RNA transcripts for I-J polypeptides are apparently not encoded between the I-A and I-E subregions of the murine major histocompatibility complex

(RNA blot/T-cell hybridoma/suppressor cell/I region genes)

MITCHELL KRONENBERG*, MICHAEL STEINMETZ*†, JOAN KOBORI*, ELLEN KRAIG*, JUDITH A. KAPP‡, CARL W. PIERCE‡, CRAIG M. SORENSEN‡, GEN SUZUKI§, TOMIO TADA§, and LEROY HOOD*

*Division of Biology, California Institute of Technology, Pasadena, California 91125; †Department of Pathology and Laboratory Medicine, The Jewish Hospital of St. Louis, and Department of Pathology and of Microbiology and Immunology, Washington University School of Medicine, St. Louis, Missouri 63110; and
§Department of Immunology, University of Tokyo, 7-3-1 Hongo Bunkyo-Ku, Tokyo, Japan

Contributed by Leroy Hood, June 6, 1983

ABSTRACT The I-J subregion of the mouse major histocompatibility complex has been reported to encode antigenic determinants expressed by suppressor T cells. Previously, cloned clones were obtained from mouse sperm DNA that contain all of the sequences between the I-A and I-E subregions, where I-J has been mapped genetically. However, hybridization of these sequences to RNA prepared from several I-J-positive suppressor T-cell hybridomas did not reveal the presence of a transcript. In addition, no rearrangements in this DNA were detected in the suppressor T cells that we have analyzed. Our results indicate that the I-J polypeptides are not encoded between the I-A and I-E subregions of the major histocompatibility complex. We discuss several hypotheses concerning the possible location and expression of I-J genes.

The I region of the murine major histocompatibility complex (MHC) encodes polymorphic cell surface molecules that are involved in lymphocyte interactions (1, 2). The I region has been divided by analysis of recombinant inbred mice into five subregions arranged in the order I-A, I-B, I-J, I-E, I-C (3). The I-A and I-E subregions code for class II molecules or Ia antigens, which are found predominantly on the surfaces of B cells and macrophages (4, 5). These molecules are composed of two polypeptides, an α chain of approximately 32,000 daltons and a β chain of approximately 28,000 daltons (6). To elicit an immune response, some T lymphocytes, mostly helper T cells, must recognize syngeneic class II molecules in addition to the stimulating antigen (7, 8). The other subregions, I-B, I-J, and I-C, are less well characterized. For example, while the I-B subregion may affect the level of the immune response to certain antigens, it does not encode serologically detectable molecules and its existence is controversial (9, 10). On the other hand, numerous alloantiseri and monoclonal antibodies have been raised against lymphocytes from strains that differ only in the I-J subregion (11–14). The antigens recognized by these antibodies have a unique distribution; they are found predominantly on the surface of suppressor T cells and on soluble factors with suppressive activity secreted by these cells (11–16). In addition, I-J-positive subsets of helper T lymphocytes (17) and macrophages (18) have been detected. Interest in the I-J subregion and its gene products has been heightened by two findings. First, I-J-encoded molecules are associated with or are actually part of the antigen-binding polypeptides made by suppressor T cells (19–21). Thus, a biochemical characterization of I-J molecules may help to define the T-cell antigen receptor. Second, in some cases I-J-encoded molecules appear to regulate interactions between T lymphocytes (22).

There are several reports indicating that the I-J determinant is present on proteins with a molecular mass of approximately 25,000 daltons (20, 21, 23). In addition, because the I-J serological determinant can be found on proteins translated in vitro by a rabbit reticulocyte lysate, the antigenic determinant is presumed to be a polypeptide rather than a carbohydrate structure (23). Little more is known, however, about the structure of I-J-encoded molecules.

Recently, a 2.0-kilobase (kb) DNA sequence has been defined that should contain the gene(s) encoding the I-J serological determinant (24, 25). Using two human cDNA probes specific for DRα, and DCβ, we have isolated 200 kb of contiguous I region sequence from BALB/c mouse (H-2d) DNA. The coding sequences for the Eα chain from the I-A subregion and the Eα chain from the I-E subregion were found to be separated by only 33 kb of DNA (24). Next, to define more precisely the boundaries of the I-A and I-E subregions, we identified polymorphic restriction enzyme sites in the I region and tested genomic DNA from parental and I-region recombinant strains for the presence of these polymorphic sites (24, 25). This analysis permitted us to correlate the genetic map of the I region with the molecular map of the clodi cluster. We discovered that the boundaries of the I-A and I-E subregions, which define I-J, are less than 2.0 kb apart (see Fig. 1). The distance separating these subregions might be smaller than 2.0 kb, but no polymorphic restriction enzyme sites have been found to permit characterization of this segment of DNA.

There are two reasons why this region seems too small to encode the I-J polypeptides. First, the 2.0-kb sequence is actually contained within the Eβ gene and includes part of the intron between the first (β1) and second (β2) major exons and probably all of the second major exon (β2) of the Eα gene (see Abbreviations: kb, kilobase(s); MHC, major histocompatibility complex. †Present address: Basel Institute for Immunology, 457 Grenzacherstrasse, Postfach 4005, Basel, Switzerland. §Present address: Laboratory of Immunology, National Inst. of Health, Bethesda, MD 20205.

5704
Fig. 1). Second, distinct I-J-encoded specificities have been detected on different T-cell subpopulations or cloned T cells, suggesting that the I-J subregion could in fact contain several genes (26-28). To resolve this problem, we have proposed several models for the location and expression of the I-J gene(s) (24). The models are summarized below.

1. The I-J gene is encoded between the I-A and I-E subregions. (a) The I-J gene is actually the $E_p$ gene. The I-J specificity is formed via special post-translational modification (e.g., glycosylation) of the $E_p$ gene product. (b) The I-J gene is formed via an alternative RNA splicing pattern that includes some of the $E_p$ exons and some exons unique to I-J. (c) The I-J gene is transcribed from the DNA strand complementary to that which encodes the $E_p$ gene. (d) The I-J gene is formed by a DNA rearrangement in suppressor T lymphocytes, which inserts some coding sequences between the I-A and I-E subregions.

2. The I-J structural gene is not encoded between the I-A and I-E subregions. (a) Because of the occurrence of multiple recombination events, the map order I-A/I-J/I-E is incorrect. (b) The I-J subregion contains a control element that regulates the expression of I-J genes encoded elsewhere. (c) The I-J serologic determinants are not encoded in the MHC. These determinants may be present on T-cell receptors for self MHC molecules.

In this paper, we report experiments that exclude models 1 a–d, which propose that the I-J gene is located between the I-A and I-E subregions. The remaining models (2 a–c), which propose that the I-J coding sequence is located elsewhere, are considered in the Discussion.

MATERIALS AND METHODS

RNA and Southern Blots. Blot hybridization experiments were carried out as described (29, 30). To reduce background caused by the hybridization of repeat nucleotide sequences in the probes, both the prehybridization and hybridization solutions contained denatured BALB/c liver DNA at 25 µg/ml. The filters were washed with 15 mM NaCl/1.5 mM sodium citrate/0.1% NaDodSO4 at 68°C and exposed to Kodak XAR-5 film in the presence of an intensifying screen for 4 days.

Hybridization Probes. Hybridization probes derived from the I region included the sequences diagrammed in Fig. 1. 3.48 and 8.47 are 7.2- and 4.8-kb HindIII restriction fragments from cosmids 8.4, which were subcloned into pBR325 by standard methods (31, 32). Probe 3, a 2-kb EcoRI restriction fragment derived from cosmids 24.2, and probe 4, a 500-base-pair Sau3A fragment purified from cosmids 39.1, have been described (24). pH211a is a 688-base-pair Pst I/Pvu II subclone from a cDNA clone encoding an H-2d transplantation antigen (33).

RESULTS

Transcripts Homologous to the Presumed I-J Subregion Are Not Detectable in Suppressor T Cells. To identify I-J transcripts, we analyzed 13 suppressor T-cell hybridomas for the presence of RNA molecules that hybridize to cloned DNA probes containing the putative I-J subregion. Some properties of these suppressor cells (nos. 1–13) are summarized in Table 1. All of the hybridomas were tested with anti-I-J alloantiserum, anti-I-J monoclonal antibodies, or both (refs. 14, 27, 34–36; unpublished data). In each of 13 cases tested, the cells produced an I-J-positive suppressor factor. Furthermore, eight out of eight hybridomas tested were positive for cell-surface I-J determinants. Because extensive sequence polymorphism might greatly reduce hybridization of probes derived from the H-2d haplotype to transcripts derived from other strains, we tested one hybridoma generated from a fusion with BALB/c lymphocytes, the same inbred strain from which the cosmid clone bank had been prepared. Blots of poly(A)+ RNA from each of the T cells were hybridized with a mixture of 32P-labeled subclones 8.48 and 8.47 (Fig. 1), which together contain the $E_p$ gene and all of the DNA between the I-A and I-E subregions. As shown in a set of characteristic blots that were hybridized in parallel (Fig. 2), there was no evidence of a transcript homologous to the probe in any of the suppressor T-cell RNA that we analyzed. Each of the RNA preparations was tested independently on at least two blots (Fig. 2; unpublished data). Since MHC-encoded transplantation antigens are expressed by most cell types, rehybridization of the same RNA blots with pH211a, a probe encoding an H-2d transplantation antigen, provided a positive control for the presence of intact RNA. Although the hybridomas transcribe various amounts of RNA homologous to this probe, a distinct 1.9-kb band was detected in each case (unpublished data). To determine the detection limit of the RNA blots, additional control hybridizations were carried out using subclones 8.48 and 8.47 and various amounts of spleen poly(A)+ RNA. We could detect hybridization of the $E_p$ coding sequence in these subclones to 0.1–0.3 µg of spleen RNA. To estimate the abundance of the $E_p$ transcript in spleen poly(A)+ RNA, we compared the $E_p$ hybridization signal to the signal obtained with a k light chain constant-region probe hybridized to various amounts of purified SI07 myeloma $k$ chain RNA. This permitted us to correlate the intensity of hybridization bands on autoradiograms with the number of picograms of complementary RNA present. In this manner, we estimated that the $E_p$ RNA represents approximately 0.03% of the spleen poly(A)+ RNA (unpublished data). Similar estimates for the abundance of class II mRNA have been reported from other laboratories (37–39). Fig. 2 shows that we could detect a RNA species that is 0.03% of 0.3 µg spleen RNA tested. Thus, we should have been able to detect sequences present at less than one part in 109 contained in the 10 µg of T-cell RNA analyzed. Since the

<table>
<thead>
<tr>
<th>No.</th>
<th>Hybridoma</th>
<th>H-2d haplotype*</th>
<th>Antigen specificity†</th>
<th>Suppressor factor</th>
<th>Cell surface</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>258C4.4</td>
<td>q GAT/GA</td>
<td>+</td>
<td>+</td>
<td>NT</td>
</tr>
<tr>
<td>2</td>
<td>342B1.11</td>
<td>s GAT/GT</td>
<td>+</td>
<td>+</td>
<td>NT</td>
</tr>
<tr>
<td>3</td>
<td>366C6.4</td>
<td>s GAT</td>
<td>NT</td>
<td>+</td>
<td>NT</td>
</tr>
<tr>
<td>4</td>
<td>367A5.2</td>
<td>k GT</td>
<td>+</td>
<td>+</td>
<td>NT</td>
</tr>
<tr>
<td>5</td>
<td>368B1.5</td>
<td>d GT/GAT</td>
<td>+</td>
<td>+</td>
<td>NT</td>
</tr>
<tr>
<td>6</td>
<td>301D4.5</td>
<td>f GT</td>
<td>+</td>
<td>+</td>
<td>NT</td>
</tr>
<tr>
<td>7</td>
<td>301A2.3</td>
<td>f GT</td>
<td>+</td>
<td>+</td>
<td>NT</td>
</tr>
<tr>
<td>8</td>
<td>469B5.5</td>
<td>a GT/GAT</td>
<td>+</td>
<td>+</td>
<td>NT</td>
</tr>
<tr>
<td>9</td>
<td>372B3.5</td>
<td>b GT/GAT</td>
<td>+</td>
<td>+</td>
<td>NT</td>
</tr>
<tr>
<td>10</td>
<td>372D6.5</td>
<td>b GT/GAT</td>
<td>+</td>
<td>+</td>
<td>NT</td>
</tr>
<tr>
<td>11</td>
<td>395A4.4</td>
<td>s GT/GAT</td>
<td>+</td>
<td>+</td>
<td>NT</td>
</tr>
<tr>
<td>12</td>
<td>7C3-13</td>
<td>k NP</td>
<td>+</td>
<td>+</td>
<td>NT</td>
</tr>
<tr>
<td>13</td>
<td>7F4-30</td>
<td>k NP</td>
<td>+</td>
<td>+</td>
<td>NT</td>
</tr>
<tr>
<td>14</td>
<td>366D3.2</td>
<td>b GT/GAT</td>
<td>+</td>
<td>+</td>
<td>NT</td>
</tr>
</tbody>
</table>

NT, not tested.

*All the hybridomas were formed by fusion of lymphocytes to the hypoxanthine/aminopterin/thymidine-sensitive AKR thymoma BW5147, and they therefore contain H-2d haplotype DNA contributed by the tumor parent.

†Antigen-binding specificities of T-hybridoma suppressor factors. GA is a synthetic poly(Glu-Ala)60; GAT is poly(Glu-Ala-Tyr)60; GT is poly(Glu-Ala-Tyr); NP is 4-hydroxy-3-nitrophenyl acetyl. In all cases, the immunogen is listed first followed by any other known specificities.

Table 1. Characteristics of suppressor T-cell hybridomas

1-J-1
Eβ RNA is about 1,300 nucleotides long and T-cell hybridomas generally contain 0.2–0.3 pg of poly(A)+ RNA per cell (40), our detection limit is roughly 3 or 4 copies per cell.

The I-J Gene Is Not Formed by Rearrangement of DNA Sequences Between the I-A and I-E Subregions. We wished to determine whether a DNA rearrangement in suppressor T cells might be necessary to form a complete I-J gene. High molecular weight DNA that was prepared from four T-cell hybridomas (342B1.11, 366D3.2, 372B3.5, and 395A4.4) was digested with one of several restriction enzymes, and the digests were hybridized with probe 3, probe 4, or a mixture of subclones 8.47 and 8.48 (Fig. 1). When the hybridization pattern obtained from the T-cell hybridomas was compared with that of liver DNA of the appropriate MHC haplotype, no DNA rearrangements were observed on any of these Southern blots. From the I-region restriction map of the liver DNA (24), we concluded that in each case the restriction fragment(s) detected should include the entire sequence between the I-A and I-E subregions. Fig. 3 shows some cases in which the enzyme digestion allowed us to distinguish between these restriction fragments derived from the normal T-cell parent and those from the BW5147 fusion partner.

**DISCUSSION**

Using a series of cosmids, it was previously determined that the boundaries of the I-A and I-E subregions are separated by only 2.0 kb of DNA (24, 25). This analysis confines the serologically assigned I-J subregion to a relatively short sequence that is actually located within the Eβ gene (Fig. 1). A number of experimental results have convinced us that the BALB/c cosmid DNA map is an accurate representation of the germ line sequences found between the Eβ and En genes. First, we have obtained seven different overlapping cosmids clones containing all or part of the DNA between Eβ and En (24). This eliminates the possibility that the map fuses noncontiguous sequences and thereby misses a gene. Second, extensive comparison of the restriction enzyme maps between BALB/c mice and other strains indicate that the BALB/c genomic DNA contains neither a deletion nor an inversion breakpoint between the Eβ and En genes (24, 25). Finally, we have obtained cosmids DNA clones from the AKR inbred strain, which has a MHC haplotype (H-2β) different from BALB/c. In this strain, the Eβ and En genes are also separated by approximately 33 kb (unpublished data).

We have found no evidence for an I-J coding sequence located between the I-A and I-E subregions. If some I-J exons were present between the I-A and I-E subregions, there could be two reasons for failing to detect them. (i) When cloned suppressor hybridomas are grown in vitro, the I-J serologic spec-

**Fig. 1.** Genetic and molecular map of the I region. The same 19-kb sequence is depicted in A, B, and C. (A) Boundaries of the I-A and I-E subregions as previously determined (24, 25). I-J has been omitted because it is confined to the same 2.0-kb sequence as I-J and codes for no serologically detectable product. (B) Positions and lengths of the hybridization probes. (C) Locations of some Eβ exons as determined by Southern blot hybridization (25) and the direction of transcription of the Eβ gene. β1 and β2 code for the first and second external domains, respectively, while TM indicates the transmembrane/cytoplasmic exon(s).

**Fig. 2.** Hybridization of subclones 8.48 and 8.47 to spleen poly(A)+ RNA (A) and to 10 μg of suppressor T-cell poly(A)+ RNA (B). T-cell hybridomas are designated as numbered in Table 1. RNA preparations on several different filters were hybridized with the subclones. Migration distances of the molecular weight markers, mouse and Escherichia coli ribosomal RNA, and their lengths in kb are indicated.
in a manner.

We have cloned cosmids probes spanning most of the 200 kb of cloned I-region DNA to blot of poly(A)⁺ RNA from two suppressor T cells and have not detected any complementary sequences (unpublished data). Therefore, any short exon between the I-A and I-E subregions would have to be present on an extremely large primary transcript originating outside the sequences we have cloned. The minimum length of such a transcript, assuming that the I-J exon is located at one of the ends of the mRNA, would be appropriately 60 kb. Alternatively, a short exon encoding the polymorphic I-J determinant could be co-expressed with other exons after a DNA rearrangement. However, gene rearrangement was not observed in Southern blots of four suppressor T cells. Because the DNA from these cells was tested after digestion with several different restriction enzymes (Fig. 3 and unpublished data), it is unlikely that the DNA was fortuitously rearranged to the same size restriction fragment as is present in the germ line. In addition, any relevant rearrangement should have been detected because the hybridizing restriction fragments include all of the DNA between the I-A and I-E subregions. Finally, although a small insertion (≈250 base pairs) would not greatly augment the size of the hybridizing restriction fragments and therefore might not have been detected, a larger insertion would be required to form an I-J structural gene. We therefore conclude that the I-J gene is not formed by a DNA rearrangement between the I-A and I-E subregions.

Our data suggest that the genes encoding I-J serologic determinants expressed by suppressor T cells do not map between the I-A and I-E subregions. There are several possibilities to explain the discrepancy between the immunogenetic and molecular maps. (i) It may be that I-J is linked to the MHC but the assigned map position is incorrect. The genetic mapping is based on analysis of a few recombinant mice and the reasonable assumption that multiple crossovers are rare. Yet in some regions of the mouse genome, multiple crossovers are relatively common (41). If multiple crossovers in the I region were the rule rather than the exception, then conceivably I-J could have an apparent map position between I-A and I-E even though it is located elsewhere in the MHC. However, it should be noted that we do not know of any documented cases of double recombinants involving the well-characterized MHC loci. (ii) Each inbred mouse strain may contain several pseudoallelic I-J genes and a regulatory locus that determines the I-J gene expressed. To explain the reported properties of I-J gene products, the regulatory locus must be polymorphic, it must map between the I-A and I-E subregions, and it must regulate expression of the pseudoallelic genes in a haplotype-specific manner. There are no proven examples of such a polymorphic regulatory sequence. (iii) I-J may map outside of the MHC, perhaps even to another chromosome. Although it is difficult to reconcile the genetic map with this hypothesis, the recombinant congenic mice used to define I-J also may differ at a non-MHC locus that is involved in suppression. There are several mechanisms that could then lead to an apparent map position for the I-J serologic determinant between the I-A and I-E subregions. For example, the I-J serologic determinants may not be present on I-region encoded molecules but instead may be located on molecules encoded elsewhere that recognize Ep. This is possible because some T cells have antigen receptors that recognize either self class II molecules or a complex of antigen with self class II peptides. If receptors that recognize the same Ep allele tend to have common serologic determinants, regardless of the non-MHC genetic background of the strain involved, then the apparent genetic map position of these receptors will be the Ep gene itself. However, it should be noted that receptors for class II molecules are best characterized for helper as opposed to suppressor T cells. In addition, some inbred mouse strains do not express an I-E molecule and therefore may not express Ep (42), while some strains that express different I-J alleles may have identical Ep genes (43).

Our experiments indicate that I-J exons are not likely to be present between the boundaries of the I-A and I-E subregions. While none of the three alternative models for the location of the I-J gene is entirely satisfactory, the possibility of a high frequency of multiple crossovers leading to an incorrect map order for the subregions of the I-region is most attractive, because this hypothesis requires the fewest ad hoc assumptions and special mechanisms to account for the reported data. We have cloned the recombination point between the I-A and I-E subregions in several I-region recombinant strains (25). Further characterization of these recombinants as well as the isolation of cloned
sequences encoding the I-J polypeptide will finally resolve the paradox of I-J gene location and expression.

We thank Dr. Donal Murphy for many helpful discussions. This work was supported by National Institutes of Health Research Grants AI 17565, AI 18095, AI 15333, by Public Health Service Grant EMB 1F32CA06693 awarded by the National Cancer Inst., Department of Health and Human Services (E.K.), and by a grant from the Ministry of Education, Culture, and Science of Japan.