly Glu Pro Leu Pro Gln Ser Gln Ser Val Thr Ser Ala Pro Met Ala Glu Asn Pro Glu
(Ch)  ACC TGC ATG GCT CAG GGC TTC AAC CCC CCC CAC CTC TTC GTC AGG TGG ATG AGA AAC GGG GAA CCC CTC CCC CAA AGC CAA TCG GTC ACA TCG GCC CCC ATG GCG GAG AAC CCC GAA
(Mo)  --- --- T-- -TG A-- --- --- TCT --T G-A G-- A-- AGT --G CAG --- C-- CAG -GA --- C-- -T- T-G --- --- GAG A-G -AT --G --C AGT --- --G --- C-A --- CCT GGG -CC
              Leu Val Lys             Ser     Ala Asp Ile Ser     Gln     Leu Gln Arg     Gln Leu             Glu Lys Tyr                     Pro     Pro Gly Ala

                              310                                     320                                     330                                 340
      Asn Glu Ser Tyr Val Ala Tyr Ser Val Leu Gly Val Gly Ala Glu Glu Trp Gly Ala Gly Asn Val Tyr Thr Cys Leu Val Gly His Glu Ala Leu Pro Leu Gln Leu Ala Gln Lys
(Ch)  AAT GAG TCC TAC GTG GCC TAC AGC GTT TTG GGG GTG GGG GCC GAA GAG TGG GGC GCC GGC AAC GTC TAC ACG TGC CTG GTG GGC CAC GAA GCT CTG CCC CTC CAG CTG GCC CAG AAG
(Mo)  CCA -GC -T- --- T-T A-- C-- --- A-C C-- ACT --- ACA -AG --G --A --- AA- T-- --A G-G AC- --T --C --T G-T --A --- --- --G --C --- --A -A- -T- G-- A-- G-- -G-
      Pro Gly Phe     Phe Thr His     Ile     Thr     Thr Glu                 Asn Ser     Glu Thr             Val                             His Leu Val Thr Glu Arg

                      Cu4   ||   C-term                                           360                         367
      Ser Val Asp Arg Ala Ser Gly Lys Ala Ser Ala Val Asn Val Ser Leu Val Leu Ala Asp Ser Ala Ala Ala Cys Tyr Stop
(Ch)  TCG GTG GAT AGG GCT TCG GGT AAA GCA AGT GCT GTC AAT GTC TCC TTG GTG TTG GCC GAC TCG GCC GCC GCC TGC TAT TAA TTAATTAACCCGCTCGTTAAGCGGCCGCTCGATTGGGATTAAAGAGC
(Mo)  A-C --- --C -A- T-C A-T --- --- C-C -CA CTG TA- --- --- --- C-- A-C A-- T-T --- A-A -G- -G- A-- --- --- -G-
      Thr         Lys Ser Thr         Pro Thr Leu Tyr             Ile Met Ser     Thr Gly Gly Thr

(Ch)  AGATGTCCATTpoly(A)
```

The chicken μ polypeptide was identified as a 70 000 Mr protein by SDS-gel electrophoresis after immunoprecipitation of 19S mRNA cell-free translation products with anti-chicken μ chain antiserum (data not shown). According to this criterium, spleen poly(A)-containing RNA was enriched by successive sucrose gradient fractionation and used as a cDNA probe to select μ-containing recombinant plasmids. A differential screening with a cDNA probe for chicken liver mRNA sequences was performed in order to eliminate sequences present in both tissues.

A clone was selected with an inserted sequence of 1.2 kbp. As an additional characterization, the size of the corresponding spleen mRNA species was determined, by using this clone as a probe in an RNA transfer experiment : after hybridization, a signal was obtained with a mRNA of 2.4 kb (fig. 1, lanes A and B), a size analogous to the one of the mouse mRNA coding for the secreted form of the μ chain, as was determined by hybridizing a mouse spleen mRNA sample with a mouse Cμ probe in a parallel track (data not shown).

2) Comparison of chicken and mouse Cμ sequences.

The chicken Cμ sequence (fig. 2) is definitively identified on the basis of its homology with the mouse Cμ carboxyterminal region and may be aligned without interruption next to the mouse fourth and third domains (fig. 3). For the second and first domains, an alignment was tempted by introducing several gaps or insertions according to internal heavy chain amino acid homologies (13). Each of the domains, as defined by such an alignment, contain the two cysteine residues attributed to the intrachain disulfide bridges (Cμ1 : 6 ; Cμ2 : 56-111 ; Cμ3 : 158-217 ; Cμ4 : 265-327). The other cysteines correspond to interchain (Cys 126 and 205) and inter-subunit (Cys 366) disulfide bridges (12).

The mouse Cμ chain has five sites of carbohydrate attachment at an Asn residue in a three amino acid recognition sequence : Asn-X-Thr/Ser

Figure 3. *Comparison between chicken and mouse Cμ nucleotide and amino acid sequences.*
The nucleic acid sequence, together with the deduced amino acid sequence of the chicken Cμ clone are aligned with the mouse Cμ nucleic and amino acid sequences (11,12). Chicken amino acids are numbered and domain boundaries are as follows : Cμ2 (28-130) ; Cμ3 (131-237) ; Cμ4 (238-348) ; COOH-terminus (349-367). The mouse amino acids identical to the chicken amino acids are omitted and the nucleotide homologous positions are indicated by dashes. The gaps introduced to maximize the homology are represented by brackets. The poly(A) addition signal in the 3' untranslated region is underlined.

Table I

domains compared	% homology chicken/mouse nucleotide	chicken/mouse amino acid	human/mouse amino acid
Cμ2 : Cμ2	37	18	59
Cμ3 : Cμ3	52	39	53
Cμ4 : Cμ4	57	45	78
C-ter : C-ter	67	42	89
Cλ : Cλ	63	58	71

Interspecies homologies between Cμ and Cλ domains. The μ chains were divided into 3 homology units and a COOH-terminal region (C-ter) according to Dayhoff (13) : Cμ2 = residues 28-130 ; Cμ3 = residues 131-237 ; Cμ4 = residues 238-348 ; COOH-terminal region = 349-367. Mouse amino acid and nucleotide μ sequences are taken from Kehry et al. (12) and Goldberg et al. (11), respectively. Human μ amino acid sequence is from ref. 15, with corrections obtained from nucleic acid sequence data (16). The Cλ sequences are taken from ref. 1 for the chicken, from ref. 17 for the mouse $C\lambda_1$ sequence and from ref. 18 for the human Cλ X sequence. The percentage of homology is expressed as : amino acid or nucleotide identities x 100 / maximum number of residues compared.

(12). Such a sequence is present at five positions in the chicken Cμ chain, (Asn residues 51, 72, 119, 303 and 354), the only identical one being Asn 354 (Asn-Val-Ser) in the COOH-terminal region. The carbohydrate moiety covalently linked to this site is highly conserved in human and mouse μ and α COOH-terminal segment, and might be involved, like the penultimate cysteine (366), in the structure of the pentameric IgM (14). The chicken Cμ4 domain contains a putative carbohydrate acceptor site (Asn 303) which is never found in the last domain of other Ig heavy chains. The attachment of a sugar moiety at this position would be contrary to the general carbohydrate-free structure of this region.

As observed within mammalian sequences, there is a decreasing homology from the Cμ4 to the Cμ2 domains between chicken and mouse μ sequences (Table I). The higher degree of conservation of the chicken Cμ4 domain may be related, as in mammals, to the structural constraint imposed by the effector functions on the Fc region of the molecule.

To align the chicken and the mouse Cμ2 domains, a deletion of seven residues was introduced (between amino acids 76 and 77), together with two single gaps (between 48-49 and 63-64) and two insertions (38 and 45) in the chicken sequence. The large deletion is located at the border

Table II

domains compared	% homology chicken/chicken amino acid	% homology mouse/mouse amino acid
Cμ2 : Cμ3	17	17
Cμ2 : Cμ4	22	23
Cμ3 : Cμ4	18	21
Cλ : Cμ2	19	25
Cλ : Cμ3	19	19
Cλ : Cμ4	23	26

Intraspecies homologies between Cμ and Cλ domains. The Cμ domains were aligned with each other and with the Cλ sequences according to Dayhoff (13). References for sequence data are as in Table I. The percentage of homology is expressed as : amino acid identities x 100 / maximum number of residues compared.

of a peculiar series of Ala and Thr residues, each being respectively coded by the same GCC and ACC codons. This alignment provides a 18 % homology at the amino acid level, a value which may not be substantially improved by the use of a computerized analysis.

Despite this interspecies divergence, the Cμ2 region retains the characteristic structural elements of an Ig domain. When intra-species comparisons are made, the homology between the Cμ2 and the other Cμ3 and Cμ4 domains is very similar in the chicken and the mouse (Table II). The same holds true with their respective Cλ light chain sequences (Table II). Moreover, it contains like other Cμ2 domains five proline residues in a stretch of 21 amino acids at its N-terminal region (Pro_{28}-Pro_{48}). These remarkable features of the chicken Cμ2 domain may be of evolutionary relevance. They may provide the molecule with the intermediate flexibility observed in chicken Ig (19). Such an increase of rotational freedom has been described for macroglobulins, when passing from lower vertebrates to mammals (20). They may also reflect, in some way, the evolutionary history of the Cμ2 region, which was most probably deleted at the time of the emergence of other heavy chain subclasses.

The chicken 3' untranslated region displays no obvious homology with the corresponding mouse sequence. It is rather short (58 bp) ; it presents three consecutive stop codons covering each reading frame and a minor type of polyadenylylation signal A-U-U-A-A-A (21,22).

The generation of the mouse μ membrane mRNA requires an alterna-

tive splicing event which joins the Cμ4 to two 3' exons, creating a longer mRNA molecule than the one for secreted form (2.7 versus 2.4 kb) (23, 24). The presence of a splicing signal G-T at the precise end of the Cμ4 domain (Gly_{348}) suggests that the chicken B cell may use a similar mechanism to generate μ membrane mRNA. Accordingly, the μ mRNA from a chicken lymphoma (TLT1) (9) expressing IgM at its surface has a slightly higher Mr than the major μ species present in the chicken spleen (fig. 1, compare lanes C and A).

It has been suggested that the μ heavy chain has been more preserved during evolution than the other Ig classes (25). Surprisingly however, the constant region of the chicken λ light chain is much more homologus to the mouse Cλ chain (58 %) than are the most conserved regions of their respective μ chains (Cμ4 or COOH-terminal : 45-42 %).

A sequential rearrangement of Ig genes has been described in the mouse pre-B cells, involving first the μ locus and then the light chain genes (26,27). In the chicken, a simultaneous expression of the IgM heavy and light chains seems to occur in the bursal lymphoid cells (28). If this observation is confirmed at the gene level, it remains to be explored whether it is due to a species-specific regulatory pathway of B cell differentiation or to a distinctive genomic organization of the chicken μ and λ loci.

ACKNOWLEDGMENTS

We thank Viviane Anquez for excellent technical assistance, Richard Schwartzman and Brigitte Charézieux for help in the preparation of the manuscript, A. Benedict for the gift of goat anti-chicken μ chain antiserum, B.W. Calnek for the TLT_1 chicken B lymphoma cell line and D. Baltimore for the mouse Cμ probe. We thank also J.F. Rodier (Institut Jacques Monod) and J.M. Claverie (Institut Pasteur) for the following computer analysis : Nucleic Acid Dot Matrix (B. Caudron) and Amino Acid Alignment of Needleman and Wunch (Smith and Fitch). This work was supported by the Centre National de la Recherche Scientifique (A.T.P. Biologie du Développement et de la Reproduction), the Institut de la Santé et de la Recherche Médicale (C.R.L. 811054) and the Fondation pour la Recherche Médicale.

REFERENCES.

1. Reynaud, C.A., Dahan, A. and Weill, J.C. (1983) Proc. Natl. Acad. USA 80, 0000-0000.

2. Bailey, J.M. and Davidson, N. (1976) Anal. Biochem. 70, 75-85.
3. Alwine, J.C., Kemp, D.J. and Stark, G.R. (1977) Proc. Natl. Acad. Sci. USA 74, 5350-5354.
4. Wahl, G.M., Stern, M. and Stark, G.R. (1979) Proc. Natl. Acad. Sci. USA 76, 3683-3687.
5. Birnboim, H.C. and Doly, J. (1979) Nucl. Acids Res 7, 1513-1523.
6. Colman, A., Byers, M.J., Primrose, S.B. and Lyons, A. (1978) Eur. J. Biochem. 91, 303-310.
7. Hoopes, B.C. and McClure, W.R. (1981) Nucl. Acids Res. 9, 5493-5504.
8. Maxam, A.M. and Gilbert, W. (1980) Methods Enzymol. 65, 499-560.
9. Hirai, K. and Calnek, B.W. (1979) Infect. Immun. 25, 964-970.
10. Sutcliffe, J.G. (1978) Nucl. Acids Res. 5, 2721-2728.
11. Goldberg, G.I., Vanin, E.F., Zrolka, A.M. and Blattner, F.R. (1981) Gene 15, 33-42.
12. Kehry, M., Sibley, C., Fuhrman, J., Schilling, J. and Hood, L.E. (1979) Proc. Natl. Acad. Sci. USA 76, 2932-2936.
13. Dayhoff, M.O. Atlas of Protein Sequence and Structure (1978) vol. 5, sup. 3, pp. 197-227, National Biomedical Research Foundation, Washington.
14. Feinstein, A. and Munn, E.A. (1969) Nature 224, 1307-1309.
15. Putnam, F.W., Florent, G., Paul, C., Shinoda, T. and Shimizu, A. (1973) Science 182, 287-292.
16. Rabbitts, T.H., Forster, A. and Milstein, C.P. (1981) Nucl. Acids Res. 9, 4509-4524.
17. Bernard, O., Hozumi, N. and Tonegawa, S. (1978) Cell 15, 1133-1144.
18. Milstein, C., Clegg, J.B. and Jarvis, J.M. (1968) Biochem. J. 110, 631-652.
19. Zabyansky, Y. and Ivannikova, E.I. (1974) Mol. Biol. Rep. 1, 301-304.
20. Holowka, D.A. and Cathou, E. (1976) Biochemistry 15, 3386-3390.
21. Jung, A., Sippel, A.E., Grez, M. and Schutz, G. (1980) Proc. Natl. Acad. Sci. USA 77, 5759-5763.
22. Hagenbuchle, O., Bovery, R. and Young, R.A. (1980) Cell 21, 179-187.
23. Rogers, J., Early, P., Carter, C., Calamme, K., Bond, M., Hood, L. and Wall, R. (1980) Cell 20, 303-319.
24. Alt, F.W., Bothwell, A.L.M., Knapp, M., Siden, E., Mather, E., Koshland, M. and Baltimore, D. (1980) Cell 20, 293-301.
25. Wasserman, R.L. and Capra, J.D. (1978) Science 200, 1159-1161.
26. Maki, R., Kearney, J.K., Paige, C. and Tonegawa, S. (1980) Science 209, 1366-1369.
27. Alt, F., Rosenberg, N., Lewis, S., Thomas, E. and Baltimore, D. (1981) Cell 27, 381-390.
28. Grossi, C.E., Lydyard, P.M. and Cooper, M.O. (1977) J. Immunol. 119, 749-756.