

Characterization of the Primary Structure of T Cell Receptor β Chains in Cells Infiltrating the Salivary Gland in the Sicca Syndrome of HIV-1 Infection

Evidence of Antigen-driven Clonal Selection Suggested by Restricted Combinations of V β J β Gene Segment Usage and Shared Somatically Encoded Amino Acid Residues

Edward Dwyer, Silviu Itescu, and Robert Winchester

Department of Pediatrics, Columbia University, College of Physicians and Surgeons, New York 10032

Abstract

Infection with HIV-1 occasionally results in a sicca syndrome, termed the diffuse infiltrative lymphocytosis syndrome, characterized by infiltration of the salivary glands with a predominance of CD8 T cells. This response is strongly associated with certain MHC class I and class II alleles. To define the salivary gland T cell receptor (TCR) repertoire, the primary structure of the TCR β -chains was determined using *in situ* cDNA synthesis followed by the "anchored" polymerase chain reaction. The sequences of 59 β -chains from five individuals with diffuse infiltrative lymphocytosis syndrome shared structural features suggesting antigenic clonal selection. Certain combinations of V β J β gene segments were selectively overrepresented in the repertoire sample, demonstrating a common restricted usage of certain V β and J β gene segments. The β -chains derived from these overrepresented V β J β combinations revealed a preference for specific amino acids at position 97 in the third complementarity-determining region, a residue postulated to contact peptide antigen. Moreover, the nucleotides encoding this position were not germline in origin. TCR β -chains in nonoverrepresented V β J β combinations did not exhibit preferential usage of selected somatically encoded residues. The pattern of TCR β -chains expressed in the salivary gland of a control person with primary Sjögren's syndrome was considerably more heterogeneous and different from that found in diffuse infiltrative lymphocytosis syndrome. (*J. Clin. Invest.* 1993. 92:495-502.) **Key words:** T cell repertoire • Sjögren's syndrome • diffuse infiltrative lymphocytosis syndrome • immunoglobulin • autoimmunity

Introduction

An unusual form of sicca syndrome termed the diffuse infiltrative lymphocytosis syndrome (DILS),¹ is a distinctive but infrequent host response to infection with HIV. DILS is characterized by infiltration of predominantly CD8 T cells into lacrimal and salivary glands and is frequently accompanied by a

Address correspondence to Edward Dwyer, M.D., Department of Pediatrics, Columbia University, College of Physicians and Surgeons, 630 W. 168th Street, New York, NY 10032.

Received for publication 12 August 1992 and in revised form 12 January 1993.

1. Abbreviations used in this paper: CDR3, third complementarity-determining region; DILS, diffuse infiltrative lymphocytes; OR, odds ratio; TCR, T cell receptor.

J. Clin. Invest.

© The American Society for Clinical Investigation, Inc.

0021-9738/93/00/0000/00 \$2.00

Volume 92, July 1993, 000-000

significant peripheral CD8 T cell lymphocytosis as well as a maintenance of relatively normal levels of circulating CD4 cells (1). The majority of peripheral CD8 and CD4 T cells exhibit a mature memory phenotype characterized by the abundant expression of β 1 and β 2 integrins, without a detectable increase in cells bearing markers of recent activation such as HLA-DR molecules or IL-2 receptors (2). HIV-1 antigens are identifiable only in monocyte lineage cells at sites of tissue infiltration and the virus is culturable from blood monocytes (3). Susceptibility to developing DILS is strongly associated with the DRB1*1102 and DRB1*1301 alleles of HLA-DR5 and DR6, respectively, while HLA-B35 is significantly decreased in frequency (4). These features contrast with the more common idiopathic autoimmune disorder of primary Sjögren's syndrome which is characterized by infiltration of the salivary glands with a predominance of CD4 T cells and susceptibility is associated with HLA-DR2 or DR3 (5).

The ability of T cells to respond to foreign antigens such as HIV-1 is determined by the clonally distributed $\alpha\beta$ T cell receptor (TCR) which is a membrane-bound 88-kD heterodimer that recognizes foreign peptides presented by self MHC molecules (6). The unique recognition structure of each clonal TCR is determined by the combinatorial pairing of distinct V, D, and J gene segments which together constitute each individual molecule. In the case of the TCR β chain, the potential repertoire is composed of 24 V β families, 2 D β segments, and 13 J β segments (7). Although several studies of the β -chain TCR repertoire in humans have demonstrated a nonrandom distribution of V β family expression in both normals and various disease states (8-14), a recent investigation specifically focusing on the primary structure of the entire β -chain variable region revealed a nonbiased pairing of individual V β families with specific J β segments (15). In addition to the use of these multiple germline elements, structural diversity of the β chain is greatly enhanced by exonucleic "nibbling" and N region diversification during junctional pairing of individual V, D, and J segments which results in random deletion or addition of individual nucleotides in generating nongermline encoded amino acids (6); this region of the molecule which exclusively emphasizes somatic diversity is referred to as the third complementarity-determining region (CDR3), and it is this region of the TCR β -chain which is postulated to directly interact with foreign peptides presented by self MHC molecules (16, 17). In support of this model of molecular interaction are several studies which have demonstrated *in vitro* that murine T cell clones responsive to a cytochrome *c*-derived peptide express TCR β chains that show not only restricted usage of specific V β J β combinations, but perhaps more importantly, also exhibit a preferential expression of particular nongermline somatically encoded amino acids in the CDR3 region, most notably at amino acid position 97 (18-20) which is five residues carboxyl

to the invariant cysteine at position 92. Further emphasizing the importance of determining complete primary structure data in obtaining evidence of antigen-specific immune responsiveness is the recent finding that single nonconservative substitutions at specific amino acid positions in the cytochrome *c* peptide are capable of transferring responsiveness to alternative clones which express TCR β chains featuring different $V\beta J\beta$ combinations as well as different non-conservative amino acid substitutions at position 97 of the CDR3 region (18).

This study analyzes the structure of the TCR β -chain to gain insight into the biologic basis of the infiltration into tissues by T cells in the DILS response to HIV-1 infection. Several alternative possibilities exist for mechanisms underlying the accumulation of lymphocytes, each of which have different predictions in terms of the character of the T cell repertoire as determined by the structure of the TCR β chains. In view of the increased frequency of B cell lymphoma in DILS (4), the lymphocytic accumulation in tissues could represent a parallel preneoplastic expansion of one or a few T cell clones which would be recognized by marked TCR β chain homogeneity at the level of nucleotide sequence. In contrast, a process driven by nominal antigen should exhibit some evidence of selection at the level of primary protein structure by demonstrating preferential usage of $V\beta$ and $J\beta$ elements, and especially, characteristic features of the CDR3 region. In the instance of superantigen drive, which has been suggested by two reports of HIV⁺ patients (12, 21), one would expect preferential deletion or enhancement of certain $V\beta$ families while allowing random structural features in the CDR3 region and the lack of any preferential usage of specific $J\beta$ elements. Since CD4 lymphocytes and monocytes which function as reservoirs of HIV-1 are systemically disseminated, this antigen or superantigen drive would be expected in both the involved tissue and elsewhere in the body. Finally, the process of tissue accumulation could be unrelated to any antigen drive or neoplastic event, perhaps reflecting secondary tissue recruitment of blood lymphocytes that have undergone activation as a result of locally secreted cytokines, and in this case, no commonality of primary structure would be evident in a comparison of TCR β chains from different individuals with DILS.

Current techniques assessing TCR structural features use either a limited set of monoclonal antibodies to measure $V\beta$ frequencies, or use potentially more informative molecular biologic techniques which may be divided into two general categories. In one, illustrated by construction of a formal cDNA library, the repertoire is determined by sequencing individual β chain clones to define CDR3 organization and obtain a frequency distribution of $V\beta$ and $J\beta$ usage by enumerating their occurrence. In the other, primers specific for each of the $V\beta$ segments are used in quantitative PCR methods which is relatively simple and has the important advantage that all known $V\beta$ gene segments can be measured (22). However, indeterminate variabilities in the efficiency of particular primers introduces uncertainties into the quantitation process and the method is presently incapable of determining $V\beta J\beta$ combinatorial frequencies or evaluating CDR3 structure. This report uses a novel approach which entails an initial stage of in situ cDNA synthesis performed on a frozen section, thereby eliminating any bias introduced by in vitro manipulation or differential priming, and is followed by amplification of cDNA by the anchored polymerase chain reaction (23) which has been

demonstrated to nonartificially amplify all expressed $V\beta$ segments (15, 24); the primary structure is obtained through nucleotide sequencing. This approach addresses the role of antigen specificity by infiltrating T cells in salivary gland tissue by concentrating on the presence of preferential $V\beta J\beta$ combinations as well as the amino acid structure of the third complementarity-determining region in TCR β -chains isolated directly from tissue.

Methods

Patients. Each individual with DILS met the diagnostic criteria of at least 6 mo of salivary gland enlargement, HIV-1 positivity confirmed by isolation of the virus and demonstration of specific antibodies by ELISA and Western blotting, elevated CD8 T cell levels in the peripheral blood, and at least two foci of lymphocytic infiltration with a predominance of CD8 lineage T cells demonstrated in lip biopsy specimens. The individual with primary Sjögren's syndrome had symptomatic xerostomia, more than two lymphocytic foci on minor salivary gland biopsy, and anti-Ro/SS-A autoantibodies. All tissue analyzed in this report was residual diagnostic lip biopsy material obtained with informed consent and under an approved protocol.

Methods. The method of in situ cDNA synthesis was according to a modification of the method of Tecott et al. (25). Briefly, all specimens were snap frozen with liquid nitrogen and stored until cut into 10- μ m sections. Each tissue section was hybridized for 12–16 h at room temperature with 50 μ l of a solution containing 50% formamide, 4 \times SSC, and a 36-bp primer (80 ng/ml) complementary to a sequence in the constant region of the TCR β -chain (CCCATTCACCCACCCAGC-TCAGCTCCACGTGGTCCGGG). After hybridization, sections were washed in 2 \times SSC at room temperature for 10 min for three washes, with a final wash in .5 \times SSC at 40°C for 2 h. cDNA was synthesized by adding to each tissue section a volume of 50 μ l which contained 10 U AMV reverse transcriptase (Promega Corp., Madison, WI; buffer supplied) and 1 mM dNTP's, and incubated at 42°C for 1 h. Each section was washed three times in 2 \times SSC at room temperature for 10 min, followed by a final wash at 40°C for 2 h. For each sample, at least five sections from each specimen were processed simultaneously according to this protocol. The cDNA was eluted from each tissue section with 0.15 M NaOH. .25 volume of 1 M Tris-HCl (pH = 6.0) was added to each specimen followed by ethanol precipitation. The cDNA was subsequently amplified by a modification of the anchored PCR procedure of Loh et al. (23). Specifically, the first strand cDNA product was G-tailed with terminal deoxynucleotidyl transferase (Promega Corp.; buffer supplied) using 1 mM dGTP for 30 min at 37°C, and then amplified using a .1- μ M 34-mer C-tailed primer which also has a restriction site for NotI (TTGACGCGGCCGCTACTAGCCCCCCCCCCCCCCC), a .8- μ M primer containing only the NotI site (TTGACGCGGCCGCTACTAGC), and a 1- μ M third primer complementary to the constant region, 5' of the original 36-mer (CCTGTGGCCAGGCAC-ACCAG). The PCR protocol using a Hybaid (Hook and Tucker Instruments, United Kingdom) thermal cyclor consisted of denaturation at 94°C for 1 min, annealing at 53°C for 2 min, and synthesis at 72°C for 3 min. The product was size-selected using a low melting agarose gel and 200–400-bp fragments were reamplified for 20 cycles (94°C for 1 min, 55°C for 2 min, and 72°C for 3 min) using a constant region primer 5' of the previously used constant region primer (GTGTGG-CCTTTTGGGTGTGG) while the other primer was the 20-mer NotI primer. The product was again size-selected on an agarose gel, digested with NotI, and ligated into Bluescript (Stratagene Inc., La Jolla, CA) which had been digested with NotI and SmaI. The β -galactosidase system was employed to select white colonies which were then screened by restriction digestion with SstI and BglII, which selectively cut a BglII restriction site present in the β -chain constant region just 5' of the last-used primer. All clones with inserts revealed by BglII and SstI digestion were sequenced using a DNA sequencer (373A; Applied Biosystems Inc., Foster City, CA) using the Taq dye deoxy terminator

sequencing kit (Applied Biosystems Inc.) in accordance with the manufacturer's protocol.

Statistics. Statistical analysis used Chi-square analysis or Fisher's exact test for which an odds ratio is given for each calculation. In each instance the null hypothesis for $V\beta$ or $J\beta$ frequency calculations was that there was an equal likelihood of the presence of each gene element.

Nomenclature. The $V\beta$ family designation is according to Wilson et al. (26) and Robinson (27); the β -chain amino acid position number is according to Kabat et al. (28). The length of the CDR3 regions is arbitrarily defined as the number of amino acid residues between the invariant cysteine at position 92 and the invariant phenylalanine at position 108 from which is subtracted the length of the shortest CDR3 region to give the absolute number of varying residues.

Results

The nucleotide sequence and inferred primary amino acid sequence of 62 TCR β -chains were obtained from lip biopsy specimens of five patients with HIV-associated DILS. Their length ranged from 100 to 250 bp and reflected extension from the constant region primer to well into the $V\beta$ segment. Each β chain had a different G-tailed terminus, indicating that they arose from separate priming events. The number of β -chains from each individual range from 10 to 17. 59 of the $V\beta$ segments sequenced could be unambiguously assigned to one of the 24 known $V\beta$ families and, with few exceptions, have sequence identity with known $V\beta$ alleles; the remaining 3 $V\beta$ sequences were too short to unequivocally assign to a specific $V\beta$ family, but were included in the analysis of $J\beta$ usage and CDR3 structure. The patterns of $V\beta$ and $J\beta$ segment usage as well as the organization of the structure of the CDR3 region was determined by enumeration of the β chain sequences. For reference purposes and to control for the possibility of intrinsic salivary gland specificity, 55 TCR β -chains derived from the lip biopsy of an individual with primary Sjögren's syndrome are also included in the analysis.

$V\beta$ Frequency distribution. Fig. 1 shows the frequency of $V\beta$ segment usage in the TCR of T cells infiltrating salivary tissue of each of the five samples from DILS as determined from the sequence obtained for each isolate. 15 of the 24 known $V\beta$ families are collectively expressed in this sample. Although relatively heterogeneous, the frequency distribution of $V\beta$ segments for the group shows a distribution biased towards the use of certain $V\beta$ families and is non-Gaussian ($\chi^2 = 63.65, P = 0.00001$). Six $V\beta$ families, $V\beta 2, 5, 6, 8, 12,$ and $13,$ are expressed at a level greater than one standard deviation above the median (median = 2.0, SD = 2.6) and collectively constitute 64% of the total $V\beta$ families expressed. With the exception of patient 3 where $V\beta 5$ was used in four of the 10 β -chains isolated from that individual, there is not a dominant usage of one specific $V\beta$ family by the entire group or by any of the other four patients. 8 of the 24 known $V\beta$ families were not encountered in any of the β chain sequences from the five samples.

Fig. 2 compares the frequency of $V\beta$ usage in the 59 β chains of the DILS group with 53 TCR β -chains obtained from the same anatomic site from one individual with primary Sjögren's syndrome. That 20 of the 24 $V\beta$ families are represented in the Sjögren's syndrome sample indicates the $V\beta$ repertoire diversity in this single person exceeds that found in the entire DILS group. Further emphasizing differences in $V\beta$ usage of these two diseases involving accumulation of lymphocytes in the same organ is the observation that there are seven

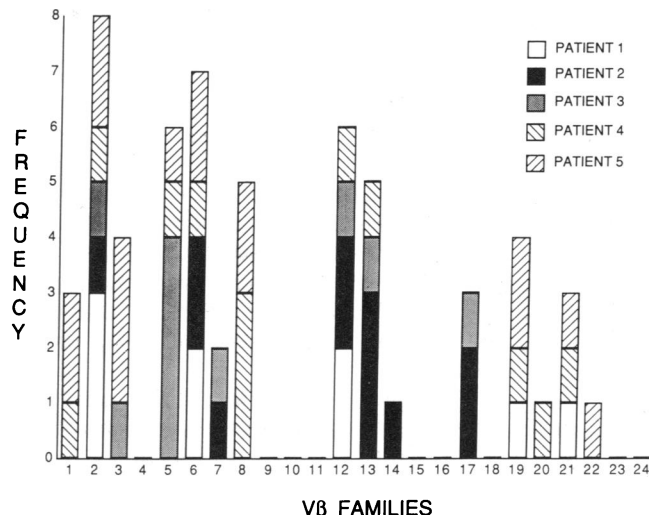


Figure 1. Frequency of TCR $V\beta$ families in DILS. 15 $V\beta$ families are expressed in 59 TCR β -chains expressed in the salivary glands of five individuals with DILS (median = 2.0, SD = 2.6). The distribution is nonrandom ($\chi^2 = 63.65, P = 0.00001$) and exhibits six $V\beta$ families ($V\beta 2, 5, 6, 8, 12,$ and 13) which are expressed at a level greater than one standard deviation above the median which collectively account for 64% of the total $V\beta$ distribution.

$V\beta$ families, $V\beta 4, 9, 10, 11, 15, 16,$ and $23,$ expressed in the Sjögren's syndrome sample that are not found in the sample of DILS tissue. The TCR repertoire observed in the Sjögren's syndrome sample contains only three $V\beta$ families, $V\beta 5, 6, 13,$ expressed at a level greater than one standard deviation above the median (median = 2.0, SD = 2.4), and these constitute 45% of the total $V\beta$ segment expression.

$J\beta$ Frequency distribution. The frequency of $J\beta$ segment expression for persons with DILS is shown in Fig. 3. All $J\beta$

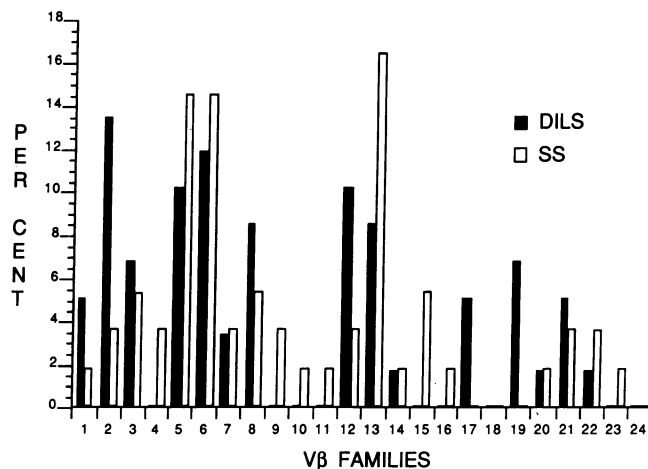


Figure 2. Comparative distribution of TCR $V\beta$ families in the salivary glands of DILS ($n = 59$) and Sjögren's syndrome ($n = 53$). One individual with Sjögren's syndrome expresses 20 $V\beta$ families which represents greater diversity than the collective $V\beta$ repertoire of five individuals with DILS; this includes seven $V\beta$ families ($V\beta 4, 9, 10, 11, 15, 16,$ and 23) not present in the $V\beta$ repertoire of any of the DILS patients. There are three $V\beta$ families ($V\beta 5, 6,$ and 13) in Sjögren's syndrome which are expressed at a level greater than one standard deviation greater than the median (median = 2.0, SD = 2.4) which collectively account for 45% of the total $V\beta$ distribution.

segments analyzed are identical to previously reported J β segment sequences. Of the 13 known J β segments, 10 are represented in the β -chain repertoire of the T cells in the salivary glands in DILS. The overall distribution is moderately heterogeneous (median = 4.0, SD = 4.5) in that there is not a singular dominance of one J β segment; however, there is a preferential skewing of J β segment usage, with four of the total of 13 J β segments, J β 1.1, 2.1, 2.5, and 2.7, found to be present at a level greater than one standard deviation above the median and collectively account for 70% of the expressed J β segments in glandular tissue ($\chi^2 = 48.2$, $P = 0.000003$). Three J β segments, J β 1.3, J β 2.4, and J β 2.6, are not represented in any of the 62 TCR β -chains analyzed.

Fig. 4 shows the frequency distribution of J β segment expression from the individual with Sjögren's syndrome as compared to the that among the five persons with DILS. There is considerable overlap between the two groups with regard to the preferential use of certain J β segments, e.g., J β 1.1, J β 2.1, and J β 2.7; however, selective differences are evident as shown by the significantly increased frequency of J β 2.5 usage in DILS ($P = 0.017$, odds ratio = 9.72) and a trend of enhanced usage of J β 2.3 in Sjögren's syndrome which approaches statistical significance ($P = 0.086$).

D β Frequency distribution. Analysis of D β segment usage reveals that there is not a significant difference in the usage of either of the two D β segments, with D β 1.1 found in 32 β -chains and D β 2.1 in 30 β -chains. One β -chain incorporated both D β 1.1 and D β 2.1 and 5 β -chain transcripts did not exhibit an identifiable D β segment. In addition, within each D β segment group, there was not a demonstrable preference for the use of one particular translational reading frame for either D β 1.1 or D β 2.1 (data not shown).

Inferred length of the CDR3 region. Fig. 5 illustrates the length of the CDR3 regions for the β -chains of DILS as well as primary Sjögren's syndrome, arbitrarily defined as the number

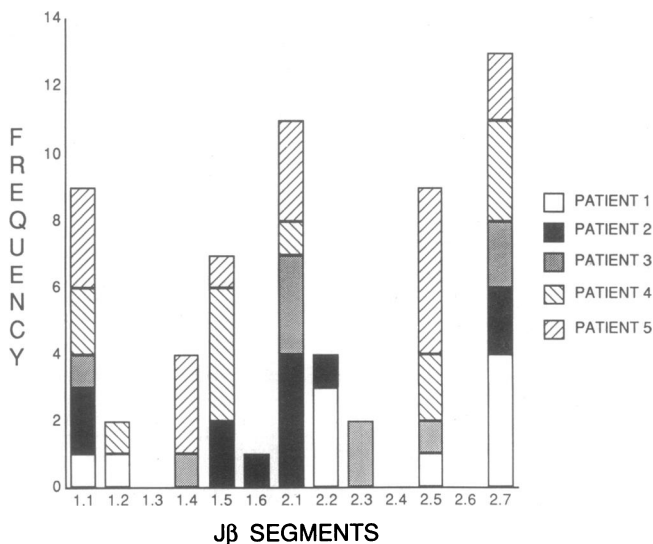


Figure 3. Frequency of TCR J β segments in DILS. 10 of the 13 J β segments are expressed in 62 TCR β -chains in the salivary glands of five individuals with DILS (median = 4.0, SD = 4.5) The distribution is nonrandom ($\chi^2 = 48.2$, $P = 0.000003$) with four J β segments (J β 1.1, 2.1, 2.5, and 2.7) expressed at a frequency greater than one standard deviation above the median which collectively account for 70% of the total distribution.

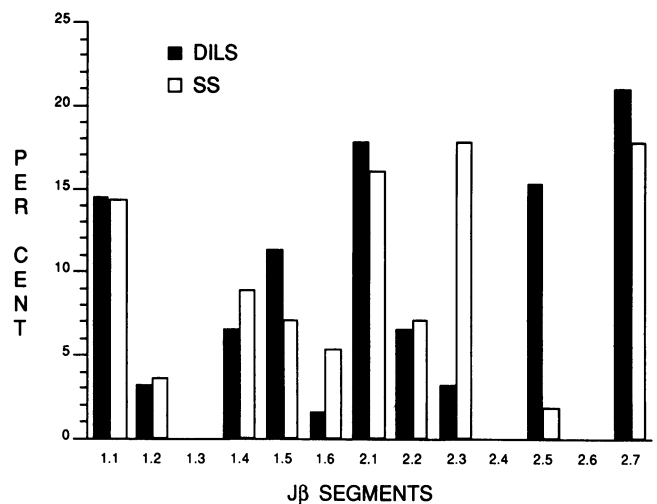


Figure 4. Comparative distribution of TCR J β segment expression in the salivary glands of DILS ($n = 62$) and Sjögren's syndrome ($n = 55$). There is considerable similarity in the skewness of J β segment usage in DILS and Sjögren's syndrome, with the only significant difference being the preferential usage of J β 2.5 by the DILS patients ($P = 0.017$).

of amino acid residues between the invariant cysteine at position 92 and the invariant phenylalanine at position 108 from which is subtracted the length of the shortest CDR3 region, 9 residues, to derive the number of inserted residues. The total number of amino acids in the CDR3 region of the DILS β -chains is significantly longer (mean = 12.4, SD = 1.72) and shows less variability as compared to the Sjögren's patient (mean = 11.7, SD = 1.93, $P = 0.04$). The CDR3 region of the DILS β -chains show more size restriction in that 73% are 11, 12, or 13 amino acids in length, or 2, 3, or 4 inserted residues, whereas the CDR3 regions of the individual with Sjögren's syndrome display more variability in size with a random distribution from 9 to 14 amino acids in length, or from zero to five inserted residues.

Association of V β with J β . Fig. 6 depicts the association that was found in a given β chain between the presence of a specific V β family and the occurrence of a particular J β segment in each of the 59 TCR β -chains from DILS tissue. In addition to the apparently background random combinatorial pairing of some V β families with individual J β segments, there is evidence for a non-random selective usage of particular V β J β combinations. For example, the combination of V β 5J β 2.1 is encountered in four separate β -chains derived from three different individuals ($P = 0.009$, OR = 13.14), accounting for 7% of all β -chains as against an expected frequency of 1.8%. The V β 6J β 2.7 combination is represented in the repertoire of four different persons ($P = 0.026$, OR = 7.33), and the combination of V β 13J β 1.1 is expressed in three separate β -chains ($P = 0.022$, OR = 12.00). For each of the V β segments expressed in these three particular V β J β combinations, the associated J β constitutes greater than 50% of the total J β segments associated with that particular V β family; V β 17J β 2.1 also fulfills this criterion but only approaches statistical significance for a V β J β combination ($P = 0.08$). The TCR β -chains containing the V β J β combinations V β 5J β 2.1, V β 6J β 2.7, V β 13J β 1.1, and V β 17J β 2.1, are designated "overrepresented" β -chains.

With the exception of two β -chains which represent examples of V β 5J β 2.1 combinations, the repertoire of Sjögren's syn-

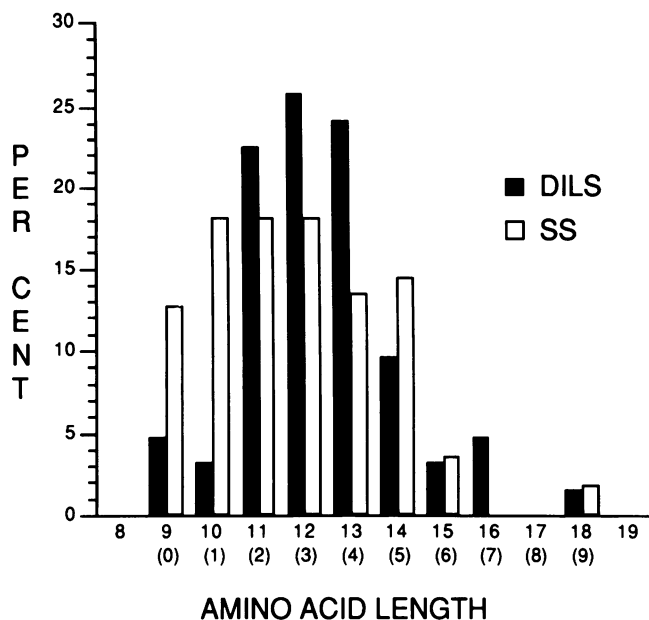


Figure 5. Amino acid length of the CDR3 region in individual TCR β -chains in DILS ($n = 62$) and Sjögren's syndrome ($n = 55$). The total length of the CDR3 region is defined as the number of amino acids between the invariant cysteine at position 92 and the invariant phenylalanine at position 108; the number of inserted amino acids, shown in parentheses, is determined by subtracting the number of amino acids in the shortest CDR3 region (e.g., 9 residues) from the total number of amino acids in each CDR3 region. The total number of amino acids in the CDR3 region of the TCR β chains of the patients with DILS (mean = 12.4, SD = 1.72) is significantly greater and shows less variability than that of Sjögren's syndrome (mean = 11.7, SD = 1.93, $P = 0.04$). 73% of CDR3 regions in DILS exhibit two to four inserted residues while the CDR3 regions of the Sjögren's syndrome patient demonstrate random variation between 0 and 5 inserted residues.

drome does not include any examples of expressed β chains which encode the overrepresented $V\beta J\beta$ combinations present in DILS (data not shown). Rather it demonstrates entirely different overrepresented $V\beta J\beta$ combinations which include $V\beta 5J\beta 2.7$ ($P = 0.042$, OR = 5.7), $V\beta 6J\beta 1.1$ ($P = 0.011$, OR = 10.75), and $V\beta 15J\beta 2.3$ ($P = 0.04$, OR = 18.8).

Association of particular $V\beta J\beta$ combinations with CDR3 structure. In view of the observation that β -chains with particular $V\beta J\beta$ combinations were found in salivary glands of patients with DILS at levels greater than expected by chance, the CDR3 regions of β -chains composed of these overrepresented combinations were compared to those where no preferential $V\beta J\beta$ combinations were observed. The primary structures of the CDR3 regions of selected β -chains present in these overrepresented $V\beta J\beta$ combinations are shown in Table I. Specifically with regard to amino acid position 97, it can be seen that there is a strong relationship between the use of specific $V\beta J\beta$ combinations and the identity of the amino acid at position 97. For example, in the combination $V\beta 13J\beta 1.1$ β -chains (Table I A), serine is present at position 97 in two of the three isolates ($P = 0.01$, OR = 54) and, as indicated by Fig. 7 which shows the total frequency of individual amino acids at position 97, this accounts for two of the four serine residues used at this position in the entire sample of 62 β -chains. Similarly, glutamine accounts for two of the four residues at position 97 in β chains

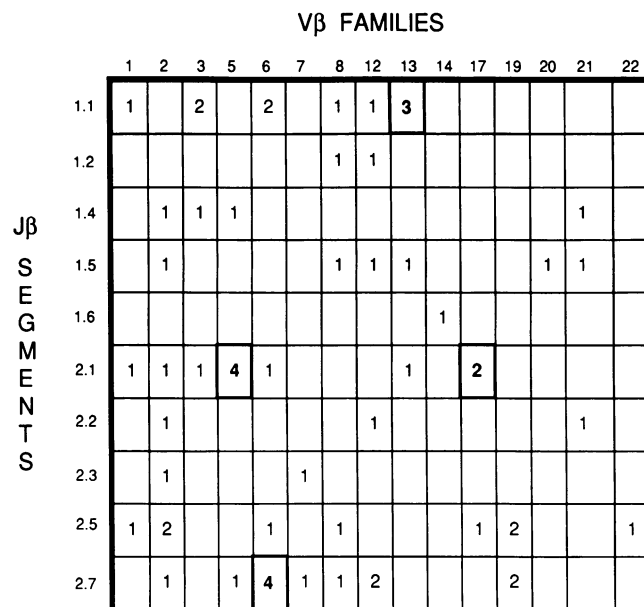


Figure 6. Coexpression of $V\beta$ families with specific $J\beta$ segments in individual TCR β chains in the salivary glands of DILS. There is a preferential association of certain $V\beta$ families with specific $J\beta$ segments as demonstrated by the pairings of $V\beta 5J\beta 2.1$ ($P = 0.009$, OR = 13.14), $V\beta 6J\beta 2.7$ ($P = 0.026$, OR = 7.33), and $V\beta 13J\beta 1.1$ ($P = 0.022$, OR = 12.00).

with the combination $V\beta 5J\beta 2.1$ (Table I B; $P = 0.01$, OR = 54), and in the case of the $V\beta 6J\beta 2.7$ combination (Table I C), aspartic acid is present at position 97 in two of the four β -chains ($P = 0.01$, OR = 54), again accounting for two of the four occurrences of this amino acid at this position. A further example of this association is found in the $V\beta 17$ family where two of the three isolates are paired with $J\beta 2.1$. Table I shows that both of these $V\beta 17J\beta 2.1$ chains are encoded with a glutamic acid at position 97 ($P = 0.006$). In addition, emphasizing the importance of similarity in residues found at position 97 is the presence of identical amino acids adjacent to position 97 in two of the four $V\beta J\beta$ categories, e.g., serine residues at position 98 in the β chain with the $V\beta 6J\beta 2.7$ combination and tyrosine residues at position 96 in the $V\beta 13J\beta 1.1$ combination. Moreover, it should be noted that with the exception of $V\beta 17J\beta 2.1$, where both of the β -chains are derived from the same individual, all the other $V\beta J\beta$ pairings are contributed by different individuals.

The significance of the relationship between $V\beta J\beta$ and amino acid residue at position 97 is further enhanced by the observation that, in most of the β -chains, the amino acids found at and around position 97 are encoded by somatically added nucleotides, rather than the individual V and J segment germline elements used elsewhere in the chain. The somatically added nucleotides are depicted in lower case letters. As can be seen in Table I different codons encode the same amino acid at these positions, suggesting that factors operating at the level of the expressed protein rather than at the level of DNA are involved in this nonrandom representation of amino acid residues. The 53 β -chains expressed in the salivary tissue of the one individual with primary Sjögren's syndrome do not exhibit a significant association between any $V\beta J\beta$ combination and a particular amino acid residue at position 97, or at any adjacent position in the CDR3 region (data not shown).

Table 1. Overrepresented VβJβ Combinations Encode Identical Amino Acids at Position 97 of the Third Complementarity-determining Region

Amino acid position:	92	97	108	Patient	P value*
A	Vβ13--TGT	C A S S GCC AGC AGT Y S A G P L Tct tct gcc GGC ccg ctc	E A F F GAA GCT TTC TTT--Jβ1.1	3	0.01
	Vβ13--TGT	C A S S GCC AGC AGT Y S S N T Tat tcc AGt AAC ACT	E A F F GAA GCT TTC TTT--Jβ1.1	2	
B	Vβ5 -- TGC	C A S S GCC AGC AGC L Q G L G S TTA CAG GGG ctg ggg agc	N E Q F AAT GAG CAG TTC TTC--Jβ2.1	5	0.01
	Vβ5 -- TGT	C A S S GCC AGC AGC Q Q T G G S D caa cag ACA GGG GGa tca gAC	N E Q F AAT GAG CAG TTC TTC--Jβ2.1	4	
C	Vβ6 -- TGT	C A S S GCC AGC AGC S D S G R L Tca gaT AGC GGG AGa ctc	H E Q Y cac gat CAG TAC TTC--Jβ2.7	5	0.01
	Vβ6 -- TGT	C A S S GCC AGC AGC R D S T CGG GAc agt aCC	Y E Q Y TAC GAG CAG TAC TTC--Jβ2.7	4	
D	Vβ17--TGT	C A S R GCC AGT AGa K E G V aaa gAG GGT gtc	D E Q F gAT GAG CAG TTC TTC--Jβ2.1	2	0.02
	Vβ17--TGT	C A S S GCC AGT AGC L E F TTa gaa ttc	D E Q F gAT GAG CAG TTC TTC--Jβ2.1	2	

Large letters designate germline nucleotides, small letters are non-germline nucleotides, and underlined segments are Dβ segments. * Fisher exact test for the association of VβJβ combinations with amino acid residue at position 97.

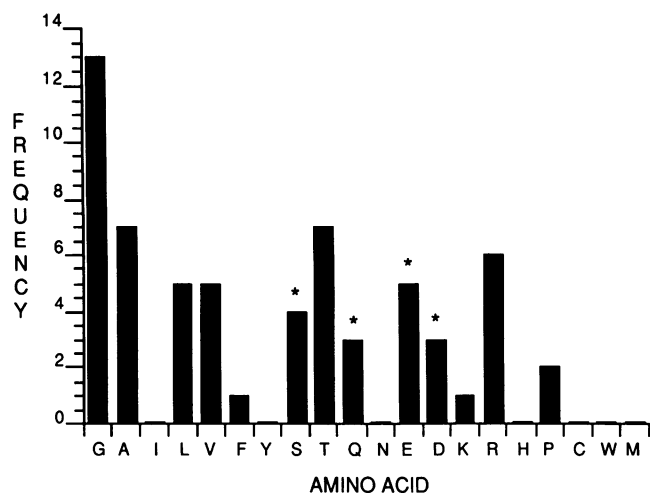


Figure 7. Distribution of amino acid residues used at position 97 in individual TCR β -chains in DILS. *Amino acids represented in preferential $V\beta J\beta$ pairings. (A, Ala, C, Cys, D, Asp, E, Glu, F, Phe, G, Gly, H, His, I, Ile, K, Lys, L, Leu, M, Met, N, Asn, P, Pro, Q, Gln, R, Arg, S, Ser, T, Thr, V, Val, W, Trp, Y, Tyr).

Discussion

The central finding of this study, using direct enumeration of the primary structural elements of each isolated TCR β chain, is that a subpopulation of T cells infiltrating the minor salivary glands of HIV⁺ individuals with DILS express a restricted β -chain TCR repertoire. This repertoire exhibits structural characteristics consistent with what would be predicted if the clonal selection of these T cells occurred through an antigen driven mechanism. The findings that support an antigen driven selection event include: the selective overrepresentation of particular $V\beta J\beta$ combinations, the presence of identical amino acid residues at position 97 in β chains included in these same $V\beta J\beta$ combinations even if the chains are derived from different individuals, and that the amino acids at this position are specified by different codons which are not specified by the germline. As a counterpoise to these nonrandom events is the observation that there is no evidence of bias in the use of the $D\beta$ elements which adds further support to the considerable evidence in favor of selective usage of the other TCR gene elements. That these β -chain structural features are specific to the immune response of DILS is supported by the finding that three of the four overrepresented $V\beta J\beta$ combinations of the DILS repertoire are not represented in any of the β -chains of the person with Sjögren's syndrome. More generally, a recent study of 250 TCR β -chains from five healthy individuals, including a pair of identical twins, did not demonstrate a preferential association between any particular $V\beta$ family and $J\beta$ gene segment in individual TCR β chains (15). These observations argue against the presence of biasing structural features, either of the genome or of the mature protein product, as being responsible for the selective $V\beta J\beta$ combinations present in DILS. Taken together with the fact that DILS is a distinctive host response to HIV-1 infection which occurs in the setting of particular MHC class I and II alleles, the present observations suggest the possibility that the unique features of the response of DILS are mediated by the recognition properties of these TCR β chains for certain HIV-encoded peptides.

The identification of a subset of β -chains characterized by overrepresented combinations of $V\beta J\beta$ segments allowed an

approach to the problem of identifying TCR's from clones that might have specifically responded to an antigen as opposed to TCR's from other clones that reflect nonspecific recruitment mechanisms. The fact that analysis of the CDR3 regions of these two categories of β -chains revealed a highly nonrandom occurrence of particular amino acids at position 97 supports the rationale of this approach. For example, each of the $V\beta J\beta$ chains from the overrepresented groups in DILS (e.g., $V\beta 5J\beta 2.1$, $V\beta 6J\beta 2.7$, $V\beta 13J\beta 1.1$, and $V\beta 17J\beta 2.1$) demonstrates a selectivity for specific amino acids at position 97 of the CDR3 region, which is the same structural residue implicated in direct interaction with the antigenic peptide in several *in vitro* and transgenic investigations (18–20). The significance of this association is enhanced by the observation that in each of the four preferred $V\beta J\beta$ classes the common amino acid is encoded by nongermline nucleotides. The existence of several structurally distinct overrepresented $V\beta J\beta$ categories suggests the possible existence of several antigenic peptides responsible for eliciting the immune response in the salivary glands of DILS.

Most studies examining either tissue-specific or peripheral $V\beta$ repertoires have demonstrated elements of nonrandomness and while the $V\beta$ repertoire of DILS identifies six of 24 $V\beta$ families ($V\beta 2$, 5, 6, 8, 12, and 13) as accounting for 64% of the total $V\beta$ segments isolated, the principal finding here is that the $V\beta$ repertoire of each individual with DILS is more similar to other individuals with DILS than it is to the individual with Sjögren's syndrome. For example, the one individual with Sjögren's syndrome exhibits greater diversity in expressing 20 of 24 $V\beta$ families than the collective repertoire of five individuals with DILS which only express 15 $V\beta$ families. Moreover, seven of the $V\beta$ families expressed in the Sjögren's syndrome patient's repertoire are absent from any of the DILS repertoires. Additionally, it should be noted that the data from the individual with Sjögren's syndrome are in general agreement with a recent study which found a preferential usage of $V\beta 2$ and $V\beta 13$ from lip biopsies in primary Sjögren's syndrome (29). In the present study $V\beta 13$ is the most frequently used $V\beta$ family, accounting for 17% of the $V\beta$ repertoire; however, $V\beta 2$ comprises only 3.7% of the total repertoire. Both DILS and Sjögren's syndrome exhibit considerable overlap in expressing skewed $J\beta$ segment usage where both diseases emphasize only 4 of 13 $J\beta$ segments in accounting for 70% of the repertoire.

While a superantigen mechanism has been postulated as an explanation for distortions in the frequency of $V\beta$ family representation in AIDS (12), it is unlikely that the present findings of skewed $V\beta$ expression in DILS are a consequence of exposure to an HIV-related superantigen; such a mechanism fails to provide an explanation for the selectivity of the $J\beta$ segments and the similarities in other features of CDR3 structure that were found in the present study. Another possible explanation for $V\beta$ gene segment restriction could be selective trafficking into the salivary gland of T cells expressing particular $V\beta$ families. Comparative analysis with primary Sjögren's syndrome, used here principally as a methodologic and biologic control, suggests that the $V\beta$ bias in DILS is not simply the result of salivary gland tropism of a select subset of T cells since the two diseases display $V\beta$ repertoires that are only partially overlapping.

The inclusion of only 8 of 59 or 13% of overrepresented TCR β chains as representative of an antigen-reactive group is consistent with previous analyses investigating the frequency of

antigen-specific T cells in pathologic lesions. For example, Bloom and colleagues demonstrated that only 2% of infiltrating T cells in cutaneous tuberculoid leprosy lesions responded in an antigen-specific manner to *Mycobacterium leprae* (30), and the relevance of these data to the current study is enhanced by the recent finding of a subset of peptide-specific human T cell clones responsive to *M. leprae* that also use TCR β -chains within a single V β family and also encode identical amino acids at position 97 of the CDR3 region (31). Similarly, in studying the antigen-induced animal model of autoimmunity, experimental autoimmune encephalomyelitis, Cohen and colleagues recently reported that only 5% of T cells derived from pathologic lesions were reactive to the stimulating antigen myelin basic protein (32). The estimation of an even higher frequency of antigen-specific T cells in DILS salivary tissue may be partially explained by the findings of several recent studies which have consistently shown that HIV⁺ individuals demonstrate markedly elevated numbers of HIV-reactive T cells even in the peripheral blood (33–35); presumably this is partly a consequence of the systemic distribution of HIV-1 coupled with the anatomical localization of the infection to the immune system itself. In fact, despite elements of organ specificity such as that exhibited in DILS, the disseminated nature of the infection predicts that many of the same structurally selective TCR's could be isolated from the peripheral blood as those isolated from the salivary gland, especially in consideration of the uniquely prominent peripheral CD8 lymphocytosis in this HIV⁺ group.

Acknowledgments

We thank Ms. Modrie Payne for excellent assistance in preparation of the manuscript.

This work was funded in part by a grant from the New York chapter of the Arthritis Foundation (E. Dwyer), a Postdoctoral Fellowship from the Arthritis Foundation (S. Itescu), National Institutes of Health grants AI19411 and AR39626, and American Foundation for AIDS Research grant 500255-12-PG.

References

- Itescu, S., L. J. Brancato, J. Buxbaum, P. K. Gregersen, C. C. Rizk, S. Croxson, G. Solomon, and R. Winchester. 1990. A diffuse infiltrative CD8 lymphocytosis syndrome in human immunodeficiency virus (HIV) infection: a host immune response associated with HLA-DR5. *Ann. Intern. Med.* 112:3–10.
- Itescu, S., J. Dalton, H. Zhang, and R. Winchester. 1993. Tissue infiltration in a CD8 lymphocytosis syndrome associated with HIV-1 infection has the phenotype of an antigen driven process. *J. Clin. Invest.* 91:2216–2225.
- Itescu, S., E. Dwyer, S. Choi, H. Zhang, Y. Huang, A. Friedman-Kien, and R. Winchester. 1992. The sicca syndrome in HIV-1 infection is dependent on the host immune response rather than the infecting viral strain. *Arthritis Rheum.* 35:S210. (Abstr.)
- Itescu, S., U. Mathur-Wagh, M. L. Skovron, L. J. Brancato, M. Marmor, A. Zeleniuch-Jacquotte, and R. Winchester. 1992. HLA-B35 is associated with accelerated progression to AIDS. *J. Acquired Immune Defic. Syndr.* 5:37–45.
- Fox, R. I., F. V. Howell, R. C. Bone, and P. Michelson. 1984. Primary Sjögren syndrome: clinical and immunopathologic features. *Semin. Arthritis Rheum.* 14:77–105.
- Matis, L. A. 1990. The molecular basis of T-cell specificity. *Annu. Rev. Immunol.* 8:65–82.
- Kronenberg, M., G. Siu, L. E. Hood, and N. Srastrri. 1986. The molecular genetics of the T-cell antigen receptor and T-cell antigen recognition. *Annu. Rev. Immunol.* 4:529–591.
- Howell, M. D., J. P. Diveley, K. A. Lundeen, A. Esty, S. T. Winters, D. J. Carlo, and S. W. Brostoff. 1991. Limited T-cell receptor beta-chain heterogeneity among interleukin 2 receptor-positive synovial T cells suggests a role for superantigen in rheumatoid arthritis. *Proc. Natl. Acad. Sci. USA.* 88:10921–10925.
- Paliard, X., S. G. West, J. A. Lafferty, J. R. Clements, J. W. Kappler, P. Marrack, and B. L. Kotzin. 1991. Evidence for the effects of a superantigen in rheumatoid arthritis. *Science (Wash. DC).* 253:325–329.

- Bucht, A., J. R. Oksenberg, S. Lindblad, A. Gronberg, L. Steinman, and L. Klareskog. 1992. Characterization of T-cell receptor alpha beta repertoire in synovial tissue from different temporal phases of rheumatoid arthritis. *Scand. J. Immunol.* 35:159–165.
- Kotzin, B. L., S. Karuturi, Y. K. Chou, J. Lafferty, J. M. Forrester, M. Better, G. E. Nedwin, H. Offner, and A. A. Vandenbark. 1991. Preferential T-cell receptor beta-chain variable gene use in myelin basic protein-reactive T-cell clones from patients with multiple sclerosis. *Proc. Natl. Acad. Sci. USA.* 88:9161–9165.
- Imberti, L., A. Sottini, A. Bettinardi, M. Puoti, and D. Primi. 1991. Selective depletion in HIV infection of T cells that bear specific T cell receptor V beta sequences. *Science (Wash. DC).* 254:860–862.
- Williams, W. V., Q. Fang, D. Demarco, J. VonFeldt, R. B. Zurier, and D. B. Weiner. 1992. Restricted heterogeneity of T cell receptor transcripts in rheumatoid synovium. *J. Clin. Invest.* 90:326–333.
- Davey, M. P., M. M. Meyer, D. D. Munkirs, D. Babcock, M. P. Braun, J. B. Hayden, and A. C. Bakke. 1991. T-cell receptor variable beta genes show differential expression in CD4 and CD8 T cells. *Hum. Immunol.* 32:194–202.
- Rosenberg, W. M. C., P. A. H. Moss, and J. I. Bell. 1992. Variation in human T-cell receptor VB and JB repertoire: analysis using anchor polymerase chain reaction. *Eur. J. Immunol.* 22:541–549.
- Chothia, C., D. R. Boswell, and A. M. Lesk. 1988. The outline structure of the T-cell alpha beta receptor. *EMBO (Eur. Mol. Biol. Organ.) J.* 7:3745–3755.
- Davis, M. M., and P. J. Bjorkman. 1988. T-cell antigen receptor genes and T-cell recognition. *Nature (Lond.)* 334:395–402.
- Jorgensen, J. L., U. Esser, B. F. De StGroth, P. A. Reay, and M. M. Davis. 1992. Mapping T-cell receptor-peptide contacts by variant peptide immunization of single-chain transgenics. *Nature (Lond.)* 355:224–230.
- Engel, I., and S. M. Hedrick. 1988. Site-directed mutations in the VDJ junctional region of a T cell receptor beta chain cause changes in antigenic peptide recognition. *Cell.* 54:473–484.
- Hedrick, S. M., I. Engel, D. L. McElligott, P. J. Fink, M. L. Hsu, D. Hansburg, and L. A. Matis. 1988. Selection of amino acid sequences in the beta chain of the T cell antigen receptor. *Science (Wash. DC).* 239:1541–1544.
- Laurence, J., A. S. Hodssev, and D. N. Posnett. 1992. Superantigen implicated in dependence of HIV-1 replication in T-cells on TCR V Beta expression. *Nature (Lond.)* 358:255–259.
- Choi, Y. W., B. Kotzin, L. Herron, J. Callahan, P. Marrack, and J. Kappler. 1989. Interaction of *Staphylococcus aureus* toxin "superantigens" with human T cells. *Proc. Natl. Acad. Sci. USA.* 86:8941–8945.
- Loh, E. Y., J. F. Elliott, S. Cwiria, L. L. Lanier, and M. M. Davis. 1989. Polymerase chain reaction with single-sided specificity: analysis of T cell receptor delta chain. *Science (Wash. DC).* 243:217–220.
- Plaza, A., D. H. Kono, and A. N. Theofilopoulos. 1991. New human VB genes and polymorphic variants. *J. Immunol.* 147:4360–4365.
- Tecott, L. H., J. D. Barchas, and J. H. Eberwine. 1988. In situ transcription: specific synthesis of complementary DNA in fixed tissue sections. *Science (Wash. DC).* 240:1661–1664.
- Wilson, R. K., E. Lai, P. Concannon, R. K. Barth, and L. E. Hood. 1988. Structure, organization and polymorphism of murine and human T-cell receptor alpha and beta chain gene families. *Immunol. Rev.* 101:149–172.
- Robinson, M. A. 1991. The human T cell receptor beta-chain gene complex contains at least 57 variable gene segments. Identification of six V beta genes in four new gene families. *J. Immunol.* 146:4392–4397.
- Kabat, E. A., T. T. Wu, H. M. Perry, K. S. Gottesman, and C. Foeller. 1991. U.S. Dept. of Health and Human Services, NIH, Bethesda, MD.
- Sumida, T., F. Yonaha, T. Maeda, E. Tanabe, T. Koike, H. Tomioka, and S. Yoshida. 1992. T cell receptor repertoire of infiltrating T cells in lips of Sjögren's syndrome patients. *J. Clin. Invest.* 89:681–685.
- Modlin, R. L., J. Melancon-Kaplan, S. M. M. Young, C. Pirmex, H. King, J. Convit, T. H. Rea, and B. R. Bloom. 1988. Learning from lesions: patterns of tissue inflammation in leprosy. *Proc. Natl. Acad. Sci. USA.* 85:1213–1217.
- van Schooten, W. C. A., J. L. Ko, N. van der Stoep, J. B. A. G. Haanen, L. Pickering, R. R. P. de Vries, and B. R. Bloom. 1992. T-cell receptor B-chain gene usage in the T-cell recognition of *Mycobacterium leprae* antigens in one tuberculoid leprosy patient. *Proc. Natl. Acad. Sci. USA.* 89:11244–11248.
- Mor, F., and I. R. Cohen. 1992. T cells in the lesion of experimental autoimmune encephalomyelitis. *J. Clin. Invest.* 90:2447–2455.
- Clerici, M., D. R. Lucey, R. A. Zajac, R. N. Boswell, H. M. Gebel, H. Takahashi, J. A. Berzofsky, and G. M. Shearer. 1991. Detection of cytotoxic T lymphocytes specific for synthetic peptides of gp160 in HIV-seropositive individuals. *J. Immunol.* 146:2214–2219.
- Lieberman, J., J. A. Fabry, M. C. Kuo, P. Earl, B. Moss, and P. R. Skolnik. 1992. Cytotoxic T lymphocytes from HIV-1 seropositive individuals recognize immunodominant epitopes in Gp 160 and reverse transcriptase. *J. Immunol.* 148:2738–2747.
- Hadida, F., A. Parrot, M. P. Kieny, B. Sadat Sowti, C. Mayaud, P. Debre, and B. Autran. 1992. Carboxyl-terminal and central regions of human immunodeficiency virus-1 NEF recognized by cytotoxic T lymphocytes from lymphoid organs. An in vitro limiting dilution analysis. *J. Clin. Invest.* 89:53–60.