Successive crystal structure snapshots suggest the basis for MHC class I peptide loading and editing by tapasin

Ida Hafstrand[a], Ece Canan Sayitoglu[b], Anca Apavaloaei[c], Benjamin John Josey[b], Renhua Sun[a], Xiao Han[a], Sara Pellegrino[d], Didem Ozkazanc[b], Renée Potens[b], Linda Janssen[c], Johan Nilvebrant[e], Per-Åke Nygren[e], Tatyana Sandalova[a], Sebastian Springer[c], Anna-Maria Georgoudaki[b], Adil Doganay Duru[a,b,1], and Adnane Achour[a,1,2]

[a]Science for Life Laboratory, Department of Medicine Solna, Karolinska Institute, and Division of Infectious Diseases, Karolinska University Hospital, Solna, SE-17176 Stockholm, Sweden; [b]Nova Southeastern University Cell Therapy Institute, Dr. Kiran C. Patel College of Allopathic Medicine, Nova Southeastern University, Fort Lauderdale, FL 33301; [c]Department of Life Sciences and Chemistry, Jacobs University Bremen, 28759 Bremen, Germany; [d]Department of Pharmaceutical Science, General and Organic Chemistry Section, University of Milano, 20133 Italy; and [e]Department of Protein Science, School of Engineering Sciences in Chemistry, Biotechnology and Health, AlbaNova University Center, Royal Institute of Technology, SE-10691 Stockholm, Sweden

Edited by Peter Cresswell, Yale University School of Medicine, New Haven, CT, and approved February 4, 2019 (received for review May 14, 2018)

MHC-I epitope presentation to CD8$^+$ T cells is directly dependent on peptide loading and selection during antigen processing. However, the exact molecular bases underlying peptide selection and binding by MHC-I remain largely unknown. Within the peptide-loading complex, the peptide editor tapasin is key to the selection of MHC-I–bound peptides. Here, we have determined an ensemble of crystal structures of MHC-I in complex with the peptide exchange-associated dipeptide GL, as well as the tapasin-associated scoop loop, alone or in combination with candidate epitopes. These results combined with mutation analyses allow us to propose a molecular model underlying MHC-I peptide selection by tapasin. The N termini of bound peptides most probably bind first in the N-terminal and middle region of the MHC-I peptide binding cleft, upon which the peptide C termini are tested for their capacity to dislodge the tapasin scoop loop from the F pocket of the MHC-I cleft. Our results also indicate important differences in peptide selection between different MHC-I alleles.

MHC-I | tapasin | peptide editing | TAPBPR

Peptide presentation by major histocompatibility class I (MHC-I) molecules is essential for recognition of intracellular pathogens and malignancies (1–6). Following their transfer by the transporter associated with antigen processing (TAP) into the endoplasmic reticulum (ER), peptides are selectively loaded on MHC-I with the help of the peptide loading complex (PLC) multiprotein machinery, which consists of tapasin, ERp57, and calreticulin (7). Until recently, the interactions of some of these proteins, especially tapasin, with MHC-I/peptide (pMHC) complexes have been mainly deduced from mutational and functional studies. However, a recent low-resolution cryo-EM structure revealed the spatial arrangement of tapasin and MHC-I molecules within the PLC (8). Furthermore, two low-resolution crystal structures (3.3 and 3.5 Å resolution, respectively) provided insights into binding and interactions between MHC-I and the tapasin homolog TAP-binding protein-related (TAPBPR) (9, 10). Thomas and Tampé described the presence of a "jack hairpin" that lodges in under the MHC-I peptide binding cleft and of a "scoop loop" that interacts with the C-terminal part of the MHC-I peptide binding cleft (9). The N-terminal domain of TAPBPR envelops the α_2 helix and stabilizes the open conformer of MHC-I molecules (9, 10), in a similar manner to how HLA-DM interacts with peptide-empty MHC class II molecules (MHC-II) (11). However, it should be noted that the overall sequence identity between tapasin and TAPBPR is only 22% (12). Importantly, the sequences of the scoop loops are also very different in tapasin compared with TAPBPR. Furthermore, in contrast to tapasin, TAPBPR does not bind ERp57 nor calreticulin and is not an integral component of the PLC (13). Altogether, these structural results indicate a potential key role in peptide editing for both the scoop loop and the jack hairpin. It should however also be noted that a large amount of additional contacts is formed between TAPBPR and the peptide binding cleft, the α3 domain, and the β_2m subunit of MHC-I alleles (9, 10), which may also play a key role in sensing the presence of peptides with adequate binding affinity.

In MHC-II antigen processing and presentation, the peptide-editing HLA-DM has been previously mentioned as equivalent to tapasin (14). A series of studies conducted by the Rötzschke research group demonstrated that small organic compounds and small peptides, usually dipeptides, can act similarly to HLA-DM, resulting in the stabilization of peptide-receptive MHC-II molecules (15–17). A similar system has also been characterized for MHC-I, demonstrating that dipeptides, including the dipeptide GL, can act as both stabilizers of peptide-empty MHC-I molecules and peptide exchange catalysts (18, 19).

Significance

We have determined a suite of crystal structures of MHC-I molecules that provide insights in the molecular basis underlying peptide binding and selection by tapasin. We demonstrate that the scoop loop of tapasin binds in the F pocket of MHC-I molecules without affecting the conformation of the rest of the peptide binding cleft, which remains empty and peptide-receptive. Furthermore, the results of our study suggest that peptides may bind first with their N termini to the peptide binding cleft and are tested thereafter for their capacities to dislodge and replace the tapasin scoop loop from the F pocket. Mutation studies revealed the importance of specific residues within the scoop loop of tapasin, with significant differences for selection between different MHC-I alleles.

Author contributions: T.S., S.S., and A. Achour designed research; I.H., E.C.S., A. Apavaloaei, B.J.J., R.S., X.H., D.O., R.P., L.J., J.N., T.S., S.S., A.-M.G., and A.D.D. performed research; S.P. contributed new reagents/analytic tools; I.H., A. Apavaloaei, R.S., X.H., S.P., J.N., P.-Å.N., T.S., S.S., A.-M.G., A.D.D., and A. Achour analyzed data; and I.H., R.S., X.H., P.-Å.N., T.S., S.S., A.D.D., and A. Achour wrote the paper.

The authors declare no conflict of interest.

This article is a PNAS Direct Submission.

Data deposition: The atomic coordinates and structure factors have been deposited in the Protein Data Bank, www.wwpdb.org (PDB ID codes 6GB7, 6GB5, and 6GB6).

[1]A.D.D. and A.A. contributed equally to this work.

[2]To whom correspondence should be addressed. Email: adnane.achour@ki.se.

This article contains supporting information online at www.pnas.org/lookup/suppl/doi:10.1073/pnas.1807656116/-/DCSupplemental.

Published online February 26, 2019.

Here, we determined a suite of crystal structures of MHC-I molecules that provide insights in the molecular basis underlying peptide binding and selection by tapasin. Mutation studies revealed the importance of specific residues within the scoop loop of tapasin, with significant differences for selection between different MHC-I alleles. Altogether, our results indicate that candidate epitopes most probably bind first with their N termini to the MHC-I peptide binding cleft before being tested for their capacity to dislodge the tapasin scoop loop from the F pocket of MHC-I molecules.

Results and Discussion

We initiated the present study by determining the crystal structure of the murine MHC-I allele H-2K^b in complex with the dipeptide GL to 1.8-Å resolution (*SI Appendix*, Table S1). The unbiased electron density revealed that the peptide loading cleft was mainly empty, besides very few water molecules, most localized in the vicinity of the F pocket. Clear density allowed for unambiguous modeling of the GL dipeptide within the F pocket (Fig. 1 *A* and *B*). Importantly, the peptide binding cleft of H-2K^b/GL takes a nearly identical conformation compared with classical pMHC complexes such as H-2K^b bound to the Sendai virus-derived epitope SEV (FAPGNYPAL) (20), indicating that GL stabilization of the F pocket is enough to keep the peptide binding cleft in the same conformation as when a full-length peptide is bound (Fig. 1*C*). Thus, binding of GL renders the MHC-I cleft fully available for peptide binding and selection and does not result in any widening of the C-terminal section of the peptide cleft as described in the structures of the TAPBPR/MHC-I complex (9, 10). Several hydrogen bonds are formed between GL and the heavy chain residues D77, Y84, T143, K146, and W147, as well as three water molecules (Fig. 1*D*). It should also be noted that the omit map demonstrates that the side chain of the MHC-I residue Y84 is pointing toward the peptide binding cleft, forming two hydrogen bonds with the C terminus of GL (Fig. 1*D*).

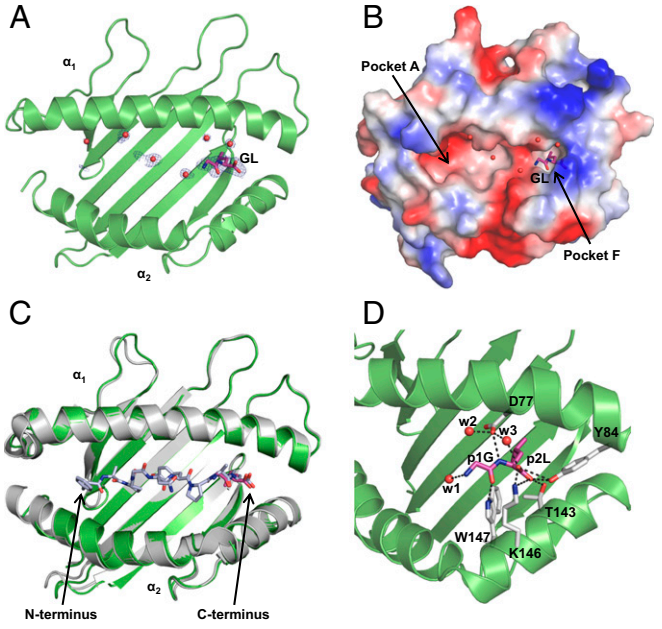

Fig. 1. The GL dipeptide binds within the MHC-I F pocket, resulting in a peptide-receptive conformation. (*A*) The 2FoFc omit map drawn at 1σ level reveals unambiguous electron density for the GL dipeptide within the F pocket of the H-2K^b peptide binding cleft. The omit map, in blue, also allows for unambiguous positioning of all of the side chains of the MHC-I peptide binding cleft. For clarity, only the omit map for the side chains of MHC residues comprising the helices is displayed. The omit map for MHC-I residues that are localized at the bottom of the peptide binding cleft is not displayed. The side chains of the MHC heavy chain residues are not displayed either. A few water molecules (red balls) are mostly grouped around the C-terminal part of the peptide binding cleft. (*B*) Surface representation of the peptide-binding cleft reveals that it is empty with three water molecules surrounding the GL dipeptide that binds in the F pocket. (*C*) The peptide binding cleft of the H-2K^b/GL complex, in white, takes the same conformation as in H-2K^b in complex with the full-length prototypic peptide SEV (FAPGNYPAL), in gray. No widening of the cleft is observed. (*D*) The side chain of the leucine residue in the GL dipeptide anchors in the F pocket while several hydrogen bonds (dashed lines) are formed between the COO$^-$ terminal of the dipeptide and the MHC-I residues Y84, T143, and K146. Additional hydrogen bonds are formed between the backbone of the dipeptide and three water molecules as well as the MHC-I residues D77 and W147.

We thereafter assessed if the GL dipeptide is present in any of the PLC protein components and, indeed, found that a GL sequence is indeed localized at the tip of a disordered loop in tapasin (*SI Appendix*, Fig. S1), which is not visible in the previously determined crystal structure of tapasin (21). The 5.8-Å cryo-EM analysis of the peptide loading complex indicated that this tapasin loop could be located close to the F pocket of MHC-I molecules (8). Hypothesizing that GL can bind simultaneously with other peptides, we refolded H-2D^b with both GL and a C-terminally truncated version of SEV corresponding to the stretch of residues FAPGNYP (FP7). The crystal structure of the H-2D^b/FP7/GL complex, determined to 2.3 Å, revealed that both GL and FP7 bound simultaneously to the MHC-I cleft. The dipeptide GL occupies the F pocket similarly to the H-2K^b/GL complex, while FP7 stretches throughout the peptide binding cleft (Fig. 2 *A*–*C*), with peptide residues p1–p5 taking the same conformation as in the crystal structure of H-2D^b in complex with the full-length SEV peptide (FAPGNYPAL) (22). In contrast, residues p6Y and p7P take significantly different conformations compared with the full-length SEV peptide (Fig. 2*D*). Thus, our structural analysis suggests that the N-terminal part of an MHC-I–restricted peptide possibly binds first to the peptide binding cleft while the C-terminal part is tested for its capacity to potentially dislodge the GL dipeptide, part of the tapasin loop, from the F pocket.

Next, we investigated if an MHC-I molecule could be refolded simultaneously with the tapasin scoop loop and the truncated peptide FP7. The crystal structure of H-2D^b in complex with both the mouse tapasin-associated loop-derived TL10 (EDAGGGGLSK) and the peptide FP7 was determined to 2.15 Å (Fig. 3 *A*, *B* and *C*). The p1–p5 section of FP7 takes a conformation identical to that in H-2D^b/FP7/GL (Fig. 3*D*). The flexibility of TL10 due to a large number of glycine residues allowed us to unambiguously define the exact conformation of only the five C-terminal residues of this tapasin-derived loop, including the leucine residue that clearly occupies the F pocket (Fig. 3 *A*–*C*). Thus, the GL tip of TL10 binds similarly to the dipeptide GL within the MHC-I F pocket while the remaining part of this loop protrudes out toward the solvent (Fig. 3*D*). Here again, while it has been demonstrated that the C-terminal part of MHC-I molecules widens upon binding to TAPBPR (9, 10), simultaneous binding of the tapasin-associated TL10 and the truncated peptide FP7 does not affect the conformation of the α_1 and α_2 helices, compared with the crystal structure of the same MHC-I molecule in complex with the full-length epitope SEV (Fig. 3 *E* and *F*). Furthermore, in the TAPBPR structures, attention was drawn to the positioning of the MHC-I heavy chain residue Y84, the side chain of which rotated away from the peptide binding cleft to interact with the TAPBPR residue E105 (*SI Appendix*, Fig. S2). The asymmetric unit of the TAPBPR comprises four complexes, two of which have the side chain of Y84 pointing as described in this previous report toward TAPBPR. However, it should be noted that the side chain of Y84 still points toward the peptide binding cleft of the MHC

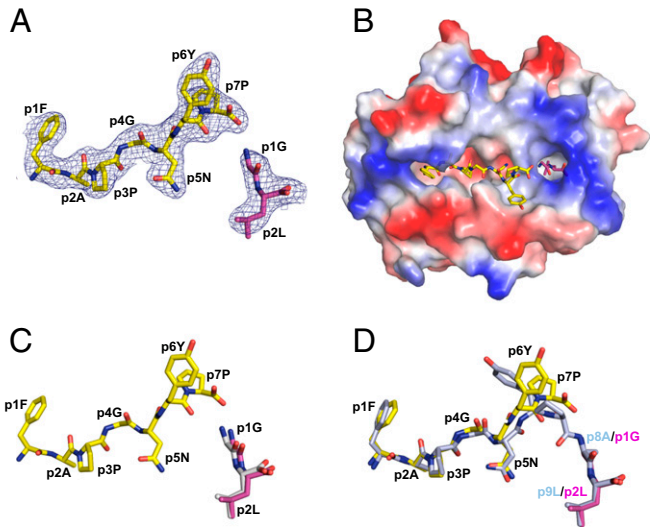

Fig. 2. A truncated version of SEV binds simultaneously with GL to the MHC-I peptide binding cleft. (*A*) The omit map for the H-2D^b/FP7/GL complex, displayed at 1σ, reveals that the truncated version of the SEV epitope, FP7 (FAPGNYP), and the dipeptide GL bind simultaneously to the MHC-I cleft. (*B*) Surface representation of H-2D^b with the two peptides displayed as sticks reveals how FP7 (in yellow) stretches throughout the peptide binding cleft, while GL (in magenta) binds to the F pocket. (*C*) Simultaneous binding of the FP7 and GL peptides (in yellow and magenta, respectively) does not affect the conformation of the dipeptide compared with H-2K^b/GL (in white). (*D*) Superposition of the H-2D^b/FP7/GL crystal structure on the crystal structure of H-2D^b in complex with the full-length SEV peptide reveals that the five N-terminal residues in FP7 (in yellow) take the same conformations as the corresponding stretch of residues in the full-length peptide SEV (in light blue). In contrast, the side chains of residues p6Y and p7P in FP7 take different conformations compared with SEV, protruding out toward the solvent. The dipeptide GL in H-2D^b/FP7/GL (in magenta) binds exactly as the two C-terminal residues p8A and p9L in SEV.

allele in the two other complexes within the same asymmetric unit (PDB ID code 5WER).

It is well established that residue Y84 forms two essential hydrogen bonds with the terminal COO^- group of any MHC-I–bound prototypic full-length peptide (23, 24). Here, in contrast to the TAPBPR structures, the crystal structure of the tapasin-derived TL10 scoop loop in complex with MHC-I keeps the same conformation for residue Y84 as in all known MHC/full-length peptide complexes. In contrast to the TAPBPR/MHC-I structures, the side chain of Y84 forms hydrogen bonds with the backbone of TL10 (PDB ID codes 5OPI and 5WER) (9, 10) (Fig. 3 and *SI Appendix*, Fig. S2). It should also be noted that since the scoop loop of TAPBPR comprises an α-helical segment, which dives toward the F pocket, it could provide a certain level of rigidity resulting in the widening of the MHC peptide binding cleft, while the scoop loop of tapasin should be more flexible due to the presence of a larger amount of glycine residue. It should therefore not be able to influence the structure of the MHC peptide binding cleft as much as observed for TAPBPR.

There are several possible explanations for this dichotomy between our and others' results regarding both the widening of the cleft and the rotation of the side chain of Y84 observed in the structures of TAPBPR in complex with pMHC molecules. Binding of the FP7 fragment at the N-terminal section of the peptide binding cleft, as revealed by the crystal structure of the H-2D^b/FP7/GL complex, could be enough for the C-terminal part of the MHC-I cleft to take a conformation similar to the final MHC-I/peptide product. However, in this case, we should also have observed a movement of the α_2 helix in the crystal structure of the H-2K^b/GL complex, which was not the case. A second possibility could be that only the jack-hairpin in both tapasin and TAPBPR is responsible for this observed movement of the α_2 helix, and that since we only have the scoop loop in our structures, the conformations of our pMHC are not affected. One possible consequence of this widening is that it could facilitate for the C-terminal section of the full-length peptide to reach within the F pocket and compete with the GL tip of the scoop loop. Finally, a third possibility is that TAPBPR and tapasin do not make use of equivalent molecular approaches to select MHC-I–restricted peptides.

To investigate if the flexible tapasin loop is essential for the cell surface expression of pMHC complexes, tapasin was knocked out in MEF cells and constructs containing mutated or truncated versions of tapasin were transduced into these cells. The cell surface expression levels of H-2D^b and H-2K^b were thereafter assessed using flow cytometry (Fig. 4). The intracellular expression

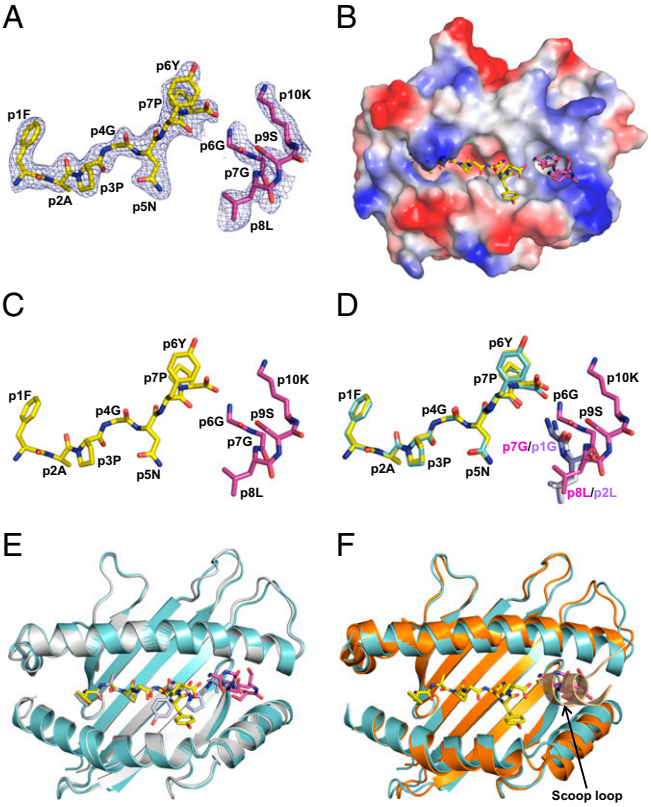

Fig. 3. The tapasin scoop loop-derived TL10 binds to the MHC-I F pocket. (*A*) The omit map for the H-2D^b/FP7/TL10 complex, displayed at 1σ, reveals that the truncated version of the SEV epitope, FP7 (FAPGNYP; in yellow), and the tapasin scoop loop TL10 (EDAGGGGLSK; in magenta) bind simultaneously to the MHC-I cleft. The side chain of the key leucine residue in TL10 is buried within the MHC-I F pocket. The quality of the electron density allowed for unambiguous positioning of the five residues GGLSK while the rest of TL10 is highly flexible. (*B*) Surface representation of the H-2D^b/FP7/TL10 complex with both the FP7 and TL10 peptides displayed as sticks reveal how the two epitopes occupy the peptide binding cleft of H-2D^b. (*C*) Side view of FP7 and TL10 reveals simultaneous occupancy of the peptide binding cleft. (*D*) The truncated peptide FP7 takes the same conformation in H-2D^b/FP7/TL10 as in the H-2D^b/FP7/GL complex. The side chain of the leucine residue in TL10 takes a similar conformation within the F pocket as for the GL dipeptide in both H-2D^b/FP7/GL and H-2K^b/GL, in white and purple, respectively. (*E*) The peptide binding cleft in H-2D^b/FP7/TL10 (in cyan) takes a similar conformation to H-2D^b in complex with the full-length peptide SEV (in white). (*F*) In contrast to the structure of TAPBPR/pMHC-I (in orange), there is no widening of the C-terminal part of the cleft (in cyan) in H-2D^b/FP7/TL10. The rigid TAPBPR α-helical scoop loop is displayed in light brown.

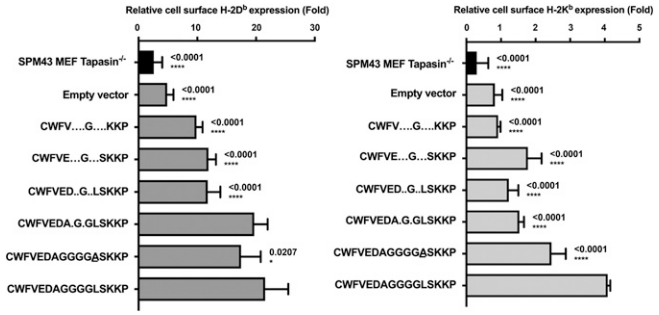

Fig. 4. Mutation of the leucine residue in the tapasin scoop loop reduces significantly H-2K^b cell surface expression levels. SPM43 MEF Tapasin-KO cells, transduced with lentiviral vectors encoding for mouse wild-type tapasin or mutated variants, were stained for cell surface expression of H-2D^b (*Upper*) and H-2K^b (*Lower*). The flow cytometry mean fluorescence intensity (MFI) for each sample was normalized to the background MFI of corresponding unstained sample. The experiment represents an average of two independent experiments in triplicates. Statistical analysis was calculated using one-way ANOVA. Asterisks indicate level of significance with $*P < 0.05$; $****P < 0.0001$. Deletion of the scoop loop abolished cell surface expression levels of both H-2D^b and H-2K^b. Mutation of the leucine residue in the tapasin scoop loop reduced significantly the cell surface expression levels of H-2K^b. In contrast, expression of H-2D^b complexes was unaffected.

levels of transduced tapasin were also verified by Western blot and confocal microscopy (*SI Appendix*, Fig. S3). While expression levels of some tapasin variants were reduced, other variants were expressed at similar levels compared with wild-type tapasin. Thus, our results demonstrate that mutation and/or removal of several residues within this loop drastically reduces the expression levels of both MHC-I alleles. Interestingly, while single mutation of the leucine residue reduced significantly the cell surface expression levels of H-2K^b, H-2D^b expression levels remained less affected. This indicates that the peptide loading and selection mechanisms steered by tapasin are most probably different with different MHC-I alleles. These results also suggest that the selection stringency based on the capacity of peptides to dislodge the loop from the F pocket is different from MHC-I allele to allele.

Previous studies of specific tapasin-dependent and tapasin-independent MHC class I alleles such as HLA-B4402/HLA-B4405, as well as HLA-B2705/HLA-B2709, respectively, revealed the key role of residue 116, which is the only residue that differs between these alleles, for tapasin dependence (25–27). It has also previously been suggested that the identity of residue 116 may result in increased or reduced conformational disorder within the F pocket, resulting in possible higher tapasin dependence for MHC-I alleles with higher conformational disorder. In the same study, the authors demonstrated that they could reduce conformational instability by mutating residue 116 and improve its interactions with residue 114, as well as reduce significantly the dependence of the substituted MHC-I allele for tapasin (28). The identity of residue 116 differs also between H-2K^b and H-2D^b (Y and F, respectively). However, the much larger amount of residues that differ between these two MHC-I alleles impairs our ability to conclude that residue 116 here again may alone account for tapasin dependence or independence (*SI Appendix*, Fig. S4).

Our mutational studies also indicate strong differences in how tapasin may select, with different levels of stringency, for epitopes that bind to different MHC alleles. Indeed, while single scoop loop mutations reduced significantly the cell surface expression levels of H-2K^b/peptide complexes, the levels of H-2D^b/peptide complexes were affected only after deletion of a large section of the tapasin loop. Thus, this result could provide an important molecular insight into the possible promiscuity or stringency of the peptide repertoire selection in different MHC-I alleles. Indeed, our results demonstrate that deletion and/or mutation of the leucine residue affects significantly the cell surface expression levels of H-2K^b complexes, while H-2D^b cell expression levels are significantly reduced only after the deletion of at least six residues in the scoop loop (Fig. 4).

As mentioned above, the sequence of the TAPBPR scoop loop is very different compared with tapasin. While the tip of the TAPBPR scoop loop comprises the stretch of residue ASSE, the corresponding loop in mouse tapasin is composed of GGLS (KGLA in human), most probably providing substantial differences in mobility and charge. Neither of the two previously published tapasin structures provides any information about the position of the key section of the scoop loop corresponding to TL10 (8, 21). Indeed, in the crystal structure of tapasin in complex with ERp57, this loop is not visible in the density and cannot be modeled (*SI Appendix*, Fig. S1) (21). This 3D model was also used as a template in the low-resolution cryo-EM structure of the PLC (8). Thus, as such, the previously published structural results do not provide any conclusive evidence for the positioning of this loop. It is however stated by Blees et al. (8) that the N-terminal β-barrel of tapasin appears very mobile in the PLC structure (*SI Appendix*, Fig. S5), meaning that the mobility of this section, including the scoop loop, extends well beyond the loop itself. Keeping this in mind, the much more flexible glycine-rich tapasin scoop loop cannot include a helical segment as found in TAPBPR, which provokes the widening of the binding groove. Instead our molecular model indicates that the GL tip of the tapasin scoop loop (G17-L18) can come close and bind into the F pocket (*SI Appendix*, Figs. S1 and S5).

Based on the results presented within this study as well as results from others (9), we here propose a hypothetical model underlying peptide loading mechanisms and selection for MHC-I molecules (Fig. 5). Our results do not provide any evidence for a widening of the peptide binding cleft of MHC-I molecules following binding to the tapasin-derived scoop loop TL10. Furthermore, the crystal structure of H-2K^b in complex with the dipeptide GL certainly demonstrates that no widening of the C-terminal part of the peptide binding cleft takes place when GL binds to the F pocket. However, this hypothesis needs to be confirmed by further structural studies involving the entire tapasin molecule and different MHC-I alleles. Altogether, our results demonstrate that the tapasin loop TL10 can bind to the F pocket, using the leucine residue as an anchor and, thereby, stabilize the peptide binding cleft. Possibly, the presence of a jack hairpin loop underneath the MHC-I peptide binding cleft may widen the C-terminal section, facilitating access for the C-terminal section of candidate epitopes to try to dislodge the GL tip of the scoop loop. The possible importance of a large amount of contacts formed between TAPBPR and MHC-I molecules for peptide editing cannot be excluded either. However, based on the analyses of our high-resolution crystal structures, our present hypothesis is that full-length peptides bind to specific MHC-I alleles with their N-terminal part binding first into the N-terminal part of the MHC-I peptide binding cleft, before being tested for their capacity to dislodge the tapasin loop from the F pocket and form the final MHC/peptide complexes (Fig. 5).

One possible explanation for these results is that the tapasin loop can be more easily dislodged in H-2D^b compared with H-2K^b, resulting in the presentation of a broader repertoire of peptides, including possibly a larger array of relatively lower-affinity epitopes by H-2D^b. Indeed, we were never able to obtain crystals of the H-2D^b/GL complex, possibly due to the fact that the dipeptide dissociates easier from this allele compared with H-2K^b. Furthermore, the binding motif requirements are much more stringent for the N-terminal and middle sections of H-2D^b–restricted epitopes compared with H-2K^b with much stricter requirements for peptide positions 2 (M, S, A), 3 (I, M, L, P, V) and most importantly position 5 (mostly N, possibly M)

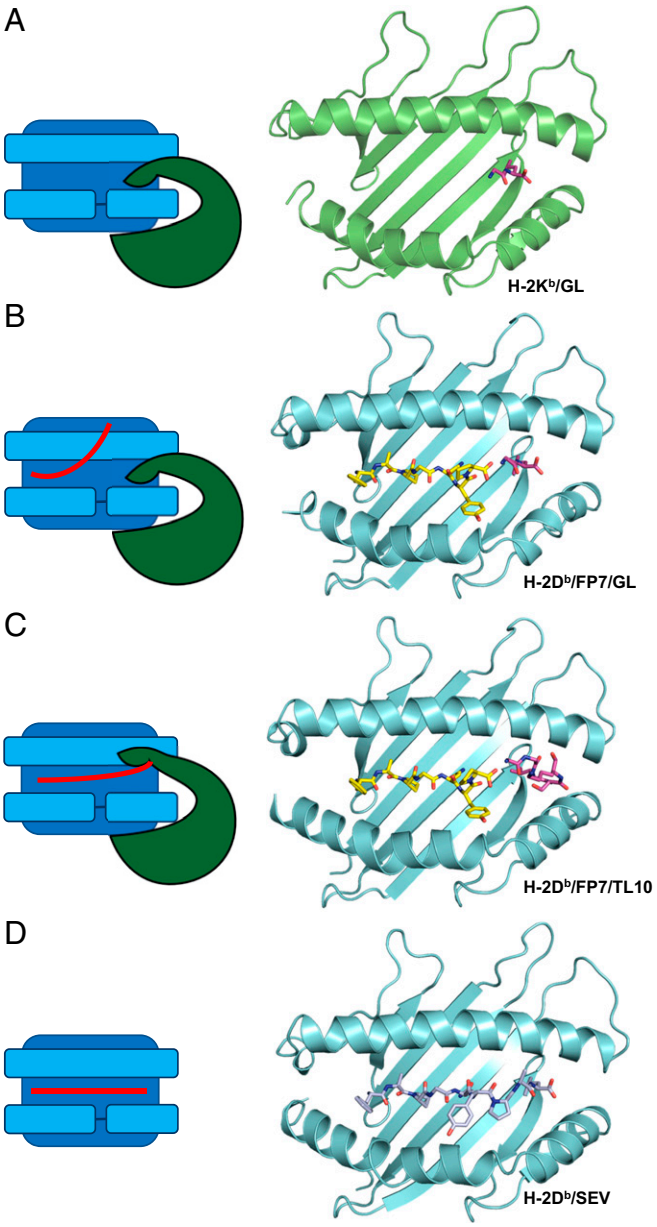

Fig. 5. Hypothetical molecular model underlying tapasin-mediated MHC-I peptide selection and binding. Based on the results obtained within this study, we propose the following hypothetical model for tapasin-mediated epitope selection and binding. (*A*) The association of the GL tip of the tapasin scoop loop (dark green) with peptide-empty MHC-I stabilizes the pMHC complex. (*B* and *C*) The candidate peptide binds to the MHC-I cleft first with its N terminus and can be released again if it displays insufficient affinity. If the peptide N-terminal section displays sufficient affinity to the cleft, the C-terminal part of the candidate peptide is tested for its capacity to dislodge the scoop loop. (*D*) If the candidate epitope is successful in dislodging the scoop loop, tapasin releases its grip from the cleft, and a full-length complex is formed.

(29, 30). Thus, a "stronger" and more stringent binding of the N-terminal and middle section in epitopes specific to H-2D^b could also reduce the potential entropic cost required for dislodging the tapasin loop compared with H-2K^b. In contrast, the more promiscuous motif for N-terminal residues in H-2K^b may thus require higher entropic cost for peptides to compete away the scoop loop, resulting potentially in a more stringent selection of adequate, often high-affinity, peptides.

Altogether, our results suggest that candidate epitopes may bind first with their N termini to the MHC-I peptide binding cleft before being tested for their capacity to dislodge the tapasin scoop loop from the F pocket of MHC-I molecules. We propose that the selection stringency based on the capacity of peptides to dislodge the loop from the F pocket is different from MHC-I allele to allele. This study together with recently published PLC structure and crystal structures of tapasin homolog protein TAPBPR could provide an important molecular insight into the possible promiscuity or stringency of the peptide repertoire selection in different MHC-I alleles. However, the impact of the tapasin scoop loop on peptide repertoire of different MHC-I alleles in health or disease has yet to be discovered.

Materials and Methods

Preparation and Crystallization of the MHC-I/Peptide Complexes. The GL peptide was purchased from Bachem. FP7 (FAPGNYP) and TL10 (EDAGGGGLSK) were prepared by microwave-assisted solid phase synthesis based on Fmoc chemistry (31) on a CEM Liberty peptide synthesizer. Peptides were purified by RP-HPLC on a Jasco BS-997-01 instrument equipped with a DENALI C-18 column from GRACE VYDAC (10 μm, 250 × 22 mm). Refolding and purification of all pMHC complexes were performed as described (3, 32–34). Refolding and purification of H-2K^b/GL and H-2D^b/FP7/GL were performed in the constant presence of 10 mM GL peptide.

Crystals for all pMHC complexes were obtained by using the hanging drop vapor diffusion method. The best H-2K^b/GL diffracting crystal appeared in pH 6.7, 2.15 M Na/K phosphate, 15% 2-Methyl-2,4-pentanediol (MPD). Two microliters of H-2K^b/GL in 10 mM Tris·HCl (pH 7.5), 150 mM NaCl, and 10 mM GL were mixed with 1 μL of well solution and equilibrated against 1 mL of well solution at 4 °C. The best crystals for H-2D^b/FP7/GL appeared in 2 M ammonium sulfate, 0.1 M Tris·HCl (pH 7.5), and 0.5 M NaCl at 4 °C. One microliter of 4 mg/mL protein solution in 10 mM Tris·HCl (pH 7.5), 150 mM NaCl, and 10 mM GL were mixed with 2 μL of crystallization reservoir solution, and thereafter equilibrated against 1 mL of well solution at 4 °C. The best crystals for H-2D^b/FP7/TL10 appeared in 1.8 M ammonium sulfate, 0.1 M Tris·HCl (pH 8), and 0.5 M NaCl at 4 °C. Two microliters of 3 mg/mL protein solution in 10 mM Tris·HCl (pH 7.5) and 150 mM NaCl were mixed with 2 μL of crystallization reservoir solution, and thereafter equilibrated against 1 mL of well solution at 4 °C.

Data Collection and Processing. Data collection for all of the pMHC complexes were performed under cryogenic conditions at both beam lines 14.1 (BESSY II) and ID30-A/MASSIF-1 (ESRF). Crystals were soaked in a cryoprotectant solution containing 25% glycerol before freezing. For H-2K^b/GL a total of 360 images were collected with 0.5° oscillation per frame. For H-2D^b/FP7/GL and H-2D^b/FP7/TL10, 2,000 and 1,800 images, respectively, were collected with 0.1° oscillation per frame. Data were processed with MOSFLM (35) and AIMLESS (36) from the CCP4 suite. Crystals for both H-2K^b/GL and H-2D^b/FP7/TL10 belong to the P21 space group, while H-2D^b/FP7/GL crystals belong to the I2 space group. Data collection statistics are provided in *SI Appendix*, Table S1.

Crystal structure determination and refinement. The crystal structures of H-2D^b/FP7/GL and H-2D^b/FP7/TL10 were determined by molecular replacement using Phaser (37) and the H-2D^b/gp33 complex with the peptide omitted (PDB ID code 1S7U) (3) as a search model. The Balbes automatic molecular replacement service was used for H-2K^b/GL. For H-2K^b/GL and H-2D^b/FP7/GL, two complexes were found in each asymmetric unit, while for H-2D^b/FP7/TL10, four complexes were found. Five percent of the total amount of reflections was set aside for monitoring refinement by R_{free}. Refinement of the crystal structure was performed using REFMAC (38) and Phenix (39). Unambiguous electron densities were observed in the peptide-binding clefts of all three pMHC complexes. Water molecules were added using COOT (40, 41) and their positions inspected manually. The stereochemistry of the final models was verified with COOT. Final refinement parameters are presented in *SI Appendix*, Table S1. Figures were created using PyMOL (The PyMOL Molecular Graphics System, Version 1.8 Schrödinger, LLC).

MHC-I cell surface expression analyses. SPM43 Tapasin-KO MEF cells were grown in DMEM, Glutamax supplemented with 10% FBS, and 1× Pen-Strep. HEK293FT cells were cultured in DMEM, Glutamax supplemented with 10% FBS, 1× nonessential amino acids, and 25 mM Hepes.

SPM43 Tapasin KO MEF cells were transduced with indicated Tapasin-iG2 Lentiviral particles (Fig. 4 and *SI Appendix, Supplementary Materials and Methods* and Fig. S4) overnight in the presence of 4 μg/mL polybrene

(Sigma-Aldrich). Transduction efficiency was checked on days 3 and 8 after transduction by flow cytometry by assessing GFP percentage (over 90% for all conditions). Tapasin expression was confirmed by Western Blot and confocal microscopy (*SI Appendix*, Fig. S4). Cell surface MHC-I levels were assessed by flow cytometry using anti-H-2K^b-PE (clone AF6-88.5; BioLegend) and H-2D^b-BV711 (clone KH95; BD) antibodies in triplicates. Data were analyzed using FlowJo V.10.5. (FlowJo LLC). Statistical analysis was performed using GraphPad Prism.

1. Trowsdale J, Knight JC (2013) Major histocompatibility complex genomics and human disease. *Annu Rev Genomics Hum Genet* 14:301–323.
2. Leone P, et al. (2013) MHC class I antigen processing and presenting machinery: Organization, function, and defects in tumor cells. *J Natl Cancer Inst* 105:1172–1187.
3. Achour A, et al. (2002) A structural basis for LCMV immune evasion: Subversion of H-2D(b) and H-2K(b) presentation of gp33 revealed by comparative crystal structure. Analyses. *Immunity* 17:757–768.
4. van Stipdonk MJ, et al. (2009) Design of agonistic altered peptides for the robust induction of CTL directed towards H-2Db in complex with the melanoma-associated epitope gp100. *Cancer Res* 69:7784–7792.
5. Hafstrand I, et al. (2016) The MHC class I cancer-associated neoepitope Trh4 linked with impaired peptide processing induces a unique noncanonical TCR conformer. *J Immunol* 196:2327–2334.
6. Hafstrand I, et al. (2018) The immunogenicity of a proline-substituted altered peptide ligand toward the cancer-associated TEIPP neoepitope Trh4 is unrelated to complex stability. *J Immunol* 200:2860–2868.
7. Blum JS, Wearsch PA, Cresswell P (2013) Pathways of antigen processing. *Annu Rev Immunol* 31:443–473.
8. Blees A, et al. (2017) Structure of the human MHC-I peptide-loading complex. *Nature* 551:525–528.
9. Thomas C, Tampé R (2017) Structure of the TAPBPR-MHC I complex defines the mechanism of peptide loading and editing. *Science* 358:1060–1064.
10. Jiang J, et al. (2017) Crystal structure of a TAPBPR-MHC I complex reveals the mechanism of peptide editing in antigen presentation. *Science* 358:1064–1068.
11. Pos W, et al. (2012) Crystal structure of the HLA-DM-HLA-DR1 complex defines mechanisms for rapid peptide selection. *Cell* 151:1557–1568.
12. Teng MS, et al. (2002) A human TAPBP (TAPASIN)-related gene, TAPBP-R. *Eur J Immunol* 32:1059–1068.
13. Boyle LH, et al. (2013) Tapasin-related protein TAPBPR is an additional component of the MHC class I presentation pathway. *Proc Natl Acad Sci USA* 110:3465–3470.
14. Schulze MS, Wucherpfennig KW (2012) The mechanism of HLA-DM induced peptide exchange in the MHC class II antigen presentation pathway. *Curr Opin Immunol* 24: 105–111.
15. Marin-Esteban V, Falk K, Rötzschke O (2004) "Chemical analogues" of HLA-DM can induce a peptide-receptive state in HLA-DR molecules. *J Biol Chem* 279:50684–50690.
16. Höpner S, et al. (2006) Small organic compounds enhance antigen loading of class II major histocompatibility complex proteins by targeting the polymorphic P1 pocket. *J Biol Chem* 281:38535–38542.
17. Gupta S, et al. (2008) Anchor side chains of short peptide fragments trigger ligand-exchange of class II MHC molecules. *PLoS One* 3:e1814.
18. Saini SK, et al. (2013) Dipeptides promote folding and peptide binding of MHC class I molecules. *Proc Natl Acad Sci USA* 110:15383–15388.
19. Saini SK, et al. (2015) Dipeptides catalyze rapid peptide exchange on MHC class I molecules. *Proc Natl Acad Sci USA* 112:202–207.
20. Fremont DH, Matsumura M, Stura EA, Peterson PA, Wilson IA (1992) Crystal structures of two viral peptides in complex with murine MHC class I H-2Kb. *Science* 257:919–927.
21. Dong G, Wearsch PA, Peaper DR, Cresswell P, Reinisch KM (2009) Insights into MHC class I peptide loading from the structure of the tapasin-ERp57 thiol oxidoreductase heterodimer. *Immunity* 30:21–32.
22. Glithero A, et al. (1999) Crystal structures of two H-2Db/glycopeptide complexes suggest a molecular basis for CTL cross-reactivity. *Immunity* 10:63–74.
23. Matsumura M, Fremont DH, Peterson PA, Wilson IA (1992) Emerging principles for the recognition of peptide antigens by MHC class I molecules. *Science* 257:927–934.
24. Lybarger L, et al. (2003) Enhanced immune presentation of a single-chain major histocompatibility complex class I molecule engineered to optimize linkage of a C-terminally extended peptide. *J Biol Chem* 278:27105–27111.
25. Serçinoğlu O, Ozbek P (2018) Computational characterization of residue couplings and micropolymorphism-induced changes in the dynamics of two differentially disease-associated human MHC class-I alleles. *J Biomol Struct Dyn* 36:724–740.
26. Abualrous ET, et al. (2015) F pocket flexibility influences the tapasin dependence of two differentially disease-associated MHC class I proteins. *Eur J Immunol* 45: 1248–1257.
27. Ostermeir K, Springer S, Zacharias M (2015) Coupling between side chain interactions and binding pocket flexibility in HLA-B*44:02 molecules investigated by molecular dynamics simulations. *Mol Immunol* 63:312–319.
28. Garstka MA, et al. (2011) Tapasin dependence of major histocompatibility complex class I molecules correlates with their conformational flexibility. *FASEB J* 25: 3989–3998.
29. Falk K, Rötzschke O, Stevanović S, Jung G, Rammensee HG (1991) Allele-specific motifs revealed by sequencing of self-peptides eluted from MHC molecules. *Nature* 351: 290–296.
30. Vita R, et al. (2015) The immune epitope database (IEDB) 3.0. *Nucleic Acids Res* 43: D405–D412.
31. Pellegrino S, Annoni C, Contini A, Clerici F, Gelmi ML (2012) Expedient chemical synthesis of 75mer DNA binding domain of MafA: An insight on its binding to insulin enhancer. *Amino Acids* 43:1995–2003.
32. Achour A, et al. (1999) Murine class I major histocompatibility complex H-2Dd: Expression, refolding and crystallization. *Acta Crystallogr D Biol Crystallogr* 55:260–262.
33. Sandalova T, et al. (2005) A structural basis for CD8+ T cell-dependent recognition of non-homologous peptide ligands: Implications for molecular mimicry in autoreactivity. *J Biol Chem* 280:27069–27075.
34. Sandalova T, et al. (2005) Expression, refolding and crystallization of murine MHC class I H-2Db in complex with human beta2-microglobulin. *Acta Crystallogr Sect F Struct Biol Cryst Commun* 61:1090–1093.
35. Leslie AG (2006) The integration of macromolecular diffraction data. *Acta Crystallogr D Biol Crystallogr* 62:48–57.
36. Evans PR, Murshudov GN (2013) How good are my data and what is the resolution? *Acta Crystallogr D Biol Crystallogr* 69:1204–1214.
37. McCoy AJ, et al. (2007) Phaser crystallographic software. *J Appl Cryst* 40:658–674.
38. Winn MD, Murshudov GN, Papiz MZ (2003) Macromolecular TLS refinement in REFMAC at moderate resolutions. *Methods Enzymol* 374:300–321.
39. Adams PD, et al. (2010) PHENIX: A comprehensive python-based system for macromolecular structure solution. *Acta Crystallogr D Biol Crystallogr* 66:213–221.
40. Emsley P, Lohkamp B, Scott WG, Cowtan K (2010) Features and development of coot. *Acta Crystallogr D Biol Crystallogr* 66:486–501.
41. Emsley P, Cowtan K (2004) Coot: Model-building tools for molecular graphics. *Acta Crystallogr D Biol Crystallogr* 60:2126–2132.