Antigen Recognition Determinants of γδ T Cell Receptors

Sunny Shin,1 Ramy El-Diwany,1,2 Steven Schaffert,1 Erin J. Adams,1,2 K. Christopher Garcia,1,2 Pablo Pereira,3 Yueh-hsiu Chien1,2*

The molecular basis of γδ T cell receptor (TCR) recognition is poorly understood. Here, we analyze the TCR sequences of a natural γδ T cell population specific for the major histocompatibility complex class Ib molecule T22. We find that T22 recognition correlates strongly with a somatically recombined TCRδ complementarity-determining region 3 (CDR3) motif derived from germ line-encoded residues. Sequence diversity around these residues modulates TCR ligand-binding affinities, whereas V gene usage correlates mainly with tissue origin. These results show how an antigen-specific γδ TCR repertoire can be generated at a high frequency and suggest that γδ T cells recognize a limited number of antigens.

1Department of Microbiology and Immunology, Stanford School of Medicine, Stanford, CA 94305, USA. 2Program in Immunology, Stanford University, and 3Unité du Développement des Lymphocytes, Centre National de la Recherche Scientifique, Unité de Recherche Associée 1961, Institut Pasteur, 75724 Paris, France.
*To whom correspondence should be addressed. E-mail: chien@stanford.edu

The γδ and αβ T cells contribute to host immune defense in distinct ways. Whereas αβ T cells are essential in pathogen clearance, γδ T cells have been implicated in the regulation of immune responses (1). Although it is clear that γδ T cells can recognize antigens directly without antigen processing and presentation requirements (2), it is unclear what the majority of γδ T cell ligands are and how they are recognized. This has made it difficult to define the precise function of γδ T cells. Previously, we found that the closely related major histocompatibility complex (MHC) class Ib molecules T10 and T22 (94% amino acid identity) are induced on activated cells and are ligands for a sizable population (~0.1% to 2%) of γδ T cells in unimmunized mice (3). This is potentially an important γδ T cell-ligand pair that could help to regulate immune responses. To understand how this antigen-specific repertoire is generated, particularly the high initial frequency of these cells, we used a T22 tetrameric staining reagent to identify and isolate T22-specific γδ T cells and determined their TCR sequences.

Most splenic γδ T cells express Vγ1 and Vγ4, whereas Vγ7-expressing γδ T cells are more prevalent in the intestinal intraepithelial lymphocyte (IEL) compartment (4–6). This bias in Vγ usage has led to the suggestion that Vγ-encoded residues enable these T cells to respond to antigens unique to their resident tissues (1, 7). Because T22-specific γδ T cells are present in both the spleen and IEL compartments, we first tested whether T22 specificity correlates with Vγ gene usage (8). We found that multiple Vγs and Vδs are associated with T22-specific γδ T cells from these two tissues; however, the majority of T22 tetramer-positive cells express Vγ1 and Vγ4 in the spleen, whereas a sizable population of these cells express Vγ7 in the IEL compartment (Fig. 1A and table S1 and S3). This result indicates that Vγ usage is more reflective of the tissue origin than of the antigen specificity for this ligand.

We then compared the TCR sequences of individual T22 tetramer-positive and -negative cells (8). Although no conserved sequences in T22-specific TCRγ chains can be identified (tables S1 to S4 and fig. S1), we found that ~90% of the tetramer-positive IELs and ~40% to 60% of the splenic tetramer-positive TCRs contained a prominent CDR3 sequence motif (Fig. 2A). This motif consists of a tryptophan (W) encoded by the Vδ3 or D61 gene segments and the sequence serine–glutamic acid–glycine–tyrosine–glutamic acid (SEGYE), followed by a P nucleotide–encoded leucine (L). Other than the motif, the CDR3 sequences are diverse, encoded by various Vδs, N and P nucleotides, and D61 in different lengths and reading frames. It is interesting that Vδ6A is the only Vδ to encode a tryptophan residue in the CDR3δ and is overrepresented in T22-specific γδ TCRs (Fig. 1B). Additionally, the CDR3δ length distribution is narrower and longer than that of γδ TCRs in general (Fig. 2, B and C).

To test whether TCRs derived from T22 tetramer-positive cells confer T22 binding specificity, we expressed several of these TCRs in the TCRβ-deficient Jurkat T cell line J.RT3-T3.5, which lacks endogenous surface TCR expression (8, 11). We found that cells expressing TCRs that have the W-(S)EGYEL motif could bind T22 tetramer, whereas those that lack this motif could not (Fig. 3 and fig.
S2). Thus, the higher rate of splenic tetramer-positive T cells without the TCR6 motif may be due to a higher false-positive rate in identifying these cells. This may be caused by the experimental limit associated with fluorescence-activated cell sorting (FACS), especially for low tetramer binders. (T22 tetramer stains IELs at a higher intensity than splenic cells.) Indeed, more recent experiments with a slightly more stringent FACS gating showed that ~70% of the splenic tetramer-positive cells have the TCR6 motif (12). Regardless, although both KN6 and 93A10 TCRs use a Vy4-V65 gene combination, only KN6 contains the W-(SEG)EYL CDR3e motif and is T22-specific. G8 (Vy4-Vx11.3), KN6 (Vy4-V65), as well as 93B7, 93D11, and 917B7 (Vy1-V66A), all bind T22 but use three different Vy-Vx pairs. This indicates that the W-(SEG)EYL CDR3e motif correlates much better than V gene usage with antigen recognition. Consistent with this is the structural analysis of the G8-T22 complex showing that the residues W and GYEL in the G8 CDR3 are the principal T22 contact residues (13).

To test whether variability in the sequences surrounding the W-(SEG)EYL motif influences ligand binding, we compared the T22 binding characteristics of cells expressing similar levels of the 93B7, 93D11, and 917B7 TCRs, which differ only in those residues. As shown in Fig. 3, these TCRs exhibit significant differences in the half-life ($t_{1/2}$) and affinity ($K_d$) of T22 tetramer binding. Thus, sequence variations around this motif can modulate the affinity and the kinetics of ligand binding.

These results indicate that, for T22 specificity, a CDR3 sequence generated by somatic rearrangement is necessary. This is similar to antibody specificities, which reside predominantly in the CDR3 of the heavy chain (14, 15). Also, in the case of mβ TCRs, peptide-MHC specificity is determined largely by CDR3e and CDR3b, but the nature of the antigen-recognition determinants of T22-specific γδ TCRs and mβ TCRs are quite different. The T22-specific CDR3e motif is encoded mainly by D62 with contributions from V6, D61, and P nucleotides, whereas in mβ TCRs the most critical residues for peptide-MHC recognition are encoded either completely or partially by N nucleotides in both CDR3α and CDR3β (15).

To determine whether a largely intact D62 is a unique feature of T22-specific TCRs or of γδ TCRs in general, we analyzed the D62 length distribution of in-frame thymocyte TCRγ sequences ($N = 431$). We found that ~23% of these sequences contain D62 in its entirety, whereas an additional ~30% retain at least 13 out of 16 D62 nucleotides (Table 1). A similar D62 length distribution was also found in nonselected TCRs ($N = 271$) consisting of out-of-frame TCRγ chains and TCRγ rearrangements from CD3ε−/− thymocytes, which cannot express surface TCR (Table 1). This indicates that TCRγ rearrangements are strongly biased toward maintaining long D62 regions. In the periphery, more than 50% of both the T22-specific and non–T22-specific splenic and IEL sequences contain D62 in its entirety, and more than 70% of the sequences have less than three nucleotides deleted (Table 1), indicating that the resulting
TCRs are further selected for full use of the Dβ8 segment. In contrast, Dβ sequences from lymph node CD4+, Vβ17+ αβ T cells (16) show that only 3 to 7% are intact and fewer than 15 to 30% have been truncated by three nucleotides or less (Table 1).

Another feature distinguishing TCRδ CDR3 sequences from those of TCRβ and IgH chains is the J region. In both the TCRβ and the IgH chains, multiple J regions (12 Jβs and 6 Jδs in mice) provide important framework residues and also contribute to antigen binding via their N-terminal residues (J5). Exonuclease digestion and the addition of N nucleotides to the J region contribute to variability and thus to antigen binding (J5). In contrast, adult murine γδ TCRs use only one Jβ, and the degree of exonuclease digestion is quite limited compared with αβ TCRs in that more than 98% of the sequences (T22-specific as well as non-specific) retain the first or second N-terminal amino acid residue encoded by Jδ1 (Table 2). This very limited J region diversity is also found among thymocytes and nonselected γδ TCRs (Table 2), revealing yet another unique feature of TCRδ gene rearrangement. This relative lack of variation suggests that, unlike Jβ, and Jβ, Jβδ1 does not play a major role in antigen recognition.

Although most γδ T cell ligands have yet to be identified, our observations indicate that rearrangements at the TCRδ locus are largely biased toward full-length D82 sequences rather than extensive D-region nucleotide deletion, as is the case for the TCRβ locus. Thus, different reading frames of D82 may contribute to the recognition of other ligands by γδ TCRs in a manner similar to that of T22-specific γδ TCRs. This would allow these germ line–encoded CDR3 sequences to coevolve with their ligands. In fact, most well-defined γδ T cell’s ligands are self-molecules that could act as indicators of physiological disturbances, such as T10 and T22 in the mouse and MICA and B, CD1, and F1–adenosine triphosphate synthase in humans (3, 17–19).

One would expect that T cell repertoire generated from somatic recombination but whose specificity is conferred by germ line–encoded amino acids (such as for T22-specific γδ TCRs) would be created much more frequently than αβ T cells whose specificity is conferred primarily by N-nucleotide additions. In fact, we find that 0.85% of nonselected TCRδ sequences (N = 353) contain this CDR3δ motif (table S5) compared to one in 105 to 106 αβ T cells specific for a given peptide-MHC before clonal expansion (20, 21). Thus, rearrangement alone could in part account for the high frequency (0.1 to 2%) of T22-specific γδ T cells in normal mice (Fig. 1A) (3, 12). If γδ TCR specificity for other ligands is determined in a similar manner, then the γδ T cell repertoire must be directed against a relatively small number of ligands but with high frequency. This could allow for a rapid and significant response without an initial need for clonal expansion.

The CDR3δ provides the TCRδ with the highest potential diversity of all antigen receptor polypeptides. The results described here show that this diversity endows T22-specific γδ TCRs with different ligand-binding affinities. Indeed, the T22-specific TCR repertoire in normal mice covers a range of affinities, as evidenced by the large range of T22 tetramer-staining intensities (Fig. 1) (3, 12). A self-reactive TCR repertoire with such diverse ligand-binding properties would enable more flexible and efficient responses to changes in self-ligand expression and at the same time allow for selection against high-affinity T cells that might respond inappropriately to basal ligand expression amounts.

Table 1. D82 length distribution in TCRδ rearrangements. Numbers represent the percentage of rearrangements with the indicated number of nucleotides removed. The lengths of D regions were analyzed in nucleotides because they can be read in all three reading frames. Sequences analyzed are functional T22 tetramer-positive and -negative TCRδ chains (tables S1 to 4); functional TCRδ chains from γδ T cell hybridomas (25) and thymocytes (26) nonselected TCRδ chains from CD3+/− thymocytes (27); and CD4+ Vβ17+ TCRδ chains from the lymph nodes of SJL mice (15) (n indicates the number of sequences analyzed).

<table>
<thead>
<tr>
<th>Dβ/DJβ nucleotides deleted</th>
<th>Spleen Tetramer+ (n = 92) (%)</th>
<th>IEL Tetramer+ (n = 93) (%)</th>
<th>Functional TCRδ chains (n = 431) (%)</th>
<th>Nonselected TCRδ chains (n = 271) (%)</th>
<th>Vββ17+ CD4+ αβTCR Dβ1 (n = 37) (%)</th>
<th>Vββ17+ CD4+ αβTCR Dβ2 (n = 37) (%)</th>
<th>Vββ17+ CD4+ αβTCR Jβδ1 (n = 57) (%)</th>
</tr>
</thead>
<tbody>
<tr>
<td>0</td>
<td>55.4</td>
<td>36.5</td>
<td>52.2</td>
<td>44.6</td>
<td>23</td>
<td>21.4</td>
<td>2.7</td>
</tr>
<tr>
<td>1–3</td>
<td>29.3</td>
<td>30.1</td>
<td>30.4</td>
<td>25.7</td>
<td>30.2</td>
<td>28.0</td>
<td>10.8</td>
</tr>
<tr>
<td>4–6</td>
<td>8.7</td>
<td>19.4</td>
<td>8.7</td>
<td>20.3</td>
<td>22.5</td>
<td>22.5</td>
<td>51.3</td>
</tr>
<tr>
<td>7–10</td>
<td>6.5</td>
<td>12.9</td>
<td>4.3</td>
<td>9.4</td>
<td>22.5</td>
<td>22.5</td>
<td>0</td>
</tr>
<tr>
<td>undetermined</td>
<td>0.1</td>
<td>1.1</td>
<td>4.3</td>
<td>0</td>
<td>1.4</td>
<td>5.5</td>
<td>35.0</td>
</tr>
</tbody>
</table>

Table 2. Jδ1 length distribution in TCRδ rearrangements. Numbers represent the percentage of rearrangements with the indicated number of amino acids (J region) removed. Sequences analyzed are functional T22 tetramer-positive and -negative TCRδ chains (tables S1 to 4); functional TCRδ chains from γδ T cell hybridomas (25) and thymocytes (26) nonselected TCRδ chains from CD3+/− thymocytes (25); out-of-frame rearrangements from γδ T cell hybridomas, and single-cell analyses from thymocytes (27); and CD4+ Vβ17+ TCRδ chains from the lymph nodes of SJL mice (15) (n indicates the number of sequences analyzed).

<table>
<thead>
<tr>
<th>Jβ/βδ amino acids deleted</th>
<th>Spleen Tetramer+ (n = 92) (%)</th>
<th>IEL Tetramer+ (n = 93) (%)</th>
<th>Functional TCRδ chains (n = 431) (%)</th>
<th>Nonselected TCRδ chains (n = 271) (%)</th>
<th>Vββ17+ CD4+ αβTCR Jβ (n = 75) (%)</th>
</tr>
</thead>
<tbody>
<tr>
<td>0</td>
<td>69.1</td>
<td>79.4</td>
<td>23.1</td>
<td>80.5</td>
<td>68.4</td>
</tr>
<tr>
<td>1</td>
<td>30.9</td>
<td>19.6</td>
<td>76.9</td>
<td>18.2</td>
<td>26</td>
</tr>
<tr>
<td>2</td>
<td>0</td>
<td>1</td>
<td>0</td>
<td>1.3</td>
<td>5.6</td>
</tr>
<tr>
<td>3 or more</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
</tr>
</tbody>
</table>

References and Notes
8. Materials and methods are available on Science Online.
12. S. Shin et al., unpublished data.
22. Single-letter abbreviations for the amino acid residues are as follows: 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; and Y, Tyr.
Consider the following situation: A child who has surreptitiously eaten the last cookies in a box sees her brother reach into the box. To make sense of his behavior, she must understand that he falsely believes the box still contains cookies. As adults, we readily understand that others may hold and act on false beliefs; this ability is widely held to be a cornerstone of social competence, and its neuronal correlates have recently begun to be examined (4). What are the origins of this ability? Within the field of psychology, there has been a longstanding controversy regarding this issue (2–4).

Some researchers have suggested that at about 4 years of age a fundamental change occurs in children’s understanding of others’ behavior, or “theory of mind”: They begin to realize that mental states such as beliefs are not direct reflections of reality, which must always be accurate, but representations, which may or may not be accurate (5–8). Part of the evidence for this change from a nonrepresentational to a representational theory of mind has come from young children’s well-documented failure at false-belief tasks (i.e., tasks that require the understanding that others may hold and act on false beliefs) (9–13). In a standard task (10), children listen to a story as it is enacted with dolls and toys: The first character hides a toy in one location and leaves the room; while she is gone, a second character hides the toy in a different location. When asked where the first character will look for her toy, 4 year olds typically say she will look in the first location and provide appropriate justifications for their answers. In contrast, most 3 year olds say she will look in the second (actual) location, thus failing to demonstrate an understanding that the first character will hold a false belief about the toy’s location.

Other researchers have suggested that a representational theory of mind is present much earlier and that young children’s difficulties with the standard false-belief task stem primarily from excessive linguistic, computational, and other task demands (14–18). Support for these claims comes in part from evidence that 3 year olds and even some 2 year olds succeed at a modified false-belief task (19, 20). In this version of the task, after listening to the story and watching it enacted, children are simply probed by the experimenter to look where the first character will search for her toy upon her return (“I wonder where she will look”). Most children look to the correct location, suggesting that they possess some implicit understanding that others may hold and act on false beliefs. We examined whether 15-month-old infants tested with a simpler, entirely nonverbal task would also show some implicit understanding of false belief.

We used the violation-of-expectation methodology, which has been used extensively to investigate infants’ understanding of others’ goals (21–23). For example, in one experiment (22), infants were familiarized with an actor reaching for and grasping one of two toys (defined as the target toy). Next, the locations of the two toys were reversed, and the actor reached for the target or the nontarget toy. The infants looked reliably longer at nontarget reaches. This and control results suggested that the infants encoded the target toy as the actor’s goal object, expected her to reach for it in its new location, and responded with increased attention when she did not. Similar results were found when the target toy was hidden rather than visible and was retrieved by means-end action sequences rather than by a simple reach (23). Our research built on these results. In our experiment, 15-month-old infants first watched an actor hide a toy in one of two locations. Next, a change occurred that resulted in the actor holding either a true or a false belief about the toy’s location. The experiment asked whether the infants would expect the actor to search for her toy based on her belief about its location, whether that belief was true or false.

---

**A** Familiarization trial 1

**B** Familiarization trials 2 and 3

Fig. 1. Events shown during (A) the first familiarization and (B) the second and third familiarization trials. The light gray box represents the yellow box; the dark gray box represents the green box.