Phylogenetic conservation of immunoglobulin heavy chains: direct comparison of hamster and mouse C_{μ} genes

K.L.McGuire, W.R.Duncan² and P.W.Tucker^{1*}

Immunology Graduate Program and ¹Department of Microbiology, University of Texas Southwestern Medical School, Dallas, TX 75235, USA, and ²Department of Surgery, Dalhousi University, Halifax, Nova Scotia B3H 4H7, Canada

Received 9 April 1985; Revised and Accepted 2 July 1985

ABSTRACT

We have analyzed the $J_{\text{H}}\text{-}C_{\mu}$ locus of the Syrian hamster by DNA cloning and sequencing. The single C_{μ} gene is highly homologous to that of the mouse. The hamster equivalents of the J_{H} and switch (S) recombination regions are arranged as in the mouse, but surprisingly are not highly conserved. Also unlike its close murine relative, the S_{μ} regions among inbred hamster strains are not polymorphic. The complete nucleotide sequence of hamster and mouse C_{μ} genes have been compared to partial C_{μ} sequences of other species. Conservation within a portion of the 3' untranslated region may signify functional requirements for 3' end processing. Mutational frequencies within exons and introns of hamster and mouse do not support the theory that the rate of DNA transitions to transversions decreases with evolutionary distance.

INTRODUCTION

The complex locus that encodes the immunoglobulin (Ig) heavy (H) chains and the events that control H chain expression have been studied extensively in the mouse (for reviews see refs. 1 and 2). The basic germline arrangement on chromosome 12 is as follows: 5'---~100 variable (V) segments, a long intervening sequence of undetermined length, ~20 diversity (D) segments clustered upstream of four functional joining (J) segments linked to a tandem of eight constant (C) region genes--3'. Somatic DNA recombination of a single V, D, and J segment in lymphoid tissue creates a functional variable region (VDJ) gene. During the primary immune response, transcription initiates upstream of VDJ and extends through the first C_{H} gene downstream, C_{μ} , leading to expression of the μ heavy chain of IgM. In the secondary (IgG and IgA) response, downstream C_{H} genes (eg, C_{γ} or C_{α}) replace C_{μ} by virtue of a second somatic DNA recombination referred to as class switching. This event is mediated by homologous recombination between repetitive switch (S) sequences positioned 5' to C_{μ} and the other C_{H} genes. Further diversification in expression of C_{H} is accomplished by alternative 3' end usage and splicing of separately encoded carboxyl terminal exons to produce either the secreted or membrane (M) form of the C_{H} gene products.

In addition to its role in primary antibody (Ab) response to antigenic stimulation, IgM is the first Ig to be produced in ontogeny and is also one of the major species of cell surface Ig (3,4). Thus, this molecule appears to serve a vital function in immune responses and has proven to be highly conserved phylogenetically (5). However, despite the importance of this molecule immunologically and evolutionarily, only one complete genomic μ sequence has been reported (6,7).

A particularly suitable model for an evolutionary comparison of the Cp gene is the Syrian hamster, which diverged from the mouse approximately 20 million years ago. The hamster is a unique immunological model as well. For example, these animals do not produce alloantibody against Class I major histocompatibility antigens and, in fact, appear to be non-polymorphic at these loci (8; McGuire, *et al.*, submitted). Additionally, hamsters do not appear to mount secondary responses to certain antigens, suggestive of a defect in class switching (9). Also, the B cells of adult hamsters appear to represent a less mature phenotype than adult mouse B cells (10,11).

Observations such as these prompted us to study the Cp gene in this species to determine if the Syrian hamster possesses a prototypic heavy chain locus. We report here the complete nucleotide sequence of the Syrian hamster Cp gene including the secreted terminus and 3' flanking region and a partial characterization of the J_H and the S μ regions. Our studies reveal many similarities as well as several unexpected differences between these closely related members of the rodent family.

MATERIALS AND METHODS

Construction of a Genomic Library from the Syrian Hamster LSH Inbred Strain.

High molecular weight DNA was isolated from liver tissue of the LSH inbred strain of Syrian hamster by the method of Blin and Stafford (12). The DNA was partially digested with Eco RI and fractionated on 5-20% NaCl gradients. Fractions containing fragments ranging from 8-20 Kilobase pairs (Kbp), as determined by agarose gel electrophoresis, were pooled, N-propanol precipitated, and ligated into arms prepared from the Ch4A cloning vector (13). Recombinant phage molecules were packaged and 1.9×10^6 plaque forming units (pfu) were obtained. The library was then amplified by plate lysis (14).

Isolation of a Recombinant Phage Clone Containing Cp.

Approximately 10^6 recombinant pfu from the amplified library were screened for Cp by *in situ* hybridization on nitrocellulose (15). The Cp probe was a 754 bp Pvu II-Pst I fragment encoding the Cp2 and Cp3 exons isolated

from a cDNA originally obtained from the NZB mouse plasmacytoma, 3741 (K. Marcu and P. Tucker, unpublished results). It was labeled with $\alpha^{32}\text{P}$ -dCTP to a specific activity of 10^8 cpm/ μg by the nick-translation method of Rigby, *et al.* (16). The filters were hybridized in 50% formamide, 5X SSC (0.75 M NaCl, 0.075 M sodium citrate, pH 7.2), 2X Denharts (0.04% bovine serum albumin, pyrrolidone, and ficoll), 100-200 $\mu\text{g}/\text{ml}$ degraded herring sperm DNA, and 30-50 $\mu\text{g}/\text{ml}$ polyadenylic acid (poly(A)) for 48 hours at 42°C after an 18 hour prehybridization under the same conditions without the poly(A). The filters were subsequently washed in 1X SSC at room temperature followed by a wash in 1X SSC at 42°C . Putative positive phage were screened and one was verified to strongly hybridize to the Cp probe. This clone, Ch4AH μ 1-1, was purified to homogeneity and DNA was isolated using the method of Leder *et al.* (17).

Genomic Blots and Southern Hybridizations.

10-20 μg of hamster (LSH strain) liver DNA or 2-3 μg of cloned DNA were digested with appropriate restriction endonucleases and electrophoresed on 0.8% agarose gels in 2X Hellings (0.08 M Tris-acetate, 0.04M sodium acetate, 0.035 M NaCl, 0.004 M EDTA, pH 8.0). The DNA was transferred to nitrocellulose by the method of Southern (18) with 10X SSC and hybridized as described above. The probes used are described in the text and figure legends and were nick-translated as described above. All filters were washed subsequent to hybridization in 2X SSC at room temperature followed by the washes indicated in the figure legends.

Subcloning and Sequencing.

Fragments containing the entire Cp gene were isolated and subcloned into the Hind III site or the Hind III-Eco RI sites of pAT 153. Sequencing was done by the chemical degradation method of Maxam and Gilbert (19) on DNA labeled with $\gamma^{32}\text{P}$ -ATP on the 5' ends or DNA labeled with the appropriate $\alpha^{32}\text{P}$ -dNTP on the 3' ends.

Alignments.

Sequence alignments between various Cp genes were done using the computer algorithms of Shroeder and Blattner (20) and Wilbur and Lipman (21).

Heteroduplex Analysis.

The cloned DNAs were hybridized overnight at 25°C in sealed capillaries containing 20 μl of a solution containing 50% deionized formamide, 0.1 M Tris-HCl, 10 mM EDTA, pH 8.5 at a final concentration of 2 $\mu\text{g}/\text{ml}$. The solution was then chilled to 4°C , spread on a hypophase of 50% formamide, picked up on parlodion-coated grids, stained with uranyl-acetate and shadowed with platinum/palladium and carbon on a rotating stage. The spreads were

examined with a Zeiss electron microscope, EM109, and measured with a Numonics graphic calculator digitizer interfaced to a digital VAX computer. The internal standard was pBR322.

RESULTS

General Structure of the J_H -C μ Locus in the Syrian Hamster.

Southern analysis of Syrian hamster liver DNA revealed only one C μ gene, as had been demonstrated in the mouse (22) (Figure 1). A clone containing this gene was isolated from an Eco RI partial library constructed in the Ch4A cloning vector. The clone, Ch4AH μ 1-1, contained two identically sized 7.5 Kbp Eco RI fragments and was shown by Southern hybridization to mouse probes to contain S μ and J H sequences as well as the C μ gene. Heteroduplex analysis

Fig. 1. Comparative Hybridization Analysis of the Cloned and Genomic C μ Gene in the Hamster. The probe used was the entire hamster C μ genomic gene which included the inserts of both pH μ 1-1a and pH μ 1-1b. L--15 μ g of hamster LSH strain liver DNA. C--1 ng of C μ clone Ch4AH μ 1- mixed with 10 μ g pAT153 DNA as a carrier. The DNAs were digested with the enzymes indicated. The filter was washed at a final stringency of 0.1X SSC at 42°C. The markers depicted are Eco RI digested Charon 10 and λ phage KH100nin5.

Fig. 2. Heteroduplex Analysis of J_H - C_μ Locus Between Hamster and Mouse. Analysis was carried out utilizing the mouse μ clone, Ch4A142.7, described previously (Newell *et al.*, 1980). Briefly, it contains J_H cluster, S_μ , the C_μ gene and the M1 and M2 membrane exons of the BALB/c C_μ gene. It is in the same orientation in the phage as Ch4AH μ 1-1. Eight molecules were examined and measured. A. Electron Micrograph. Variable hybridization was seen among the hybrid molecules as discussed in the text, and this molecule was chosen as a representative molecule because of its completeness. B. Interpretive Drawing. a) deletion loop of ~ 3.7 Kbp corresponding to the membrane exons of the mouse μ not represented in the hamster clone, b) ~ 3.1 Kbp of conserved sequence which corresponds to the μ gene and about 1.0 Kbp of 5' flanking sequences, c) the area of variable hybridization that is different for each hybrid molecule analysed and extends through most of the J_H complex as well as the switch recombination sequences. This area also includes the J - S_μ region that encodes the Ig enhancer, d) ~ 2.0 Kbp of homologous DNA seen in most of the molecules analyses (5/7) extending just up to and including the region encoding J_H , e) ~ 1.5 Kbp of consistently non-hybridizing DNA present on the mouse molecule, f) 1.5 Kbp of non-homologous hamster sequences plus additional 5' flanking DNA not represented on the mouse μ clone.

Fig. 3. Physical Map of the Syrian Hamster C_μ Locus. A. Restriction map of the genomic C_μ gene including the J_H and S_μ regions. Eco RI fragments cloned into Ch4AHμ1-1 are indicated above the map. Black boxes indicate C_μ exons while the white box indicates the μS terminus. B. Extended map of the C_μ gene. The narrow extension of the white box is the 3' untranslated region with the hexanucleotide AATAAA indicated by a dot. Sequencing strategies are indicated below the map. Bolder arrows indicate 3' end labeling. The entire region has been determined on both strands. Enzymes shown are: E-Eco RI, H-Hind III, K-Kpn I, B-Bam HI, X-Xba I, P-Pst I, N-Nco I, A-Ava I, Bg-Bgl II, R-Rsa I, Xn-Xmn I.

(Figure 2) as well as hybridization and restriction mapping allowed the characterization of the locus as shown in Figure 3. As indicated, the locus shares structural homology with the corresponding regions in mouse and human DNA.

Joining (J_H) Segments Reside 5' to the C_μ Locus.

To determine if the Syrian hamster possesses a J_H region, the insert of Ch4AHμ1-1 was used as a probe and hybridized to plasmids containing various J_H regions of the mouse. Hamster DNA hybridized to J_H1 and J_H4, less to J_H3, and not at all to J_H2 (Figure 4A, B). Using probes to the various mouse J segments it was determined that the Syrian hamster J_H region is localized to the Kpn I-Xba I fragment shown in Figure 3. A fragment containing J_H1 and J_H2 hybridizes to the 3' Kpn I fragment and to 5' Bam HI fragment suggesting that a mouse J_H1-like segment is localized to the 100 bp Kpn I-Bam HI fragment. The hybridization of J_H3 and J_H4 maps 3' to J_H1 suggesting that the homologous J regions lie in the same orientation as in the mouse and that they span the region noted in Figure 3. These results suggest to us that the hamster

Fig. 4. Analysis of the Hamster J_H Region. A. Hybridization of cloned hamster DNA to cloned mouse J_{H1} and J_{H2} . The Eco RI insert of Ch4AHu1.1 is hybridized to plasmid digests (2 μ g each) at a final wash condition of 3X SSC, 55°. With Bam HI (left), J_{H1} and J_{H2} co-migrate on a 1005 bp fragment; vector sequences show low, nonspecific hybridization. Hind III digestion (right) separates the 944 J_{H1} -containing fragment from the J_{H2} plus vector sequences. B. Hybridization of cloned hamster DNA to cloned mouse J_{H3} and J_{H4} . The same hamster probe hybridizes equally well to the 2390 bp J_{H4} -containing fragment and to the 4494 bp J_{H3} plus vector-containing fragment. C. Mouse J_{H4} fragment does not hybridize to hamster genomic DNA. A 2190 bp Hind III-Eco RI fragment spanning J_{H4} was hybridized to 15 μ g of mouse (left) and hamster (right) liver DNA digested with Eco RI. Final wash was 1.0X SSC, 25°. Identical results were obtained with individual fragments spanning J_{H1} , J_{H2} and J_{H3} .

possesses three J_H segments but it can not be ruled out that other J segments exist that are not homologous enough with the mouse J segments to be detected in this manner.

To further characterize this region we attempted Southern analysis of genomic DNA. Surprisingly, fragments containing just J_{H1} and J_{H2} or J_{H4} (Figure 4C) do not hybridize at all to hamster genomic DNA under normal hybridization conditions using low stringency washes. This was unexpected considering the conservation seen in the $C\mu$ gene (see below). The heteroduplex analysis (Figure 2) confirmed this limited homology. Variable hybridization was seen in this region with the possible exception of the area of homology which includes and extends 5' to J_{H1} . It is interesting that the IVS-2 of the $C\mu$ gene showed consistent hybridization in the heteroduplex analysis and yet is only 65% conserved between the two species. This suggests

that the variable hybridization seen in the J_H portion represents homologies less than 65% between the two loci. Although the hamster possesses three J_H segments similar to those in the mouse, this portion of the locus does not share extensive sequence homology with the corresponding mouse region. However, short regions of homology may exist that are undetectable by these methods.

Deletion 5' to the Hamster $C\mu$ Gene Suggests a Switch Recombination Region.

Southern hybridization analysis of liver DNA demonstrated that the germline $C\mu$ gene of the LSH hamster strain was contained on a 10.0 Kbp Eco RI fragment. However, the Eco RI fragment containing the $C\mu$ gene in Ch4AH μ 1-1 is only 7.5 Kbp suggesting that a 2.5 Kbp deletion took place upon cloning (Figure 1). This is a common problem encountered when cloning the mouse $C\mu$ gene because of rearrangements occurring in the $S\mu$ region (22,23). To determine where the deletion occurred, DNA was digested with Bam HI and Eco RI. This digest showed that the μ fragment in the clone comigrated with the genomic gene, placing the deletion area 5' to the Bam HI site. The 5.0 Kbp hybridizing fragment in the cloned DNA is the Bam HI fragment 5' to the $C\mu$ gene. Although not evident in Figure 1, a faint corresponding genomic fragment was seen at 7.5 Kbp in the original blot. The Xba I-Eco RI digest of genomic DNA yields a 7.3 Kbp fragment while the clone contains a 4.8 Kbp hybridizing band. These data support the conclusion that a 2.5 Kbp deletion has taken place in Ch4AH μ 1-1, suggesting that hamsters do possess $S\mu$ sequences. This was further supported by the fact that the clone hybridizes to a mouse $S\mu$ probe (data not shown). The data shown in Figure 1 also strongly suggest that the $S\mu$ lies entirely between the Xba I site and the Hind III site 5' to $C\mu$ 1 as depicted in Figure 3. Interestingly enough, this area also falls into the region of variable homology seen in the heteroduplex analysis. Although variable deletions in the clones may partially account for these results, they suggest to us that again the two species may share only limited sequence homology in this region. Sequence analysis of an undeleted $S\mu$ region is required to say more.

Genomic variation in the length of $S\mu$ has been demonstrated among inbred mouse strains (22). Because hamsters appear to possess a $S\mu$ region and it appears to delete similarly as that of the mouse upon cloning, an analysis of several inbred hamster strains and F1 hybrids was performed to test for genomic length variation. As can be seen in Figure 4, no variation in the region spanning $S\mu$ and $C\mu$ is observed either in the Bam HI (Figure 4A) or in the Eco RI (Figure 5B) digests among the animals tested. (The upper band in

Fig. 5. Lack of DNA Polymorphism of the Switch Recombination Region Among Inbred Hamster Strains. The probe used was the same as in Fig. 1. Enzymes used were: A, *Bam* HI and B, *Eco* RI. The lanes represent strains: 1, LSH; 2, UT4; 3, (MHA.UT2 x MHA)F1; 4, MHA; 5, (CB.BUN x CB)F1; 6, (CB x LSH)F1 and 7, MIT. The filters were washed at a final stringency of 0.1X SSC at 55°C. The congenic animals used are congenic for the hamster MHC, Hm-1, not Ig. (For a review of the origin of these strains see ref. 8.)

lane 6 is a partial digest. It was not seen consistently in other digests.) It is interesting that rearrangements and deletions occur upon cloning but genomic polymorphism is not seen in strains that were isolated from the wild forty years apart (lanes 1 and 4 versus lanes 2 and 7) (24).

Analysis of the C μ Gene.

The entire C μ gene was located on two small restriction fragments directly adjacent to the left arm of the λ cloning vector. The fragments were subcloned into pAT 153: pH μ 1-1b contains the 1.4 Kbp *Hind* III-*Hind* III-*Eco* RI fragment that encodes C μ 3 and C μ 4. These two subclones were used in subsequent sequence analysis. The strategies used are shown in Figure 3.

Figure 6 shows the DNA sequence for the C μ gene of the Syrian hamster and the predicted amino acid sequence for the secreted protein. It is compared directly to the sequence of the BALB/c C μ gene (6). The sequences are extremely similar with only 23, 22 and 21 amino acid substitutions in the first three domains, respectively. The fourth domain is slightly more conserved with only 18 amino acid substitutions. Especially conserved is the μ S exon adjacent to the fourth domain which encodes the C-terminus of the secreted protein. This exon is 95% conserved at the amino acid level between the two species, containing only one amino acid substitution.

The predicted amino acid sequence reveals three additional N-linked glycosylation sites in the hamster sequences as compared to the mouse. It also indicates that the tryptophan and cysteine residues are conserved between the two molecules, the latter suggesting disulfide bonding and secondary structure is conserved.

The 3' untranslated region is highly conserved up to the canonical hexanucleotide AATAAA (end site) that occurs some 20 nucleotides 5' to the

H 100
 M
 H 200
 M
 H 300
 M
 H 400
 M
 H 500
 M
 H 600
 M
 H 700
 M
 H 800
 M
 H 900
 M
 H 1000
 M
 H 1100
 M
 H 1200
 M
 H 1300
 M
 H 1400
 M
 H 1500
 M
 H 1600
 M
 H 1700
 M
 H 1800
 M
 H 1900
 M
 H 2000
 M
 H 2100
 M
 H 2200
 M
 H 2300
 M
 H 2400
 M
 H 2500
 M
 H 2600
 M
 H 2700
 M
 H 2800
 M
 H 2900
 M
 H 3000
 M
 H 3100
 M
 H 3200
 M
 H 3300
 M
 H 3400
 M
 H 3500
 M
 H 3600
 M
 H 3700
 M
 H 3800
 M
 H 3900
 M
 H 4000
 M
 H 4100
 M
 H 4200
 M
 H 4300
 M
 H 4400
 M
 H 4500
 M
 H 4600
 M
 H 4700
 M
 H 4800
 M
 H 4900
 M
 H 5000
 M
 H 5100
 M
 H 5200
 M
 H 5300
 M
 H 5400
 M
 H 5500
 M
 H 5600
 M
 H 5700
 M
 H 5800
 M
 H 5900
 M
 H 6000
 M
 H 6100
 M
 H 6200
 M
 H 6300
 M
 H 6400
 M
 H 6500
 M
 H 6600
 M
 H 6700
 M
 H 6800
 M
 H 6900
 M
 H 7000
 M
 H 7100
 M
 H 7200
 M
 H 7300
 M
 H 7400
 M
 H 7500
 M
 H 7600
 M
 H 7700
 M
 H 7800
 M
 H 7900
 M
 H 8000
 M
 H 8100
 M
 H 8200
 M
 H 8300
 M
 H 8400
 M
 H 8500
 M
 H 8600
 M
 H 8700
 M
 H 8800
 M
 H 8900
 M
 H 9000
 M
 H 9100
 M
 H 9200
 M
 H 9300
 M
 H 9400
 M
 H 9500
 M
 H 9600
 M
 H 9700
 M
 H 9800
 M
 H 9900
 M
 H 10000
 M

site of poly(A) addition (25). DNA 3' to the poly(A) addition site is extremely conserved as well (see Table I).

The overall structure of the gene is very similar to the mouse, in that exon and intron lengths are highly conserved and no gaps are required to align the coding regions. Heteroduplex analysis (Figure 2) revealed complete complementarity between the two Cp genes including all the introns and the 3' flanking region. The homology extends ~1.0 Kbp 5' to the Cp gene in all hybrid molecules analysed, suggesting that the high degree of conservation between the two genes applies to the 5' flanking region as well. However, as previously noted, this extensive homology does not extend into the areas encoding the J_H and Sp regions.

In Table I, the nucleotide sequence of the hamster Cp gene was compared with available sequences from other species. In addition to the complete comparisons between the hamster and mouse genes, alignments were done with the cDNA sequences for chicken (26) and rabbit (27), as well as the limited genomic sequence available for the human gene (28-30). The mouse and the hamster genes are equally homologous to the chicken gene with the exception of the avian μ S exon, which is markedly more homologous with the mouse than with the hamster. It is interesting that the rabbit constant region domains are consistently more homologous with the hamster gene, while the μ S exon is again more related to the mouse. This holds true for the limited analysis possible for the human gene as well.

When comparing the 3' untranslated region of the rabbit and human genes to the corresponding regions in hamster and mouse (Figure 7), the homology appears to divide into two sections. The 5' side of the 3'-untranslated region, immediately adjacent to the termination codon of the μ S exon, is not highly conserved. This is especially evident in the rabbit sequence in which three gaps (including 29 bp) and ten single base changes occur. While the

Fig. 6. Sequence of the Syrian Hamster Cp Gene. The coding strand sequence of the hamster (H) Cp gene is compared directly to the murine (M) BALB/c Cp gene (6). Nucleotide substitutions in the mouse gene are indicated below the hamster sequence. The predicted hamster amino acid sequence is shown in single letter code directly above the first base of the corresponding codon in the DNA sequence while substitutions in the mouse protein are indicated below the mouse DNA sequence. The dashed lines indicate identity and open spaces indicate gaps. The end site, AATAAA, is in lower case letters and the putative N-linked glycosylation sites are underlined. The nucleotide numbering for the hamster gene is given to the right. The one letter amino acid code is as follows: A, alanine; C, cysteine; D, aspartic acid; E, glutamic acid; F, phenylalanine; G, glycine; H, histidine; I, isoleucine; K, lysine; L, leucine; M, methionine; N, asparagine; P, proline; Q, glutamine; R, arginine; S, serine; T, threonine; V, valine; W, tryptophan and Y, tyrosine.

Table I. Comparisons of Available Cp Sequences with Hamster and Mouse.

	hamster/ mouse	hamster/ human	mouse/ human	hamster/ rabbit	mouse/ rabbit	hamster/ chicken	mouse/ chicken
5' flanking region	74.5 6,40						
Cu1	86.3 0,0			76.3 1,3	73.4 1,3		
IVS1	78.8 2,9						
Cu2	84.4 0,0			70.8 0,0	69.2 1,5	41.2 2,27	43.6 3,27
IVS2	65.2 3,25	47.5 3,11	42.0 4,22				
Cu3	85.8 0,0			73.3 0,0	69.5 0,0	53.8 0,0	52.3 0,0
IVS3	70.6 2,7						
Cu	87.4 0,0	80.1 1,3	78.2 1,3	83.5 0,0	81.1 0,0	58.6 0,0	55.4 1,7
uS	90.3 0,0	85.5 0,0	87.1 0,0	83.9 0,0	88.7 0,0	58.0 0,0	63.3 0,0
3'UT	88.5 0,0	67.8 4,10	70.9 4,7	77.9 6,38	78.9 6,34		
3' flanking region	78.1 2,20	52.9 2,25	44.4 1,8				

The partial sequences for the Cp genes of human (29), rabbit (27), and chicken (26) were compared to the Cp genes of the hamster and mouse (6). The numbers on the top represent percentage agreements; the numbers below designate the number of gaps required to best align the sequences followed by the number of bases included in the gaps.

human sequence contains no gaps, a lack of sequence conservation is evident. Even the mouse sequence shows more divergence in this area than in the downstream region, which is marked on its 5' and 3' ends with arrows in Figure 6 (the 3' arrow also denotes the site of poly(A) addition). Homologies with the hamster gene for the 5' region are 81.1%, 50.9%, and 59.4% for mouse, human, and rabbit, respectively. The corresponding homologies for the 3' region are 93.8%, 81.0%, and 83.1%. These data suggest that sequence conservation of the downstream portion of the 3' untranslated region is important, perhaps for recognition in termination and polyadenylation.

mouse and the human M1 exons (29 and Word *et al.*, unpublished) encompassed by our probe. These data suggest that the sequences diverge 3' to the C μ gene similarly to the divergence seen 5' to the gene.

DISCUSSION

We present here a characterization of the C μ heavy chain locus of the Syrian hamster. Overall the hamster and mouse Ig heavy chain loci share several structural features. The hamster has one genomic C μ gene which is extremely homologous with the mouse, and this conservation extends approximately 1.0 Kbp 5' to C μ 1 as well as at least 0.2 Kbp 3' to the μ S end site. These animals also possess J $_H$ regions and a switch region similar to the mouse, based on cross-hybridization of cloned DNAs. However, these portions of the locus are not highly conserved based on genomic Southern and heteroduplex analysis. Additionally, switch region polymorphism was not observed in this species as has been reported for the mouse. Surprisingly, hamster and mouse sequences appear to diverge 3' to the C μ coding region in that a probe encompassing the mouse μ membrane region did not cross-hybridize definitively with hamster genomic DNA even under low stringency wash conditions.

It has been reported by Eidels *et al.* (11) that the IgM immunoprecipitated from the surface of hamster lymphoid cells is consistently larger than the published apparent molecular weight for murine cell surface IgM (84 Kd vs. 82 Kd). The structural genes for C μ are identical in length between the two species so this difference in size should reflect post-translational modification (unless the membrane exons vary in size). This is supported by the presence of three additional putative N-linked glycosylation sites in our sequence. Another interesting aspect of hamster Ig previously reported by Eidels *et al.* (11) was that two subpopulations of proteins migrating at 84 Kd were immunoprecipitated by a rabbit anti-hamster immunoglobulin antiserum; one which was also immunoprecipitated by a rabbit anti-mouse μ reagent whereas the other was not. The authors suggested two possibilities for the identity of the "non-IgM" subpopulation: Either the molecules are IgM but lack the determinants necessary to be recognized by the rabbit anti-mouse μ reagent or they represent another class of Ig that comigrates with IgM. Our finding of only one C μ gene rules out the first hypothesis. Alternatively, these molecules may be non-reactive because of differential glycosylation or they represent, as previously proposed, another class of Ig that comigrates with IgM.

We observed significant interspecies homology in the 3' untranslated

region around the end and polyadenylation sites. This has been noted for many other conserved genes, including the actin family (32), globin, insulin, and growth hormone (33). It is interesting that the C μ sequences begin to diverge again immediately 3' to the poly(A) addition site. This suggests the sequences important in recognition of the poly(A) addition site lie upstream and that downstream sequence conservation is not essential for C μ gene processing. This is in contrast to a recent report by McDevitt *et al.* (34) which suggests sequences immediately downstream are important in generating a poly(A) addition site for an adenovirus E2A transcription unit. If this is the case for all eukaryotic transcription units, it is difficult to explain the lack of conservation in the downstream sequences unless they are species specific or secondary structure is conserved despite the lack of sequence homology.

The exact evolutionary relationship between the hamster and the mouse is not known. Simpson (35) places the divergence point somewhere in the late Miocene or early Pliocene period, approximately 20 million years ago. The C μ sequence homology suggests that these animals are closely related. Kimura (36) has developed a model to calculate rates of evolutionary distance for each codon position based on transition and transversion frequency and divergence time. He determined K values, or base substitution rates, for several genes including rabbit and mouse β -globin genes. These values were 0.16, 0.13, and 0.43 for the three codon positions, respectively. We calculated the corresponding K values for the hamster and mouse C μ genes as 0.10, 0.12, and 0.26. It is interesting that the values for the first two codon positions are essentially equal in the two comparisons even though the divergence time of mouse and rabbit (~80 million years) is approximately four times that of mouse and hamster. Even the third base changes are not four times as great, suggesting to us that these conserved gene sequences will only tolerate a certain amount of alteration and that other constraints placed on these sequences limits the divergence despite the time of separation.

Tucker *et al.* (37) in comparing the mouse γ 2b and γ 1 Ig genes found a higher percentage of third base changes in replacement codons than in nonreplacement ones. We have determined that the same holds true for the mouse C μ gene with respect to the hamster. Of the 370 conserved amino acids, 70 have silent third base mutations for a rate of 18.9%, slightly higher than the overall nucleotide divergence of 14%. However, 24 of the 85 non-conserved amino acid codons have third base alterations for a rate of 28.2%. This re-emphasizes that once a codon has changed, there is a selective pressure to

alter the third base to a more preferred codon.

Holmquist (38) has reported that the ratio of transitions to transversions decreases with an increase in evolutionary distance. To test this, we determined the transition to transversion ratio in the two C μ genes. This ratio is 1.4:1 (197/141) overall but 1.1:1 (77/72) in the introns and 1.7:1 (120/69) in the exons. The ratio is 2.8:1 for silent third base changes of which over half are transitions in conserved codons (53/92). The alteration to preferred codons does not account for this totally in that the ratio in non-conserved codons is 1.9:1. This suggests that even with evolutionary pressure to use a more preferred codon, there is still a preponderance of transitions. Therefore, our data do not support the Holmquist hypothesis but instead suggest that exons and introns fix transitions at different rates. The exact meaning of this observation is not clear. It is obvious, however, that mutations do not collect in these genes randomly as a function of time, but that they are directed, probably by mechanisms not yet fully understood.

If the inter-species homology demonstrated in the hamster J $_H$ -C μ region is generally reflected at other loci, it is improbable that many of the immunological curiosities of this animal are structurally derived. For example, there appears to be no simple molecular explanation, based on our finding of an S μ -like region, for the defect in class switching in certain immune responses. Perhaps the functional relevance of this finding draws on the structural uniformity of S μ and not on its presence or absence. A lack of DNA polymorphism in hamsters is demonstrable not only in the μ switch region but also in Class I MHC genes (McGuire *et al.*, submitted). It is highly improbable that this is due to a breeding artifact since the strains studied were independently derived from animals isolated from the wild 40 years apart. The theoretical basis that we favor is that the hamster underwent an evolutionary bottleneck; i.e., that all existing members of the species arose from a limited number of progenitors. But regardless of the mechanism, the lack of selective pressure to generate polymorphism must be linked to the hermit-like lifestyle of this species and ultimately must modulate the functional expression, if not the structure, of the immunological loci.

ACKNOWLEDGEMENTS

We gratefully acknowledge Dr. M. Sullivan for the heteroduplex analysis, Ms. C. Pramanik for technical assistance, Ms. M. Gardner and Ms. M. Allan for secretarial assistance, and Drs. D. Yuan, J. Guise and J.D. Capra for critically reviewing the manuscript. This work was supported by the Cancer

Immunology Training Grant CA 09082 and National Institutes of Health grants AI 18106, CA 31534 and RR01133. K. McGuire is a Robert A. Welch Fellow.

*To whom reprint requests should be sent

REFERENCES

1. Leder, P. (1982) *Scientific American*, 246, 102-115.
2. Tonegawa, S. (1983) *Nature*, 302, 575-581.
3. Vitetta, E.S., Baur, S. and Uhr, J.W. (1971) *J. Exp. Med.*, 134, 242-264.
4. Marchalonis, J.J., Cone, R.E. and Atwell, J.L. (1972) *J. Exp. Med.*, 135, 956-971.
5. Clem, I.W. and Leslie, G.A. (1971) *Proc. Natl. Acad. Sci. USA*, 68, 139-141.
6. Goldberg, G.I., Vanin, E.F., Zrolka, A.M. and Blattner, F.R. (1981) *Gene*, 15, 33-42.
7. Kawakami, T., Takahashi, N. and Honjo, T. (1980) *Nuc. Acids Res.*, 8, 3933-3945.
8. Streilein, J.W. and Duncan, W.R. (1981). *Transpl. Proc.*, 15, 1540-1545.
9. Streilein, J.W., Duncan, W.R. and Thompson, S.A. (1977) *Transpl. Proc.*, 9, 1229-1233.
10. Coe, J.E. (1971) *Immunology*, 21, 175-191.
11. Eidels, L., Proia, R.L., Streilein, J.W. and Hart, D.A. (1979) *Mol. Immunol.*, 16, 541-546.
12. Blin, N. and Stafford, D.W. (1976) *Nuc. Acids Res.*, 3, 2303-2308.
13. Blattner, F.R., Blechl, A.E., Denniston-Thompson, K., Faber, H.E., Richards, J.E., Slightom, J.L., Tucker, P.W. and Smithies, O. (1978) *Science*, 202, 1279-1283.
14. Tucker, P.W. (1984) In: "Handbook of Experimental Immunology", 4th edition, Weir & Herzenberg, eds., Blackwell publications, Ltd., Edinburgh, England, in press.
15. Benton, W.D. and Davis, R.W. (1977) *Science*, 196, 180-182.
16. Rigby, P.W.J., Dieckmann, M., Rhodes, C. and Berg, P. (1977) *J. Molec. Biol.*, 113, 237-251.
17. Leder, P., Tiemeier, O. and Enquist, L. (1977) *Science*, 196, 175-178.
18. Southern, E.M. (1975) *J. Molec. Biol.*, 98, 503-517.
19. Maxam, A.M. and Gilbert, W. (1980) *Meth. Enzymol.*, 65, 499-560.
20. Schroeder, J.L. and Blattner, F.R. (1982) *Nuc. Acids Res.*, 10, 69-84.
21. Wilber, W.J. and Lipman, D.J. (1983) *Proc. Natl. Acad. Sci. USA*, 80, 726-730.
22. Marcu, K.B., Banergi, J., Penncavage, N.A., Lang, R. and Arnheim, N. (1980) *Cell*, 22, 187-196.
23. Newell, N., Richards, J.E., Tucker, P.W. and Blattner, F.R. (1980) *Science*, 209, 1128-1133.
24. Streilein, J.W. and Duncan, W.R. (1979) *Immunogenetics*, 9, 563-573.
25. Proudfoot, N.J. and Brownlee, G.G. (1976) *Nature*, 272, 211-215.
26. Dahan, A., Reynaud, C.-A. and Weil, J.-C. (1983) *Nuc. Acids Res.*, 11, 5381-5389.
27. Bernstein, K.E., Alexander, C.B., Reddy, E.P. and Mage, R.C. (1984) *J. Immunol.*, 132, 490-495.
28. Dolby, T.W., Devuono, J. and Croce, C.M. (1980) *Proc. Natl. Acad. Sci. USA*, 77, 607-6031.
29. Milstein, C.P., Deverson, E.V. and Rabbitts, T.H. (1984) *Nuc. Acids Res.*, 12, 6523-6535.

30. Takahashi, N., Nakai, S. and Honjo, T. (1980) Nuc. Acids Res., 8, 5983-5991.
31. Richards, J.E., Gilliam, A.C., Shen, A., Tucker, P.W. and Blattner, F.R. (1983) Nature, 306, 483-487.
32. Gunning, P., Mohun, T., Ng, S.-Y., Ponte, P. and Kedes, L. (1984) J. Mol. Evol., 20, 202-214.
33. Miyata, T., Yasunaga, T. and Nishida, T. (1980) Proc. Natl. Acad. Sci. USA, 77, 7328-7332.
34. McDevitt, M.A., Imperiale, M.J., Ali, H. and Nevins, J.R. (1984) Cell, 37, 993-999.
35. Simpson, G.G. (1959) Cold Spring Harbor, Symp. Quant. Biol., 24, 255-271.
36. Kimura, M. (1978) Proc. Natl. Acad. Sci. USA, 78, 454-458.
37. Tucker, P.W., Marcu, K.B., Slightom, J.L. and Blattner, F.R. (1979) Science, 206, 1299-1303.
38. Holmquist, R. (1983) J. Mol. Evol., 19, 134-144.