Multi-level Strategy for Identifying Proteasome-Catalyzed Spliced Epitopes Targeted by CD8+ T Cells during Bacterial Infection

Graphical Abstract

Authors
Anouk C.M. Platteel, Juliane Liepe, Kathrin Textoris-Taube, ..., Peter M. Kloetzel, Michele Mishto, Alice J.A.M. Sijts

Correspondence
michele.mishto@kcl.ac.uk (M.M.), e.j.a.m.sijts@uu.nl (A.J.A.M.S.)

In Brief
Proteasomes both degrade proteins and ligate generated products, creating "spliced peptides" composed of distant protein parts. Platteel et al. now describe a multi-level strategy for identifying proteasome-generated spliced T cell epitopes. This work suggests ways of defining spliced epitopes within any antigen of interest and to determine their immunological relevance.

Highlights
- Development of in-silico-based, multi-level strategy for identifying spliced epitopes
- Developed strategy identifies two spliced bacterium-derived CD8+ T cell epitopes
- Proteasome-catalyzed peptide splicing increases the pathogen-derived peptide pool
Multi-level Strategy for Identifying Proteasome-Catalyzed Spliced Epitopes Targeted by CD8+ T Cells during Bacterial Infection

Anouk C.M. Platteel,1,8 Juliane Liepe,2,3,8 Kathrin Textoris-Taube,4,5,6 Christin Keller,4,5 Petra Henklein,4 Hanna H. Schalkwijk,1 Rebeca Cardoso,1 Peter M. Kloetzel,4,5 Michele Mishto,4,5,7,8,* and Alice J.A.M. Sijts1,9,10,*

1Division of Immunology, Faculty of Veterinary Medicine, Utrecht University, 3571 EK Utrecht, the Netherlands
2Centre for Integrative Systems Biology and Bioinformatics, Department of Life Sciences, Imperial College London, SW7 2AZ London, UK
3Max-Planck-Institute for Biophysical Chemistry, 37077 Göttingen, Germany
4Berlin Institute of Health, 10117 Berlin, Germany
5Berlin Institute of Health, 10117 Berlin, Germany
6Shared Facility for Mass Spectrometry, Charité - Universitätsmedizin Berlin, 10117 Berlin, Germany
7Centre for Inflammation Biology and Cancer Immunology (CIBCI) & Peter Gorer Department of Immunobiology, King’s College London, SE1 1UL London, UK
8These authors contributed equally
9Senior author
10Lead Contact
*Correspondence: michele.mishto@kcl.ac.uk (M.M.), e.j.a.m.sijts@uu.nl (A.J.A.M.S.)
http://dx.doi.org/10.1016/j.celrep.2017.07.026

SUMMARY

Proteasome-catalyzed peptide splicing (PCPS) generates peptides that are presented by MHC class I molecules, but because their identification is challenging, the immunological relevance of spliced peptides remains unclear. Here, we developed a reverse immunology-based multi-level approach to identify proteasome-generated spliced epitopes. Applying this strategy to a murine Listeria monocytogenes infection model, we identified two spliced epitopes within the secreted bacterial phospholipase PlcB that primed antigen-specific CD8+ T cells in L. monocytogenes-infected mice. While reacting to the spliced epitopes, these CD8+ T cells failed to recognize the non-spliced peptide parts in the context of their natural flanking sequences. Thus, we here show that PCPS expands the CD8+ T cell response against L. monocytogenes by exposing spliced epitopes on the cell surface. Moreover, our multi-level strategy opens up opportunities to systematically investigate proteins for spliced epitope candidates and thus strategies for immunotherapies or vaccine design.

INTRODUCTION

CD8+ T cell responses play an important role in the clearance of intracellular pathogens and in protection from subsequent infections. CD8+ T cells react to epitopes presented by major histocompatibility complex class I (MHC class I) molecules at the cell surface. Epitope generation usually starts through proteasomal processing of pathogen-derived intracellular proteins. Peptides released by the proteasome are translocated by the transporter associated with antigen processing (TAP) into the endoplasmic reticulum (ER), where they may undergo N-terminal trimming by ER-resident aminopeptidases (ERAPs) (Cascio et al., 2003; Rock et al., 1994; Sijts and Kloetzel, 2011) and are loaded onto MHC class I molecules for presentation at the cell surface (Groettrup et al., 2010).

Given the central role of proteasomes in MHC class I antigen processing, their catalytic activity plays a fundamental role in shaping the pathogen-derived peptide repertoire toward which CD8+ T cells may react. The catalytic activity of proteasomes is displayed by three subunits, β1, β2, and β5, which are replaced by the induced catalytic subunits β1i/LMP2, β2i/MECL1, and β5i/LMP7 when cells are exposed to an inflammatory milieu (Groettrup et al., 2010; Sijts and Kloetzel, 2011). The changes in the catalytic subunit composition induce subtle modifications in conformation of the active sites, which affect proteasome proteolytic dynamics, and have consequences for the quantities of peptides produced (Arciniega et al., 2014; Liepe et al., 2015; Mishto et al., 2014; Ruschak and Kay, 2012; Sijts et al., 2000). Proteasomes can simply hydrolyze the peptide bonds of the antigen, thereby releasing canonical non-spliced peptides, or splice two distal fragments of the antigen, thereby generating peptides with a novel sequence (Vigneron et al., 2004). The latter mechanism is called proteasome-catalyzed peptide splicing (PCPS) (Mishto et al., 2012).

Despite the well-established crucial role of proteasomes in MHC class I antigen processing, the immunological relevance of PCPS is still a matter of debate, in part because only a few proteasome-generated spliced epitopes, mainly derived from tumor-associated antigens (Dalet et al., 2011; Ebstein et al., 2016; Hanada et al., 2004; Michaux et al., 2014; Vigneron et al., 2004; Warren et al., 2006), have been described so far. Some recent studies, however, hint at the importance of PCPS in MHC class I antigen presentation. Indeed, based on mass spectrometry (MS) analyses, approximately one-third of the peptides presented by MHC class I molecules on human lymphoblastoid
and fibroblast cells are proteasome-generated spliced peptides, and one-third of self-antigens presented by MHC class I on these cells are represented by spliced peptides only (Liepe et al., 2016). Such a significant increase in the variety of peptides potentially recognized by CD8+ T cells due to PCPS is also strengthened by their amount, which has been calculated in the order of one fourth of the MHC class I-restricted self-peptides (Liepe et al., 2016). This is in agreement with the amount of tumor-associated spliced epitopes presented by human MHC class I molecules at the cell surface, as shown for the gp100<sub>217-226</sub> antigen (Ebstein et al., 2016).

The relevance of PCPS in the cell-mediated immune response during infections is still unknown. We recently described cross-reactivity of CD8+ T cells, primed during <i>L. monocytogenes</i> by the dominant listeriolyisin O (LLO)-derived epitope LLO<sub>296-304</sub> toward the spliced epitope LLO<sub>294-297-304</sub> (Platteel et al., 2016). Nevertheless, so far, no clear evidence that PCPS leads to spliced peptide-specific CD8+ T cell responses during infection has been found. The scarcity of knowledge on the role of PCPS in immune recognition illustrates the difficulty of identifying immunologically relevant spliced epitopes, which lies in adopting a reverse immunology approach, requiring a complex in silico-in vitro workflow. To propel further investigation into the role of PCPS in antigen processing, we have developed a multi-level spliced epitope identification strategy, which we tested in the murine <i>L. monocytogenes</i> infection model. <i>L. monocytogenes</i> is a Gram-positive bacterium that primarily infects phagocytes and then mobilizes the host cell cytoskeleton to spread to neighboring cells. To enter the cytosol of infected cells, the bacteria secrete LLO and the phospholipases PlcA/phosphatidylinositol-specific phospholipase C (PI-PLC) and PlcB/phosphatidylcholine-prefering phospholipase C (PC-PLC) (Bielecki et al., 1990; Pamer, 2004; Portnoy et al., 1992). Bacterial clearance during <i>L. monocytogenes</i> infection is mediated by CD8+ T cells specific for the secreted bacterial proteins. Therefore, to test and validate our spliced epitope identification approach, we decided to use the secreted PlcA and PlcB proteins as model antigens.

**RESULTS**

**Identification of <i>L. monocytogenes</i>-Derived Spliced Epitope Candidates**

LLO<sub>296-304</sub> and its spliced variant LLO<sub>294-297-304</sub> are the main, known, targets of CD8+ T cells responding to <i>L. monocytogenes</i> infection in C57BL/6 mice (Geginat et al., 2001; Platteel et al., 2016). The PlcA and PlcB antigens used in this study to develop a reverse immunology-based spliced epitope prediction method possess 317 residues and 292 residues, respectively. These two proteins lack known H-2K<sup>b</sup> epitopes. To identify potential spliced epitopes, we computed all possible 8-mer and 9-mer non-spliced and spliced peptides from the two antigens (Figure 1). Here, the peptide length was restricted to reduce the number of potential epitopes since 8-mer and 9-mer represent the majority of the mouse MHC class I ligandome. The resulting database contained in total 1.56 × 10<sup>8</sup> sequences (Figures 2A and 2B). We then reduced the number of sequences by introducing further restrictions, which were based on previous studies. In particular, among the peptide candidates, we (1) set a maximal distance of 40 residues between the two splice reactants (i.e., the intervening sequence) and (2) excluded spliced peptides that were generated by binding of splice-reactants derived from two molecules of the same antigen (i.e., trans PCPS), as suggested by Dalet et al. (2010). Thus, only cis-spliced peptides, i.e., peptides generated from fragments of the same molecule, were included. This reduced database contained 4.67 × 10<sup>3</sup> entries for potential 8–9-mer spliced peptides and 1,188 entries for the potential 8–9-mer non-spliced peptides (Figures 2A and 2B). Because we were interested in antigenic peptides efficiently presented in the H-2K<sup>b</sup> cleft, we further reduced the database by selecting only those peptides that were predicted to bind H-2K<sup>b</sup> molecules with an IC<sub>50</sub> < 100 nM using the algorithm stabilized matrix method (SMM) (Peters and Sette, 2007) (Figures 1, 2C, and 2D). This selection step reduced the PlcA and PlcB candidate list to 989 spliced peptides and five non-spliced peptides. From these, we selected the 22 peptides with the lowest predicted IC<sub>50</sub>, i.e., lower than 16 nM, which were all spliced peptides (Table 1; Figures S1 and S2), for further testing in ex vivo CD8+ T cell recognition assays (Figure 1).

In summary, we applied several reduction steps in silico based on previously collected experimental evidence in order to obtain a candidate list of 22 predicted PlcA- and PlcB-spliced epitopes, starting from 1.56 × 10<sup>8</sup> potential spliced peptides. This led to progressive emergence of a peptide sequence pattern in the reduced databases, which included the known H-2K<sup>b</sup> anchor sites (Falk et al., 1991) (Figures 2E and 2F).

**Specific CD8+ T Cells Are Activated by PlcB-Derived and Proteasome-Generated Spliced Epitopes in <i>L. monocytogenes</i> Infection**

To test the immunogenicity of the selected antigenic spliced epitope candidates, C57BL/6 mice were infected intravenously (i.v.) with 2,000 colony forming units (CFUs) <i>L. monocytogenes</i>. At day 7, the peak of the response (Busch et al., 1998), their splenocytes were isolated and frequencies of CD8+ T cells producing interferon (IFN)-γ upon recognition of the spliced peptides were measured ex vivo, by intracellular cytokine staining and flow cytometry (Figure 3A). Two spliced PlcB epitope candidates, i.e., PlcB<sub>189-191/163-167</sub> and PlcB<sub>189-192/164-167</sub>, as well as the known immunodominant non-spliced epitope LLO<sub>296-304</sub> stimulated IFN-γ production in a significant number of CD8+ T cells (Figure 3B), indicating that these peptides are recognized by CD8+ T cells responding to <i>L. monocytogenes</i> infection.

To confirm that the PlcB-derived spliced epitopes were generated by the proteasome, we performed in vitro digestions of the synthetic substrate PlcB<sub>159-171/185-196</sub> with 20S proteasomes purified from the spleens of <i>L. monocytogenes</i> infected or uninfected mice. The substrate sequences were derived from the original PlcB protein, although part of the intervening sequence (PlcB<sub>172-184</sub>) was removed to facilitate the in vitro reaction as previously demonstrated (Dalet et al., 2010). MS analysis of the digests demonstrated the proteasome-mediated generation of the spliced epitopes PlcB<sub>189-191/163-167</sub> and PlcB<sub>189-192/164-167</sub> (Figure S3). By applying the quantitative method QME (quantification with minimal effort) to samples collected at different time intervals following in vitro digestion of PlcB<sub>159-171/185-196</sub> we measured the generation kinetics of the spliced epitopes.
PlcB₁₈₉⁻₁₉₁/₁₆₃⁻₁₆₇ and PlcB₁₈₉⁻₁₉₂/₁₆₄⁻₁₆₇ (Figure S4). No remarkable differences in generation kinetics of the two spliced epitopes were detected between reactions carried out by proteasomes purified from either uninfected or infected mouse spleens.

Thus, we confirmed that the spliced epitopes PlcB₁₈₉⁻₁₉₁/₁₆₃⁻₁₆₇ and PlcB₁₈₉⁻₁₉₂/₁₆₄⁻₁₆₇ can be generated by the proteasome, both early in infection when proteasome composition has not yet been changed as result of the inflammatory milieu (see results with proteasomes of uninfected mouse spleens) and in later stages of infection.

CD₈⁺ T Cells Specific for the PlcB-Derived Spliced Epitopes Do Not Cross-React against Non-spliced Peptides in L. monocytogenes Infection

CD₈⁺ T cell responses to PlcB₁₈₉⁻₁₉₁/₁₆₃⁻₁₆₇ and PlcB₁₈₉⁻₁₉₂/₁₆₄⁻₁₆₇, detected ex vivo, may have been primed by non-spliced epitope variants rather than by the spliced epitopes. To rule out this possibility, we compared the frequencies of CD₈⁺ T cells in L. monocytogenes infected mouse spleens specific for the spliced epitopes PlcB₁₈₉⁻₁₉₁/₁₆₃⁻₁₆₇ and PlcB₁₈₉⁻₁₉₂/₁₆₄⁻₁₆₇, to the response elicited by the two non-spliced peptides PlcB₁₆₀⁻₁₆₇ and PlcB₁₉₆⁻ which shared part of the N or C terminus with the spliced epitopes (their generation kinetics are shown in Figure S4). Only the spliced epitopes were specifically recognized by CD₈⁺ T cells primed during L. monocytogenes infection (Figure 4A).

To further investigate the possibility that CD₈⁺ T cells, specific for the two spliced epitopes, were primed by Listeria-derived non-spliced peptides, we adapted an in silico analysis method, described by Calis et al. (2012). We first computed all spliced epitope-related peptides that could potentially prime cross-reactive CD₈⁺ T cells (i.e., CD₈⁺ IFN-γ⁺ T cells) were targeted by the L. monocytogenes-induced CD₈⁺ T cell response in vivo.

Figure 1. Reverse-Immunology Approach to Identifying Immunogenic Spliced Peptides In Vivo

The systematic identification of immunogenic peptides requires a combined in silico-in vitro approach. Illustrated is the identification of non-spliced peptides (pink arrows) and identification of spliced peptides (blue arrows). For the latter, we first computed the complete list of theoretically possible spliced peptides from a given antigen and reduced this step by step to a number of spliced peptides predicted to bind to the H-2Kb molecule. The final candidate list of epitopes then was tested ex vivo. C57BL/6 mice were infected with L. monocytogenes. After 7 days, the splenocytes were isolated and incubated with the spliced peptides from the final candidate list. Peptides that triggered IFN-γ production in CD₈⁺ T cells ex vivo (i.e., CD₈⁺ IFN-γ⁺ T cells) were targeted by the L. monocytogenes-induced CD₈⁺ T cell response in vivo.
sites (3, 5, and 8 for the H2-K\(^b\) molecule) as well as the positions 1, 2, and 8 (for 9-mer) and 2, 3, and 8 (for 8-mer) are often irrelevant for efficient TCR recognition (Falk et al., 1991; Matsumura et al., 1992). We therefore computed all theoretical peptide sequences that followed the patterns ISXPXYKX, XXYPFYKX, and IXXPFYKX for the peptide PLCB\(^{189–192/164–167}\) (ISYPFYKL).
and ISXAXYKK, XXYAFYKK, and IXXYFKY for the peptide
PLCb189–191/163–167 (ISYAFLKL), where X could be any of
the 20 amino acids. This resulted in 7,999 related peptides per
candidate, which could theoretically prime CD8+ T cells re-
once also to the spliced epitope-specific, if any of these pep-
tides were generated in vivo. However, none of these theoretical
peptide sequences could be found in the L. monocytogenes
non-spliced proteome (represented by 2,844 protein entries)
(data not shown), thereby further suggesting that a hypothetical
cross-reactivity of CD8+ T cells against non-spliced Plc-
derived epitopes to the spliced epitopes, or the other way
around, was unlikely.

To confirm that the specific CD8+ T cell-mediated response
toward the spliced epitopes in infected mice was the outcome
of the infection, we compared the frequencies of spliced
epitope-specific IFN-γ+ CD8+ T cells among splenocytes of
infected mice to those of uninfected mice. For both the spliced
(PlcB189–191/163–167 and PlcB189–192/164–167) and non-spliced
(LO296–304) epitopes, we observed a significantly larger prevalence
of IFN-γ+ CD8+ T cells in samples from infected mice
compared to those from uninfected mice (Figure 4B), suggesting
that the identified PlcB-derived spliced epitopes participate in
the priming of CD8+ T cell responses upon L. monocytogenes
infection.

To prove that the identified spliced epitopes are processed
from the PlcB protein and presented in vivo, EL4 cells (H-2Kb)
were transduced with retroviral constructs expressing either
full length PlcB or a mutated form of PlcB (PlcB-delta), in which
the spliced epitopes’ anchor residues were substituted for serine
residues (F164S and L167S: PlcB189–192/164–167 mutated to
ISYPYSKLS and PlcB189–191/163–167 mutated to ISYASYKLS),
to interfere with H-2Kb binding (Table 1). These EL4 cells, as well
as control EL4 cells transduced with an empty vector (Figure 5A)
or EL4 cells expressing LLO1–415, were co-cultured with purified
CD8+ splenocytes derived from PlcB-spliced epitope-respon-
sive infected (Figure 3) or uninfected mice. Activation-induced
IFN-γ release by CD8+ T cells was quantified. As shown in Figure
5B, CD8+ T cells derived from infected mice were activated
by co-culture with EL4-LLO and EL4-PlcB cells, or by incubation
with identified spliced epitopes (shown in Figure 3), but not by
co-culture with PlcB-delta-expressing or mock transduced cells.
Thus, disruption of spliced epitope presentation on H-2Kb
molecules in EL4-PlcB-delta cells abrogated the ability of these
cells to trigger IFN-γ production in PlcB-specific CD8+ T cells
derived from infected mice (Figure 5B). This indicates that these T cells
carry the spliced epitope sequences on EL4-PlcB cells
and are not directed to peptides derived from another part
of PlcB or from a different Listeria protein. As expected, CD8+
splenocytes derived from non-infected mice failed to respond to any
transduced EL4 cell line.

To exclude a theoretical cross-reaction of the spliced epitope-
specific CD8+ T cells against other spliced and non-spliced pep-
tides generated by proteasomes from the epitope surrounding
sequences, we analyzed PlcB189–191/185–196 digestion products
(Figure S3) by MS for the presence of other peptide products
that carried one of the two residues substituted in the PlcB-
delta, i.e., F164 or L167. Three non-spliced and ten spliced
peptides containing PlcB F164 and/or L167 were identified
(Table S1). We already demonstrated that the spliced epitopes
PlcB189–191/163–167 and PlcB189–192/164–167 activated the CD8+
T cells of infected mice, whereas the non-spliced peptides
PlcB189–192/164–167 did not (Figure 4A). Among the other peptide
products, only the spliced peptide PlcB189–191/163–167 was predicted
to bind the H-2Kb complex with an IC50 lower than 100 nM, i.e.,
73 nM, which is much larger than the IC50 of the overlapping
and also than the IC50 of the spliced epitope candidates tested nega-
tive for recognition by CD8+ T cells of Listeria-infected mice.

Table 1. Predicted Spliced Epitope Candidates

<table>
<thead>
<tr>
<th>Peptide a</th>
<th>Sequence</th>
<th>IC50 (nM) b</th>
<th>T1/2 (hr) c</th>
</tr>
</thead>
<tbody>
<tr>
<td>LO296–304</td>
<td>VAYGROVYL</td>
<td>8 ± 0.2</td>
<td></td>
</tr>
<tr>
<td>PlcA70–72/92–96</td>
<td>MSYLQQL</td>
<td>2.5 –</td>
<td></td>
</tr>
<tr>
<td>PlcA73–73/92–96</td>
<td>MSYNYQL</td>
<td>3.2 –</td>
<td></td>
</tr>
<tr>
<td>PlcA262–266/273–275</td>
<td>TSLTFAAL</td>
<td>3.5 –</td>
<td></td>
</tr>
<tr>
<td>PlcA262–267/285–286</td>
<td>TSLTFTNL</td>
<td>5.7 –</td>
<td></td>
</tr>
<tr>
<td>PlcA262–268/254</td>
<td>TSLTFTPL</td>
<td>6.7 –</td>
<td></td>
</tr>
<tr>
<td>PlcA70–72/57–61</td>
<td>MSYNLAL</td>
<td>5.3 –</td>
<td></td>
</tr>
<tr>
<td>PlcA262–266/284–286</td>
<td>TSLTFLNL</td>
<td>7.3 –</td>
<td></td>
</tr>
<tr>
<td>PlcA112–114/92–96</td>
<td>KIYLQQL</td>
<td>8.4 –</td>
<td></td>
</tr>
<tr>
<td>PlcA19–23/59–61</td>
<td>CFFTFIAL</td>
<td>3.8 –</td>
<td></td>
</tr>
<tr>
<td>PlcA19–23/56–58</td>
<td>CFFTFNL</td>
<td>6.1 –</td>
<td></td>
</tr>
<tr>
<td>PlcA19–23/50–52</td>
<td>CFFTSAL</td>
<td>8.7 –</td>
<td></td>
</tr>
<tr>
<td>PlcA19–24/57–58</td>
<td>CFFTPNL</td>
<td>7.0 –</td>
<td></td>
</tr>
<tr>
<td>PlcB189–191/163–167</td>
<td>ISYAFYKL</td>
<td>1.1 5.25 ± 0.92</td>
<td></td>
</tr>
<tr>
<td>PlcB189–192/164–167</td>
<td>ISYPFYLK</td>
<td>1.9 4.05 ± 1.07</td>
<td></td>
</tr>
<tr>
<td>PlcB171–173/163–167</td>
<td>IHAYFYLK</td>
<td>6.1 2.39 ± 0.71</td>
<td></td>
</tr>
<tr>
<td>PlcB65–69/93–96</td>
<td>VNTHYANL</td>
<td>7.5 4.78 ± 1.69</td>
<td></td>
</tr>
<tr>
<td>PlcB189–191/181–185</td>
<td>ISYMHANN</td>
<td>8.1 0.17 ± 0.12</td>
<td></td>
</tr>
<tr>
<td>PlcB171–174/164–167</td>
<td>IHYYFYLK</td>
<td>8.1 1.66 ± 0.63</td>
<td></td>
</tr>
<tr>
<td>PlcB251–255/270–272</td>
<td>KSYLVARL</td>
<td>8.6 3.13 ± 0.56</td>
<td></td>
</tr>
<tr>
<td>PlcB118–121/137–140</td>
<td>STFLFANA</td>
<td>9.9 1.54 ± 0.25</td>
<td></td>
</tr>
<tr>
<td>PlcB189–192/182–185</td>
<td>ISYPHANN</td>
<td>11.6 0.22 ± 0.14</td>
<td></td>
</tr>
<tr>
<td>PlcB177–179/163–167</td>
<td>ISQAFYKL</td>
<td>15.4 1.43 ± 0.73</td>
<td></td>
</tr>
<tr>
<td>PlcB160–167 d</td>
<td>FDTAFYKL</td>
<td>2949 1.07 ± 1.07</td>
<td></td>
</tr>
<tr>
<td>PlcB189–196 d</td>
<td>ISYPYGRH</td>
<td>10840 1.07 ± 1.07</td>
<td></td>
</tr>
<tr>
<td>PlcB189–191/163,164S–167S</td>
<td>ISYASYKLS</td>
<td>288 1.07 ± 1.07</td>
<td></td>
</tr>
<tr>
<td>PlcB189–192/164S–167S</td>
<td>ISYPSYKLS</td>
<td>473 1.07 ± 1.07</td>
<td></td>
</tr>
</tbody>
</table>

aSpliced epitope candidates PlcA and PlcB antigens were predicted by
applying the in silico computation method shown in Figure 1.

bThe predicted IC50 for binding to the H-2Kb MHC class I molecule was
calculated using SMM algorithm (Peters and Sette, 2007).

cThe H-2Kb peptide complex stability was empirically measured; mean
half-life (T1/2) and SD of two independent experiments are shown.

dThe half-life of LLO296–304 was in agreement with previous studies
(Platteel et al., 2016).

eThe two non-spliced peptides listed here were used as negative con-
trols, to exclude potential cross-reactivity of spliced epitope-specific
CD8+ T cells.

fPeptides failed to stabilize H-2Kb molecules.

gIn these spliced PlcB sequences, residue 164F and 167L were replaced
for S residues, to decrease H-2Kb binding affinity.
(Figures S2 and 3 and see correlation between the frequency of the IFN-γ+ CD8+ T cells and the IC50 described below).

From these diverse control experiments, and the in vitro digestion experiments (Figure S4), we infer that cells expressing the Listeria PlcB protein generate the two PlcB-derived spliced epitopes in similar amounts (Figure S4) and present at least one of them to CD8+ T cells, triggering T cell activation. Given the absence of a CD8+ T cell response (1) toward the spliced PlcB epitopes in uninfected mice, (2) toward a non-spliced peptide sharing part of the spliced epitope sequence in infected mice, and (3) toward EL4 cells expressing a PlcB variant in which the MHC class I anchor sites of the two spliced epitopes are mutated, the possibility of cross-reactivity of the spliced PlcB epitope-specific CD8+ T cells is unlikely. This is further supported by the absence of other peptides produced by proteasomes and predicted to efficiently bind to H-2Kb molecules.

Figure 3. CD8+ T Cells from Infected Mice Specifically Recognize PlcB Spliced Epitopes

C57BL/6 mice were infected with L. monocytogenes, and 7 days later the frequency of CD8+ T cells specific for synthetic peptides among the splenocytes of infected mice was measured ex vivo by intracellular staining of IFN-γ in CD8+ T cells.

(A) Representative FACS plots of the staining in presence or absence of the target peptide LLO296–304.

(B) Frequency of IFN-γ+ CD8+ T cells, upon ex vivo stimulation with the PlcA- or PlcB-spliced peptides and the non-spliced LLO296–304 epitope, among splenocytes derived from infected mice (n = 5–43). Accumulated data of five independent experiments are shown, corrected for IFN-γ background measured in control samples incubated without peptide. The violin plots indicate the density of the measurements on each side, with all single mice indicated as dots. Red lines indicate the median of the measurements. Significant differences between the splenocytes incubated with or without the peptides are marked as * (ANOVA with Bonferroni and Welch’s corrections; LLO296–304, PlcB189–191/163–167, and PlcB189–192/164–167, p < 0.001).
Measured H-2Kb-Spliced Peptide Stability and the Predicted Binding Affinity Correlate with the Relative Frequency of Specific CD8+ T Cells in Infected Mice

During the development of our in silico strategy for selecting the best peptide candidates for this study, we assumed, based on earlier studies (Sijts and Pamer, 1997; van der Burg et al., 1996; Watson et al., 2012), a correlation between the predicted IC50 (i.e., H-2Kb binding affinity) and the recognition of peptides by specific CD8+ T cells. Therefore, we restricted the list of spliced peptides that might be recognized during L. monocytogenes infection by their predicted IC50. The validity of this approach was evaluated on the results of the ex vivo stimulation of splenocytes derived from infected mice with the PlcB-spliced peptides (Figure 6). In agreement with our assumption, we observed a significant inverse correlation between the predicted IC50 of the spliced peptides and the specific response (measured as frequency of IFN-γ+ CD8+ T cells) in infected mice (Figure 6A; Table 1).

As proof of principle, we experimentally measured another parameter depicting the affinity between MHC class I and peptide, i.e., the stability of H-2Kb-peptide complexes at the cell surface of RMA-S cells (Figure 6B). Five out of ten spliced PlcB epitope candidates tested, including PlcB189–191/163–167 and PlcB189–192/164–167, upregulated H-2Kb levels on RMA-S cells (Figure 6C, left panel), indicating that these peptides bound to and stabilized RMA-S-expressed H-2Kb molecules. The other spliced epitope candidates tested (Figure 6C, right panel) as well as the non-spliced peptides PlcB160–167 and PlcB189–196 (data not shown), only weakly upregulated or failed to upregulate H-2Kb on RMA-S cells. This difference in peptide binding capacity was illustrated by the half-lives that we computed based on chase analyses of peptide-pulsed RMA-S cells (Table 1), which varied from 5.25 hr for H-2Kb-complexed with “strong” binders to 0.17 hr for weak binders. Of note, the two identified spliced epitopes PlcB189–191/163–167 and PlcB189–192/164–167 were among the three highest affinity H-2Kb binders (Table 1), with computed half-lives exceeding these of H-2Kb molecules bound to the control epitope LLO296–304. Comparing the data for all spliced peptides tested, we found a significant inverse correlation between the predicted IC50 and the measured half-lives of H-2Kb-peptide complexes at the cell surface (Figure 6D). Accordingly, we also found a significant direct correlation between the half-lives of H-2Kb-peptide complexes and the relative frequency of IFN-γ+ CD8+ T cells among infected mouse splenocytes for the spliced peptides derived from PlcB (Figure 6E). Thus, selection of spliced epitope candidates based on predicted IC50 is a valid approach for focusing ex vivo analyses on the most promising epitope candidates.

DISCUSSION

We here present a multi-level spliced epitope identification approach that will be instrumental in uncovering the role of PCPs in immune recognition, both of pathogen-derived and self-antigens. Application of this approach enabled us to gain insight into the contribution of PCPs to MHC class I antigen processing. Focusing on two L. monocytogenes proteins as model antigens, we show that PCPs generates two overlapping spliced epitopes that participate in CD8+ T cell priming following infection. The precise contribution of the two respective peptides to CD8+ T cell activation will be unraveled in future experiments aimed at examining the MHC class I presentation kinetics of the spliced epitopes in relation to each other as well as to other, non-spliced Listeria epitopes.
The method developed for this study to identify spliced epitope candidates uses an in silico-guided approach for the identification of spliced peptides, combined with experimental outcomes and the SMM prediction algorithm. As part of this approach, in silico-predicted epitope candidates are selected based on their low predicted IC50, which showed to be the right approach since we observed a correlation between predicted IC50 (and H-2Kβ peptide complex stability) and the expansion of specific CD8+ T cells during infection, in agreement with earlier published works (Sijts and Pamer, 1997; van der Burg et al., 1996; Watson et al., 2012).

The demonstration that proteasome-generated spliced epitopes trigger a specific CD8+ T cell response during L. monocytogenes infection in mice could represent a milestone in the investigation of T cell immunity against pathogens. The ability of PCPS to expand the epitope variety could be a key means by which the immune system tackles pathogens.

Thus, the so-far neglected possibilities for epitopes generated by PCPS may be a tremendous asset for vaccination approaches focusing on single pathogen-derived proteins. To illustrate this, we performed a preliminary in silico analysis to identify PlcA and PlcB epitope candidates for the most frequent human MHC class I haplotypes (Figures S5 and S6). We calculated that the number of potential spliced peptides predicted to bind with high affinity to the most frequent HLA-A or HLA-B variants largely exceeds that of the non-spliced peptides. Thus, no non-spliced epitope candidates with a predicted IC50 below 100 nM were found in PlcB or in PlcA in combination with HLA-A*01:01, HLA-A*03:01, HLA-B*44:02, or HLA-B*44:03, while four and three non-spliced candidates were predicted for the PlcA antigen in combination with HLA-A*02:01 and HLA-B*07:02, respectively. In contrast, there are hundreds of potential spliced epitopes for both antigens predicted to bind to each of the investigated MHC class I haplotypes (Figures S5 and S6).

In conclusion, we suggest that spliced peptides might provide immunological targets as well as a key means by which the immune system tackles pathogens.

**EXPERIMENTAL PROCEDURES**

**In Silico Generation of Spliced Peptide Target List**

Given the PlcA (ACE05212.1) or PlcB (ACE05216.1) sequences, we first computed all potential non-spliced and spliced peptides of length L = 8 and
\[ L = 9, \text{ resulting in } N \text{ peptides, denoted as } [n-p1]/[p2-c], \text{ where } N \text{ can be obtained via Equation 1 derived from Liepe et al. (2010).} \]

\[ N = \sum_{n=1}^{3} \sum_{p1=1}^{S} \sum_{p2=1}^{S-n} S + 1, \quad \text{(Equation 1)} \]

where \( S \) is the length of the parental antigen (PlcA and PlcB, respectively). This list was then reduced by excluding all spliced peptides generated through trans PCPS. This means that for a given spliced peptide denoted as \([n-p1]/[p2-c]\) all peptides with \( p1 < p2 \) (for PCPS in same order as in parental antigen) or \( c < n \) (for PCPS in reverse order compared to the parental antigen) are maintained in the reduced database. We next removed all peptides with intervening sequence length \( l \) larger than 40 amino acids, where \( l = p2-p1-1 \) for splicing in normal order, and \( l = c-n-1 \) for splicing in reverse order. For all remaining 8- and 9-mer spliced peptides, we computed in silico the IC50 as a measure for binding strength to the murine H-2Kb or the human HLA-A*01:01, HLA-A*02:01, HLA-A*03:01, HLA-B*07:02, HLA-B*44:02, and

\[
\begin{align*}
\text{Strong binders} & : \quad \text{PlcB}_{189-191}/163-167 \\
& \quad \text{PlcB}_{189-192}/164-167 \\
& \quad \text{PlcB}_{65-69}/88-90 \\
& \quad \text{PlcB}_{251-255}/270-272 \\
& \quad \text{PlcB}_{118-121}/137-140 \\
\text{Weak binders} & : \quad \text{PlcB}_{171-173}/163-167 \\
& \quad \text{PlcB}_{189-191}/181/185 \\
& \quad \text{PlcB}_{171-174}/164-167 \\
& \quad \text{PlcB}_{189-192}/182-185 \\
& \quad \text{PlcB}_{177-179}/163-167
\end{align*}
\]

Figure 6. Correlation between Predicted IC50, H-2Kb-Peptide Stability, and Frequency of Specific CD8+ T Cells in Infected Mice

(A) Correlation between the predicted IC50 of PlcB-spliced peptides and the frequency of IFN-\( \gamma \) CD8+ T cells specific for the peptides among splenocytes derived from \( L. \) monocytogenes infected mice (for all PlcB peptides tested on five to 37 mice). There is a significant inverse correlation (Spearman correlation test \( p < 0.001, C \text{ value} = -0.559 \)). Values are the mean and bars the SEM of tested mice for each peptide \((n = 5-37)\). IC50 of the peptides for the H-2Kb complex was predicted by the SMM algorithm (Peters and Sette, 2007).

(B) H-2Kb-peptide stability was measured with RMA-S cells incubated with synthetic peptides and chased in the absence of peptide. At different time points the remaining H-2Kb-peptide complexes were measured using FACS and measured in MFI (mean fluorescence index).

(C) The stability of the H-2Kb complexes bound to each PlcB-spliced peptide is shown. Peptides were divided in strong binders (with MFI > 500 at t = 0; left panel) or weak binders (with MFI < 200 at t = 0; right panel). Incubation without peptide resulted in a background MFI level of ~40. Values are the mean MFI ± SD of two independent experiments.

(D) Correlation between the predicted IC50 of the PlcB-spliced peptides \((n = 10)\) and the H-2Kb-peptide half-life at the cells surface is shown. Values are the mean of two independent experiments. There is a significant inverse correlation (Spearman correlation test \( p = 0.022, C \text{ value} = -0.709 \)).

(E) Correlation between the H-2Kb-peptide stability and the frequency of IFN-\( \gamma \) CD8+ T cells responsive against the spliced peptides among splenocytes derived from infected mice (for all peptides tested on five to 37 mice). There is a significant correlation (Spearman correlation test \( p < 0.001, C \text{ value} = 0.498 \)). Values are the mean and bars the SEM of mice tested for each peptide \((n = 5-37)\).
H-2Kb haplotypes using the offline version of the SSMM algorithm (Peters and Sette, 2007). We considered all spliced peptides with predicted IC50 < 100 nM and a further IC50 cutoff of 16 nM.

**Peptide Synthesis and 20S Proteasome Purification**

The PlcB159–171/185–196 (KFDTAFYKLGLAINFTAISYPPGYH) polypeptide substrate and all non-spliced and spliced PlcA and PlcB peptides were synthesized using Fmoc solid phase chemistry. 20S proteasomes were purified from five pooled spleens of C57Bl/6 mice infected or uninfected by L. monocytogenes (Platteel et al., 2016). The purity of the proteasome preparation is shown in Figure S7. The differences in proteasome subunit composition were described previously (Platteel et al., 2016).

**Identification and Quantification of Peptide Products from In Vitro Digestions by Proteasomes**

PlcB159–171/185–196 (20 μM) was digested by 2 μg 20S proteasomes in 100 μL TEAD buffer (Tris 20 mM, EDTA 1 mM, NaCl 1 mM, DTT 1 mM [pH 7.2]) over time, at 37°C. Identification of the polypeptide digestion products was performed by liquid-chromatography mass spectrometry (LC-MS) analyses: 15–μL digested samples were analyzed directly by nanoscale LC-MS/MS using an Ultimate 3000 and LTQ Orbitrap XL mass spectrometer (Thermo Fisher Scientific). The system comprises a 5-mm x 300-μM, 100-Å trapping column (PepMap C18, 5 μm; Dionex) and a PiccoChip analytical column (Replosil-pur; 3 μm; New Objective). The mobile phase (A) was 0.1% (v/v) formic acid in water and (B) was 80:20 (v/v) acetonitrile/water containing 0.1% (v/v) formic acid. Elution was carried out using a gradient 15–50% B in 35 min with a flow rate of 300 nL/min. Full MS spectra (m/z 300–2,000) were acquired in an Orbitrap instrument at a resolution of 60,000 (full width at half maximum [FWHM]). At first, the most abundant precursor ion was selected for either data-dependent collision-induced dissociation (CID) fragmentation with parent list (1st, 2nd charge state included). Fragment ions were detected in an Ion Trap instrument. Dynamic exclusion was enabled with a repeat count of 2- and 60-s exclusion duration. Additionally, the theoretically calculated precursor ions of the expected spliced peptides were pre-elected for two Orbitrap CID (m/z 350–2,000) and higher energy collisional dissociation (HCD; m/z 100–1500) fragmentation scans. The maximum ion accumulation time for MS scans was set to 200 ms and for MS/MS scans to 500 ms. Background ions at m/z 391.2943 and 445.1200 act as lock mass. Peptides were identified by PD1.4 software (Thermo Fisher Scientific) based on their merged tandem mass spectra (MS/MS) of CID and HCD. In addition, for spliced peptides we compared the retention time and the merged MS/MS of CID and HCD with the fragmentation pattern of their synthetic counterparts. The database used for the LC-MS/MS analyses was generated by applying the SpliceMet’s ProteaJ algorithm thereby allowing the identification of non-spliced and spliced peptides (Liepe et al., 2010).

The polypeptide digestion kinetics were analyzed with the ESI-ion trap instrument DECA XP MAX (Thermo Fisher Scientific) as described (Liepe et al., 2010). The quantification of peptides produced in the in vitro digestion kinetics was carried out by applying the QME method to the LC-MS analyses (Mishito et al., 2012). QME estimates the absolute content of spliced and non-spliced peptide products based on their MS peak area measured in the digestion probe. The QME algorithm parameters were empirically computed in our previous study (Mishito et al., 2012) and here applied.

**Cell Culture**

RMA-S cells were cultured in IMDM (Invitrogen Life Technologies), supplemented with 10% fetal calf serum (FCS; LONZA), 2 mM L-glutamine, 30 μM 2-ME, penicillin/streptomycin. Transduced EL4 cells were cultured in RPMI (Invitrogen Life Technologies), supplemented with 10% FCS, 2 mM L-glutamine, 30 μM 2-ME, penicillin/streptomycin, and 5 μg/mL puromycin (Sigma-Aldrich) as appropriate.

**MHC Class I Peptide Stability Assays**

RMA-S off rate assays were performed as described (Deol et al., 2007). In short, TAP-deficient RMA-S cells were incubated overnight (o/n) in the presence or absence of 100 μM synthetic peptide, at 37°C. Cells were harvested, washed three times with PBS, and chased in the absence of peptide, at 37°C, for IC50 /2KCS peptide ligand expression with a down-regulation-sensitive, biotin-conjugated mouse antibody (Ab) (AF6-88.5; BD Bioscience) and PE-conjugated SA (eBioscience) and analyzed using fluorescence-activated cell sorting (FACS) Canto II (BD Biosciences) and FlowJo software (Tree Star). The T1/2 of the H-2Kb-peptide complex at the cell surface was computed as described (Textoris-Taube et al., 2015), based on the mean fluorescence intensities of peptide-pulsed cells corrected for background levels from cells that were not incubated with peptide.

**Retroviral Transduction of EL4 Cells**

LLLOI1 was amplified from L. monocytogenes 10403S using forward primer AGATCTGTTGACCCATGAAAAATAATGCTAG and reverse primer TTAAATCTGTATAATGGTTAGTTGTC. The product was verified by sequencing and cloned into a pMSCVires-GFP vector (Addgene). Synthetic DNA sequences of PlcB and PlcB-delta (with the substitutions F164S and L167S) were purchased (Geneart Invitrogen) and cloned into pMSCV-Puro-ires-GFP, pMSCV constructs and pc10A1, encoding the envelope proteins, were transfected into 293T cells using Lipofectamine 2000 reagent (Thermo Fisher Scientific) as instructed. After 24 hr, the medium was removed and cells were co-cultured with EL4 cells, added in 10 mL medium, for 48 hr at 37°C. EL4 cells were then harvested, resuspended in 10 mL DMEM medium without (LLO-transduced cells) or with puromycin (PlcB, PlcB-delta and mock cells). Expression of the constructs was verified by GFP expression using FACS and found to be ~99% (Figure 5A) for PlcB, PlcB-delta, and mock cells. LLO-expressing cells were sorted by FACS twice, which resulted in expression levels of ~20%.

**Mice and Infection**

*L. monocytogenes* strain 10403S was grown in brain-heart infusion medium (Sigma-Aldrich) and harvested while in log phase. C57Bl/6 mice were purchased from Charles River. For infection, 6 to 8-week-old female mice were inoculated intravenously in the tail vein with 2000 bacteria (0.1 LD50) in 200 μL PBS. All animal experiments were approved by the Animal Ethics Committee from Utrecht University (DEC 2014.I.11.081 and DEC 2014.I.11.003).

**Intracellular IFN-γ Staining**

Erythrocyte-deplete single spleen cell suspensions were prepared and 0.5 × 107 splenocytes were incubated with or without 1 μg/mL synthetic peptide for 6 hr in 1 mL RPMI medium, containing 50 μg/mL gentamycin (Gibco), and 10 μM monensin (eBioscience), at 37°C. Subsequently, cells were stained with an antigen-presenting cell (APC)-conjugated anti-mouse CD8 antibody (53-6.7; eBioscience) in the presence of anti-mouse CD16/CD32 antibody (clone 2.4G2; made in house). Cells were fixed with 2% paraformaldehyde and then stained with phycocerythrin (PE)-conjugated anti-mouse IFN-γ antibody (XM1G1.2; eBioscience) in the presence of 0.5% saponin and analyzed on a FACS Canto II (BD Biosciences) using FlowJo software (Tree Star). Percentages of specific IFN-γ–CD8+ splenocytes were calculated by subtracting the background IFN-γ–CD8+ splenocytes incubated without peptide per individual mouse.

**Detection of IFN-γ Release**

CD8+ splenocytes were enriched by negative selection for CD4 (GK1.5), B220 (RA3-6B2), CD11b (M1/70), and MHCII (M5/114)-expressing cells using Dynal beads (Invitrogen). 4 × 106 CD8+ cells (purity ~ 80%) were co-cultured with 2 × 105 transduced EL4 cells for 48 hr, and supernatants were then collected for analysis in a monoplex IFN-γ assay using the Magpix (Luminex XMAP) system, as instructed. Briefly, supernatant were incubated o/n with magnetic IFN-γ capture beads (clone AN-18), in 96-well flat-bottomed plates (Greiner bio-one, 655096). Captured IFN-γ was detected with biotin-conjugated anti-IFN-γ mAb (clone XM1G1.2) and Streptavidin-PE. Cytokine concentrations in the tested samples were calculated from a standard curve generated with rIFN-γ and MFI data were analyzed using a 5-parameter logistic method (XPONENT software, Luminex).
Statistical Analysis

Data were tested for normality distribution and homoscedasticity by Kolmogorov-Smirnov, Shapiro-Wilk, and Levene tests. To test for significant differences between CD8+ T cell responses to different epitope candidates, we applied an unpaired ANOVA test with Bonferroni post hoc correction for multiple comparisons and Welch’s correction. To compare frequencies of peptide-specific CD8+ T cells between infected and uninfected mice, we applied an unpaired t test with Welch’s correction. The same test was performed in Figure 5 to compare INF-γ release by CD8+ T cells toward EL4 cells expressing LLO, PlcB, PlcB-delta as compared to those expressing the empty vector (mock). For the correlations, we applied a Spearman test. p < 0.05 was considered to be significant.

ACCESSION NUMBERS

The FACS files and mass spectrometry RAW files reported in this paper are available in Mendeley Data at DOI http://dx.doi.org/10.17632/983y6nncrx.1.

SUPPLEMENTAL INFORMATION

Supplemental Information includes seven figures and one table and can be found with this article online at http://dx.doi.org/10.1016/j.celrep.2017.07.026.

AUTHOR CONTRIBUTIONS


ACKNOWLEDGMENTS

Support was by European Union’s Seventh Framework Programme (FP7/2007-2013) grant no. 280873; ADITEC to A.J.A.M.S., Berlin Institute of Health (BIH, CRG1-TP1) and Einstein Stiftung Berlin (A2013-174) to P.M.K., and NC3Rs through a David Sainsbury Fellowship to J.L. (NC/K001949/1). We thank P. Kunert and B. Brecht-Jachan (Charité Berlin) for technical assistance. We thank the Shared Facility Mass Spectrometry of the Charité for support in data acquisition.

Received: October 31, 2016
Revised: June 26, 2017
Accepted: July 12, 2017
Published: August 1, 2017

REFERENCES


