Volume 8 Number 14 1980 Nucleic Acids Research

Structure of the constant and 3' untranslated regions of the murine Balb/c γ2a heavy chain messenger RNA

Jean-Louis Sikorav, Charles Auffray and François Rougeon

Unite de Genetique et Biochimie du Developpement, E.R.A. C.N.R.S. 851, Institut Pasteur, 28, rue du Docteur Roux, 75724 Paris Cedex 15, France

Received 27 May 1980

ABSTRACT

The complete sequence for the constant and 3' untranslated regions of a mouse γ2a immunoglobulin heavy chain mRNA is reported. The sequence is 1093 nucleotides long coding for the CH1 (amino-acids 118-214), the Hinge (215-230), the CH2 (231-340) and the CH3 (341-447). The 3' untranslated region is 103 nucleotides long preceding the poly(A). The nucleotide sequence predicts as in the case for γ1 and γ2b heavy chains an additional lysine residue before the termination codon. This sequence has been compared to the corresponding sequences of γ1 and γ2b heavy chain mRNAs. These sequences are respectively 75% and 84% homologous. The CH2 domains of γ2a and γ2b are 95% homologous at the nucleotide level. The cross-over point of a γ2a - γ2b heavy chain variant is located in a segment of 73 perfectly matching nucleotides. The 3' non coding regions of γ2a and γ2b are 89% homologous.

INTRODUCTION

Myeloma proteins and antibodies have been shown to have a basic structure consisting of two identical light (L) chains (MW 22000) and two identical heavy (H) chains (MW 50000). The light and heavy chains are divided into consecutive homology regions having approximately 110 residues. These domains consist of a disulfide loop of about 65-70 residues and a connecting segment. Each domain is associated with a different function. Sequence studies and X-Ray cristallography studies have shown that an antibody combining site is formed by non-covalent interactions between the NH2 terminal domains (V regions) of the light chain and the heavy chain (for review, see reference 1). Immunoglobulins are divided into five different classes (IgM, IgD, IgG, IgA and IgE) which are determined by the amino sequence of the three of four COOH terminal domains of the heavy chains (1). Analysis of the structure of the constant regions of light and heavy chains suggests that heavy chains have evolved by gene duplication from a common ancestor with light chain (for review, see reference 2). The constant domains of light and heavy chains are encoded by distinct exons (3-11).

© IRL Press Limited, 1 Falconberg Court, London W1V 5FG, U.K.

In the mouse, the IgG class can be divided into four subclasses : γ1, γ2a, γ2b, γ3, each of them containing three domains : CH1, CH2, CH3 and the Hinge region. Amino-acid sequence studies of mouse and human heavy chains and nucleic acid hybridization with cDNAs complementary to heavy chain mRNA of different γ subclasses demonstrated that the γ subclass chains are more closely related to each other than to μ or α chains and that γ2a and γ2b proteins share considerable homology (12,14). The results suggest that the γ chain class has evolved by gene duplication of the three domain structure (13,15). In the present report we present the complete sequence for the constant region and for the 3' untranslated region of the mouse Balb/c γ2a immunoglobulin heavy chain mRNA. This sequence has been determined from a bacterial plasmid (pG2a-10-21) containing a full length transcript of a γ2a heavy chain mRNA (16). This sequence has been compared to the corresponding sequences of γ2b and γ1 heavy chain mRNAs (7,9).

MATERIALS AND METHODS

1 - Chemicals and Enzymes

(γ - ^{32}P) ATP and 3'-(α - ^{32}P) dATP (Cordycepin triphosphate) were respectively obtained from the Radiochemical Centre Amersham (England) and from New England Nuclear (Boston, Massachusetts, U.S.A.). T4 polynucleotide kinase and calf thymus terminal deoxynucleotidyl transferase were purified as previously described (17,18). BamHI, PstI and AluI were purified according to the published procedures (19). All other restriction endonucleases were obtained from New England Biolabs.

2 - DNA

The construction and the characterization of the E.coli strain carrying the hybrid plasmid (pG2a-10-21) containing a full length transcript of a γ2a heavy chain mRNA has been previously reported (16). The supercoiled DNA was purified according to Katz et al. (20).

3 - DNA sequencing

DNA fragments were labelled either at 5' ends using (γ - ^{32}P) ATP and polynucleotide kinase in the exchange reaction (21) or at 3' ends by incorporation of 3'-(α - ^{32}P) dATP with terminal transferase (22). Partial chemical degradation was performed according to Maxam and Gilbert (23) (Four base reactions were used : G, G+A, C+T, C). The products were analysed on 20% and 8% 0.35 mm thick urea-polyacrylamid gels according to Sanger and Coulson (24).

RESULTS AND DISCUSSION

1 - Nucleotide sequence of the constant and 3' untranslated regions of the γ2a heavy chain mRNA

The strategy for sequencing the plasmid [pG2a-10-21] containing a full length transcript of a γ2a heavy chain mRNA is presented in Fig.1. The DNA fragments endlabelled as described in Methods were either cut with another enzyme or subjected to strand separation (23). The exact boundaries of the domains have been assigned from the protein sequence data (12) and by the alignment maximizing homology with the domains of the γ2b protein, established by location of intervening sequences (8,9,25). The complete sequence of the constant and untranslated regions (1093 residues) of the γ2a heavy chain mRNA is presented in Fig. 2. The sequence starts with Ala at position 118 of the γ2a protein sequence (12) and ends with the poly(A) segment of the mRNA. The nucleotide sequence predicts the complete amino-acid sequence of the constant region of the Balb/c γ2a heavy chain consisting of the CH1 (amino-acids 118-214), the Hinge (215-230), the CH2 (231-340) and the CH3 (341-

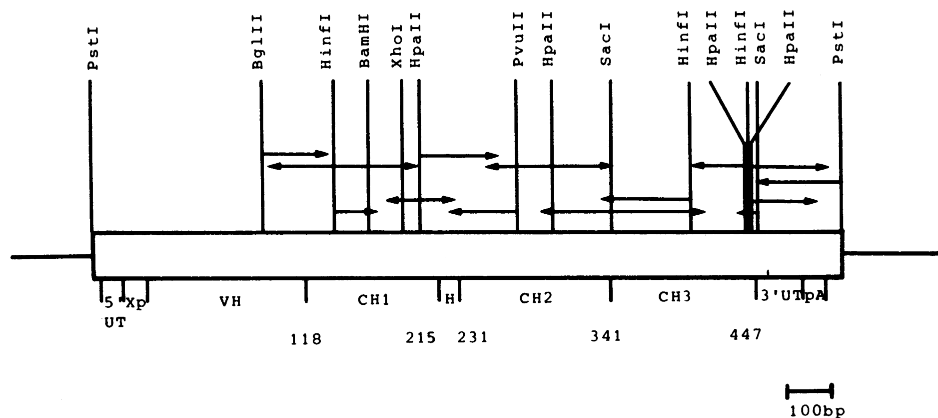

Fig. 1 : Sequencing strategy. Restriction sites used for sequencing are indicated in the upper part of the figure.
3' end labelling was used only for the two SacI sites (see Methods). The horizontal arrows represent the direction and range of the sequence. The different segments of the mRNA are delimited by vertical bars on the lower part.
5' UT and 3' UT : untranslated regions of 5' and 3' extremities, CH1, CH2, CH3 : regions coding for the different domains of the protein.
Xp = extra-piece segment ; H = Hinge ; pA = poly-adenylic acid.
The limits of constant domains are indicated by the first amino-acid position (see also Fig. 2).

```
118 120                                   130                                      140                                      150
 -  -  -  -  -  P  -  -  -  -  -  -  -  -  G  -  -  -  -  -  -  -  -  -  -  -  -  -  -  -  -  -  -  -  -  -
------------C-------A------------------G------------------U-----C--------G-----------------C--C---------
                                                                                    *
GCCAAAACAACAGCCCCAUCGGUCUAUCCACUGGCCCCUGUGUGUGGAGAUACAACUGGCUCCUCGGUGACUCUAGGAUGCCUGGUCAAGGGUUAUUUCCCUGAG
 A  K  T  T  A  P  S  V  Y  P  L  A  P  V  C  G  D  T  T  G  S  S  V  T  L  G  C  L  V  K  G  Y  F  P  E
                        160                                      170                                      180
 -  -  -  V  -  -  -  -  -  -  -  -  -  S  -  -  -  -  -  -  L  -  -  -  G  -  -  -  M  -  -  -  -  -  -
--------UG----U------------------------CA--------------------C-------------GA--------UA-G-----------------C
               *
CCAGUGACCUUGACCUGGAACUCUGGAUCCCUGUCCAGUGGUGUGCACACCUUCCCAGCUGUCCUGCAGUCUGACCUCUACACCCUCAGCAGCUCAGUGACUGUA
 P  V  T  L  T  W  N  S  G  S  L  S  S  G  V  H  T  F  P  A  V  L  Q  S  D  L  Y  T  L  S  S  S  V  T  V
       190                                 200                                  210         214
 P  -  -  -  -  -  -  -  T  V  -  -  S  -  -  -  -  -  -  -  -  -  T  -  -  -  -  -  L      γ2b protein
C----C-----U-----A--U---A--G----------GC--U--U-----A--C-----------C---------A---C--           γ2b mRNA
            *
ACCUCGAGCACCUGGCCCAGCCAGUCCAUCACCUGCAAUGUGGCCCACCCGGCAAGCAGCACCAAGGUGGACAAGAAAAUU             γ2a mRNA
 T  S  S  T  W  P  S  Q  S  I  T  C  N  V  A  H  P  A  S  S  T  K  V  D  K  K  I              γ2a protein

                                              - CH1 -


215        219       220                  227            228   230
 -  -  S  -  -  I  S  -  -  N  -  -  -  -  -  K  E  C  H  -  -  -              γ2b protein
--------C------ AUUUCA --------C--------------- AAGGAGUGUCAC ---------           γ2b mRNA

GAGCCCAGAGGGCCC(      )ACAAUCAAGCCCUGUCCUCCAUGC(           )AAAUGCCCA          γ2a mRNA
 E  P  R  G  P          T  I  K  P  C  P  P  C                K  C  P           γ2a protein

                                  - HINGE -
```

```
231                                     240                                     250                                     260
 -   -   -   -   E   -   -   -   -   -   -   -   -   -   -   N   -   -   -   -   -   -   -   -   -   T   -   K   -   -   -
--U---------GA--------------------------------U----------------------------CA----AG-----G---
GCACCUAACCUCUUGGGUGGACCAUCCGUCUUCAUCUUCCCUCCAAAGAUCAAGGAUGUACUCAUGAUCUCCCUGAGCCCCAUAGUCACAUGU
 A   P   N   L   L   G   G   P   S   V   F   I   F   P   P   K   I   K   D   V   L   M   I   S   L   S   P   I   V   T   C
                                270                                     280                                     290
 -   -   -   -   -   -   -   -   -   -   -   -   -   -   -   -   -   -   -   -   -   -   -   -   -   -   -   -   -   -   -
----------------------------C---------------------------------------------------------------
GUGGUGGUGGAUGUGAGCGAGGAUGACCCAGAUGUCCAGAUCAGCUGGUUUGUGAACAACGUGGAAGUACACACAGCUCAGACACAAACCCAU
 V   V   V   D   V   S   E   D   D   P   D   V   Q   I   S   W   F   V   N   N   V   E   V   H   T   A   Q   T   Q   T   H
                              300                                     310
 -   -   -   -   -   -   -   I   -   -   -   -   T   -   -   -   -   -   -   -   -   -   -   -   -   -   -   -   -   -   -
---------------------A--------------CA------------------------------------------------------
                                                                      *
AGAGAGGAUUACAACAGUACUCUCCGGGUGGUCAGUGCCCUCCCCAUCCAGCACCAGGACUGGAUGAGUGGCAAGGAGUUCAAAUGCAAGGUC
 R   E   D   Y   N   S   T   L   R   V   V   S   A   L   P   I   Q   H   Q   D   W   M   S   G   K   E   F   K   C   K   V
                          330                                     340
 -   -   -   -   -   -   S   -   -   -   -   -   -   -   -   -   -          γ2b protein
------------------U-A------------------------------                         γ2b mRNA
AACAACAAAGACCUCCCAGCGCCCAUCGAGAGAACCAUCUCAAAACCCAAA                         γ2a mRNA
 N   N   K   D   L   P   A   P   I   E   R   T   I   S   K   P   K          γ2a protein
```

- CH2 -

```
341                                 350                                      360                                      370
 -   L   -   -   -   -   -   -   -   T   -   -   -   -   A   -   Q   L   S   R   -   D   -   S   -   -   -   L   -   V   G
---CU---C-----------A-----CACU-----G-------C---GC--U--U-C-G----G-U----G---C--U---C-----GUG-G-
GGGUCAGUAAGAGCUCCACAGGUAUAUGUCUUGCCUCCACCAGAAGAAGAGAUGACUAAGAAACAGGUCACUCUGACCUGCAUGGUCACAGAC
 G   S   V   R   A   P   Q   V   Y   V   L   P   P   P   E   E   E   M   T   K   K   Q   V   T   L   T   C   M   V   T   D
                                380                                      390                                      400
 -   N   -   G   -   -   S   -   -   -   -   -   S   -   -   H   -   -   E   -   -   -   D   -   A   -   -   -   -   -   -   -
----AC----G------CAGU------------G---U---C-U------GAG---------G----C-C------U--U--------C---
                                                                                                    *
UUCAUGCCUGAAGACAUUUACGUGGAGUGGACCAACAACGGGAAAACAGAGCUAAACUACAAGAACACUGAACCAGUCCUGGACUCUGAUGGU
 F   M   P   E   D   I   Y   V   E   W   T   N   N   T   K   T   E   L   N   Y   K   N   T   E   P   V   L   D   S   D   G
                            410                                      420                                     430
 -   -   -   L   -   -   -   -   N   M   K   T   S   K   -   E   K   T   D   -   F   -   -   N   -   R   -   -   -   -   K
-----------A--U--------C-AUA--A---CA-GC--G----A-A---C-G--UC--U---A--CAAC---AGA------------A-A
UCUUACUUCAUGUACAGCAAGCUGAGAGUGGAAAAGAAGAACUGGGUGGAAAGAAAUAGCUACUCCUGUUCAGUGGUCCACGAGGGUCUGCAC
 S   Y   F   M   Y   S   K   L   R   V   E   K   K   N   W   V   E   R   N   S   Y   S   C   S   V   V   H   E   G   L   H
                         440                           447
 -   Y   Y   L   K   -   T   I   -   -   S   -   -   -               γ2b protein
---U--U--CU--AG----C-A--------U-----------                           γ2b mRNA
AAUCACCACACGACUAAGAGCUUCUCCCGGACUCCGGGUAAA                           γ2a mRNA
 N   H   H   T   T   K   S   F   S   R   T   P   G   K               γ2a protein
```

- CH3 -

```
------------------G-----------U---------UGG---C---A-------A-------------------U------
UGAGCUCAGCACCCACAAAACUCUCAGGUCCAAAGAGACACCCACACUCAUCUCCAUGCUUCCCUUGUAUAAAUAAAGCACCCAGCA
STOP


-A------U---------                                    γ2b mRNA
     *
AUGCCUGGGACCAUGUAA...........AACC...CC                γ2a mRNA
```

- 3' UNTRANSLATED REGION -

Fig. 2 : Nucleotide sequence of the constant and 3' untranslated regions of the γ2a heavy chain mRNA : The nucleotide sequence has been divided into five parts : CH1, Hinge, CH2, CH3 and 3' UT regions. The two lower lines give the nucleotide sequence and the predicted amino-acid sequence of the γ2a chain. *: 5' methyl-cytosines resulting from the methylation of the EcoRII sites in E.coli C600 (34). The two upper lines give the nucleotide and amino-acid sequences of the γ2b chain (9). Only the nucleotides or amino-acids which differ from the γ2a chain have been mentioned. Homologous-positions are indicated by dashes. The brackets in the Hinge region indicate the deletions required to maximize the homology. The single letter code for amino-acid is given in ref. 33.

447). The untranslated region consists of 103 residues, starts with UGA and ends with GU at the poly-adenylation site. Like other mRNAs from eucaryotic organisms,the 3' untranslated region contains the hexanucleotide AAUAAA, located 26 nucleotides from the poly(A) (26,27).

The amino-acid sequence deduced from DNA sequencing has been compared with the amino-acid sequence of the γ2a protein of MOPC 173 established by Fougereau et al. (12). Twenty-five differences have been observed, ten of them are exchanged Asn $\rightleftarrows$ Asp or Gln $\rightleftarrows$ Glu. Fifteen are substitution or displacement of amino-acids. The DNA sequence also predicts an additional lysine at the carboxy terminal end which is not found in the secreted serum protein. The same evidence for a post translational processing of a lysine has been reported in the case of γ2b and γ1 heavy chains (7,9,25).

2 - Comparison of the coding sequences of γ2a and γ2b mRNAs

γ2a and γ2b sequences have been aligned in order to maximize the homology between the two closely related genes (Fig. 2). Amino-acid or nucleotide matches have been counted for each segment of the coding sequence. The CH1 of γ2a and γ2b exhibit considerable homology since their sequences can be aligned without deletion or insertion (Fig. 2). Thirty-seven base changes have been found accounting for 13 amino-acid changes and 21 silent mutations.

The Hinge of γ2a is shorter than the Hinge of γ2b. Two deletions have to be introduced in order to maximize the homology : Ile - Ser at position 219-200 and Lys - Glu - Cys - His at position 227-228.

The sequences of the 48 matching nucleotides differ only by 2 residues (96% homology). As expected from protein sequence data, the CH2 of γ2a and γ2b are highly conserved. 94% of the amino-acids are identical corresponding to 95% homology at the nucleotide level. The most striking feature of the comparison of the CH2 domains is the total identity of 73 nucleotides between position 305 (Ala in γ2a and Thr in γ2b at position 307), and position 330 (Ala in γ2a and Ser in γ2b at position 332). This segment corresponds to the cross-over region in the CH2 domain in a variant of the MPC 11 myeloma characterized by Birshtein et al. that synthesizes a γ2b - γ2a hybrid heavy chain (25,28). Because of the identity of the two sequences in this region, it is difficult to know exactly the cross-over point. If hybrid proteins are generated by cross-over involving homologous recombination, it can be supposed that other variants could be generated by recombination in regions completely homologous such as the region between positions 271 and 299 where 81 nucleotides match perfectly. The highest variability is found in the CH3 domain. Out of 107 amino-acids only 65 are conserved ; this corresponds to

about 61% of homology at the DNA level. Nevertheless, the silent mutation rate is only slightly higher than in the CH1. Thus, the alteration rates of triplet coding for similar amino-acid at the same position are respectively : CH1 (7.9%), Hinge (0%), CH2 (1.3%) and CH3 (9.2%). This relatively low rate of codon alteration is also illustrated by the pattern of codon utilisation (Tab. 1). It appears clearly that the codon usage for the constant regions of γ2a and γ2b is very similar. This suggests selective pressure for more preferred codons.

3 - Comparison of 3' untranslated regions of γ2a and γ2b mRNAs

The 3' non coding regions of γ2a and γ2b have exactly the same length of 103 residues and the same dinucleotide GU at the poly-adenylation site. The two sequences exhibit an unexpected high degree of homology since out of 103 aligned nucleotides 90 match perfectly (89% of homology). For comparison, it can be noted that β^{min} and β^{maj} globin mRNAs have 94% homology in their coding regions and only 70% in their 3' non coding regions(29). The observation that the 3' non coding region is as conserved as or more conserved than some of the coding regions suggests that some parameters such as secondary structure, mRNA stability or translational efficiency could exert selective pressure on the primary structure of the mRNA.

4 - Comparison of C regions of murine γ1, γ2a and γ2b heavy chains

The comparison of the amino-acid and nucleotide sequences of the constant regions of γ1, γ2a and γ2b is presented in Tab. 2 and Fig. 3. We compa-

		γ2a	γ2b			γ2a	γ2b			γ2a	γ2b			γ2a	γ2b
F (Phe)	UUU	1	1	S (Ser)	UCU	4	5	Y (Tyr)	UAU	3	2	C (Cys)	UGU	4	4
	UUC	8	8		UCC	8	10		UAC	7	8		UGC	6	7
L (Leu)	UUA	0	0		UCA	4	7	(Ter)	UAA	0	0	(Ter)	UGA	1	1
	UUG	3	2		UCG	3	0		UAG	0	0	W (Trp)	UGG	6	6
L (Leu)	CUU	0	1	P (Pro)	CCU	7	6	H (His)	CAU	1	2	R (Arg)	CGU	0	0
	CUC	7	8		CCC	8	8		CAC	8	6		CGC	0	0
	CUA	2	2		CCA	14	15	Q (Gln)	CAA	1	2		CGA	0	0
	CUG	9	9		CCG	2	2		CAG	8	7		CGG	2	2
I (Ile)	AUU	2	2	T (Thr)	ACU	9	9	N (Asn)	AAU	3	4	S (Ser)	AGU	4	6
	AUc	9	12		ACC	12	10		AAC	12	11		AGC	12	14
	AUA	1	1		ACA	9	10	K (Lys)	AAA	10	13	R (Arg)	AGA	6	4
M (Met)	AUG	6	4		ACG	1	2		AAG	14	11		AGG	0	1
V (Val)	GUU	0	1	A (Ala)	GCU	3	5	D (Asp)	GAU	7	7	G (Gly)	GGU	6	5
	GUC	13	13		GCC	5	3		GAC	8	9		GGC	2	3
	GUA	5	3		GCA	2	2	E (Glu)	GAA	7	1		GGA	4	6
	GUG	18	16		GCG	1	0		GAG	10	14		GGG	3	4

Table 1. Codon usage for the C region of the γ2a and γ2b (9) heavy chain mRNAs.

COMPARISON	CH1	HINGE	CH2	CH3	3' UT	Entire CH
γ2a/γ2b	87% / 88% (0,0)	64% / 70% (2,18)	94% / 95% (0,0)	61% / 74% (0,0)	89% (0,0)	79% / 85% (2,18)
γ2a/γ1	85% / 88% (0,0)	38% / 44% (1,9)	67% / 77% (1,9)	63% / 72% (0,0)	70% (2,2)	69% / 76% (2,18)
γ2b/γ1	84% / 88% (0,0)	18% / 32% (2,27)	67% / 74% (1,9)	54% / 68% (0,0)	70% (2,2)	65% / 73% (3,36)

Table 2. Comparison of mouse γsubclass C region amino-acid and nucleotide sequences. The sequences were aligned to maximize the percent homology (amino-acid or nucleotide identities x 100 / total number of amino-acids or nucleotides compared). The upper portion of the square gives the amino-acid percent homology, the lower portion the nucleotide percent homology. In brackets is the number of gaps required followed by the total number of nucleotides included in the gaps. The data for γ2b are from (9), those for γ1 from (7) and (30) for the 3' UT region.

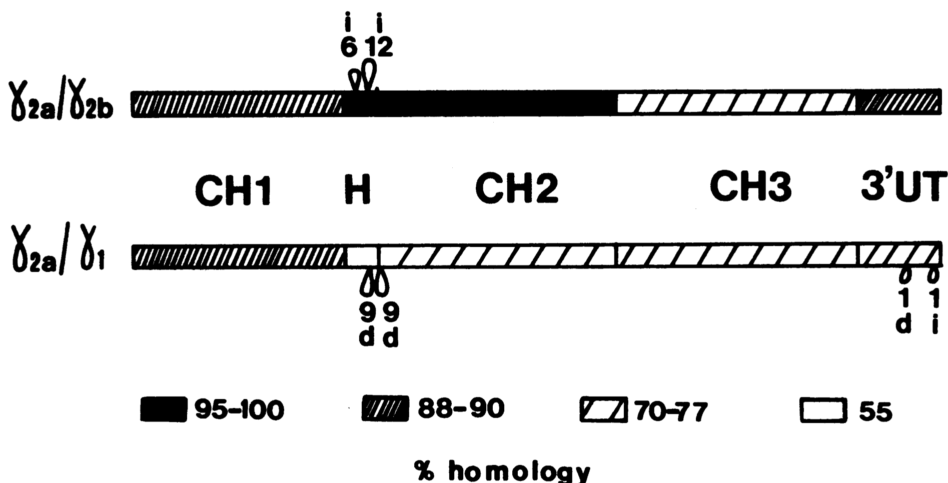

Fig. 3 : Homology maps of γ2a, γ2b and γ1 mRNAs :
Increased homology is shown by increasing density of hatching. Insertions (i) or deletions (d) are shown as loops with the number of nucleotides indicated.

red the γ2a sequence presented here with γ1 and γ2b sequences published by Honjo et al. (7) and Yamawaki-Kataoka et al. (9).

It appears that the CH1 and the CH3 of the 3 proteins can be aligned without introducing a gap. These conclusions are different from those derived from the protein sequence. Comparing corresponding domains two by two, it appears that the rates of divergence of one CH1 relatively to another are very similar, since they exhibit 88% of homology at the nucleotide level. In contrast, the maximum of divergence is found in the CH3 domain with about 70% of homology while an intermediate rate of divergence is observed when comparing the CH2 from γ2a and γ2b with γ1. In contrast, the CH2 of γ2a and γ2b are highly conserved with 95% of homology at the nucleotide level. The maximum of divergence is found in the Hinge where deletion or insertion have to be introduced to maximize the homology.

The 3' non coding region appears to be at least as conserved than some of the coding regions. It is worth noting that the three 3' non coding regions of γ1 (30), γ2a and γ2b (25) heavy mRNAs have exactly the same length of 103 nucleotides. Nevertheless, the poly-adenylation site is not the same for similar classes of proteins : GU has been found for γ2a and γ2b (9,25), GC for γ1 (30), whereas UG has been found for κ chains (31,32).

Comparative sequences of IgG subclasses led to the hypothesis that subclasses resulted from recent tandem duplication (13,15). Studies on the homology of domains from IgG subclasses reported here show that the evolution of corresponding domains occurs with variable rates of divergence.

It is generally assumed that preservation of homology correlates with function. Nevertheless, at the present time the biological signification of the codon usage pattern and the conservative structure of the 3' non coding region remain unclear.

The determination of the corresponding flanking and intervening sequences of the C γ2a gene would permit to obtain a more complete knowledge on the molecular mechanisms of the evolution of the γ subclass.

AKNOWLEDGEMENTS

We thank B. CHAMBRAUD and R. NAGEOTTE who provided excellent technical assistance.

This work was supported by grants from the C.N.R.S. (A.T.P. 3663 and 4247) and the Fondation pour la Recherche Médicale Française.

REFERENCES

1 - Putnam, F.W. (1977) in The Plasma Proteins, Putnam, F.W., Ed., New-York Academic Press, 1-153.
2 - Putnam, F.W. (1977) ibid, 223-284.
3 - Brack, C., Hirama, M., Lenhard-Schuller, R. and Tonegawa, S. (1978) Cell 15, 1-14.
4 - Seidman, J.G. and Leder, P. (1978) Nature 276, 790-795.
5 - Early, P.W., Davis, M.M., Kaback, D.B., Davidson, N. and Hood, L. (1979) Proc. Natl. Acad. Sci. USA 76, 857-861.
6 - Sakano, H., Rogers, J.H., Hüppi, K., Brack, C., Traunecker, A., Maki, R., Wall, R. and Tonegawa, S. (1979) Nature 277, 627-632.
7 - Honjo, T., Obata, M., Yamawaki-Kataoka, Y., Kataoka, T., Kawakami, T., Takahashi, N. and Mano, Y. (1979) Cell 18, 559-568.
8 - Tucker, P.W., Marcu, K.B., Newell, N., Richards, J. and Blattner, F.R. (1979) Science 206, 1303-1306.
9 - Yamawaki-Kataoka, Y., Kataoka, T., Takahashi, N., Obata, M. and Honjo, T. (1980) Nature 283, 786-789.
10 - Gough, N.M., Kemp, D.J., Tyler, B.M., Adams, J.M. and Cory S. (1980) Proc. Natl. Acad. Sci. USA 77, 554-558.
11 - Calame, K., Rogers, J., Early, P., Davis, M., Livant, D., Wall, R. and Hood, L. (1980) Nature 284, 452-455.
12 - Fougereau, M., Bourgois, A., de Preval, C., Rocca-Serra, J. and Schiff, C. (1976) Ann. Immunol. (Inst. Past.) 127 C, 607-631.
13 - Adetugbo, K. (1978) J. Biol. Chem. 253, 6068-6075.
14 - Yamawaki-Kataoka, Y., Sato, K., Shimizu, A., Kataoka, T., Mano, Y., Ono, M., Kawakami, M. and Honjo, T. (1979) Biochemistry, 490-494.
15 - Milstein , C. and Pinck, J.R.L. (1970) Progr. Biophys. Mol. Biol. 21, 209-263.
16 - Auffray, C., Nageotte, R., Chambraud, B. and Rougeon, F. (1980) Nucleic Acids Research 8, 1231-1241.
17 - Chang, L.M.S. and Bollum, F.J. (1971) J. Biol. Chem. 246, 909-916.
18 - Panet, A., VandeSande, J.H., Loewen, P.C., Khorana, H.G., Raae, A.J., Lillehaug, J.R. and Kleppe, K. (1973) Biochemistry 12, 5045-5049.
19 - Greene, P.J., Heyneker, H.L., Bolivar, F., Rodriguez, R.L., Betlach, M.C., Covarrubias, A.A., Backman, K., Russel, D.J., Tait, R. and Boyer, H.W. (1978) Nucleic Acids Research 5, 2373-2380.
20 - Katz, L., Kingsbury, D.T. and Helinsky, D.R. (1973) J. Bact. 114, 557-591.
21 - Berkner, K.L. and Folk, W.R. (1977) J. Biol. Chem. 252, 3176-3184.
22 - Roychoudhury, R. and Wu, R. (1979) Methods in Enzymology 65, 43.
23 - Maxam, A.M. and Gilbert, W. (1977) Proc. Natl. Acad. Sci. USA 74, 560-564.
24 - Sanger, F. and Coulson, A.R. (1978) F.E.B.S. Letters 87, 107-116.
25 - Tucker, P.W., Marcu, K.B., Slightom, J.L. and Blattner, F.R. (1979) Science 206, 1299-1302.
26 - Proudfoot, N.J. and Brownlee, G.G. (1976) Nature 263, 211-214.
27 - Ullrich, A., Shine, J., Chirgwin, J., Pictet, R., Tischer, E., Rutler, W. and Goodman, H.M. (1977) Science 196, 1313-1319.
28 - Kenter, A.L. and Birshtein, B.K. (1979) Science 206, 1307-1309.
29 - Konkel, D.A., Maizel, J.V. and Leder, P. (1979) Cell 18, 865-873.
30 - Dunnick, W., Rabbitts, T.H. and Milstein, C. (1980) Nucleic Acids Research 8, 1475-1484.
31 - Hamlyn, P.H., Brownlee, G.G., Cheng, C.C., Gait, M.J. and Milstein, C. (1978) Cell 15, 1067-1075.
32 - Gough, N.M., Cory, S. and Adams, J.M. (1979) Nature 281, 394-396.

33 - Dayhoff, M.O. (1976) Atlas of Protein Sequence and Structure (National Biomedical Research Foundation, Silver Spring, MD) 5, 189-190.

34 - Ohmori, H., Tomizawa, J.I. and Maxam, A.M. (1978) Nucleic Acids Research 5, 1479-1485.