Molecular Immunology 85 (2017) 66–80

Contents lists available at ScienceDirect

Molecular Immunology

journal homepage: www.elsevier.com/locate/molimm

A new mechanism shapes the naïve CD8⁺ T cell repertoire: the selection for full diversity

Pedro Gonçalves[a,b,*,1], Marco Ferrarini[c], Carmen Molina-Paris[c], Grant Lythe[c], Florence Vasseur[a,b], Annik Lim[a], Benedita Rocha[a,b,*], Orly Azogui[b]

[a] Lymphocyte Population Biology Unit, CNRS URA 196, Institut Pasteur, Paris, France
[b] INSERM, U1151, CNRS, UMR8253, Faculté de Médecine Paris Descartes, Paris, France
[c] Department of Applied Mathematics, University of Leeds, Leeds LS29JT, UK

ARTICLE INFO

Article history:
Received 4 September 2016
Received in revised form 16 January 2017
Accepted 28 January 2017
Available online 8 March 2017

Keywords:
CD8⁺ T cells
Mice TCR repertoires
CDR3 sequences
TCR diversity
TCR cross-reactivity

ABSTRACT

During thymic T cell differentiation, TCR repertoires are shaped by negative, positive and agonist selection. In the thymus and in the periphery, repertoires are also shaped by strong inter-clonal and intra-clonal competition to survive death by neglect. Understanding the impact of these events on the T cell repertoire requires direct evaluation of TCR expression in peripheral naïve T cells. Several studies have evaluated TCR diversity, with contradictory results. Some of these studies had intrinsic technical limitations since they used material obtained from T cell pools, preventing the direct evaluation of clonal sizes. Indeed with these approaches, identical TCRs may correspond to different cells expressing the same receptor, or to several amplicons from the same T cell. We here overcame this limitation by evaluating TCRB expression in individual naïve CD8⁺ T cells. Of the 2269 *Tcrb* sequences we obtained from 13 mice, 99% were unique. Mathematical analysis of the data showed that the average number of naïve peripheral CD8⁺ T cells expressing the same TCRB is 1.1 cell. Since TCRA co-expression studies could only increase repertoire diversity, these results reveal that the number of naïve T cells with unique TCRs approaches the number of naïve cells. Since thymocytes undergo multiple rounds of divisions after TCRB rearrangement and 3–5% of thymocytes survive thymic selection events the number of cells expressing the same TCRB was expected to be much higher. Thus, these results suggest a new repertoire selection mechanism, which strongly selects for full TCRB diversity.

1. Introduction

The immune system is known to have Promethean properties, i.e., to be able to recognize all types of natural and artificial antigens introduced in the organism. How much this remarkable capability depends on the diversity or on the cross-reactivity of peripheral T cell repertoires is still a subject of debate.

The antigen specificity of CD8⁺ T cells is determined by a dimer of TCRB and TCRA chains, which binds peptides presented by the major histocompatibility class I complex (pMHC). The TCRB and TCRA chains have three regions of hypervariability, the complementarity determining regions (CDR). The CDR1 and CDR2 loops are encoded by the germline V gene segment, while the CDR3 loop is created by V(D)J recombination (Von Boehmer, 2004). The *Tcrb* rearrangements begin at the $CD44^{low}CD25^{+}CD4^{-}CD8^{-}TCR^{-}$ (triple negative 3-TN3) thymocyte differentiation stage by recombining one of each of 35 TRBV, 2 TRBD and 12 TRBJ genes in mice (Lefranc, 2001). A semi-random cleavage of the recombination hairpins intermediates results in nibbling at the V-D-J junctions. These events, and the further addition of N and P nucleotides, result in a major increase in CDR3 diversity. TCR diversity studies are often focused on *Tcrb* CDR3 region because this region is the most diverse portion of the TCR and functional/crystallographic analysis shows that the interaction between the pMHC complex and the TCR is predominantly mediated via this region (Das et al., 2015). The expression of an in-frame TCRB induces allelic exclusion, a burst of

Abbreviations: BM, bone marrow; CDR3, complementarity determining region 3; "clonotypes", cells expressing identical *Tcrb* chains; HP, homeostatic proliferation; LCMV, Lymphocytic Choriomeningitis Virus; LN, lymph node; MHC, major histocompatibility complex; MoAbs, monoclonal antibodies; Ms., manuscript; SP, spleen; SPF, specific-pathogen-free mice; TCR, T cell receptor.

* Corresponding authors at: Lymphocyte Population Biology Unit, CNRS URA 196, Institut Pasteur, Paris, France.
E-mail addresses: pedro.goncalves@pasteur.fr (P. Gonçalves), benedita.rocha@inserm.fr (B. Rocha).
[1] Present address: Innate Immunity Unit, INSERM, U668, Institut Pasteur, Paris, France.

http://dx.doi.org/10.1016/j.molimm.2017.01.026
0161-5890/

6–8 cell divisions (Kreslavsky et al., 2012; Penit et al., 1995; Penit and Vasseur, 1997; Von Boehmer, 2004), and the transition to the ISP (immature single positive) and the $CD4^+CD8^+$ (DP) thymocyte differentiation stages. Thymocytes divide at the ISP and at least 3 times at the DP stage (Kreslavsky et al., 2012).

A diversity of the peripheral T cell pools also depends on *Tcra* rearrangements and TCRB/TCRA pairing. Indeed, in DP cells *Tcra* chains are rearranged by recombining one of each of 132 TRAV and 60 TRAJ fragments genes in mice (Genolet et al., 2012; Lefranc, 2001). The same events occurring during TCRB V-D-J recombination also occur during *Tcra* V-J rearrangements, inducing a major variability of the *Tcra* CDR3 region. Because the first in-frame TCRA chain may not pair efficiently to the expressed TCRB chain, DP cells have the ability to rearrange multiple *Tcra* chains until a compatible TCRAB dimer is formed. In theory, these events could generate a potential repertoire of more than 10^{15} different TCRs (Von Boehmer, 2004).

It is estimated that in the mouse thymus around 5×10^7 $TCR\alpha b^+$ thymocytes are generated each day (Shortman and Jackson, 1974; Von Boehmer, 2004). These cells undergo negative, positive and agonist selection, or may die by neglect. Lastly, thymic egress is also restricted (Von Boehmer, 2004). Each of these selection events, as well as peripheral T cell survival, is not solely dependent on the TCR-ligand interactions of each individual cell, but is strongly influenced by competition between different T cells (Freitas and Rocha, 2000; Hao et al., 2006). Because it cannot be deduced, in a non-manipulated mouse, from current understanding of these multiple selecting events, the diversity of peripheral T cell repertoires, must be measured directly.

Several studies addressed this issue, with contradictory results. In the mouse, it was claimed that 10% (Carey et al., 2016), 28% (Casrouge et al., 2000), 55% (Peaudecerf et al., 2012) or 68% (Quigley et al., 2010) of *Tcrb* chains were unique. In humans, the reported number of unique TCRs ranges from 10^6 to 2×10^7 (Arstila et al., 1999, 2000; Qi et al., 2014; Robins et al., 2009). These differences could be partially due to differences in the T cell populations studied, or/and by bias introduced by the methodology used to evaluate diversity (PCR amplification, followed by spectrotyping and cloning *versus* PCR amplification followed by next-generation sequence analysis). However, these approaches had common as well as specific technical limitations. None allows the evaluation of sequencing efficiency, i.e., to identify the number of T cells that had their TCR actually amplified. None can exclude bias introduced by primer competition or/and, in next-generation sequencing, by filters used to eliminate potential PCR errors. In most studies these selection filters are not defined. In one study where selected filters were fully described reported that up to 50% of the TCR sequences were eliminated (Nguyen et al., 2011). Apart from these biases, the successive preparation steps may reduce putative diversity by preferentially selecting more abundant TCRs. Importantly, bulk studies are unable to identify the number and the size of different clones. After PCR amplification, it is uncertain if identical TCR sequences correspond to multiple cells sharing the same TCR or to multiple amplicons from the TCR of a single cell. Lastly, in several cases it is not clear how representative the sample was, with respect to the total number of T cells belonging to the same population. To overcome these limitations, TCR expression must be determined in single cells.

Several recent studies used single-cell approaches to determine TCR expression of total naïve or naïve antigen-specific cells from non-immunized mice (Cukalac et al., 2015; Eltahla et al., 2016; Quinn et al., 2016; Stubbington et al., 2016). All these studies report a higher diversity of *Tcrb* expression than that determined by bulk studies. In particular, Quinn et al. studied over 300 T cells (15–72 cells mouse) specific for a peptide of the influenza virus using tetramers (Quinn et al., 2016). They observed that this naïve repertoire is almost completely diverse but did not determine overall TCRB repertoires. Rigorous analysis of repertoire diversity will require studies in which multiple mice are studied and the number of cells sequenced in each mouse is representative of the total population from that mouse. Over more than a decade, we have developed and validated the parameters required for quantifying the expression of multiple mRNAs in single cells (Peixoto et al., 2004), including primer design and concentrations required to prevent primer competition and the conditions of amplification allowing the detection of as little as 2mRNA/cell, while preventing saturation. We now used this experience to develop a single-cell approach allowing evaluating the *Tcrb* expression in single-cells. Here, we describe the evaluation of the repertoire diversity in single $CD8^+$ T cells from specific-pathogen-free (SPF) adult mice. In contrast to bulk cell approaches, this single-cell analysis allows evaluation of the sequencing efficiency since we directly determined the number of cells where an in-frame *Tcrb* chain was sequenced. Our approach prevents primer competition, since a single primer pair is used for the PCR amplification of the *Tcrb* in each individual cell. It allows direct evaluation of PCR errors, by sequencing simultaneously the *Tcrb* of monoclonal TCR-Tg single-cells expressing known *Tcrb* chains. Of the 2269 *Tcrb* chains we sequenced, 99% were unique. Mathematical analysis of representative samples indicate that, solely based on *Tcrb* expression, the average number of naïve $CD8^+$ T cells expressing the same *Tcrb* is 1.1 cells. This average "clonotype" size is unexpected, taking into account the number of divisions of TCRB expressing immature thymocytes. At the DN3 thymocyte differentiation stage, the expression of an in-frame TCRB induces allelic exclusion, a burst of 6–8 cell divisions (Kreslavsky et al., 2012; Penit et al., 1995; Penit and Vasseur, 1997; Von Boehmer, 2004), and the transition to the ISP (immature single positive) and the $CD4^+CD8^+$ (DP) thymocyte differentiation stages. Thymocytes divide at the ISP and at least 3 times at the DP stage (Kreslavsky et al., 2012). Since immature thymocytes undergo at least 11 divisions after *Tcrb* rearrangement (Kreslavsky et al., 2012; Penit et al., 1995; Penit and Vasseur, 1997; Von Boehmer, 2004), if all these divisions were productive up to 2^{11} cells expressing the same *Tcrb* could be generated. Even if only 3–5% survive thymus selection events (Huesmann et al., 1991), an average 62–102 cells should express the same *Tcrb*. Therefore, our results suggest that, superimposed on the known mechanisms of repertoire selection (negative, positive, agonist and death by neglect), a remarkable selection for full TCRB diversity also occurs.

2. Material and methods

2.1. Mice

Specific-pathogen-free C57BL/6 (B6) mice expressing the CD45.2 allotype marker, and CD45.1 $Rag2^{-/-}$ P14 (P14) mice expressing a transgenic TCR specific for LCMV epitope GP_{33-41} (GP33) backcrossed onto the $Rag2^{-/-}$ C57BL/6 (B6) background, were obtained from our breeding colonies at the Centre de Distribution, Typage et Archivage (CDTA, Orleans, France). All animal experiments were performed in accordance with National and European Commission guidelines for the care and handling of laboratory animals and were approved by the site ethical review committee.

2.2. The evaluation of T cell numbers in each $CD8^+$ T cell subpopulation

To eliminate blood-derived T cells, SPF B6 mice ($CD45.2^+$) were exsanguinated before organ removal. From each mouse the spleen (SP), 2 femurs and 40 lymph nodes (LNs) were removed from each