Structure-Function Relationships among Highly Diverse T Cells that Recognize a Determinant from Influenza Virus Hemagglutinin

By Alexander H. Taylor, Ann M. Haberman, Walter Gerhard, and Andrew J. Caton

From the Wistar Institute of Anatomy and Biology, Philadelphia, Pennsylvania 19104

Summary

We have analyzed the structural and genetic basis for T cell recognition of the complex formed between antigen and class II products of the major histocompatibility complex by performing sequence analysis of T cell receptors (TCRs) induced in response to the helper T cell site 1 of the influenza virus hemagglutinin. The results demonstrate, first, that structurally highly diverse TCRs can be utilized in recognition of the same antigen/1-Ed complex: 12 of 13 TCRs utilize unique Vα/Vβ gene segment combinations, suggesting that ~70 different Vα/Vβ combinations are available to BALB/c mice in response to this determinant. Second, comparison of these sequences with the ability of each hybridoma to recognize a panel of peptide analogues suggests that α and β chains of these TCRs frequently determine specificity for the NH2-terminal and the COOH-terminal portions, respectively, of the site 1 determinant.

The specificity of T cell responses is mediated through the clonally distributed, surface membrane-bound α/β heterodimer of the TCR, and is directed toward recognition of complexes formed between partially degraded protein antigens and self-MHC molecules (1, 2). Molecular characterization has identified three basic mechanisms, termed germline, combinatorial, and junctional diversity, by which TCRs with diverse specificities and unique structures can be generated (reviewed in references 3–5). Previous structural analyses of TCRs generated in response to defined antigen/class II MHC complexes have generally identified a limited number of different V region/germ-line gene segments among TCRs with the same specificity (6–14). Since TCRs that utilized the same V gene segments in recognition of a common determinant often express very similar junctional sequences, these studies have suggested an important role for particular junctional sequences in the formation of specificity for individual antigen/class II MHC complexes (9, 10, 15, 16). However, because of the structurally restricted nature of these responses, how the V gene repertoire contributes to the specificity of TCRs for different antigen/class II MHC complexes has been unclear.

Previous studies of the antigenicity of the influenza virus A/PR/8/34 (PR8)1 hemagglutinin (HA) have identified several immunodominant Th determinants recognized by BALB/c mice (17). One such determinant, termed site 1, comprises amino acids 110–120 of the HA1 subunit, and elicits Th responses that are I-Ed restricted (17–19). Each site 1–specific Th hybridoma analyzed thus far has been found to exhibit a distinct pattern of fine specificity for a panel of 20 different analogue peptides of the site 1 determinant, demonstrating that there exists extensive functional diversity in the Th response to the site 1/I-Ed complex (20). We wished to investigate how the germ-line gene segments and somatically assembled junctional regions interact to form TCRs with these distinct specificities for the site 1 determinant. Sequence analysis indicates that a substantial proportion of the available germline V and J gene repertoires are expressed by TCRs that recognize the site 1/I-Ed complex. Comparison of the TCR sequences with the ability of the hybridomas to recognize the panel of peptide analogues demonstrates that the expression of different gene elements forms the basis for the functional diversity of the response to the site 1 determinant. This comparison further suggests that the TCR α chain and β chain play important roles in recognition of the NH2-terminal and COOH-terminal portions, respectively, of the site 1 determinant.

Materials and Methods

Nucleotide Sequence Analysis. The hybridomas were generated by a variety of protocols after immunization either with whole virus or purified HA protein as described previously (17–21). The following pairs of oligonucleotide primers, which are complementary to adjacent positions at the 5' end of α and β chain C region mRNA, were used to direct sequence analysis of the hybridoma

1 Abbreviations used in this paper: HA, hemagglutinin; PR8, influenza virus A/PR/8/34.
An Influenza Antigen/MHC Complex Induces Diverse T Cell Receptors

P148.1 2011

Published December 1, 1990
Figure 1. Legend appears on following page.
TCR, variable regions: Cα1, 5'dATTTTTTTACTGGTACGC'3', Cα2, 5'dGTTCTGGCTCTGGAGCCC'3'. In some cases, sequence analysis was performed using a modification of the PCR (22) as follows. Primers Cα2 and Cβ2 were 5' end labeled with γ[32P]-ATP and used to prime the synthesis from 10 μg of poly(A)-selected hybridoma RNA of full-length cDNA copies of α and β chain V regions, which were synthesized and isolated from 5% polyacrylamide-urea gels as described (23). Approximately 10-20 G residues were added to the 3' end of the cDNAs by incubation with terminal transferase (24), and the products were purified by phenol extraction and ethanol precipitation. Tail DNA was annealed to 1 ng of the oligonucleotide 5'GATCATGGTCCTAC-ATTGGC3' by heating to 65°C followed by slow cooling to 10°C in a 20 μl volume of 20 mM NaCl, 10 mM Tris- HCl, 1 mM EDTA, pH 7.5. Second-strand synthesis was carried out by the addition of 1 μl 0.2 M MgCl2, 2 μl 0.1 M dithiothreitol, 2 μl 0.5 mM each dNTP and 1 μl Sequenase (United States Biochemical Corp., Cleveland, OH) followed by incubation at room temperature and then 37°C for 15 min each. Reaction mixtures were boiled, 100 ng each of the oligonucleotide 5'dGATCATGGTCCTACTATTGGC'3' and either the Cα1 or Cβ1 primer was added, and reactions were diluted to 100 μl and subjected to 30 cycles of amplification using Taq DNA polymerase (Perkin Elmer Cetus, Norwalk, CT) under standard conditions (22). Full-length reaction products were isolated by electrophoresis in a 3% Nusieve GTG (FMC Bioproducts, Rockland, ME) gel and purified by phenol extraction and ethanol precipitation. 1/10 of the product was then subjected to 10 further rounds of amplification under the same conditions, and the reaction products were once more gel purified. A quarter of this material was then used for dsDNA sequence analysis using Sequenase and either the Cα1 or Cβ1 oligonucleotide to prime synthesis. In cases where a mixed sequence that included the α or β chain V region sequence of the BWS147 fusion partner was obtained, the reaction products obtained after the first round of amplification were digested with a restriction enzyme (usually PstI or PvuII) that cleaves the BWS147-specific dsDNA but does not cleave the T cell-specific DNA. This prevented reamplification of the BWS147-specific DNA and allowed a single sequence to be obtained. The remaining TCR sequences were determined after cloning of the V region sequences in λ GT11. The Cα2 and Cβ2 primers were phosphorylated and used to prime the synthesis of cDNA from 10 μg of poly(A)-selected hybridoma RNA. This was converted to dsDNA and packaged into λ GT11 using standard methods (25). Replicate filters were screened using 32P-labeled Cα1 and Cβ1 oligonucleotides, and also with the oligonucleotides BWα5'dGCCCTAATTGGCTACTG3' and BWB5'dGTTTACTAGTTATTCTGCCC'T3' which identified clones derived from the α and β V regions of the fusion partner.

Lymphokine Secretion Assay. Lymphokine secretion assays were performed as described previously (20). Briefly, 2 x 104 T cell hybridomas/well were cultured with 4 x 105 irradiated syngeneic splenocytes/well (2,200 rad) and various concentrations of synthetic peptides (0.00025-25 μM) in a total volume of 0.2 ml. These cultures were incubated in flat-bottomed microtiter plates at 37°C, 7% CO2. Supernatants were harvested after 48 h, and the presence of IL-3 was determined by bioassay using the IL-3-dependent cell line DA-1. DA-1 viability was determined by a colorimetric assay using the dye 3-(4,5-dimethylthiazol-2-yl)-2,5-diphenyltetrazolium bromide, as originally described by Mosmann (26). For every T cell, the concentration of each analogue required for 50% of the maximal stimulation achieved with the parental site 1 peptide (SFEREFIPKPE) was determined from the dose-response curves. These values were then normalized against that obtained with the parental site 1 peptide to give the relative potency of the peptides in stimulating individual T cells.

Results and Discussion. Sequence analysis of TCR α and β V regions from 13 site 1-specific hybridomas was performed after either molecular cloning into λ GT11 (25) or one-sided amplification in the PCR (27-29). Several of the Vα and Vβ chains were analyzed by both methods, and in all such cases, identical sequences were obtained. The nucleotide and deduced amino acid sequences of the site 1-specific hybridoma Vα and Vβ chains are shown in Fig. 1 A and B, respectively. Where possible, Vα, Jα, Vβ, and Jβ sequences have been assigned to previously described families. The TCR gene segment usage, α and β chain junctional amino acid sequences, and site 1 analogue fine specificity analysis are summarized in Fig. 2.

A large number of different Vα, Vβ, Jα, and Jβ gene segments were identified among the site 1-specific hybridomas. 10 distinct Vα gene segments drawn from six different families and five different Vβ gene segments drawn from four different families were found. Similarly, 11 different Jα gene segments and seven different Jβ gene segments were isolated. These different gene segments were used in many unique combinations. For example, the Vβ8.3 gene segment was used in conjunction with six different Vα genes, as well as with four different Jβ gene segments, and each of the Vα4 family members associated with a different Vβ gene segment. With the exception of 7/6AH1 and 1E10 (both of which express Vα10b and Vβ8.3), each of the site 1-specific hybridomas expresses a unique combination of Vα and Vβ gene segments. Despite this overall diversity, particular α and β chain gene segments and gene segment combinations were repetitively observed. For example, the Vβ8.3/Jβ13 combination was expressed by 4 of the 13 hybridomas (Fig. 2). Three members...
of the Vβ8.3/Jβ1.3-expressing group (7/6AH1, 1E10, and LD1) display identical entire β chains, including junctionally encoded amino acid residues. The fourth member, PID3A6, expresses a β chain that differs only by two conservative junctional amino acids. These examples of identical, and in one case highly related, Vβ chains include junctional amino acids that must, in part, be encoded by N nucleotide addition, suggesting an important role for the β chain junctional sequences in the formation of specificity for the site 1/Eβ complex. The coselection of specific junctional sequences in the context of common V gene elements has previously been observed in the responses to cytochrome c (15, 16) and myelin basic protein (9, 10). Because of the large number of different gene segment combinations that were utilized in the site 1 response, however, the repetitive isolation of specific junctional sequences is not a prominent feature of formation of specificity for this determinant.

The structural diversity of these TCRs is reflected by the unique pattern of reactivity that each hybridoma displays toward peptide analogues of the site 1 determinant (Fig. 2). Among the hybridomas from the Vβ8.3/Jβ1.3 group that express identical (or in one case, very similar) β chains, the differences in specificity for the peptide analogues occur because of differences in the TCR α chains. Moreover, since these hybridomas differ in their specificity only for analogues at positions 110–113, the results demonstrate that the α chain is making important contacts with the NH2-terminal portion of the site 1 determinant. It is, furthermore, interesting that the identical specificities of these hybridomas for analogues at positions 114–118 are unique to and characteristic of hybridomas that express the Vβ8.3/Jβ1.3 combination; none of the other hybridomas show this reactivity pattern, including those that express Vβ8.3 in conjunction with other Jβ gene segments. As noted above, the Vβ8.3/Jβ1.3 hybridomas utilize identical (or nearly identical) β chain junctional sequences, the presence of which implies that these β chains have precise and closely related structural requirements for recognition of the site 1/Eβ complex. In conjunction with the clear contribution made by the α chain in recognition of positions 110–113, the association between the expression of identical β chains and shared specificity for analogues at positions 114–118 suggests that these TCR β chains play an important role in recognition of the COOH-terminal half of the site 1 determinant.

A second major group of four hybridomas shares in common the expression of one of two genes from the Vα4 family, and utilizes highly diverse β chains derived from four different Vβ gene families. This group is characterized by shared specificity for analogues at positions 110–112; each of the hybridomas displays a unique specificity for analogues at positions 114–118. Since this group does not contain any examples of hybridomas whose TCRs differ only in the α or β chain, it is not possible to make precise assignments of the role of each chain in determining specificity for the different ana-
logues. It is nevertheless noteworthy that this group of hybridomas is consistent with the Vβ8.3/Jβ1.3 group, in as much as there is an association between the sharing of common α or β chain sequences, and shared specificity for the NH2-terminal and COOH-terminal portions, respectively, of the site 1 determinant.

The extensive V and J gene segment diversity of the site 1-specific TCRs is in contrast to many previous studies (6-14), which demonstrated that TCRs induced in response to certain antigen/class II MHC complexes utilized a limited number of germ-line gene elements. Based on the sample obtained here, the total number of different gene segments that are expressed in recognition of the site 1/1-Ecomplex can be estimated from the frequency with which individual Va, Vβ, Jα, and Jβ gene segments were repetitively observed among the site 1-specific hybridomas (Table 1). For the Vβ and Jβ gene segments, the most likely estimate is that five different Vβ and eight different Jβ gene segments are utilized in the formation of TCRs that recognize the site 1/1-E complex, closely resembling the actual number observed. For Vα and Jα gene segments, the most likely estimates are that 22 different Vα and 35 different Jα gene segments are utilized in TCRs with this specificity. This analysis suggests, then, that approximately one-fifth each of the total Vα and Vβ, and greater than one-half of the total Jα and Jβ germ-line repertoires, are utilized by the BALB/c mouse in response to this determinant. Because both positive and negative selection act on developing thymocytes to influence the repertoire that is expressed by peripheral T cells (30, 31), the available repertoire does not constitute a random expression of the germ-line repertoire. These projections are based on the assumption that each species is randomly drawn from the available repertoire; accordingly, the proportion of the available, peripheral germ-line gene repertoire that is utilized cannot be assessed.

It is also possible to estimate the number of different Vα/Vβ gene segment combinations that are utilized to recognize the site 1/1-E complex (Table 2). Among 13 TCRs, there is one pair of hybridomas (7/6AH1 and 1E10) that expresses the same Va/Vβ gene segment combination. Based on this repeat expression of a single Va/Vβ pair, the most likely estimate is that the BALB/c mouse utilizes 74 different Va/Vβ combinations in recognition of the site 1/1-E complex. Since this estimate is based on the identification of only a single Va/Vβ pair, there is a considerable amount of error associated with it. Although the most likely estimate is that 74 different combinations are utilized, the lower 95% confidence bound predicts that 20 different combinations are used. Even this lower limit greatly exceeds similar estimates of the number of different Va/Vβ combinations that are utilized by TCRs specific for other Th determinants (Table 2). Interestingly, this number of potential Va/Vβ combinations is approximately equal to the number of different Va/Vβ combinations that can similarly be estimated to be utilized in the BALB/c antibody response to the antigenic site Cb of the PR8 HA. Site Cb-specific antibodies recognize overlapping epitopes on the HA, which are abolished by amino acid substitutions in a linear stretch of six amino acids on the surface of the HA (32). The structure of this antigenic site resembles models of the structure of antigen/class II MHC complexes, which propose that T cells recognize the antigenic peptide found in a cleft on the surface of the class II molecule (3, 33). Comparison of the response to these determinants is interesting in view of the requirement as noted above that T cells undergo processes of positive and negative selection that eliminate many thymocytes and restrict the available TCR repertoire to those that can recognize foreign antigens in the context of the host's particular MHC molecules (30, 31). There is no equivalent bias introduced into the generation of the B cell repertoire; nevertheless, similar numbers of V gene combinations are estimated to be utilized by B and T cells in recognition of these presumably comparable antigenic structures.

Table 1. Estimated TCR Gene Segment Repertoires Utilized for Recognition of Site 1/1-E Complex

<table>
<thead>
<tr>
<th>Site 1-specific hybridomas</th>
<th>Most likely site 1-specific repertoire</th>
<th>Estimated total germ-line repertoire</th>
<th>Most likely site 1-specific repertoire/estimated total germ-line repertoire</th>
</tr>
</thead>
<tbody>
<tr>
<td>No. of species/sample size</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Va</td>
<td>10/13</td>
<td>22</td>
<td>100</td>
</tr>
<tr>
<td>Vβ</td>
<td>5/13</td>
<td>25</td>
<td>25</td>
</tr>
<tr>
<td>Jα</td>
<td>11/13</td>
<td>35</td>
<td>50</td>
</tr>
<tr>
<td>Jβ</td>
<td>7/13</td>
<td>8</td>
<td>12</td>
</tr>
</tbody>
</table>

The most likely number of different species (n) is calculated by finding the value that maximizes the formula (40): S(r,d)/r!\cdot d!; where S(r,d) are Stirling's numbers of the second kind, and d is the number of different species observed among a sample of size r. This statistic estimates the most likely number of different species that exist in a population from which the observed number of different species was randomly drawn. The predicted number of different α and β chain gene segments that are used in the recognition of the site 1/1-E complex is shown. The estimated total germ-line repertoire (3) and the ratio of the predicted number of species to the estimated germ-line repertoire are also shown.

1648  An Influenza Antigen/MHC Complex Induces Diverse T Cell Receptors
Table 2. Estimated TCR $\alpha$/$\beta$ Combinations Utilized for Recognition of Defined Antigenic Determinants

<table>
<thead>
<tr>
<th>Sample</th>
<th>No. of species/sample size</th>
<th>Estimated total $\alpha$/$\beta$ repertoire</th>
</tr>
</thead>
<tbody>
<tr>
<td>a. Site 1</td>
<td>12/13</td>
<td>74</td>
</tr>
<tr>
<td>b. Cytochrome c</td>
<td>6/10</td>
<td>8</td>
</tr>
<tr>
<td>c. Cytochrome c*</td>
<td>4/15</td>
<td>4</td>
</tr>
<tr>
<td>d. $\lambda$ cl*</td>
<td>6/12</td>
<td>7</td>
</tr>
<tr>
<td>e. Myelin basic protein</td>
<td>2/8</td>
<td>2</td>
</tr>
<tr>
<td>f. Myelin basic protein</td>
<td>4/33</td>
<td>4</td>
</tr>
<tr>
<td>g. $\text{Cb}$</td>
<td>20/23</td>
<td>77</td>
</tr>
</tbody>
</table>

The estimated number of different $\alpha$ gene combinations that are used in recognition of individual determinants is indicated. The number of different $\alpha$/$\beta$ combinations that are expressed in recognition of (a) the site 1/1-Ed complex; (b) the COOH-terminal determinant of pigeon cytochrome c induced in B10.A mice (7); (c) the COOH-terminal determinant of pigeon cytochrome c induced in B10.A mice (7); (c) the COOH-terminal determinant of pigeon cytochrome c induced in B10.A mice, from Sorger et al. (6); (d) residues 12-26 of the $\lambda$ virus repressor protein cl induced in A/J mice (8); (e) the NH$_2$-terminal determinant of myelin basic protein (10); (f) the NH$_2$-terminal determinant of rat myelin basic protein induced in B10.PL mice (9). Also shown (g) is the estimated number of different $\alpha$/$\beta$ combinations that are used in recognition of the antigenic site Cb on the HA (32).

Based on Southern blot analysis, two hybridomas from this panel, E2 and 2C2, could not be unambiguously assigned to a particular $\alpha$11 family member, and therefore were not included in the present analysis; the $\alpha$11 family member assignments for hybridomas B10, C557, and 2B4 are from reference 7; the $\alpha$11 family member assignment for C8 is from reference 41.

It is not clear why the degree of structural diversity of the site 1-specific TCRs differs so dramatically from that observed in response to other antigen/class II MHC complexes. It is possible that there is substantial variability in the number of different gene segments and gene segment combinations that can be expressed in response to individual antigen/MHC complexes. This might in part account for differences in the immunodominance of particular antigenic determinants (20, 34). Alternatively, the considerable structural diversity of the response to the site 1 determinant might reflect the phylogenetic distance between the HA and the proteins of the mouse. The more structurally restricted responses to determinants such as cytochrome c and myelin basic protein might be a consequence of the requirement for self-tolerance and the close relationship between these proteins and the corresponding self-determinants of the responding animal. By this model, however, it is unclear why the murine response to $\lambda$ phage repressor protein cl also shows limited structural diversity. The analysis of the genetic basis for the generation of TCRs that recognize other determinants of the HA might allow these issues to be resolved.

We thank Joshua Kavaler for discussion and for help with the statistical analysis, and Sue Stark for excellent assistance.

This work was supported by grants AI-24541 and AI-13989 from the National Institutes of Health.

Address correspondence to Andrew J. Caton, The Wistar Institute, 3601 Spruce Street, Philadelphia, PA 19104.

Received for publication 26 June 1990 and in revised form 31 August 1990.

References

324:679.


T cell receptor junctional regions and the MHC molecule affect the recognition of antigenic peptides by T cell clones. J. Immunol. 144:1127.