ELSEVIER

Seminars in Immunology xxx (2007) xxx–xxx

seminars in
IMMUNOLOGY

www.elsevier.com/locate/ysmim

Review

Polyspecificity of T cell and B cell receptor recognition

Kai W. Wucherpfennig [a,*], Paul M. Allen [b], Franco Celada [c], Irun R. Cohen [d], Rob De Boer [e], K. Christopher Garcia [f], Byron Goldstein [g], Ralph Greenspan [h], David Hafler [i], Philip Hodgkin [j], Erik S. Huseby [k], David C. Krakauer [l], David Nemazee [m], Alan S. Perelson [n], Clemencia Pinilla [o], Roland K. Strong [p], Eli E. Sercarz [q,**]

[a] *Dana-Farber Cancer Institute, Harvard Medical School, Boston, MA 02115, USA*
[b] *Department of Pathology and Immunology, Washington University School of Medicine, St. Louis, MO 63110, USA*
[c] *Department of Oncology, Biology & Genetics, University of Genova, Genova, Italy*
[d] *Department of Immunology, The Weizmann Institute of Science, Rehovot 76100, Israel*
[e] *Theoretical Biology Department, Utrecht University, Utrecht, The Netherlands*
[f] *Howard Hughes Medical Institute, Departments of Microbiology & Immunology, and Structural Biology, Stanford University School of Medicine, Stanford, CA 94305-5124, USA*
[g] *Theoretical Biology and Biophysics Group, Los Alamos National Laboratory, Los Alamos, NM 87545, USA*
[h] *The Neurosciences Institute, San Diego, CA 92121, USA*
[i] *Department of Neurology and Center for Neurologic Disease, Brigham and Women's Hospital, Harvard Medical School, Boston, MA 02115, USA*
[j] *The Walter and Eliza Hall Institute of Medical Research, Parkville, Victoria, Australia*
[k] *Howard Hughes Medical Institute and Integrated Department of Immunology, National Jewish Medical and Research Center, Denver, CO 80206, USA*
[l] *Santa Fe Institute, Santa Fe, NM 87501, USA*
[m] *Department of Immunology, The Scripps Research Institute, La Jolla, CA 92037, USA*
[n] *Department of Theoretical Biology and Biophysics, Los Alamos National Laboratory, Los Alamos, NM 87545, Mexico*
[o] *Department of Immunochemistry, Torrey Pines Institute for Molecular Studies, San Diego, CA 92121, USA*
[p] *Division of Basic Sciences, Fred Hutchinson Cancer Research Center, Seattle, WA 98109, USA*
[q] *Division of Immune Regulation, Torrey Pines Institute for Molecular Studies, San Diego, CA 92121, USA*

Abstract

A recent workshop discussed the recognition of multiple distinct ligands by individual T cell and B cell receptors and the implications of this discovery for lymphocyte biology. The workshop recommends general use of the term polyspecificity because it emphasizes two fundamental aspects, the inherent specificity of receptor recognition and the ability to recognize multiple ligands. Many different examples of polyspecificity and the structural mechanisms were discussed, and the group concluded that polyspecificity is a general, inherent feature of TCR and antibody recognition. This review summarizes the relevance of polyspecificity for lymphocyte development, activation and disease processes.

Keywords: T cell receptor; Peptide; MHC; Polyspecificity; Recognition

1. Purpose of the workshop

A recent workshop at the Santa Fe Institute organized by E. Sercarz, I. Cohen and A. Perelson on "Degeneracy and Complexity in the Immune System" discussed the emerging realization that lymphocyte receptors recognize multiple distinct ligands and the implications of this discovery for many different aspects of T cell and B cell biology. Participants described their work on the characterization of peptide/MHC ligands (E. Sercarz, H. Eisen, C. Pinilla and D. Hafler), structural mechanisms of recognition by TCRs and other immune receptors (P. Allen, K. C. Garcia, R. Strong and K. W. Wucherpfennig), T cell development (E. Huseby), antibody degeneracy and B cell tolerance (I. Cohen and D. Nemazee), quantitative aspects and mathematical models of specificity and degeneracy (A. Perelson, R. de Boer, F. Celada,

* Corresponding author at: Dana-Farber Cancer Institute, 44 Binney Street, Boston, MA 02115, USA. Tel.: +1 617 632 3086; fax: +1 617 632 2662.
** Corresponding author.
E-mail addresses: kai_wucherpfennig@dfci.harvard.edu (K.W. Wucherpfennig), esercarz@tpims.org (E.E. Sercarz).

1044-5323/$ – see front matter © 2007 Elsevier Ltd. All rights reserved.
doi:10.1016/j.smim.2007.02.012

Table 1
Definitions of terminology

Polyspecificity	This term emphasizes two important features of TCR recognition: the ability to recognize multiple distinct peptide/MHC ligands as well as the specificity with which each of these ligands is recognized. The workshop recommends general usage of this term.
Degeneracy	Emphasizes the finding that some of the peptide/MHC ligands recognized by a particular TCR can be distinct in both primary sequence and structure. In general, the term degeneracy is a better fit for peptide binding to MHC molecules than for TCR recognition. Peptides that are highly diverse in sequence can be bound by MHC molecules and even the requirements for anchor residues are typically not very strict. In contrast, subtle changes in a peptide can result in loss of TCR recognition.
Molecular mimicry	Polyspecificity in the context of autoimmune diseases: a self-reactive T cell is stimulated by microbial peptide(s) successfully processed from the microbe and bound to MHC molecules, leading to activation and expansion of autoreactive T cell populations. A large fraction of the early literature used this term.
Plasticity and flexibility	One structural explanation for polyspecificity: in crystal structures of TCRs with and without peptide/MHC ligands, substantial movements in TCR loops were documented, in particular in the CDR3 loops. However, not all cases of polyspecificity may be caused by flexibility in TCR loops.
Cross-reactivity	Similar to polyspecificity, but this term is not as explicit in emphasizing the existence of multiple peptide/MHC ligands. This term was originally used to indicate unexpected reactivity to targets that differed from those used to initially define the clone.

Table 2
Major conclusions from workshop

1. Many different terms are currently used to describe polyspecificity of TCR recognition. The field would benefit from usage of one well-defined term. The workshop recommends polyspecificity because it emphasizes two key aspects, the ability of TCRs to recognize multiple peptide/MHC ligands and the specificity with which each ligand is recognized.
2. Polyspecificity is an inherent, general property of TCR recognition and relevant to many aspects of T cell biology.
3. T cells are specific because they recognize a *small fraction* of all ligands. However, they recognize a *substantial number* of ligands because the total number of potential ligands is very large. Thus, T cells are both specific and degenerate.
4. Mathematical models support the conclusion that polyspecificity ensures that a sufficient number of T cells are recruited into an immune response against a particular pathogen.
5. Motif analysis and structural studies demonstrate that limited sequence similarity can be sufficient for recognition of distinct MHC-bound peptides. However, sequence combinations are important because even small changes can affect peptide conformation.
6. Polyspecificity is relevant for T cell development, in particular during positive and negative selection. When negative selection is impaired, a large fraction of mature T cells are highly degenerate. T cell clones appear to differ in their degree of degeneracy.
7. Weak ligands contribute to TCR triggering. T cell activation can thus be described as an integration of signals resulting from multiple ligands, such as one or several strong and multiple weak ligands.
8. Autoreactive T cells can be activated by a number of microbial peptides that have limited sequence similarity with the self-peptide.
9. Alloreactive T cells can recognize other peptide/MHC combinations and the peptides recognized in the context of different MHC molecules can have little primary sequence similarity.

B. Goldstein and P. Hodgkin) as well as complexity in other biological systems (R. Greenspan and D. Krakauer). This report represents a synthesis of both individual presentations as well as in-depth discussions of key issues in this field (summaries and PowerPoints of individual presentations can be found at http://www.santafe.edu/events/workshops/index.php/Degeneracy_and_Complexity_in_the_Immune_System) (Tables 1 and 2).

2. Perspective

One of the implications of the clonal selection theory was that a given T cell is highly specific for a single "cognate" peptide/MHC ligand and that recognition of alternative ligands is a rare event. This view was based on the observation that a given T cell clone typically does not respond when tested against several other antigens and that small changes in a peptide can result in a loss of T cell recognition. T cell activation was thus thought to be the result of recognition of a single MHC-bound peptide present at a sufficient density at the cell surface. However, several examples emerged in the early 1990s in which T cell clones or hybridomas were found to recognize alternative peptide/MHC ligands that were remarkably different from the "cognate" antigen in their primary sequence [1–3]. At the same time, the structural requirements for peptide binding to MHC molecules were elucidated, demonstrating that peptide-binding motifs were quite degenerate, in particular for MHC class II molecules [4–7]. Peptides with minimal sequence homology to the original peptide were identified by considering the requirements for both MHC binding and TCR recognition. These studies demonstrated for the first time that a single T cell could respond to a variety of different peptides that were quite distinct from each other [1–3]. The presentations at the workshop showed that the interaction of T cells with a multiplicity of endogenous and exogenous ligands is central to many immunological processes. Processes such as positive selection in the thymus, survival of naïve T cells, the high frequency of allorecognition and the induction of autoimmunity by exogenous microbial ligands are just some of the distinctive features of the system, explicable from this viewpoint.

3. Examples of polyspecificity from diverse murine and human T cell systems

3.1. Systematic identification of peptide ligands with limited sequence similarity

Polyspecificity was first systematically studied by considering the structural requirements for both MHC class II binding

Please cite this article in press as: Wucherpfennig KW, et al., Polyspecificity of T cell and B cell receptor recognition, Semin Immunol (2007), doi:10.1016/j.smim.2007.02.012

and TCR recognition of a myelin basic protein (MBP) peptide recognized by T cell clones isolated from multiple sclerosis (MS) patients [3,7]. The analysis was based on the realization that peptide-binding motifs by MHC class II molecules were quite degenerate because multiple amino acid substitutions were tolerated at the two major anchor residues of the MBP peptide for HLA-DR2 (DRA, DRB1*1501). Furthermore, a high degree of specificity of TCR recognition was typically limited to two to three peptide side chains. A large number of viral and bacterial peptides could thus be identified that activated these human CD4 T cell clones. The peptide sequences were quite distinct from the MBP peptide and from each other. Functional dissection of the peptide sequences demonstrated that the residues occupying pockets of the MHC molecule were particularly diverse, while more sequence similarity was observed for residues involved in TCR recognition. Depending on the T cell clone, these peptides had sequence identity either at the P2, P3 positions (–HF–) or at the P3, P5 positions (–F–K) in the central VHFFK peptide segment [3,7]. The degree of cross-reactivity was even larger when simultaneous substitutions were made at positions P3 and P5 [8]. These studies thus demonstrated that a single T cell clone could recognize a number of different peptides with limited sequence similarity.

3.2. Identification of diverse peptide ligands for T cell clones with combinatorial peptide libraries

Synthetic combinatorial peptide libraries have emerged as a powerful tool for the identification of multiple distinct peptide ligands for human/murine CD4 and CD8 T cell clones [9,10]. For each position, a set of 20 libraries is synthesized in which one of the 20 naturally occurring amino acids occupies the defined position while all other positions are synthesized with amino acid mixtures. All positions of a 9-mer peptide can thus be interrogated with a set of 180 mixtures (9×20). Incorporation of defined amino acids in the active mixtures permitted synthesis of super-agonist peptide ligands for a variety of T cell clones that are active at substantially lower peptide concentrations than the peptide used to isolate the T cell clone. Furthermore, protein databases searches led to the identification of alternative self or microbial peptides for these clones [11].

These libraries are composed of complex mixtures. Individual peptides are present at concentrations that are far too low to induce T cell activation. The estimated number of stimulating ligands differs by several orders of magnitude depending on the activity assumed for an individual peptide, for example, 2×10^3 versus 2×10^6 stimulatory peptides for a 9-mer library, depending on whether peptides are assumed to have an average activity of 10 pM or 10 nM. Since both weak and strong ligands are likely to contribute to activation, the precise number cannot be determined with certainty. It is, however, evident that the number of peptides contributing to T cell activation by such libraries is large. The specificity of TCR recognition is also reflected by this analysis because the *fraction* of stimulatory peptide ligands in these highly complex libraries is still likely to be small.

3.3. An example of degeneracy: The 2C TCR

The 2C TCR has been characterized in great detail, and these studies revealed a diverse array of peptide/MHC ligands for this TCR that induce a wide range of different responses [12,13]. The TCR originated from an allogeneic CD8 cytotoxic T cell clone and naturally processed peptides bound to L^d (allogeneic) and K^b (syngeneic) MHC class I molecules have been identified that have very limited sequence similarity. Also, the 2C TCR can recognize multiple MHC molecules (several MHC class I, a non-classical MHC class I and a MHC class II) during positive selection in the thymus. The various peptide/MHC complexes induce a wide range of different functional responses, ranging from positive to negative selection, to potent cytotoxic responses and antagonism. Despite the ability of 2C TCR to recognize such a wide variety of peptide/MHC ligands, this receptor can exhibit exquisite specificity because single conservative substitutions in peptides can abolish activity.

4. Structural mechanisms of specificity and polyspecificity

4.1. General features of TCR recognition

Crystal structures of TCR–peptide/MHC complexes have provided a molecular understanding of the mechanisms of TCR specificity [14–16]. The 2C TCR structure demonstrated a diagonal binding mode in which the TCR covers almost the entire MHC-embedded peptide [15]. Four TCR loops can contribute to peptide recognition, the centrally located hypervariable CDR3 loops and the germline-encoded CDR1 loops. Most TCRs exhibit exquisite specificity for one or a few peptide side chains at peptide positions directly contacted by TCR, such as the P5 position in the peptide center. Loss of a single hydroxyl group on a peptide side chain can abrogate TCR recognition if a critical hydrogen bond to a TCR loop is absent [16,17].

4.2. Recognition by autoimmune TCRs

More recently, the structures of several autoimmune TCRs have been determined that show some unusual features of TCR binding to self-peptide/MHC. The 172.10 TCR originated from a T cell clone that induces experimental autoimmune encephalomyelitis (EAE) and recognizes the N-terminal acetylated peptide of MBP (Ac1–11) bound to I-A^u [18]. All T cell clones that recognize this peptide/MHC complex use the Vβ8.2 segment, and the Vβ dominates the interaction with the MHC molecule both in terms of the number and specificity of contacts. Contacts made by the Vβ8.2 CDR1 and CDR2 loop to the MHC helix are strikingly similar to D10 TCR that uses the same Vβ. The MBP peptide only partially fills the binding site, limiting the potential peptide interaction surface with TCR. Peptide contacts are only made by the CDR3 loops, and there is only a single hydrogen bond between the TCR and a peptide side chain. Scanning of combinatorial peptide libraries demonstrated that this clone had a preference for native MBP residues at positions P1 Ala, P3 Gln, P5 Arg and P6 Pro, with the highest

Please cite this article in press as: Wucherpfennig KW, et al., Polyspecificity of T cell and B cell receptor recognition, Semin Immunol (2007), doi:10.1016/j.smim.2007.02.012

degree of specificity localized to the P5 Arg side chain that forms the only hydrogen bond to the TCR (CDR3β loop) [19]. These results are not consistent with generalized TCR degeneracy, but rather suggest a model in which TCR engagement of alternative peptide/MHC ligands results from highly specific, alternative structural solutions.

The crystal structure of a human TCR isolated from a patient with MS demonstrated a binding topology that differs substantially from the central binding mode described above [20]. This TCR (Ob.1A12) originated from a patient with relapsing-remitting MS and recognizes a MBP peptide (res. 85–99) bound to HLA-DR2 (DRA, DRB1*1501) [7]. Transgenic mice that express this TCR and HLA-DR2 can develop spontaneous EAE, demonstrating that this TCR can be pathogenic *in vivo* [21]. Compared to other structures, this TCR is shifted towards the peptide N-terminus and the DRβ chain helix. The shift towards the peptide N-terminus is considerable because the CDR3 loops are not centered over the P5 peptide side chain, but rather over P2. Similar to the EAE TCR described above, peptide contacts are largely limited to the CDR3 loops. The Ob.1A12 T cell clone can be activated by several microbial peptides that have limited sequence similarity with the MBP peptide. In all of these peptides, P2 His and P3 Phe, the primary TCR contact residues contacted by the TCR CDR3 loops, are conserved. This TCR thus exhibits exquisite specificity (substitutions of P2 His and P3 Phe are not tolerated) but can nevertheless recognize a number of alternative ligands. The unusual topology may have contributed to escape from negative selection.

4.3. Influence of a subtle change in the MHC binding of a ligand on TCR recognition

Specificity of TCR recognition is not only conferred by peptide side chains that directly contact TCR loops but also indirectly by peptide side chains that occupy pockets of the MHC molecule, as illustrated by a hemoglobin peptide (Hb, res. 67–76) bound to I–E^k and an altered peptide ligand (APL) of the hemoglobin epitope in which P6 Glu is substituted by Asp [22]. The P6 side chain occupies a pocket of I–E^k and the Glu to Asp change is subtle because it shortens the side chain by one carbon atom. The two peptides bound with the same affinity to I–E^k, but the APL was an antagonist/weak agonist for T cells from 3.L2 TCR transgenic mice. Immunization with these two peptides demonstrated that they stimulated essentially distinct bulk populations of T cells, and the crystal structure demonstrated that the single amino acid change in the P6 pocket resulted in minor differences in the conformation of the P6, P7 and P8 peptide segment. Small changes in MHC anchor residues can thus contribute to the specificity of TCR recognition through subtle changes of peptide conformation.

In a second study, T cells expressing an engineered high-affinity 3.L2 TCR (25 nM) remarkably had unchanged antigen sensitivity and retained antigen specificity [23]. Functional testing of a large set of single substituted TCR contact residue Hb peptides revealed a large increase in the number of stimulatory peptides. This apparent discrepancy between overall antigen specificity of the high-affinity TCR and the large number of recognized single substituted peptides was resolved when chimeric peptides between Hb and non-stimulatory moth cytochrome peptides were tested, showing that MHC anchor residues significantly affected TCR recognition. Together, these two studies highlight how small changes in MHC anchor residues can contribute to the specificity of TCR recognition through subtle changes of peptide conformation.

4.4. Receptor binding of multiple distinct ligands without substantial conformational changes

Several crystal structures have demonstrated substantial conformational differences between free and MHC-bound TCR, indicating that flexibility of TCR loops may contribute to recognition of distinct ligands [24,25]. The CDR3 loops are longer than the germline encoded CDR1 and CDR2 loops and can contain multiple glycine residues, enhancing rotational freedom. However, such flexibility/plasticity may not be the only structural explanation for polyspecificity.

An example of degenerate recognition in which the receptor represents a largely rigid structure is NKG2D, an activating receptor expressed by NK cells and subsets of T cells. Human NKG2D binds to MIC-A, MIC-B and several UL16 binding proteins (ULBP) which have a MHC-like fold, but do not bind peptide or β_2-microglobulin. Degenerate recognition is observed at several different levels. First, each NKG2D monomer binds to one of the MIC-A α-helices and the MIC-A residues contacted by the two NKG2D monomers are largely distinct. Comparison of structures of human NKG2D bound to MIC-A or ULBP3 demonstrates a second level of degeneracy. Two tyrosine residues of NKG2D make important contacts at each of the NKG2D monomer interfaces with MIC-A and ULBP3, but recognize an almost completely different set of amino acids on their interaction partners. Nevertheless, the conformation of the NKG2D binding surface is similar between these two structures [26].

5. Thymic selection determines the degree of TCR specificity and degeneracy

To address the impact of thymic negative selection on TCR specificity, $IA^b + 3\,K$ reactive T cells were isolated from conventional C57BL/6 mice and mice severely deficient in negative selection. T cells from normal C57BL/6 mice were very sensitive to amino acid substitutions at many positions of the peptide and MHC. In contrast, T cells from negative selection limited mice had a much wider range of specificity requirements [27]. Some T cells were biased in needing specific amino acids of the $IA^b\alpha$ chain and not the peptide or $IA^b\beta$ chain, while some were biased in needing specific amino acids of the $IA^b\beta$ chain and not the peptide or $IA^b\alpha$ chain. In addition to these "off center" T cells, some T cells required very few side chains, while another set had similar specificity requirements as T cells from C57BL/6 mice. Many of the T cells from negative selection limited mice were also highly alloreactive with some responding to most allogeneic MHC they were challenged with. Furthermore, transgenic mice expressing two of the TCRs derived from

Please cite this article in press as: Wucherpfennig KW, et al., Polyspecificity of T cell and B cell receptor recognition, Semin Immunol (2007), doi:10.1016/j.smim.2007.02.012

negative selection limited mice underwent positive selection on MHC class I and class II proteins, and CD8 T cells from both of these mice were activated by peptides bound to the MHC class I molecule H2-K^b. Thus, in the absence of proper negative selection, TCRs can react with MHC proteins in a class- and allele-independent fashion.

Finding extremely polyspecific TCRs led to the question of what makes a TCR specific or degenerate. Detailed binding studies demonstrated that highly degenerate TCRs do not differ from conventional TCRs in either binding kinetics or equilibrium binding affinity to their peptide/MHC ligand. Rather, the major indicator of TCR cross-reactivity is the number of peptide and MHC side chains that contribute ≥1.5 kcal/mol of binding energy. For example, a MHC class, allele and peptide cross-reactive TCR used two side chains of the MHC + peptide to contribute ≥1.5 kcal/mol of binding energy, while at least seven side chains of the MHC + peptide complex contributed this much energy for conventional TCRs [28]. The loss of 1.5 kcal/mol of binding energy would convert a high affinity 10 μM affinity interface to a 125 μM affinity interface, approximately the limit of affinity for T cell activation. These quantitative studies strongly suggest that a larger fraction of the binding energy for degenerate TCRs is derived from conserved structural features of MHC and peptide. In addition, it suggests that negative selection functions to eliminate T cells that have the highest degree of MHC and peptide degeneracy and thus biases the repertoire towards recognition of peptide side chains.

6. Receptor editing by autoreactive B cells

Similar to immature T cells, a large fraction of immature B cells is self-reactive, suggesting that most randomly generated antigen receptors are autoreactive. This high degree of autoreactivity was predicted several years ago based on mathematical modeling [29]. Cloning of antibodies from immature human B cells and analysis of their reactivity has shown that ∼55–75% of developing B cells exhibit reactivity to several autoantigens. If tolerance were maintained solely by clonal elimination mechanisms, massive cell loss would result. However, self-tolerance can be brought about indirectly by receptor editing, in which secondary rearrangements lead to the generation of B cells carrying non-autoreactive receptors [30,31]. There is a division of labor between antibody heavy and light chain genes: the heavy chain locus rearranges first and generates substantial diversity, while the absence of diversity (D) segments in the light chain locus facilitates editing by permitting direct joining of new variable (V) and joining (J) segments. Editing displays a low threshold affinity in anti-MHC transgenic mice (3–83) [32], a finding that may explain why editing appears to occur in substantial numbers of immature B cells. Editing may also occur during T cell development, but has not been studied as extensively as in B cells [33].

7. Polyspecificity and disease—Autoimmunity and viral infections

Several examples of self-reactive T cells that recognize a series of microbial peptides were already described above. These T cell clones were selected using a particular antigen and thus represent only one facet of the self-reactive repertoire. CD4 T cells that proliferated in response to self-peptide/MHC complexes in the absence of deliberate addition of antigen were therefore cloned from normal donors. T cells that proliferated based on loss of CFSE fluorescence ($CSFE^{low}$) were single cell sorted and the resulting clones were tested for their ability to respond to a panel of self (MBP, GAD65, insulin, proinsulin) or microbial (tetanus toxoid, mumps virus) antigens. In total, ∼0.04% of CD4 T cells proliferated to self-peptide/MHC and entry into cell cycle was dependent upon CD28 co-stimulation. A substantial subset of the $CFSE^{low}$ clones (14/545, 2.5%) responded to three of these antigens and thus exhibited a high degree of polyspecificity. For example, such clones responded to GAD65, insulin, proinsulin and mumps virus or other combinations of tested antigens. Proliferation of $CFSE^{low}$ clones recognizing multiple antigens was in almost all instances blocked by an HLA-DR antibody. In contrast, none of the non-proliferating cells ($CFSE^{high}$ 0/73 clones) exhibited such polyspecificity. These results demonstrate that CD4 T cells capable of recognizing multiple distinct antigens are present in the T cell repertoire of normal individuals [34].

In aggregate, the results discussed at the workshop demonstrated many examples of TCR cross-reactivity between foreign and self-antigens, and the results described above illustrate that such T cells are present in the repertoire of normal subjects. Given such degeneracy of TCR recognition, which factors limit the development of autoimmunity? Studies with the N-terminal MBP epitope demonstrated that peptide length is an important factor. Even though all major MHC and TCR contacts are contained in the Ac1–6 peptide, a longer peptide (Ac1–11) is substantially more effective in inducing EAE. Increasing the affinity of the Ac1–6 peptide by substitution of position 4 (Lys to Met) does not compensate for the loss of the C-terminal 7–11 SQRSK peptide segment and no disease is observed in Ac1–6 (M4) immunized mice. This may be owing to an activation of Ac1–6 (M4) specific T cells that tend to competitively exclude the pathogenic clones (E. Maverakis et al., submitted for publication). Studies with combinatorial peptide libraries discussed above also demonstrated a significant contribution of the C-terminal segment of the Ac1–11 peptide to TCR recognition [19], but the structural mechanisms are not yet understood. A second consideration is that even though different self- and microbial peptides may stimulate the same T cell clones, immunization with these peptides may lead to the expansion of T cell populations that are largely distinct, similar to the results with the Hb peptide and its APL described above.

Polyspecificity is also relevant in the context of viral infections, due to extensive T cell cross-reactivity between antigens, even from apparently unrelated viruses. Prior infections can thus shape the T cell repertoire and reshuffle the clonal hierarchy, affecting which T cell specificities become dominant in subsequent encounters with other pathogens [46]. Over a lifetime of viral encounters, it can be expected that the polyspecific memory clones will preferentially survive.

Please cite this article in press as: Wucherpfennig KW, et al., Polyspecificity of T cell and B cell receptor recognition, Semin Immunol (2007), doi:10.1016/j.smim.2007.02.012

8. Quantitative aspects of polyspecificity and mathematical models

8.1. Quantitative assessment of the peptide repertoire presented to CD8 T cells

The availability of complete genome sequences and knowledge of the requirements for peptide processing by the proteasome, peptide transport by TAP and peptide binding by MHC molecules now permits a quantitative assessment of the number of peptides in human and microbial proteomes that can be presented to CD8 T cells. MHC class I molecules present 9-mer peptides (8–11 mers) and analysis of the human proteome (~30,000 proteins) demonstrated that the majority of 9-mers (76%) are unique in sequence, resulting in a total of $\sim 10^7$ unique 9-mer peptides. The majority of amino acids in human proteins are thus the starting point for a novel 9-mer. Processing and transport reduce the number of available peptides to ~25% (2.5×10^6 9-mers), with the proteasome being more selective than the TAP transporter. Particular alleles of MHC class I molecules are estimated to present $\sim 10^5$ distinct human peptides (in the aggregate, not in one particular cell type). The total number of 9-mers is smaller in any microbial proteome (ranging from $\sim 3.5 \times 10^5$ to 1.46×10^6 9-mers for different bacterial organisms that were studied, and from only 677 to 4946 9-mers for the analyzed viruses). Tolerance is thus induced by sets of self-peptides that are considerably more diverse than the set of microbial peptides encountered during a particular infection.

An important question is whether 9-mers are of sufficient length to permit discrimination between human and microbial peptides. The probability that a foreign peptide also occurs in the human self is about 0.2% for 9-mers. In contrast, 4-mers have no discriminating ability, and the overlap decreases to 30% for 6-mers and 3% for 7-mers. Not all residues of a 9-mer are directly contacted by the TCR, but peptide binding to MHC class I molecules contributes to specificity of recognition. Due to specificity of MHC molecules for the two primary anchor residues, the overlaps between exposed 7-mers remain well below 1%. The 9-mers used in MHC class I presentation thus tend to carry sufficient information to detect non-self among self-peptides [35].

This analysis also indicates that there are a substantial number of microbial peptides that are identical in their 9-mer sequence to human peptides (a total of ~2340 9-mer peptides for a bacterium with 1.46×10^6 9-mers; proteasome and TAP transporter constraints are expected to reduce the number to ~405 peptides). Only a small fraction of such peptides is expected to be presented by a given MHC molecule (a few percent), but due the diversity of MHC class I molecules in the population, a subset of such peptides is likely to be presented to T cells.

8.2. Detection of all epitopes requires polyspecificity of recognition

Recognition in the immune system is based on shape complementarity between B cell and T cell receptors and their ligands. The set of ligands that stimulate a B cell or T cell can be represented in shape space, and the region of space shape sensed by a particular receptor is called a recognition region. Because the shape, charge and hydrophobicity of epitopes and immune receptors are all finite, the shape space is bounded. If each receptor can only recognize a single complementary shape, then the immune system would in principle need one receptor for every possible epitope, which would require orders of magnitude more receptors than there are cells in the immune system. However, if each receptor recognized a set of nearby shapes (meaning that it is cross-reactive or degenerate) then a considerably smaller number of receptors would be required to give an essentially complete repertoire. The fraction of shape space that a recognition region covers is the equivalent of the probability that a B cell or T cell recognizes and responds to an epitope, the precursor frequency, which is about 10^{-5}. Thus, repertoires with 10-fold coverage, i.e. of size 10^6, may be sufficient to recognize the majority of epitopes. The major conclusion of this theory is that for clonal selection to work, each receptor/antibody must be degenerate and recognize a region of shape space. Current estimates suggest that each receptor responds to approximately 10^{-5} of all epitopes. The number of 9-mer peptides is $20^9 = 5 \times 10^{11}$, implying that each T cell may recognize a large number of different peptides. Thus, polyspecificity must be the rule, not the exception in immune recognition [36,37].

In biology more generally, degeneracy is related to three fundamentally important issues: one is uncertainty about the structure of a target sequence, another is limited storage capacity for the recognition elements, and the third is noise in the recognition process. Uncertainty about structure relates to high levels of environmental variability. Limiting coding relates to the fact that host genomes are finite. And noise means that even with the perfect receptor–epitope pairing, recognition and activation might not take place. Degeneracy, by promoting overlap among receptors, provides partial solutions to all three problems. A specificity/degeneracy trade off is also observed in the genetic code. The genetic code is degenerate as multiple codons are associated with the same amino acid [38].

9. Specificity versus degeneracy

There are now numerous examples of polyspecificity of TCR recognition, as discussed above, both for human and murine T cells, CD4 and CD8 T cell populations and for T cells initially selected with a microbial or self-antigen. At the same time, there is abundant evidence for an exquisite degree of TCR specificity because minor changes in a peptide can result in loss of TCR recognition. Such changes can be localized either to peptide residues that directly contact TCR loops or to peptide residues that occupy pockets of the MHC binding groove. There is also considerable evidence that particular combinations of peptide residues are required because many peptides assembled from motifs based on single amino acid substitutions are not active.

Thus, the question arises as to whether specificity and degeneracy are contradictions. The workshop concluded that both specificity and degeneracy are inherent properties of TCR recognition. T cells are specific because they recognize a *small fraction* of all possible ligands (i.e. 1 per 10^5–10^6 peptides),

Please cite this article in press as: Wucherpfennig KW, et al., Polyspecificity of T cell and B cell receptor recognition, Semin Immunol (2007), doi:10.1016/j.smim.2007.02.012

but are degenerate because the number of potential ligands is very large (5×10^{11} possible 9-mer sequences for CD8 T cells) [39].

10. Complexity in T cell biology and other biological systems

A T cell recognizes multiple different peptide/MHC ligands during its lifetime: these include self-peptide/MHC complexes during positive selection in the thymus, an outcome illustrative of the utility of degeneracy; endogenous self-peptide/MHC complexes that contribute to survival of a naïve T cell in the periphery; and potentially, microbial peptides during the course of an infection. In the latter case, the microbial peptides may activate a self-recognizing cell and cause autoimmunity if several criteria are met. Furthermore, there is considerable evidence that endogenous self-peptide/MHC complexes contribute to TCR triggering and that such weak ligands reduce the number of agonist peptide/MHC complexes required for initiation of signaling [40,41]. Given this complexity of peptide/MHC recognition, it is evident that T cell selection, survival in the competitive milieu, and activation are the sum of signaling events induced by multiple, distinct peptide/MHC complexes.

Specificity of discrimination can thus not be solely achieved at the level of single receptor–ligand interactions. The functional specificity of an immune response emerges from the integration of multiple signals by a collective of interacting cells (i.e. recognition of multiple peptides from a pathogen, recognition of pathogen-derived structures by TLR) and the interactions between multiple cell types of the immune system (i.e. T cell–B cell collaboration, effector and regulatory T cells) [42].

There is also a growing realization that biological systems are not adequately described by widely used concepts, such as linear pathway models in which a series of proteins execute a biological program in a stepwise fashion [43]. Greenspan reported an experiment in which he asked whether any behavioral phenotype could be affected from anywhere within the genome. In a matrix experiment, eight genes were examined that were expressed in the central nervous system and that covered a wide range of different functions (ranging from a G-protein coupled receptor to transcription factors). These eight genes were tested in pairwise combinations in a *heterozygous* state (so that the experiment would not simply reflect a complete loss of function), and animals were tested for eight different behaviors that had not been previously associated with these genes (such as circadian rhythm, sleep, locomotor activity, etc). Many different combinations yielded large phenotypic effects, a result that challenges the linear pathway model. Such complexity may be an inherent feature of biological systems.

11. The nomenclature discussion

Many different terms have been used to describe the finding that T cells can recognize multiple distinct peptide/MHC ligands, including molecular mimicry, flexibility, plasticity, cross-reactivity and degeneracy (Table 1). *Molecular mimicry* was widely used during the early stages of research in this field. However, it only relates to the specialized case in which T cells recognize both peptides from self- and microbial antigens and can thus not serve as a general term that covers the many other biologically relevant aspects (thymic selection, alloreactivity, etc). *Flexibility and plasticity* suggest particular structural mechanisms in which movement of TCR loops permits recognition of alternate ligands. Examples of such TCR flexibility have been reported, but it is not known whether this is the only structural mechanism. Examples from other receptor systems discussed above (i.e. NKG2D) suggest that binding of structurally diverse ligands does not always require large conformational changes of the receptor. The term *cross-reactivity* does not imply a particular structural mechanism and mainly refers to relationships among ligands. At the T cell recognition level it can be misunderstood to imply that there is a single primary ligand for a T cell (the "cognate" ligand) and that all other ligands merely represent cases of cross-reaction. In fact, all available evidence suggests that there is not a single "real" peptide ligand for a TCR, but rather a more or less diverse group of ligands.

Several workshop participants expressed concerns regarding the term *degeneracy*, in part because it is used in many different contexts. The term degeneracy is appropriate for describing peptide binding to MHC molecules because a large variety of peptides can be bound. The term degeneracy is particularly suitable for MHC class II proteins because many changes in the peptide are tolerated, even at the major anchor residues. When applied to TCR and antibody recognition, degeneracy can be misunderstood to imply that recognition is "sloppy", even though TCRs and antibodies discriminate among a larger number of ligands than any other known receptor systems. In the course of the discussion, it thus became apparent that an alternative term is required that simultaneously captures both key aspects: the ability to recognize multiple ligands as well as the specificity with which each of these ligands is recognized. The term *polyspecificity* captures the essence of both concepts and reinforces the major conclusion of the workshop that specific recognition of multiple distinct peptide/MHC ligands is an inherent property of TCR recognition (Table 1). This term has already been used to describe recognition of distinct ligands by the same antibody [44,45].

12. Summary

The substantial progress in the field reported at this meeting report leads to the conclusion that recognition of multiple peptide/MHC ligands by each TCR is a general, inherent property of this receptor system relevant for many different aspects of T cell biology (Table 2). The workshop recommends general use of the term polyspecificity because it simultaneously captures the two essential features, the recognition of multiple peptide/MHC ligands as well as the remarkable ability of TCRs to distinguish among many structurally related ligands. The number of potential peptides with unique sequence far exceeds the number of T cells in the immune system and the ability to recognize multiple

Please cite this article in press as: Wucherpfennig KW, et al., Polyspecificity of T cell and B cell receptor recognition, Semin Immunol (2007), doi:10.1016/j.smim.2007.02.012

ligands is thus required for reasonably complete coverage of all potential pathogen-derived peptides. Negative selection eliminates those T cells with the highest level of degeneracy and thus imparts T cells with the required degree of specificity.

Acknowledgments

The authors would like to thank the Santa Fe Institute and Wayne Cote for hosting this workshop and Lindsey Harvey for help in organizing the workshop and this manuscript. The work in this paper was supported by NIH Grant # R13 AI068354-01, Human Frontiers Science Program Grant # RGP0010/2004 and the Santa Fe Institute.

References

[1] Bhardwaj V, Kumar V, Geysen HM, Sercarz EE. Degenerate recognition of a dissimilar antigenic peptide by myelin basic protein-reactive T cells. Implications for thymic education and autoimmunity. J Immunol 1993;151:5000–10.
[2] Evavold BD, Sloan-Lancaster J, Wilson KJ, Rothbard JB, Allen PM. Specific T cell recognition of minimally homologous peptides: evidence for multiple endogenous ligands. Immunity 1995;2:655–63.
[3] Wucherpfennig KW, Strominger JL. Molecular mimicry in T cell-mediated autoimmunity: viral peptides activate human T cell clones specific for myelin basic protein. Cell 1995;80:695–705.
[4] Gautam AM, Lock CB, Smilek DE, Pearson CI, Steinman L, McDevitt HO. Minimum structural requirements for peptide presentation by major histocompatibility complex class II molecules: implications in induction of autoimmunity. Proc Natl Acad Sci USA 1994;91:767–71.
[5] Hammer J, Valsasnini P, Tolba K, Bolin D, Higelin J, Takacs B, et al. Promiscuous and allele-specific anchors in HLA-DR-binding peptides. Cell 1993;74:197–203.
[6] Stern LJ, Brown JH, Jardetzky TS, Gorga JC, Urban RG, Strominger JL, et al. Crystal structure of the human class II MHC protein HLA-DR1 complexed with an influenza virus peptide. Nature 1994;368:215–21.
[7] Wucherpfennig KW, Sette A, Southwood S, Oseroff C, Matsui M, Strominger JL, et al. Structural requirements for binding of an immunodominant myelin basic protein peptide to DR2 isotypes and for its recognition by human T cell clones. J Exp Med 1994;179:279–90.
[8] Ausubel LJ, Kwan CK, Sette A, Kuchroo V, Hafler DA. Complementary mutations in an antigenic peptide allow for crossreactivity of autoreactive T-cell clones. Proc Natl Acad Sci USA 1996;93:15317–22.
[9] Nino-Vasquez JJ, Allicotti G, Borras E, Wilson DB, Valmori D, Simon R, et al. A powerful combination: the use of positional scanning libraries and biometrical analysis to identify cross-reactive T cell epitopes. Mol Immunol 2004;40:1063–74.
[10] Pinilla C, Martin R, Gran B, Appel JR, Boggiano C, Wilson DB, et al. Exploring immunological specificity using synthetic peptide combinatorial libraries. Curr Opin Immunol 1999;11:193–202.
[11] Hemmer B, Gran B, Zhao Y, Marques A, Pascal J, Tzou A, et al. Identification of candidate T-cell epitopes and molecular mimics in chronic Lyme disease. Nat Med 1999;5:1375–82.
[12] Chen J, Eisen HN, Kranz DM. A model T-cell receptor system for studying memory T-cell development. Microbes Infect 2003;5:233–40.
[13] Udaka K, Tsomides TJ, Eisen HN. A naturally occurring peptide recognized by alloreactive CD8+ cytotoxic T lymphocytes in association with a class I MHC protein. Cell 1992;69:989–98.
[14] Garboczi DN, Ghosh P, Utz U, Fan QR, Biddison WE, Wiley DC. Structure of the complex between human T-cell receptor, viral peptide and HLA-A2 [comment]. Nature 1996;384:134–41.
[15] Garcia KC, Degano M, Stanfield RL, Brunmark A, Jackson MR, Peterson PA, et al. An alphabeta T cell receptor structure at 2.5 A and its orientation in the TCR–MHC complex. Science 1996;274:209–19.
[16] Garcia KC, Adams EJ. How the T cell receptor sees antigen—a structural view. Cell 2005;122:333–6.
[17] Hennecke J, Wiley DC. T cell receptor–MHC interactions up close. Cell 2001;104:1–4.
[18] Urban JL, Kumar V, Kono DH, Gomez C, Horvath SJ, Clayton J, et al. Restricted use of T cell receptor V genes in murine autoimmune encephalomyelitis raises possibilities for antibody therapy. Cell 1988;54:577–92.
[19] Maynard J, Petersson K, Wilson DH, Adams EJ, Blondelle SE, Boulanger MJ, et al. Structure of an autoimmune T cell receptor complexed with class II peptide–MHC: insights into MHC bias and antigen specificity. Immunity 2005;22:81–92.
[20] Hahn M, Nicholson MJ, Pyrdol J, Wucherpfennig KW. Unconventional topology of self peptide-major histocompatibility complex binding by a human autoimmune T cell receptor. Nat Immunol 2005;6:490–6.
[21] Madsen LS, Andersson EC, Jansson L, Krogsgaard M, Andersen CB, Engberg J, et al. A humanized model for multiple sclerosis using HLA-DR2 and a human T-cell receptor. Nat Genet 1999;23:343–7.
[22] Kersh GJ, Miley MJ, Nelson CA, Grakoui A, Horvath S, Donermeyer DL, et al. Structural and functional consequences of altering a peptide MHC anchor residue. J Immunol 2001;166:3345–54.
[23] Donermeyer DL, Weber KS, Kranz DM, Allen PM. The study of high-affinity TCRs reveals duality in T cell recognition of antigen: specificity and degeneracy. J Immunol 2006;177:6911–9.
[24] Garcia KC, Degano M, Pease LR, Huang M, Peterson PA, Teyton L, et al. Structural basis of plasticity in T cell receptor recognition of a self peptide–MHC antigen. Science 1998;279:1166–72.
[25] Housset D, Malissen B. What do TCR-pMHC crystal structures teach us about MHC restriction and alloreactivity? Trends Immunol 2003;24:429–37.
[26] McFarland BJ, Strong RK. Thermodynamic analysis of degenerate recognition by the NKG2D immunoreceptor: not induced fit but rigid adaptation. Immunity 2003;19:803–12.
[27] Huseby ES, White J, Crawford F, Vass T, Becker D, Pinilla C, et al. How the T cell repertoire becomes peptide and MHC specific. Cell 2005;122:247–60.
[28] Huseby ES, Crawford F, White J, Marrack P, Kappler JW. Interface-disrupting amino acids establish specificity between T cell receptors and complexes of major histocompatibility complex and peptide. Nat Immunol 2006;7:1191–9.
[29] Nemazee D. Antigen receptor 'capacity' and the sensitivity of self-tolerance. Immunol Today 1996;17:25–9.
[30] Tiegs SL, Russell DM, Nemazee D. Receptor editing in self-reactive bone marrow B cells. J Exp Med 1993;177:1009–20.
[31] Gay D, Saunders T, Camper S, Weigert M. Receptor editing: an approach by autoreactive B cells to escape tolerance. J Exp Med 1993;177:999–1008.
[32] Lang J, Jackson M, Teyton L, Brunmark A, Kane K, Nemazee D. B cells are exquisitely sensitive to central tolerance and receptor editing induced by ultralow affinity, membrane-bound antigen. J Exp Med 1996;184:1685–97.
[33] McGargill MA, Derbinski JM, Hogquist KA. Receptor editing in developing T cells. Nat Immunol 2000;1:336–41.
[34] Cai G, Hafler DA. Multispecific responses by T cells expanded by endogenous self-peptide/MHC complexes. Eur J Immunol 2007;37:602–12.
[35] Burroughs NJ, de Boer RJ, Kesmir C. Discriminating self from non-self with short peptides from large proteomes. Immunogenetics 2004;56:311–20.
[36] Perelson AS, Oster GF. Theoretical studies of clonal selection: minimal antibody repertoire size and reliability of self-non-self discrimination. J Theor Biol 1979;81:645–70.
[37] de Boer RJ, Perelson AS. How diverse should the immune system be? Proc R Soc Lond Ser B 1993;252:171–5.
[38] Krakauer DC, Jansen VA. Red queen dynamics of protein translation. J Theor Biol 2002;218:97–109.
[39] Borghans JAM, de Boer RJ. Crossreactivity of the T-cell receptor. Immunol Today 1998;19:428–9.

Please cite this article in press as: Wucherpfennig KW, et al., Polyspecificity of T cell and B cell receptor recognition, Semin Immunol (2007), doi:10.1016/j.smim.2007.02.012

[40] Stefanova I, Dorfman JR, Germain RN. Self-recognition promotes the foreign antigen sensitivity of naive T lymphocytes. Nature 2002;420:429–34.
[41] Krogsgaard M, Li QJ, Sumen C, Huppa JB, Huse M, Davis MM. Agonist/endogenous peptide–MHC heterodimers drive T cell activation and sensitivity. Nature 2005;434:238–43.
[42] Cohen IR. Tending Adam's garden: evolving the cognitive immune self. San Diego, CA: Academic Press; 2000.
[43] van Swinderen B, Greenspan RJ. Flexibility in a gene network affecting a simple behavior in *Drosophila melanogaster*. Genetics 2005;169:2151–63.
[44] Keitel T, Kramer A, Wessner H, Scholz C, Schneider-Mergener J, Hohne W. Crystallographic analysis of anti-p24 (HIV-1) monoclonal antibody cross-reactivity and polyspecificity. Cell 1997;91:811–20.
[45] Shoenfeld Y, Rauch J, Massicotte H, Datta SK, Andre-Schwartz J, Stollar BD, et al. Polyspecificity of monoclonal lupus autoantibodies produced by human–human hybridomas. N Engl J Med 1983;308:414–20.
[46] Selin LK, Cornberg M, Brehm MA, Kim SK, Calcagno C, Ghersi D, et al. CD8 memory T cells: cross-reactivity and heterologous immunity. Semin Immunol 2004;16:335–47.

Please cite this article in press as: Wucherpfennig KW, et al., Polyspecificity of T cell and B cell receptor recognition, Semin Immunol (2007), doi:10.1016/j.smim.2007.02.012