The EMBO Journal vol.8 no.6 pp.1749–1759, 1989

Organization and structure of the Qa genes of the major histocompatibility complex of the C3H mouse: implications for Qa function and class I evolution

Suzanne Watts, Ann Cranmer Davis, Brandon Gaut, Christopher Wheeler, Lisa Hill and Robert S.Goodenow

Department of Genetics, University of California, Berkeley, CA 94720, USA

Communicated by P.A.Peterson

We have determined the structure and organization of the entire Qa family of class I genes from the major histocompatibility complex of the C3H mouse. Restriction maps of overlapping lambda and cosmid clones reveal that there are only five Qa^k genes: Q1^k, Q2^k, Q4^k, Q10^k and a Q5/9 hybrid, presumably generated by unequal homologous recombination. The resulting deletion of Q6–Q9 is consistent with the Qa-2^{null} phenotype of this mouse strain. We have sequenced the Qa^k genes, and predict that each may encode a class I molecule with a structure comparable with that proposed for the transplantation antigens. Furthermore, these Qa products should be able to bind peptides and interact with appropriate T-cell receptors. Interestingly, in comparing Qa^k and H-2^k sequences, we find limited evidence of interlocus gene conversion between Qa and H-2 loci, suggesting that the Qa genes are not likely to serve as a reservoir of genetic information for the generation of H-2 diversity within this haplotype.

Key words: Qa^k genes/class I/major histocompatibility complex/immune regulation/gene conversion

Introduction

The class I genes of the murine major histocompatibility complex (MHC) comprise a 30–40 member multigene family of structurally related glycoproteins (Hood *et al.*, 1983; Klein, 1986). The classical transplantation antigens, encoded by the H-2K and H-2D loci of the H-2 complex, are integral membrane proteins which are found on virtually all somatic cells and function in the presentation of peptide antigens to T-cells (Hood *et al.*, 1983). Telomeric to the H-2D region is the Qa subregion of the Tla complex which contains between 1 and 10 loci in the different inbred mouse strains (Weiss, 1987). Twelve to 20 class I genes map to the Tla region (Steinmetz *et al.*, 1982; Weiss *et al.*, 1984; Chen *et al.*, 1987). Several characterized cell surface or secreted class I products with restricted tissue distribution are encoded by Qa and Tla genes, although their function is not yet understood (Chen *et al.*, 1987; Robinson, 1987a).

Crystallization of the human HLA-A2 class I molecule (Bjorkman *et al.*, 1987a,b) has revealed that the amino (N)-terminal α_1 and α_2 domains fold into two α-helices on top of an antiparallel β-sheet, forming a cleft. Experimental evidence supports a model of antigen presentation in which peptides derived from processed foreign or self proteins are inserted into this cleft during intracellular trafficking for presentation to the T-cell receptor (Maryanski *et al.*, 1986; Townsend *et al.*, 1986; Clayberger *et al.*, 1987; Song *et al.*, 1988). In addition, the extraordinary polymorphism displayed by the transplantation antigens is concentrated at amino acid positions which contribute to this putative antigen-binding cleft (Klein, 1986; Bjorkman *et al.*, 1987b). This diversity, reflected in the ~50 H-2K and H-2D alleles characterized by serological methods, may function to ensure the ability of the population to survive infectious disease (Klein, 1986). Thus, the evolutionary mechanisms which drive diversification of the transplantation antigens may be important in enabling the immune system to respond to a vast array of pathogens (Klein, 1986; Hughes and Nei, 1988).

Although the Qa antigens share the same overall domain structure as the H-2 molecules, associating noncovalently with β_2-microglobulin, there is little evidence for their participation in classical MHC-restricted antigen presentation to T-cells (Robinson, 1987a). Three Qa gene products have been characterized to date (reviewed in Robinson, 1987a). The Qa-2 molecule, expressed on hemopoietic cells in a variety of adult tissues as well as at specific stages in embryonic development (Harris *et al.*, 1984; David-Watine *et al.*, 1987; Robinson, 1987a,b; Warner *et al.*, 1987), appears to be attached to the membrane by a phospholipid linkage (Steirnberg *et al.*, 1987; Stroynowski *et al.*, 1987; Soloski *et al.*, 1988a; Waneck *et al.*, 1988). Qa-2 is encoded by genes Q7^b and Q9^b in Qa-2^{high} C57BL/6 (Mellor *et al.*, 1985; Waneck *et al.*, 1987; Sherman *et al.*, 1988; Soloski *et al.*, 1988b), and by gene Q7^d (27.1) in Qa-2^{low} BALB/c (Mellor *et al.*, 1985; Soloski *et al.*, 1988a) mice. Qa-2^{null} strains are presumed to lack the structural genes (Flaherty *et al.*, 1985; Weiss, 1987). A developmental role has been suggested for Qa-2 based on correlations between Qa-2 expression and the rate of cell division in preimplantation mouse embryos (Warner *et al.*, 1987). The Q10 gene product, expressed in the liver, has a truncated transmembrane domain and is secreted directly (Cosman *et al.*, 1982; Kress *et al.*, 1983a,b; Maloy *et al.*, 1984; Mellor *et al.*, 1984; Devlin *et al.*, 1985a). Because it is a soluble molecule, theoretically capable of entering lymphoid tissues as well as encountering circulating T-cells, a role for Q10 in generating self tolerance has been proposed (Cosman *et al.*, 1982; Kress *et al.*, 1983a,b; Maloy *et al.*, 1984). Qb-1 (Robinson, 1985), a product of the Q4 gene in C57BL/6 mice (Robinson *et al.*, 1988), is also directly secreted from a wide variety of tissues (Palmer and Frelinge, 1987). The murine Qa alleles appear to be relatively conserved in comparison with those of H-2K^k and H-2D^k (Weiss, 1987).

Recently, several groups have described human non-classical class I molecules with limited tissue distribution (Geraghty *et al.*, 1987; Srivastava *et al.*, 1987; Koller *et al.*, 1988; Mizuno *et al.*, 1988; Shimizu *et al.*, 1988). However,

©IRL Press

molecules, but may correspond instead to murine Tla products.

There are two major questions concerning the Qa molecules. The first involves their function, especially with regard to the immune system and murine development. The second addresses the potential role of the Qa genes in the concerted evolution of the class I multigene family, and, more specifically, the generation and maintenance of H-2 diversity.

The question regarding Qa function is complicated by both the relative lack of polymorphism among alleles (Weiss, 1987), which can imply conservation of a structure uniquely suited for biological function, as well as the absence of multiple Qa genes in some inbred mouse strains (O'Neill *et al.*, 1986). In fact, the A.CA strain apparently lacks genes Q1−Q9 (O'Neill *et al.*, 1986), and does not express Q10 (Lew *et al.*, 1986). Nevertheless, the structural homology of Qa and H-2 molecules has lead to speculation that Qa products might be involved in a ligand-presentation role important for cell−cell interaction during development or immune surveillance (Warner *et al.*, 1987; Janeway, 1988).

A more indirect biological role has also been suggested for Qa genes in the concerted evolution of the class I multigene family (Pease *et al.*, 1983; Ritzel *et al.*, 1984; Pease, 1985; Nathenson *et al.*, 1986). Analysis of the H-2K^{bm} (Nathenson *et al.*, 1986) and the H-2K^{km2} (Vogel *et al.*, 1988) meiotic mutants has revealed multiple clustered nucleotide and amino acid changes. Moreover, potential donors for the altered sequences could often be detected among Qa genes (Nathenson *et al.*, 1986). It was hypothesized that gene conversion might be instrumental in the generation of extraordinary H-2 diversity (Ritzel *et al.*, 1984; Pease, 1985; Nathenson *et al.*, 1986).

We have characterized the organization and structure of the Qa-region genes of the C3H mouse (H-2^k) in order to address questions regarding Qa expression and function, as well as their role in the concerted evolution of the class I multigene family. We find that each of the five Qa genes potentially encodes a structurally intact class I molecule. With respect to class I evolution, comparison of Qa and H-2 sequences fails to reveal overwhelming evidence of gene conversion between H-2 and Qa genes of the H-2^k haplotype.

Results

Organization of the C3H Qa region

In order to clone and characterize the Qa genes from the C3H mouse, we screened C3H genomic lambda and cosmid libraries with the class I probes pH-2IIa (Steinmetz *et al.*, 1981a) and p$K^b\Delta$i3. Overlapping clones containing five Qa genes were identified on the basis of restriction map homology to corresponding alleles from BALB/c (Stephan *et al.*, 1986) and C57BL/10 (Weiss *et al.*, 1984) (shown in Figure 1). The restriction maps of the C3H $Q1^k$−$Q5^k$ genes more closely resemble those of the BALB/c than C57BL/10 loci. In particular, $Q5^k$ and $Q5^d$ maps are distinct from that of $Q5^b$ which instead resembles Q6−Q9, suggesting that the B10 gene may not be an allele of $Q5^k$. Previous Southern blot data had indicated that the Qa-2^+ phenotype, controlled by genes Q6−Q9, was lacking in certain inbred mouse strains, e.g. C3H, due to a deletion of these genes (O'Neill *et al.*, 1986). Consistent with this, we isolated genes $Q5^k$ and $Q10^k$ on a single cosmid. Restriction sites between the $Q5^k$ and $Q10^k$ loci are conserved with those upstream of $Q10^d$, implying that the recombination point for the deletion of Q6−Q9 in this strain resides very near the 3′ end of $Q5^k$ (see below).

DNA sequence analysis of the C3H Qa genes

In order to determine whether the Qa^k genes might potentially encode functional products, we sequenced all of the C3H Qa-region genes. An alignment of the Qa^k structural gene sequences relative to $Q7^d$ (from clone 27.1) (Steinmetz *et al.*, 1981b) is presented in Figure 2. The previously uncharacterized $Q1^k$, $Q2^k$ and $Q5^k$ genes are not homologs of the recently described human nonclassical class I genes (Geraghty *et al.*, 1987; Srivastava *et al.*, 1987; Koller *et al.*, 1988; Mizuno *et al.*, 1988; Shimizu *et al.*, 1988).

None of the Qa^k genes contain sequences which would appear to preclude their expression. For example, there are no frame shifts or termination codons within exons encoding the external domains. However, potential premature polyadenylation signals are found in intron 3 of both $Q1^k$ and $Q5^k$. The >99% similarity between $Q4^k$ and $Q4^b$ (Robinson *et al.*, 1988) sequences implies that $Q4^k$ encodes the Qb-1^a antigen (Robinson, 1985). Likewise, $Q10^k$ is an allele of $Q10^q$ (Kress *et al.*, 1983a,b), $Q10^b$ (Mellor *et al.*, 1984; Devlin *et al.*, 1985a) and $Q10^d$ (Lalanne *et al.*, 1985), and is presumed to encode the Q10 molecule in C3H mice. In contrast, $Q5^k$ is 18% different from the partial $Q5^b$ pseudogene sequence (Robinson *et al.*, 1988), confirming the nonallelic relationship of these loci suggested by their restriction map differences (see Figure 1). To date, transcripts or protein products derived from $Q1^k$, $Q2^k$ and $Q5^k$ or their putative alleles have not been detected.

Comparison of the Qa^k sequences yields certain insights into the relationship of these genes to one another and the H-2 genes. The Qa sequences in Figure 2 are as similar to each other as to the H-$2K^k$ and H-$2D^k$ genes (e.g. 83−90%). $Q1^k$ and $Q5^k$ are the most divergent, with greater numbers of scattered nucleotide differences throughout the gene in contrast to the localized polymorphism of the H-2 genes. This could be interpreted as evidence that these loci are accumulating mutations under different selective pressures than those which constrain the evolution of the other class I gene products. Interestingly, the scattered nucleotide differences of $Q1^k$ and $Q5^k$ in and of themselves do not seem to be detrimental to the structural integrity of the predicted class I molecules (see below). Rather, they could reflect the differential function of these particular Qa gene products.

Although $Q5^k$ has the most divergent sequence relative to the other genes, it is quite similar to $Q7^d$ downstream of exon 5. Given the >99% similarity among Q7 and Q9 sequences in the H-2^b haplotype (Devlin *et al.*, 1985b) and the restriction map homology of the $Q5^k$ and $Q10^k$ intervening region to the 5′ flanking map of $Q10^d$, we propose that genes Q6−Q9 were deleted from the H-2^k haplotype by recombination of Q5 with Q9 to generate the present Q5/Q9 hybrid gene.

We also observed that each of the Qa sequences is interrupted by a unique pattern of murine repetitive elements (Singer, 1982) or insertions, primarily within intron 3. Most of the first half of $Q5^k$ intron 3 has been replaced by an ~1.1 kb nonrepetitive sequence which lacks homology to

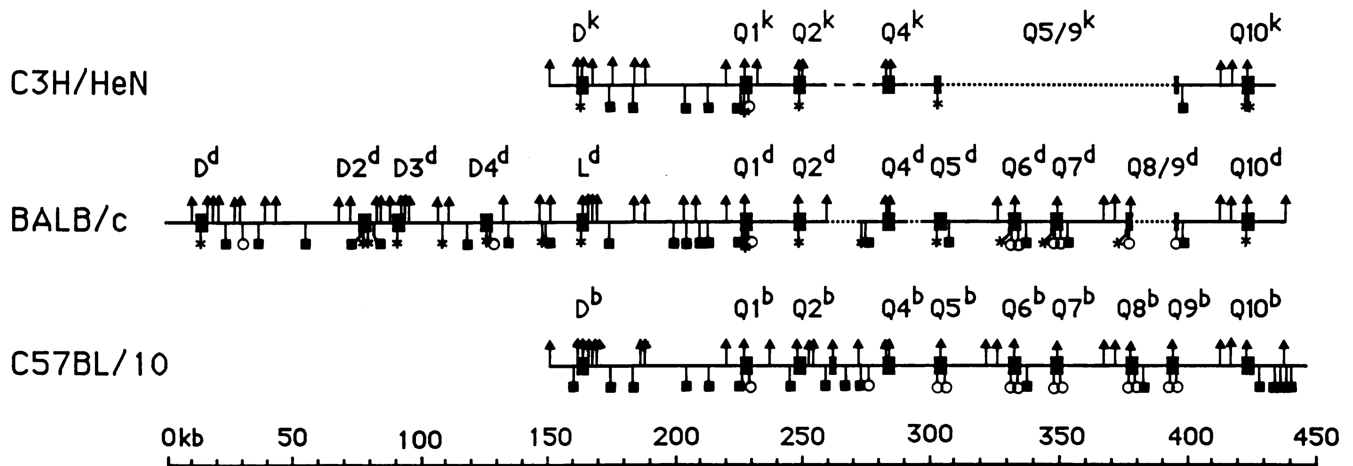

Fig. 1. Comparison of Qa region organization from C3H, BALB/c (Stephan *et al.*, 1986) and C57BL/10 (Weiss *et al.*, 1984) inbred mouse strains. Restriction sites are: *KpnI* (↑), *SacII* (↓), *XhoI* (↓) and *SalI* (↓). Solid lines indicate regions spanned by overlapping cosmid clones. The map between H-2D^k and Q1^k was taken from Figure 3 in Stephan *et al.* (1988). Gaps introduced to align the maps are designated with fine dashes. The C3H clones did not overlap between Q2^k and Q4^k (indicated by heavy dashes), so we cannot exclude the possibility that other genes might map in this region.

any sequence catalogued in the GenBank database. Like the H-2 genes (Ronne *et al.*, 1985), Q2^k, Q4^k, Q10^k and Q7^d contain a B1 repeat flanked by CT-rich tracts. Q1^k has an extensive CT tract, but lacks the B1 element. More than 2.1 kb of unidentified sequence beginning with poly(A) and ending with poly(T) has been inserted into Q2^k immediately downstream of this CT tract. Another B1 repeat in Q5^k, flanked by a 14 bp direct repeat, contains one of the polyadenylation signals in this gene. There is precedent for use of mRNA termination signals within B1 elements (Singer, 1982), so O5^k transcripts may indeed be truncated. Close to the beginning of exon 4, two copies of a B1 repeat are found in Q4^k, while there are two copies of a B2 repeat in Q10^k and Q7^d at exactly the same place, all flanked by duplications of attt or gttt. This suggests that both of these murine Alu-like sequences may integrate at similar sites within the genome. Q5^k contains a B2 element flanked by a perfect 14 bp direct repeat in intron 5. Finally, a sequence with no counterpart in GenBank interrupts Q1^k and Q2^k in the 3′ untranslated region.

We compared the Qa^k and H-2 promoters in Figure 3. Q4^k–Q10^k 5′ flanking sequences are homologous to those of the H-2 genes (Figure 3a). It may be possible that some of the DNA binding proteins, which have been detected in association with putative *cis* transcriptional control elements upstream of H-2 genes (Friedman and Stark, 1985; Vogel *et al.*, 1986; Israel *et al.*, 1987; Shirayoshi *et al.*, 1987; Sugita *et al.*, 1987; Korber *et al.*, 1988), could be involved in regulation of these Qa loci as well as the H-2 genes. Q1^k and Q2^k are not homologous to other class I genes upstream of position −82 (Figure 3b). We were unable to find similar sequences by comparing this region to the GenBank database, or to published sequences of other members of the immunoglobulin supergene family. This novel upstream sequence could represent a differentially regulated class I promoter, as is the case for the Tla^c 5′ flanking sequence which also lacks homology to the H-2 promoters, yet regulates Tla-specific expression. Alternatively, it might simply be the vestige of a recombination event(s) which resulted in duplicated structural Qa genes without their promoters.

Structure of the Qa^k molecules

The amino acid sequences of the Qa^k molecules were derived using consensus splice signals conserved with H-2 genes and are aligned to H-2K^k and H-2D^k in Figure 4. Hydrophilicity plots for all of the Qa gene products look comparable to the H-2 antigens (data not shown). There are no amino acid replacements which would obviously disrupt the class I structure. Q1^k, Q2^k, Q4^k and Q10^k potentially encode molecules with two N-linked glycosylation sites at positions 86 and 256. Q5^k has four possible glycosylation sites at amino acids 42, 86, 176 and 238. The carboxyl (C)-terminal sequences of the putative Qa^k molecules could regulate intracellular trafficking, thus determining whether these potential proteins are expressed on the cell surface, secreted, or remain in an intracellular compartment. Distinct frameshifts in exons 5 of Q4^k and Q10^k result in the same 10 amino acid C-terminus on Qb-1 and Q10, which might be important in determining the processing which leads to their secretion (Robinson, 1987b). Q1^k and Q2^k could both encode molecules with appropriately hydrophobic transmembrane domains and cytoplasmic segments comparable in length to the classical transplantation antigens. Q5^k, however, contains several hydrophilic residues within exon 5 which might exclude it from the membrane.

The Qb-1^a allele from C3H differs from the Qb-1^b of C57BL/6 (Robinson *et al.*, 1988) by only three amino acids at positions 236 (A to S), 251 (L to P) and 262 (Y to H) in the α-3 domain. The P and H substitutions in Qb-1^b could account for its more basic phenotype (Robinson, 1985). Q10^k differs from Q10^b (Mellor *et al.*, 1984) by only one amino acid, Y to H at position 84, and from Q10^q (Kress *et al.*, 1983a,b) at position 5 (M to T) (Cosman *et al.*, 1982; Kress *et al.*, 1983a,b).

The amino acid substitutions in Qa relative to H-2 molecules in the presumed peptide-binding cleft and on the tops of the α-helices are of particular interest, especially since the Qa products do not appear to act as classical restriction elements (Robinson, 1987a). Based on the assumption that the murine class I structure will be comparable to HLA-A2 (Bjorkman *et al.*, 1982a,b), we have indicated the amino

```
                Exon 1                                                                                    Intron 1
Q1   agat-ggggcaatggcactgggaaggctgctcctgctgctggcagccgccctgaccctgaccaaaaccggagcgg  gtgagtgcggggtccgaagagaaacgg------------gctctgagaa
Q2       g           g       c            g          g    g     c g a c    c          a      g g g     a aactttccaaacag     c gg
Q4   gg--- a        gtcaac   t       t     g      t      a      t g g t c c        a c        g g       ------------c       g
Q5   tg --     g     c c c c             t        c g   c c   t                      g g g       ------------c    g  gg
Q10  tgg--     g     g c c c c           g  t     g   c   c g   cag  a               g g g      a ------------ct     cag
Q7d  ggg-- a        t   aac   t      t     g     g       t g g    c c                g g g      a ------------c c    gt--

Q1   gagggtcgcggccggcacctgggatgctgcgtccccgcgtcgcccaccagaccc-tccgcccccttctccacc-------cacgtcccgcgccctgctcccctcccggaccacgga-ccgccggga
Q2     g g tg       g    a   c      g                 -     t  t    ta   c---gcgtcc ta c                          t tc   g   c t      c  g
Q4   a    ga g   g      a  t  a   c        g          t     -                  ------- ta c                              c  g tc c        g
Q5      a    g   g   g   a    g  c  a  c      t     g  t t   g     -a               agagtcc ga c       aa          tg   c  g  c c      c  g
Q10     g   g   g        a    a  c                  g        -          t        ------- ga c                              c  g  c c      c  g
Q7d  ---------------------------              -              g       c                 ------- ga c      a          a         c   g   t  c    a  c  g

                                                Exon 2
Q1   gtctcgggaggaggtccgagtttcaccgcgcgccgcccccag  gctcacactcgctgcggtatttcgagacctcggtgtcccggcccggcttcgggaagccccggttcatctctgtcggctacg
Q2     c     a     g g  c                                   a                 gtc              g    c     g        a g
Q4     c   a    t  g g  c          c t    t        c     t   a t    t c    c             g   c t    g                        t
Q5     ct c    t   g g  c     t       t    t                      t c   g c            aa       g       a
Q10    cg a        g   c     t                           ca   a        a      c  t        g   c t    g              tat       t
Q7d    c           g g  c                          ca          aa      c c    g t              c     g      t

Q1   tggacgacacgcagtttgtgcgcttcgacagcgacgcgaagaatccgagatatgagccgcgggcgccgtggatggagcaggaggggccggagtattgggagcggaacacacggagagtcaagggc
Q2                 g     c                   g    a      g            g                                t    g a   tc c
Q4        (a)     (g)    c                   g                 a                                       g a      a a   c
Q5      a          a     cca            c    g     g a   tgcga      t c          a       a   aa    c      ga c g    aa   tc c     aaa
Q10                      c                   g    c     gatg            t                                 g g     a     c
Q7d                      c             t    g  a        gatg           g                                  g g     a  tg c

                                                                                     Intron 2
Q1   agtgagaagagattccaagagagcctgagcaccctgctcagctactacaaccagagcaagggcg  gtgagtga-ccccgggtcggaggtcacgacccctcgacgtccctacacag-gggccaga
Q2   ca    c   t  g      g  t        g aa    ag               gc                 -                       - g       c  t     g       --  a gct
Q4   a     c     tt     g  t   a      g                    gc                 -          a            -                          a a     -     c
Q5   aca   c   t  g    g  g          g   gc    g  a              tgca  a          t -                              c     t tg     g gaata
Q10  a     c    t      t  t          g       g                    g  a             -    -----                    c          g  a   g  -a     g  t
Q7d  ca    c    t     g  g          g   tgca  ag                                   g                              c          g  a      -a     g  t

Q1   gacctcctgcttctcaagtcccaggttcaggagcagaactgacccgggaccggattccctttcagtttggaggagtc-gcgcgtgggcggggccgggg---gagcgaggcgctgacccagggttc
Q2      g      g    c       tg       a    a       g          t                         c  g                       ------------------        gc    g
Q4      g    c ggc  c       t  g       -t         g    g      t   t                      c  g                  cgt  t----- g              gc    g
Q5      g    a gg  c        tg        ca       c ca             t          ca        a   ----a a          ga t aac      ------            gc    g
Q10     g    c gg    ca          g      g                    a                            c    -a    g    a --a    ggc g g      g      tgca    a
Q7d    gg    c gg    ca          g  t                         a                            c   gtg    g  a gg     ggc gag       ga      act    g

         Exon 3
Q1   ctgcag   gcattcacaccttccagaagttgtctggctgtgatctggggtcagatgggcgccttcaaagcgggtacctgcagttcgcctatgatggccttgattacatcgccctgaatgaagacc
Q2   c         ttc         ac       g   a    a            g     a   g   c           c tcc              g       c                tgc        t        c
Q4   c         tc         ta t      gt a c               ag        c    c             c tcc                a                  c   c   gc                  c
Q5   c         tc                   c    a    a           ct          g    c     a     c tcc        t    tg      c t       c   c   gc                  c
Q10  c         tc         ga        tg  a    a         a ag          c    c         t c tcc                a a     a   c   c   gc          t          c
Q7d  c--       tc         ac        atg  a    a          ca          c    c             c tcc                                  a   gc                  c

Q1   tggaaacctggacagcagcagatgtggcagctcaggaaacccgacacaagtgggagcaggctggtgctgctgagaaacacaggacctacctggagggcaagtgcctgatgtggctccacagatac
Q2     a          t   gaag               g t atc     a     g               a              a    t tt    a g t     a           g      g  ca    c       g
Q4     a     g        g  g  g    ca          g    a    atc        g                          a a         ag   a g                  gc     g  ca   cc      g
Q5     a    tg        t   t  t   cag        t   g     a   c    g     t          a            a    t ttgtga ga   t       a  a    g  a   tg    a  a       tggat
Q10    a     g        g      g   c           g    gattatc          g                           a    t tt        g               c g       g  ga          t
Q7d    a     g        g   g tg   ca          g    a    atc         g                       at    a         g  ca g                 c     a  ca    c       g

                                            Intron 3
Q1   ctggagctcgggaaggagatgctgctgcgcacag     ---------------------------------gtctctggctggggctcagtcctggggaagaagaaaccctcagctggggtgatgcccc
Q2                           c              gtgcaggggccgcaggcagatcctccctctgcc    g                                                         -
Q4                           c                                           cc   ga                                                      a   g
Q5                     t     c                     ----------------------     g                      {begin insert}..................
Q10                          c                          g       c        cc   g
Q7d      c                   c                          g       c        cc   g   -
```

Insert

```
Q5   aatcttggaattgcttttttcttttttttctatttcttctttcttcctttttttcttttactgccgcagaccatctgggtccctggtctgcgaggaacggggatttcttaggcagatgggcaaaagagtc
     ggatgtagtgacagacagccaaacacacaagagaggtgtttaatctgaatgtaattttccagtcgagcctcagactttatataatacagaataaattgcaaatcataacagaataaggcacattga
     atttttccaggtgcaaaaggagtgacttcaaaaggaaccagataaatgttaaacactagagatagcacagatgcaaagggagtgactacaaagggaatttgtagggaccagataaatgtttacat
     tagagattagcattttctgccaggcaaagagtttacacttagagttaatagcctcagggtagttaactcttgatacgcctcaagactaatgtagtgtcactgaccccaaaattctctaggactcg
     ttaattcttatctatatctggagactatccattttcagtatagtatacaaatttgctcttggcatggaatgtagtctgtaataagaatttctatctcagtgagaatttcagactctatttgcaga
     catctgaattaatggctgtgctttattgttttatggtagtgcttaattagttactgaatgggttaagcctcttgctgctatctagaatctaagaacactgggaaaaggctttagctatgttataa
     tacaatcttaaaaggcatttgctataaggtaatactatatagagtgcacgtgaatccatacactagactaatgtggtggaaattaattttatgggttagaggaatagccaaagatccgggagtaa
     gtttccttgaaactcttttctcatgagaaagctttcagactttcatctgtcaagctgactccaccggagtccctggca
```

```
Q1   tgtctcagaggggagacagtgtccctgggtctcctgatccctcatcacagtgactgcactgactctcccagggctcagccttctccctggacagtgcccaggctgtctcaggagggaaggagaga
Q2                     g     a                 g   t g                                                          --
Q4                     g     a                                             g
Q10                    g     a
Q7d                    g     a

Q1   atttccctgaggtaacaacagctgctcccttcagttcccctgtagcctctgtcagccatggcctctcccaggctgggttctctgcccacacccactgtgtgtggacactgactcctgtcctgctg
Q2                                            a                  a                                -                 cct                      -
Q4                                                                                          c          a ------      c   a
Q10                                                c                    a -            a                       -----------------     a
Q7d          c                                     c   a                            cc     g                   -----------------     a
```

```
Q1   agtgtgtcagtccttacacctcaggaccggaagtctcctttccct-------gacatggacttccctacactaggctggttgccccagtttctagaattttctaaagaatacattctcccagatc
Q2           c                t        g    -a   gataagaa        c                                                 a               a
Q4           c                    a       g      aact   gg       a                  a      t       -       a
Q10          c            t   t             a  c    gg           at                     t               c    c              a
Q7d          c            t   t             a  -    gg           at                 - c                 c    c          g   at

Q1   cctccctgtctgctggggtttgcaccccttcgacaacccaattctctttattcctacactggtgaatggtcacatgaggccttattgggaaccctggaggaatataaatcatggaatttcttttt
Q2    ac           -          t      a      g          c         a a         c     g    -     g   t
Q4                 -     c                     t       c        --- a    -g          c  c      g   t                  c    ag             c
Q10                -          t-    t     -          a c      g   a                   c a     g   tt       aa ca      ---------------     c
Q7d     g          -          t-    t     -          a c      g   a                   c a     g   tt       aa ca      ---------------     c

Q1   c[t]7ctctctttc[t]5cttctttctttc[t]6ccttac[t]14
Q2   tcttc[tc]24ttttcagaa[t]5c
Q4   [ctt]24ct[ctt]5c[ct]22tcga[t]6
Q10  ttgtg[t]6cccctctcgttttctttatatctttac[t]8aagggtattatgttgcttataatcgg[t]5cttcggcactggaatgatattg[ct]4ccccaccac[acccc]29
Q7d  ttgtg[t]5cccctctcgttttctttttatctttac[t]7aagggtattatgttgcttataatcgc[t]5cttcggcactggaatgatgttg[ct]4ccccaccat[acccc]2a

     B1 repeat
Q2   gagacagggtttctctgtatatccctggctgtcctggagctcactttgtagacca-cttggcctcgaactcagaaatctgcctgcctctgcctcccgagtgctgggattaaaggcgtgtgccacc---
Q4    t              g                 a          gga        c        c                c                                            ---
Q10  cct t tca   g a cag  g            a      c      ggc      t        a               tgcc --                              g     acc
Q7d  cct t tca   g a cag  g        g   a      c      ggc      t      - a             ..[cctctg]16...                      - g     acc

Q1   .............................                 gtggaaggca-ttattttgcttgtagtca-------tcttttgtctgcactggaatgatcctgtttctccctgcccttgtatt
Q2   a-tgcccagcctttctttttttctttac[t]5  {insert}    ----g   gcc       a c g    -------gt  c   t        g          c    a
Q4   t-               a          [t]18              --   g  gg- a      t  c     tcttttg                g                           a
Q5   .........{end insert}       [t]26              ---  gt    -       t  c     -------gt               g   g           a
Q10   c  gg   aag-----------------                  --------------------------------------------------------------------------------------
Q7d   c  ggaggag------------------                  --------------------------------------------------------------------------------------

     Insert
Q2   aaaaaaaaattgtacaagctatttattaacattataacatgaagcccatttcctcacaaagaatctacaagatttaccaccttaggactcttgagtacaaccaagcaagtgctgtagattaggaa
     aaaccttcacagtctggacaagcactgaagctcagagactgactttgacttaactaatttgaaatatatttaaaactactctcagcgcgctcatacagtgatgctctaacagtcaggtccaatgc
     ctggctcgaggac {gap in sequence} acagccacaggggcctctcgtgctccctgcccttcctcgtggcctttccctctgaccagtctgtgttgtgttgacagtgatgtacagaatgag
     aggcttcccgttgttcccctttgtctttctctgtgtttcgtgctttggctggatctttcctggtcaagatagtcctttggtactagaccgtgggctgctttcccttctgtctcgatgtctttcta
     aagacgccacgtccttctggcttctccctctctctgtgctgaccactgtcttcacatctttccctccttgctttgtgtgctcttctctttctttgcttttctttcctttctcactgtattggtc
     gtgatcattccattgtctatctctttcttgttcccttgaggaatatttctccctcttccctcctctatcctgtctttctccctgttctcatcccttcagtttatcttcttgctcatgcttcttcc
     agtagagatctttatgactggcctctctgtgcctatgagaactcttttcttcccggtggctgcggtccttgtgacagctctcttggtctctggactgcctgtcctggtgctcatgtcctcttgac
     tgggaacctcgtgagtggtgcttggtcccgagcctcctttcctcactagatgatcatgagtgggcgcggttcttgtgatgcttgggatatttcatggaggtgccggtgtcgtctcctctggactt
     tatggtctcccctttcaccgtcctcacagctcttatcatcaaaacctctgttagcatctgggttttcctcggcttcatctgcaccccttgcagaatgagttctgctgtggtggtctgttctctg
     aatttgtttcatcagggtaacccctgagcttcttttattacagtcctggcttcacgaaagctggatttaggaactgcctcctcaccaactgcttgatttaatagatgccgataaaatccactgag
     atctttggaggcaagaattgccgctgctgtcgccgggctttctggcttgataaccccaaccaacaatgttcttaattgcactttgccagaatctacccacgggcgcgcctgctgccagaactgtt
     tacaatatgggcagaacggatcggcgaagacgtagacaatgaccggcgcatctttttaccgtcgaggagccagtgggattgttccatccgttgccacatttcgcgtccggctggtgcgtaaatt
     tctttttcgataagtgtgttactcaggttttcacctttctcgttgtacatgtaaccagagatagcgtgcttaccatctggagtcaggtagatggtgacgcccatatcctgatactttccgagata
     acctttcattcctccgggggcatcgaatgtttgatgattgtaatgccctgttttcaatcgctttactggagcaggaagtcct {gap in sequence} cattgtctgttcagtcccgctcta
     gcagctccctctgaaggcttcgctcacggaactctcatcatcctcagaatcactcccaaacactgacagtgtttgtaaaactcggagcaatggctgttttctttagcaaaataagcccatactgc
     ctgcccaggatcgccatcttgttccca[t]15

Q1   atcatgtgtatcagtctacacaggtgccagggaaggctgctaacctggataatttgacaagttaaatcagggttctcattaaaggagagattcctgtgaacttagagtgtttcct-gtcagaact
Q2                   cg            a g  a               atg                                   t                        -
Q4                   c                  ----------------t aa                        t                       c          -
Q5      t            c                                  a                           a            t           c          -
Q10  ---------------------------------------------------------------------------------------- a  g          gt   a a    t  - g    t
Q7d  ---------------------------------------------------------------------------------------- a  g          g    a a    t  g g    t

Q1   aaacacccagaagcctcctgctcttcctctgcccccca-caagttacagtg  [c]11       agtgaatcaggacttggactctgag-agacagggtctactgcaatccagggctggagtga
Q2        tg         c         c        --t    -          [c]6a[c]11                g              -         t
Q4        t                             --t    -          ct[c]13tcc           t                   -         t
Q5        t          c               -----------------    ct[c]7                                   a         t
Q10    c  t    tca     a a c   g  --     c                ctcctcct     g                           -         t                    -
Q7d    c  t    tca     a a c   g  --     c                ctcccccct    g          a                -     -   t               c    -

Q1   gagggaagacatcacaccccgtga-gcccactgtgttccagtgagtgctgcactggggtccacagctcactccagggatcctgtgtgacacacctgtaccttgtcccccagagtcaggggctggg
Q2               ca     t    -                                      a                                               g
Q4               ca     t    -         g                            at                                                                   a
Q5               ca     t    acttgcagaattctt        gg              a      t   ga       a           t          a    g
Q10              ca     t    -                                      a
Q7d              ca     t    -                          -           a                          t

Q1   agtcatttttctctggct----------------------------catcagcttagtgatggctgttcacttgaactgacagttaatgttg-tcgacgagatgacaacagttgtag--tctcagt
Q2                   ----------------------------- cata   t                 g                   g  ag a        c     g     at
Q4                   gagtgtcagaggttgacacatttctgct  caca   ct         c      g      g g    -  c ag  tag a     ac     g   ttag
Q5                      g              c   ca       t cacaac t     a     c      g               a    g    a    a         c     g   t ag      ta
Q10                                    c   ca       t caca   ccc         t      g                    g  ag a         c     g   ttag       a
Q7d                                    c   ca       t caca   ccc         t      g                    g  ag a         c     g   ttag       a

Q1   tttagcaccctttcagtagcatatgtgcctaatatctgatttgatatgaactcaaacacataataaattacttattttc........cattccctcttccattct-ttgactacgtctctcatg
Q2   gg t                    gt   g  tc       c                   t                    ........      t    t                 c
Q4    g ca       c     g     gct      t   ac          c         t  t          g    g     ........      a           ag -- c t-----
Q5   gg   a     at c       c  gc      t   a    a c              t         t    g       {repeat}      a           -g         c a        t
Q10  gg gt    t   c          gt   g   t   a    a      c         tg     t               ........      a        a -a ag      c a      g
Q7d  gg gt    t   c          gt   g   t   a    a      c         --g    t               ........      a        a -a ag   g c a      g

     B1 repeat
Q5   tgcctggcagtggtggcacacgcctgtaatcccagcacttgggaggcagaagcaggtggatttttgagttcgaggccagcctggtctagagagtgagttccagagacagccaggactatacaga
     gaaaccctgtcttgaaaataaaaatgtagtttc
```

```
Q1   ctatctaacattacataaggatgaccatggtgatccactggctcatgtggattccatcatagcttctgagtcccaa-aagaaaatgtgcagttgtgtgctcaggg-accagctctgcctactggt
Q2     ag      c    g         at t c                g   c  t                                  cc    g   g          t --
Q4     aag     c          ctg   g a  c                  c  t                cc- ga          c  a gaaa-------              g a
Q5     ag      c                  t   c      t     a   t c t            t  c-- g      ------- c    a   g    c        gaa a t
Q10    tg      c                  t   c     t             c  t               cg- g              c    g a  g             g a
Q7d    tg      c       -          t t c                   c  t               ctc g            cc     g    g             g a

Q1   cactagtgtgatgacagttgaagtgt-caaacggacacgtagttcagtgtcatcattgttttacctgagtcttgtgtagatttcagtttgttttgttaaccctgtg-----------------
Q2    g    aca        t    -    a       ---------          a    a                   gt   t c      ttg  ga-----------------
Q4      c    cc                 -     a      a t           g    a      a   t c               agg  c      ttg agtac[t]8g[t]14atttt
Q5               g       a -     a   g   a                a    a      a     c     at              c   c      ttg  g -----------------
Q10          cc           t    g    a     a          a   t c  a      at tt      a- c  t             c   t    atttg t [gttt]5...........
Q7d          cc      a      t    t    a      a      t    a   t c  a      a gtt       - c  t          ca   t[attt]9.....................

     B1 repeat
Q4   gttttgttttcaagacaggattttctgtatagccctggctgtcctggaactcacattgtggaccaggctgaccacgaactcagaaatctgcctgcctctgcctcccgagtgctgggattaaagg
     cgtgtgccaccaccaaccagctaatcgtgagatttctttctttttcttttcttttcttc[t]20[gtt]4ttgtttttttgagacagggttctctgtgtatccctggctgtccttgaactcactt
     tgtagaccaggctggccttgaacttagaaatttgcctgcctctgcctccggagtgctgggattaaaggcatgtgccactaccaaccagctaattgtggg

     B2 repeat
Q10  tatgtatggaagtacactgtagctgtactgatggttatgagccttcgtgtggttgttgggaattgaattttttttaggacctctgtttgctctgatcggccccgctcactccggtcaactccta
Q7d  a   c            t              g tt      t                                                 c         g    a   t

Q10  agggtcaactctgctcactcagctcagtccctgcttgttctggcccaaagatttatttattatt----gtacataagtacactgtagctgacttcgtatgcaccagaagagggcgtcagatctca
Q7d  t                     -----                                     tatta       a                       ag

Q10  ttacagatggttgtgagccaccatgtggttgctggggcttgaactcaggaccttcaaaagagcattcagtgctctttaccctctaagccatctccccagtcctcagtttgtcttcttaattgtgcg
Q7d                   t                     t                      g           -          g                                  a a

                               Exon 4
Q1   atttcttaaa-tct-ttcacacag  atcccccaaaggcacatgtgactcatcaccccagatctcaaggtgatgtcaccctgaggtgctgggccctgggcttctaccctgctgacatcaccctgac
Q2             - g - c                           c            g    a a  a  t                   a
Q4             -  - c              t             atg    a     g c                                               a     t
Q5         g   -  - c              a   g    a a  c g    a     g c                  a
Q10        g   -  - c                  a         c     ag     g              t
Q7d        g   a   c c  a                        c        g   t t    c

Q1   ctggcagttgaatggggaggagctgacccaggacatggagcttgtggagaccaggcctgcaggggatggaaccttccagaagtgggcatctgtggtggtgcctcttgggaaggagcagaattaca
Q2                                                                                          g
Q4
Q5                                                                  a a                                                     g
Q10                                         a  g
Q7d  t                               c                                t

                                                                  Intron 4
Q1   catgccatgtgtaccatgaggggctgcctgagcccctcaccct---------gagatggg  gtaaggagggtgtgggtgcagagccggggtcagggaaagctgtagcctttgcagacc--ctg
Q2              c                                ---------                            t                   g     c        --
Q4                                               ---------                            t        a          g                --
Q5              g a                              ---------          c        a  c     t                   g     c          --
Q10                                              ---------                     a      t t                 g     c a        --
Q7d             a                  t             gagatgggg                            t t                 g   a  gc g   g tg

                                                                         Exon 5
Q1   agctgctcagggctgagagctggggtcataaccctcacctttatttcctgtacctgtccttcccag  agcctcctccatacactgtttccaacatggtaatcatagctgttctggttgtccttg
Q2     t  g                                  c        c t       t                  c      c                  t
Q4                            g              c                                     c      c          cg a g
Q5    g   g                   t              c                                t  c    caaga t    ca        t       c
Q10    t  g                   g              ct                         a      t c    ac  t t    tc ca    t    a   -----------
Q7d                c    t    --- g           c    a          -                        c          cg c    t      g     a

                                                                                  Intron 5
Q1   gagctgtgatagtcattggagctgtggtcatcattggagttatggtgtcttttgtgatgaagaggaggagaaacaaag  gtagaaaaggccagggtctgagttttctcatagccttctttaga-
Q2           ca                ------------------        g        a  ag          c       g   g                    tc     g      -
Q4        -  gcca      c       ------------------        g          a       c    c       g   g    a               tc     c      -
Q5           g ca  c           ------------------        g                       c       tg  g                    tc     c      a
Q10  --      gcca      aa     c------------------        gta        a            c       g   g           c        tc  t  c      -
Q7d          gcca              ------------------        g        t     t        c       g   g           c        tc  t  c      a

Q1   ........gtgtcctctgctcattaaatggaaacacagccacacccctcattgctactgtctctaactgggtcagctgtcagttccgggaacttcctagagtcaagatcttccttgaactctcac
Q2   ........ a  g          tg      t                               gt          t    t                        t
Q4   ........    g        c  tg  g    g        a           g    t   t    c       t           t     c
Q5   (repeat)    g   at      tg      c  t t     a             c       ct          t      t ------
Q10  ........    g           tg                 a         a    a       t          t      -  t            t
Q7d  ........    g   aa      tg      c  t t     a         ct   c       ct          t      -------               ---------

     B2 repeat
Q5   gtgtgctctagggctggcgagatggctcagtgggtaagagcactgactgctcttccgaaggtcctgagttcaaatcccagcaaccacacagcaacttacaaccaatctgtaatgagatctgacac
     cctcttctggtgcatctgaagacagttagagtgtacttatgtataatagtaaatgaatctttaaaaaaaaaaagaa

                         Exon 6                         Intron 6
Q1   agcttttcttctca-cag  gtggacaaggagaggactgtgctctggctccaa  gttagtgt-gtggggacag-gattgtcctg-aggccgattagagtgaagctggagatgtt-aggagctctg
Q2           a    a           a  t aga   aacc   a  a    g          - -          -        -    -                              -
Q4          c     -                g     a            g          a -    -      a t -     --   t -   g             t   -g  t
Q5          c     -                g     a     ca     g   ca     - -           - g      -t   a -   gc                a -g
Q10   t       c   -       a  a     g     a     cc     g          - -    -    -  -  -    gg   a -   gc                  -g
Q7d         c     -                g           ca     g          - -    -    -    -   -t   a -   gc                  gg
```

```
Q1   ggaatccataatagctcctccagagaaatcttctccagaaaaatcttgtggcagcctgagttgtactgcgatattaatacatacatatacatgtgcatatacatttgtttt-atttaccctag
Q2                 t                          --------------- g   c     t    g        t       a          -g
Q4      c           t   gtt t   c   --------------- tg   t  cag tctt c  ag  -------------- atc t gt  a at------  t
Q5      c    g  a   t                        ----------------cagg   c       c--a   g      t  g    a          --------        t
Q10             a   t                         --------------- gg  t  g   t  gctt c  g   ------------ a ctc    t   ---------c  a
Q7d  a       g  a   t                        ----------------cagg   c       c--a   g      t  g    a          --------      - t

     Exon 7                                                     Intron 7
Q1   gcag-ggctgcgcccagagctctgatatctctctc-ctagattgtaaag  gtgacactctggtgcctgattggggaggggcaatgtgcatatgattgggtttcagggactcccagaa--tcct
Q2      a a a                          -                                                    g                              ac
Q4      g a a t                  g t   t a          t   atg       ca      c    g c   g   t                         c  tcc
Q5      g a a t  t               a     - c   g           g                     a                                    tcc
Q10     g ata t                  gc    - c    t          g   a         a a   g   g                          g       tcc
Q7d     g a a t  t              ga     - c   g           g         a               g                            ac   tcc

                                                                  Exon 8
Q1   gtgagtg----gtgggttgttgggatgttgtcttcacagtgatggttcatgactctctttctctag  catga  agacagctgcctggagtagactgagtgatagccagtgtgttcaggtctc
Q2         ----
Q4         ----           a         tt   g     c    c    a      t                               a     c
Q5   c     agct        a  a                              a                             g    t     c              ac tt-----
Q10     t  agtga         g     a     t                   a      gg                     g          c  a ga               g
Q7d  c     agtg        a                                 a      tg         c   a       g          c a
                                          g  t  t c  a                              c gc          c  a ga         t

Q1   tcctgtgacatcaagagtctttagacatcagctttctttagaaaatggtg {insert} ......gttccctgtgatcctatggtctcagtgtgaagaattttggagcccagtctgcccctgca
Q2                                                       {insert} .....cag                                  c              gt     cg
Q4   ------------- g      c -------tt           g  ca c           cctgat        g       g     a          g             c  t g   a
Q5      a         c     c c -------gt c a       c  ca             t  c t        g       a          g  c g             c      g-
Q10               c   c c c -------gt c         c  ca             t  t    t   t    g    g     a       c g  t          c      -
Q7d               c   c c c -------gt c         c  ca             t  c t        g       a     a     g  c g             c      -

     Insert
Q1   tctcatgttctctgcca-caccaaccccagtagcagcca-tgtgacaagttgcacagactgaggtatacttcaaaagaaggaggcctcctggaccagttgtggaactcttggccagagtattatat
Q2                  a                      a    {gap in sequence}

Q1   tgagtcttagttctcatggcaattgatttattttgtttctcttattccttttcttatttacatgttcctcgtaacctaggctaagatctcatgtgaaagtaatgggatagaaacatctagaatggg
     tcactaacatgacttatctgggaaac {gap in sequence}

Q1   caacag-accctgtccctgccctgccctgtcttcccttccacagccaaccttgctggtccagccaaacactagggggcatctgcatcctgccagctccatgctaccctgagctgcagctcctcac
Q2       g                     g
Q5     c t g          a        g                                       t   a   a            t           gt        a
Q10    c   g          at   t--- gc                          c            g  ta     a   t    t           g   c     c         at
Q7d    c t g          a        --         t    t  g      c aac           g  a a              tg       c       a
                                                         c   t       c    -  ta     a       t            g   c     c         at
```

Fig. 2. Aligned DNA sequences of the Qa^{k} genes. The Q7^{d} sequence is from clone 27.1 ((Steinmetz *et al.*, 1981b). Spaces indicate identity to the Q1^{k} sequence. Dashes are inserted to align the sequences. Dots fill space to avoid confusion with regions of identity to Q1^{k}. Short direct sequence repeats are bracketed ($[]_n$). Possible initiation codons and premature polyadenylation signals are underlined. Nucleotide positions which reveal putative gene conversion events are marked as follows: Q5/Q10 (●), Q4/Q7 (↓) and Q4/H-2D^{k} (circled).

acids with side chains which point into the antigen-binding cleft or toward the T-cell receptor in Figure 4. For the most part, substitutions at these positions are conservative. When hydrophilic residues are replaced by hydrophobic ones, and *vice versa*, most often there are known class I proteins which contain either type. For example, at position 72, H-2D^{k}, Q1^{k}, Q7^{d} and Q10^{k} share hydrophilic amino acids, while H-2K^{k}, Q2^{k}, Q4^{k} and Q5^{k} have hydrophobic ones. Both the H-2 and Qa proteins are polymorphic at the same amino acid positions, and, furthermore, the same two to five amino acids are utilized at these polymorphic positions within molecules encoded by either region. Thus, there is nothing obvious to distinguish the composition of the cleft of H-2 and Qa molecules, implying that they might bind similar kinds of peptides, and that the Qa products could potentially interact with appropriate T-cell receptors.

We reasoned that one might find Qa- or H-2-specific amino acids on the tops of the α-helices if the Qa gene products interact with T-cell receptors distinct from those recognizing the H-2 molecules. There are a few positions at which such Qa- or H-2-specific residues could be identified. For example, many H-2 antigens have an R at positions 83 and 155, distinct from every Qa and Tla gene sequenced to date. Some HLA, RLA and ChLA molecules (Klein, 1986; Lawlor *et al.*, 1988; Mayer *et al.*, 1988), as well as the human non-classical HLA-E (Koller *et al.*, 1988) also share R at 83; only RLA products share R at 155 (Klein, 1986). There are murine Qa-exclusive substitutions at positions 83 (S in Q1^{k} and Q4^{k}, N in Q5^{k}), 138 (V in Q1^{k}, Q2^{k} and Q10^{k}, R in Q5^{k}), 155 (K in Q1^{k}, Q4^{k} and Q7^{d}), and 166 (Q in Q2^{k}, Q4^{k} and Q7^{k}, K in Q5^{k}). Intriguingly, site-specific mutagenesis of amino acids at positions 138 and 166 on H-2K^{b} can affect CTL recognition of alloantigenic target cells (Ajitkumar *et al.*, 1988), confirming that these residues do influence T-cell recognition. These H-2- and Qa-specific amino acids could reflect some unique attributes of H-2 and Qa function distinguished by interaction with distinct T-cell subsets or, alternatively, might merely represent evolutionary polymorphism among groups of closely related sequences.

Evolution of class I genes within the H-2^{k} haplotype

The analysis of the H-2K^{bm} mutants revealed clustered nucleotide substitutions that could have been donated in some cases by Qa genes. As a result, it was proposed that the Tla complex might serve as a reservoir of polymorphism for generating H-2 diversity (Ritzel *et al.*, 1984; Pease, 1985; Nathenson *et al.*, 1986). Therefore, we looked for evidence of past gene conversion events among Qa^{k} and H-2^{k} loci. To do this, we compared all available class I sequences, identifying regions where H-2K^{k} or H-2D^{k} differ from their alleles, and then searched for such divergent regions among the Qa^{k} sequences. While we found some clusters of nucleotide similarity between class I sequences that could be indicative of past interlocus gene conversion, the best examples were between pairs of Qa sequences, rather than

A)

```
Q4      ctccaggcagtgagatcaggggtgggg----------------atttcccatctccccagtttcacttctgctcctaacttgggtcaggtc-----cttctgcccggacactgatgacttactggc
Q5              g     g            aagcccagggctaggg   c                     t       ----- a    t   -----         a              ga
Q7d                   g   a    a          t    g      ct      t   ca       -            c        t  -----a       a             g g- aa
Q10          c  g     g        t               ga   c c                                c    c      attcggc  g  t           t-gagg  at
H-2Kk   tg      g     g                        g      c         a                      c           -----  c   t          t    gcg a t
H-2Dk    g            g        a               g      c         t               a      c           -----        g      a      gcg
                  bind A                                              IFNγ                                                bind B
                     Region II                 KBF1 site, Region I
                                            Region III

Q2      ......  ggt        a g ag tctcatc   ctg ggaa        c  g tc t       c           t            t ag   c     t    t t
Q1      ......  ggt          g ag -------   ctg ----        c  g ta t t t   c  t               g    t  agg cc     t    t t  t
Q4      agttctcacccccattgggtgacgtg-------atcaccg----ggaaccaatcatcgtcgccgcgggagcaggttataaagtccacacagcccgccagagtcagacgcccagaatcccag
Q5        c               c  gt c -------   ctg ---- t        cg     - tac  cc t         a              gg  c     g a c  g
Q7d        ctc     -t         g   -------  t   c-----          cg    g         g    g          g t    a ag       t  a  g       cc
Q10       g            - a    g c -------t  ctg ----        c  g          cc - t                       agg  c          c g
H-2Kk     c.  t               g c -------      a----a          gt          ac  t  c            g        gga c   a t tc       g c
H-2Dk     g      tat          g a -------   -  a----   g       g           ac  t               t        gg  c tt   a   g g
                                                         CAAT                      TATAA
```

B)

```
Q1                                                                                     gatcagtatggaagggctcagcccactgcaggtggtcttggctggtg
Q2      ggatcctgaattgggaaaatgcctccattcaggacgcttttggaaaagtctataagacattttaaaattattattaata    a  a             -- aa g       t

Q1      atcagggattgtataagaaagcaggctgagcaagccatgagctacaaagctattagcactggtccccattctctctgcttcagttcccacttagaatcctctctcaagttcctgccttgaatctct
Q2          t                     g     -------                              gg                t ------------------------------------

Q1      tcaatgatagactgagatgtggaagtgtaaactgaaataaacccttcttcctcaagttgcttttggtcatggggtcttatcacagcaatagaaaccctcactaagacacagctgcacaagggcct
Q2                 a                          g                                                   g          g          g          t

Q1      cactctgttccctgcagacaccaccctctcctgcagctcagctcctgcccctcccttgccctggagcagcaccactttatagctactggggatctgatgccctctctg-acctcctgaggtagcag
Q2                                                                                          c                       a          a

Q1      gaattatctggcaggtc-------cattgggtggcaggatcctgg----ggaaccaaccagctactctgtgggcgctggttataaagtccacgcagctcgagggcctcagatgccctgtattccag
Q2                       tcatctc        a            ggaa     ____      c   c c     a   _____t      a    ct ca  a                c
```

Fig. 3. (A) Aligned DNA sequence of the H-2^k (Arnold *et al.*, 1984; Watts *et al.*, 1987) and Qa^k promoters. All sequences are compared to $Q4^k$ as described in legend to Figure 2. Potential *cis*-regulatory sequences are underlined; bind A and bind B (Korber *et al.*, 1988), Regions I, II and III (Shirayoshi *et al.*, 1987), interferon responsive sequence (IFNγ) (Korber *et al.*, 1988), KBF1 binding factor (Yano *et al.*, 1987). (B) Novel upstream sequence of $Q1^k$ and $Q2^k$.

between Qa and H-2^k genes. The most striking case involves eight nucleotides within 57 bp between $Q4^k$ and $Q7^d$ in exon 1 (see Figure 2). There is a background of sequence divergence between $Q4^k$ and $Q7^d$ within this region, but this can be attributed to possible point mutations which could have occurred subsequent to the time of the putative gene conversion event. The next most obvious potential event, supported by six nucleotide positions, involves $Q5^k$ and $Q10^k$ in the beginning of exon 1. Interestingly, we found evidence of only two potential gene conversions involving an H-2 and a Qa gene in the H-2^k haplotype. In the first, $Q4^k$ and H-$2D^k$ share two nucleotides within a span of seven bases in exon 2, resulting in amino acid replacements at positions 30 and 32 (see Figure 4). A comparison of this sequence to other H-2 genes (data not shown) reveals that H-$2D^b$, H-$2D^p$, H-$2D^d$ and H-$2L^d$ match H-$2D^k$, while H-$2K^d$ is like H-$2K^k$ and the other Qa genes. This suggests that this potential conversion event was donated by an H-2D allele into the Q4 locus. It is not clear that the interlocus gene conversion occurred in the H-$2K^k$ haplotype, since $Q4^b$ also shares this substitution (Robinson *et al.*, 1988). The second event apparently involved H-$2K^k$ and $Q5^k$, with two nucleotide substitutions within 5 bp altering amino acid 56. H-$2K^q$ (Kress *et al.*, 1983a,b) shares this sequence, again implying an H-2 to Qa sequence transfer.

In addition to clusters reminiscent of gene conversion, we looked for patterns of sequence similarities along the lengths of the genes which might reveal phylogenetic relationships among the loci. The most parsimonious explanation for the presence of a shared B1 repeat in the CT tract of intron 3 among all H-2 and Qa genes, except for $Q1^k$, is that the genes of these class I regions have diverged from a common progenitor subsequent to branching from the Tla lineage, which has a different pattern of murine repetitive elements in intron 3 (Ronne *et al.*, 1985). H-$2K^k$ and H-$2D^k$ are distinctly more related to each other than to any of the Qa genes. However, there are 16 positions at which H-$2D^k$ shares amino acids with the Qa products while differing from H-$2K^k$, but only four positions where H-$2D^k$ has unique residues with respect to the others (Figure 4), suggesting that H-$2D^k$ is more Qa-like than H-$2K^k$. $Q1^k$ and $Q2^k$ are most similar to each other taking into account not only nucleotide identity within the structural genes, but also their novel 5′ flanking sequence and the shared insertion in the 3′ untranslated region. Furthermore, these Qa genes closest to the H-2D locus are more diverged from the other Qa sequences than from the H-2 genes. Scattered point differences make $Q5^k$ the least conserved, but when we consider only positions at which $Q5^k$ shares a polymorphic substitution with some other locus, then $Q5^k$ is most closely related to $Q4^k$. $Q10^k$ is more similar to the $Q7^d$ sequence than it is to the other k haplotype genes. Thus, the Qa^k genes seem to be arranged in duplicate gene pairs, consistent with similar observations in other haplotypes (Devlin *et al.*, 1985b).

Discussion

We have cloned and characterized the Qa genes from the C3H (H-2^k) mouse. The C3H Qa region has fewer loci ($Q1^k$, $Q2^k$, $Q4^k$, $Q5/9^k$ and $Q10^k$) than BALB/c (Stephan

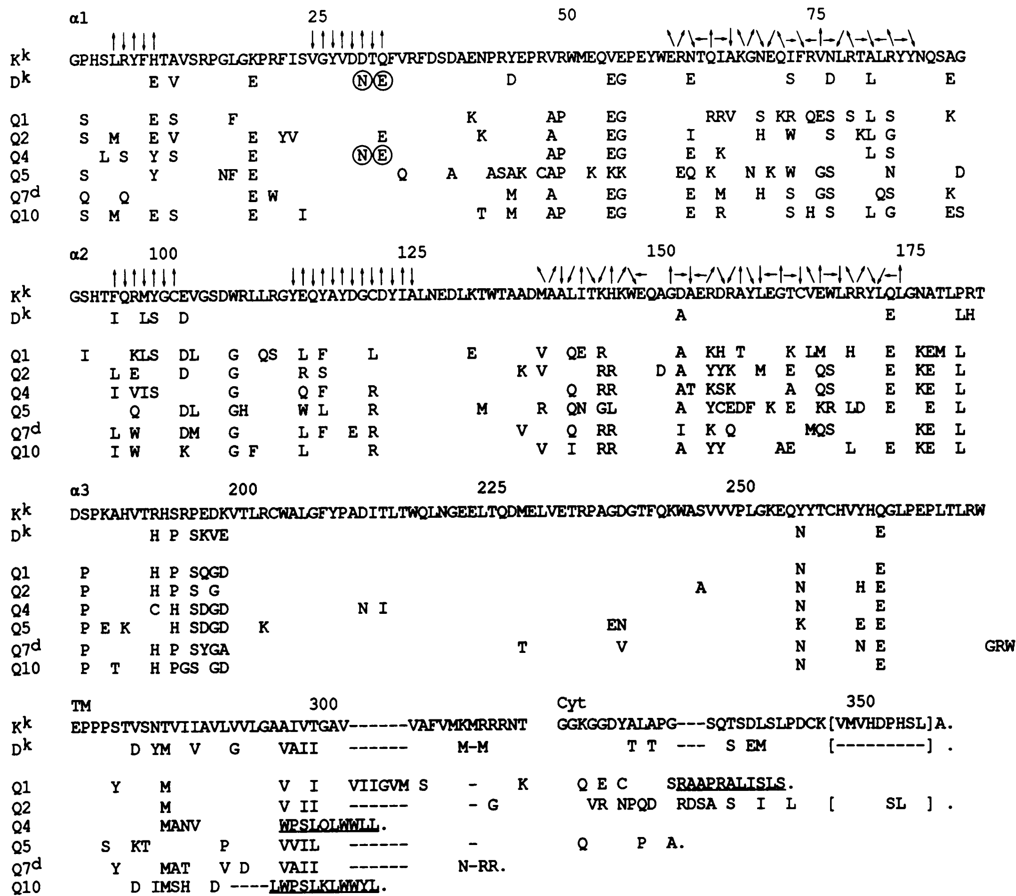

Fig. 4. Aligned amino acid sequence of H-2^k and Qa^k gene products. Spaces indicate identity to H-2K^k. Frameshifts relative to H-2K^k are underlined. Periods mark termination codons. Amino acids possibly introduced by alternative splicing at exon 8 are bracketed. Residues affected by potential Q4^k/H-2D^k gene conversion are circled. Arrows indicate the orientation of side chains of amino acids which comprise the β-sheets and α-helices of the model class I structure: pointing up or down (↑ or ↓), pointing toward or away from the cleft (→ or ←), pointing up or down and toward the cleft (↗ or ↘), pointing up or down and away from the cleft (↘ or ↗).

et al., 1986) or C57BL/10 (Weiss *et al.*, 1984). The deletion of the Q6–Q9 genes via recombination of Q5 with Q9 to form a Q5/9 hybrid accounts for the Qa-2^{null} phenotype of this strain. Each of the five Qa^k genes appears able to potentially encode a class I molecule which might be capable of binding peptides for presentation to an appropriate T-cell receptor. In comparing class I sequences, we found little evidence for extensive gene conversion among Qa and H-2 loci within the k haplotype which might contribute to the generation of H-2 polymorphism.

Several theories have been proposed to explain the generation and maintenance of the extraordinary polymorphism found among the transplantation antigen genes (reviewed in Klein and Figueroa, 1986). It is unlikely that a high mutation rate for class I genes can explain class I variability since these loci do not appear to be evolving particularly rapidly. In fact, nucleotide substitutions at synonymous positions seem to accumulate at a lower rate than for most other mammalian genes (Hayashida and Miyata, 1983). Another argument against a relatively high MHC mutation rate centers around the presence of populations of mice isolated for a known period of time in which one or very few H-2 alleles exist (Klein and Figueroa, 1986). Furthermore, recent comparisons of class I alleles among humans and chimpanzees, as well as rodents, have demonstrated that these molecules can be strikingly conserved between species (Figueroa *et al.*, 1988; Lawlor *et al.*, 1988; Mayer *et al.*, 1988). This suggests that the multiple divergent alleles may be very old, rather than evolving quickly (Mayer *et al.*, 1988).

A gene conversion mechanism was invoked from the analysis of the H-2K^{bm} meiotic mutants (Pease *et al.*, 1983; Nathenson *et al.*, 1986) as well as pairwise comparison of H-2 (Jaulin, 1985; Pease, 1985) and HLA (Jaulin, 1985) sequences which revealed clusters of amino acid differences which could be accounted for in other class I loci, often from the Qa region in mice (Nathenson *et al.*, 1986). Furthermore, because of the ratio of clustered to nonclustered events among the H-2K^{bm} mutants, interlocus gene conversion was hypothesized to be responsible for most of the class I diversity (Pease, 1985; Nathenson *et al.*, 1986). Our analysis is the first in which all of the sequences within the H-2 and Qa regions of a murine haplotype have been compared. While we see two possible examples of exchange between H-2 and Qa genes, sharing of the clustered sequences among H-2 alleles suggests that the direction of sequence transfer was from H-2 to Qa, rather than Qa to H-2. We would have expected to find many more examples of residual gene conversions with apparent Qa donors and H-2 recipients if a directionally biased recombination mechanism had been largely responsible for the generation of H-2 diversity. This does not exclude a role for interlocus gene conversion in

the concerted evolution of this multigene family, but it seems likely that other mechanisms, e.g. point mutation and selection, are also influential in the generation of H-2 variability.

Involvement of the class I molecules in antigen presentation to the variable T-cell repertoire led to the hypothesis that positive selection might be the driving force behind fixation of polymorphism in H-2 genes (Klein and Figueroa, 1986; Hughes and Nei, 1988). A role for selection is supported by a statistical study which showed that the number of nonsynonymous nucleotide substitutions outnumbered synonymous ones for nearly every amino acid within the predicted active site of the class I molecule in both humans and mice (Hughes and Nei, 1988). Although we cannot yet ascribe a biological function to Qa molecules, the observation that they do not act as classical restriction elements suggests that the Qa genes are likely to be under different selective constraints from the H-2 genes. Indeed, we found a few H-2- and Qa-specific amino acids at positions presumed to interact with the T-cell receptor (Ajitkumar *et al.*, 1988). If these particular amino acids are important in distinguishing H-2 or Qa function, then certain H-2/Qa exchanges might be deleterious. This would be expected to limit severely any rare gene conversion event between Qa and H-2 sequences involving these residues.

The function of Qa antigens remains a mystery, and is further obscured by their apparently dispensable nature in inbred mouse strains (O'Neill *et al.*, 1986), as well as their ostensible absence from other well-characterized species, e.g. humans (Klein, 1986). Large deletions of Qa genes in some inbred mouse strains have been interpreted as evidence against their having a biologically significant function. However, if Qa products interact with immune cells to fine-tune the immune response, it is conceivable that they might be expendable depending on the avidity of the T-cell receptor population to the respective H-2 molecules within a particular haplotype. This could allow deletions of Qa products to become fixed in the population in association with certain transplantation antigens. Alternatively, the function of a given Qa gene product might be assumed by other members of the class I family in a system which utilizes several molecules to modulate the immune response.

There is not unequivocal data suggesting that the Qa products function as restriction elements for $\alpha\beta$ T-cell receptors in the classical sense (Robinson, 1987a). It is becoming increasingly clear that minor histocompatibility antigens are recognized in processed form bound to transplantation antigens (Roser, 1986), and even allograft responses may involve restricted recognition of allo-peptides presented by self H-2 (Maryanski *et al.*, 1986; Clayberger *et al.*, 1987; Song *et al.*, 1988). Yet some CTLs generated against Qa-2 and a genetically engineered cell-surface Q10 seem to recognize their targets in an H-2 unrestricted manner (Forman *et al.*, 1982; Mann *et al.*, 1987; Mann and Forman, 1988), implying that the Qa structure is capable of interacting directly with an appropriate T-cell receptor. Accordingly, it has been hypothesized that the more conserved class I molecules might be recognized by $\gamma\delta$ T-cell receptors in immune surveillance of epithelial tissue (Janeway *et al.*, 1988).

The theory that Qa gene products that are secreted could circulate and interact with immune cells to help regulate the immune response is an attractive one. Qa-2 expression appears to be increased and the molecule processed to a secreted rather than a cell surface form in some subsets of activated T-cells (Soloski *et al.*, 1986), implying its potential involvement in T-cell function. However, the mechanism by which these conserved Qa molecules could mimic the large potential of self class I plus peptide structures to which T-cell receptors can exist in order to effect positive or negative immune regulation is not obvious.

The availability of Qa gene sequences from which to generate locus-specific nucleotide probes, as well as derive peptides which could be used to raise anti-Qa monoclonal antibodies, should greatly facilitate future studies of Qa structure, expression and potential function. For example, the presence of Qa-specific amino acids on the tops of the α-helices which are presumed to interact with residues on the T-cell receptor could reflect the specificity of Qa products for a unique family of T-cell receptors. Accordingly, the theory of an immune modulatory role for Qa molecules might be addressed by systematically searching for Qa and Tla gene transcripts and products, analyzing the effects of transfected Qa gene expression on the specificity of distinct T-cell populations, and comparing the T-cell receptor repertoire in inbred mouse strains bearing Qa gene deletions, such as A.CA and C3H.

Materials and methods

Isolation of Qa^k clones

C3H/HeN spleen DNA was partially digested with *Mbo*I and cloned into the *Bam*HI site of λ EMBL3 (Frischauf *et al.*, 1983). Five genome equivalents of recombinant phage were screened using the cDNA probe pH-2IIa (Steinmetz *et al.*, 1981a,b), and 150 class-I-containing clones identified. C3H/HeJ liver DNA was prepared for construction of a Stratagene Custom Library (11099 North Torrey Pines Road, La Jolla, CA 92037, USA) in the cosmid vector pWE15 (Evans and Wahl, 1987). More than six genome equivalents of cosmid library were screened with a genomic H-2K^b clone from which intron 3 had been deleted, pK$^b\Delta$i3, a gift of Richard Flavell. About 50 cosmid clones contained class I inserts in sizes ranging from 15 to 35 kb were isolated. Lambda clones were restriction mapped with *Bam*HI, *Hpa*I, *Kpn*I, *Hin*dIII, *Xba*I, *Bgl*II, *Pst*I, *Xho*I and *Sal*I, while cosmids were mapped with *Bam*HI, *Hpa*I, *Kpn*I, *Hin*dIII, *Sac*II, *Xho*I and *Sal*I (enzymes from New England Biolabs or Boehringer Mannheim Biochemicals). H-2- and Qa-containing clones were identified by comparing their restriction maps to those published for BALB/c (Steinmetz *et al.*, 1982; Stephan *et al.*, 1986) and C57BL/10 (Weiss *et al.*, 1984). Representative restriction sites were chosen for Figure 1 based on having both a complete map across the Qa region and published data for the other strains.

DNA sequencing

The Qa^k genes were subcloned into M13mp18 or mp19 from the λ clones, with the exception of Q2^k, which was obtained from a cosmid. Sequencing was by the dideoxy chain termination method using class-I-specific oligonucleotide primers as previously described (Linsk *et al.*, 1986). Most sequencing reactions were carried out using Sequenase (United States Biochemical Corporation), although Klenow (Boehringer Mannheim Biochemicals) was also used. All clones were sequenced at least twice in both orientations.

Sequence analysis

DNA sequences were analyzed using the Bionet sequence analysis package from Intelligenetics (Bion Workstation Version 5.2 for UNIX, Licensed to UCB, No. 7224.000154, 1988, Intelligenetics, 124 University Avenue, Palo Alto, CA 94301), or the UCSF Sequence Analysis Programs (Hugo Martinez, University of California, San Francisco, CA 94143).

Acknowledgements

We thank Dori Nakahara for excellent technical assistance, as well as Roxanne Morse, Mary Kuhner, Minnie McMillan and Iwona Stroynowski for informative discussions. This work was supported by NIH grant CA37099

to R.S.Goodenow. S.Watts was supported by Pre-graduate Training Grant NIH/GM07232-10.

References

Ajitkumar,P., Geier,S.S., Kesari,K.V., Borriello,F., Nakagawa,M., Bluestone,J.A., Saper,M.A., Wiley,D.C. and Nathenson,S.G. (1988) *Cell*, **54**, 47–56.

Arnold,B., Bugert,H.-G., Archibald,A.L. and Kvist,S. (1984) *Nucleic Acids Res.*, **12**, 9473–9487.

Bjorkman,P.J., Saper,M.A., Samraoui,B., Bennett,W.S., Strominger,J.L. and Wiley,D.C. (1987a) *Nature*, **329**, 506–512.

Bjorkman,P.J., Saper,M.A., Samraoui,B., Bennett,W.S., Strominger,J.L. and Wiley,D.C. (1987b) *Nature*, **329**, 512–518.

Chen,Y.-T., Obata,Y., Stockert,E., Takahashi,T. and Old,L.J. (1987) *Immunol. Res.*, **6**, 30–45.

Clayberger,C., Parham,P., Rothbard,J., Ludwig,D.S., Schoolnik,G.K. and Krensky,A.M. (1987) *Nature*, **330**, 763–765.

Cosman,D., Kress,M., Khoury,G. and Jay,G. (1982) *Proc. Natl. Acad. Sci. USA*, **79**, 4947–4951.

David-Watine,B., Transy,C., Gachelin,G. and Kourilsky,P. (1987) *Gene*, **61**, 145–154.

Devlin,J.J., Lew,A.M., Flavell,R.A. and Coligan,J.E. (1985a) *EMBO J.*, **4**, 369–374.

Devlin,J.J., Weiss,E.H., Paulson,M. and Flavell,R.A. (1985b) *EMBO J.*, **4**, 3203–3207.

Evans,G.A. and Wahl,G.M. (1987) *Methods Enzymol.*, **152**, 604–610.

Figueroa,F., Gunther,E. and Klein,J. (1988) *Nature*, **335**, 265–267.

Flaherty,L., DiBiase,K., Lynes,M.A., Seidman,J.G., Weinberger,O. and Rinchik,E.M. (1985) *Proc. Natl. Acad. Sci. USA*, **82**, 1503–1507.

Forman,J., Trial,J., Tonkonogy,S. and Flaherty,L. (1982) *J. Exp. Med.*, **155**, 749–767.

Friedman,R.L. and Stark,G.R. (1985) *Nature*, **314**, 637–639.

Frischauf,A.-M., Lehrach,H., Poustka,A. and Murray,N. (1983) *J. Mol. Biol.*, **170**, 827–830.

Geraghty,D.E., Koller,B.H. and Orr,H.T. (1987) *Proc. Natl. Acad. Sci. USA*, **84**, 9145–9149.

Harris,R.A., Hogarth,P.M., Penington,D.G. and McKenzie,I.F.C. (1984) *J. Immunogenet.*, **11**, 265–281.

Hayashida,H. and Miyata,T. (1983) *Proc. Natl. Acad. Sci. USA*, **80**, 2671–2675.

Hood,L., Steinmetz,M. and Malissen,B. (1983) *Annu. Rev. Immunol.*, **1**, 529–568.

Hughes,A.L. and Nei,M. (1988) *Nature*, **335**, 167–170.

Israel,A., Kimura,A., Kieran,M., Yano,O., Kanellopoulos,J., Le Bail,O. and Kourilsky,P. (1987) *Proc. Natl. Acad. Sci. USA*, **84**, 2653–2657.

Janeway,C.A.,Jr, Jones,B. and Hayday,A. (1988) *Immunol. Today*, **9**, 73–76.

Jaulin,C., Perrin,A., Abastado,J.-P., Dumas,B., Papamatheakis,J. and Kourilsky,P. (1985) *Immunogenetics*, **22**, 453–470.

Klein,J. (1986) *Natural History of the Major Histocompatibility Complex.* John Wiley and Sons, New York.

Klein,J. and Figueroa,F. (1986) *CRC Crit. Rev. Immunol.*, **6**, 295–386.

Koller,B.H., Geraghty,D.E., Shimizu,Y., DeMars,R. and Orr,H.T. (1988) *J. Immunol.*, **141**, 897–904.

Korber,B., Mermod,N., Hood,L. and Stroynowski,I. (1988) *Science*, **239**, 1302–1306.

Kress,M., Cosman,D., Khoury,G. and Jay,G. (1983a) *Cell*, **34**, 189–196.

Kress,M., Glaros,D., Khoury,G. and Jay,G. (1983b) *Nature*, **306**, 602–604.

Lalanne,J.-L., Transy,C., Guerin,S., Darche,S., Meulien,P. and Kourilsky,P. (1985) *Cell*, **41**, 469–478.

Lawlor,D.A., Ward,F.E., Ennis,P.D., Jackson,A.P. and Parham,P. (1988) *Nature*, **335**, 268–271.

Lew,A.M., Maloy,W.L. and Coligan,J.E. (1986) *J. Immunol.*, **136**, 254–258.

Linsk,R., Vogel,J., Stauss,H., Forman,J. and Goodenow,R.S. (1986) *J. Exp. Med.*, **164**, 794–813.

Maloy,W.L., Coligan,J.E., Barra,Y. and Jay,G. (1984) *Proc. Natl. Acad. Sci. USA*, **81**, 1216–1220.

Mann,D.W., Stroynowski,I., Hood,L. and Forman,J. (1987) *J. Immunol.*, **138**, 240–245.

Mann,D. and Forman,J. (1988) *J. Immunol.*, **141**, 1813–1818.

Maryanski,J.L., Pala,P., Corradin,G., Jordan,B.R. and Cerottini,J.-C. (1986) *Nature*, **324**, 578–579.

Mayer,W.E., Jonker,M., Klein,D., Ivanyi,P., van Seventer,G. and Klein,J. (1988) *EMBO J.*, **7**, 2765–2774.

Mellor,A.L., Weiss,E.H., Kress,M., Jay,G. and Flavell,R.A. (1984) *Cell*, **36**, 139–144.

Mellor,A.L., Antoniou,J. and Robinson,P.J. (1985) *Proc. Natl. Acad. Sci. USA*, **82**, 5920–5924.

Mizuno,S., Trapani,J.A., Koller,B.H., Dupont,B. and Yang,S.Y. (1988) *J. Immunol.*, **140**, 4024–4030.

Nathenson,S.G., Geliebter,J., Pfaffenbach,G.M. and Zeff,R.A. (1986) *Annu. Rev. Immunol.*, **4**, 471–502.

O'Neill,A.E., Reid,K., Garberi,J.C., Karl,M. and Flaherty,L. (1986) *Immunogenetics*, **24**, 368–373.

Palmer,M.J. and Frelinger,J.A. (1987) *J. Exp. Med.*, **166**, 95–108.

Pease,L.R. (1985) *Transplantation*, **39**, 227–231.

Pease,L.R., Schulze,D.H., Pfaffenbach,G.M. and Nathenson,S.G. (1983) *Proc. Natl. Acad. Sci. USA*, **80**, 242–246.

Ritzel,G., McCarthy,S.A., Fotedar,A. and Singh,B. (1984) *Immunol. Today*, **5**, 343–345.

Robinson,P.J. (1985) *Immunogenetics*, **22**, 285–289.

Robinson,P.J. (1987a) *Immunol. Res.*, **6**, 46–56.

Robinson,P.J. (1987b) *Proc. Natl. Acad. Sci. USA*, **84**, 527–531.

Robinson,P.J., Bevec,D., Mellor,A.L. and Weiss,E.H. (1988) *Immunogenetics*, **27**, 79–86.

Ronne,H., Widmark,E., Rask,L. and Peterson,P.A. (1985) *Proc. Natl. Acad. Sci. USA*, **82**, 5860–5864.

Roser,B. (1986) *Immunol. Today*, **7**, 76–77.

Sherman,D.H., Waneck,G.L. and Flavell,R.A. (1988) *J. Immunol.*, **140**, 138–142.

Shimizu,Y., Geraghty,D.E., Koller,B.H., Orr,H.T. and DeMars,R. (1988) *Proc. Natl. Acad. Sci. USA*, **85**, 227–231.

Shirayoshi,Y., Miyazaki,J.-I., Burke,P.A., Hamada,K., Appella,E. and Ozato,K. (1987) *Mol. Cell. Biol.*, **7**, 4542–4548.

Singer,M.F. (1982) *Int. Rev. Cytol.*, **76**, 67–112.

Soloski,M.J., Vernachio,J., Einhorn,G. and Lattimore,A. (1986) *Proc. Natl. Acad. Sci. USA*, **83**, 2949–2953.

Soloski,M.J., Lattimore,A., Hereld,D., Krakow,J.L., Low,M.G. and Einhorn,G. (1988a) *J. Immunol.*, **140**, 3858–3866.

Soloski,M.J., Hood,L. and Stroynowski,I. (1988b) *Proc. Natl. Acad. Sci. USA*, **85**, 3100–3104.

Song,E., Linsk,R., Olson,C.A., McMillan,M. and Goodenow,R.S. (1988) *Proc. Natl. Acad. Sci. USA*, **85**, 1927–1931.

Srivastava,R., Chorney,M.J., Lawrance,S.K., Pan,J., Smith,Z., Smith,C.L. and Weissman,S.M. (1987) *Proc. Natl. Acad. Sci. USA*, **84**, 4224–4228.

Steinmetz,M., Frelinger,J.G., Fischer,D., Hunkapillar,T., Perera,D., Weissman,S.M., Ushara,H., Nathenson,S. and Hood,L. (1981a) *Cell*, **24**, 125–134.

Steinmetz,M., Moore,K.W., Frelinger,J.G., Sher,B.T., Shen,F.-W., Boyse,E.A. and Hood,L. (1981b) *Cell*, **25**, 683–692.

Steinmetz,M., Winoto,A., Minard,K. and Hood,L. (1982) *Cell*, **28**, 489–498.

Stephan,D., Sun,H., Lindahl,K.F., Meyer,E., Hammerling,G., Hood,L. and Steinmetz,M. (1986) *J. Exp. Med.*, **163**, 1227–1244.

Stiernberg,J., Low,M.G., Flaherty,L. and Kincade,P.W. (1987) *J. Immunol.*, **138**, 3877–3884.

Stroynowski,I., Soloski,M., Low,M.G. and Hood,L. (1987) *Cell*, **50**, 759–768.

Sugita,K., Miyazaki,J.-I., Appella,E. and Ozato,K. (1987) *Mol. Cell. Biol.*, **7**, 2625–2630.

Townsend,A.R.M., Rothbard,J., Gotch,F.M., Bahadur,G., Wraith,D. and McMichael,A.J. (1986) *Cell*, **44**, 959–968.

Vogel,J., Kress,M., Khoury,G. and Jay,G. (1986) *Mol. Cell. Biol.*, **6**, 3550–3554.

Vogel,J.M., Davis,A.C., McKinney,D.M., McMillan,M., Martin,W.J. and Goodenow,R.S. (1988) *J. Exp. Med.*, **168**, 1781–1800.

Waneck,G.L., Sherman,D.H., Calvin,S., Allen,H. and Flavell,R.A. (1987) *J. Exp. Med.*, **165**, 1358–1370.

Waneck,G.L., Sherman,D.H., Kincade,P.W., Low,M.G. and Flavell,R.A. (1988) *Proc. Natl. Acad. Sci. USA*, **85**, 577–581.

Warner,C.M., Gollnick,S.O., Flaherty,L. and Godlbard,S.B (1987) *Biol. Reprod.*, **36**, 611–616.

Watts,S., Vogel,J.M., Harriman,W.D., Itoh,T., Stauss,H.J. and Goodenow,R.S. (1987) *J. Immunol.*, **139**, 3878–3885.

Weiss,E.H. (1987) *Immunol. Res.*, **6**, 179–191.

Weiss,E.H., Golden,L., Fahrner,K., Mellor,A.L., Devlin,J.J., Bullman,H., Tiddens,H., Bud,H. and Flavell,R.A. (1984) *Nature*, **310**, 650–655.

Yano,O., Kanellopoulos,J., Kieran,M., Le Bail,O., Israel,A. and Kourilsky,P. (1987) *EMBO J.*, **6**, 3317–3324.

Received on January 24, 1989; revised on March 7, 1989