Immunogenetics (2016) 68:157–165
DOI 10.1007/s00251-015-0883-9

ORIGINAL PAPER

# A combined prediction strategy increases identification of peptides bound with high affinity and stability to porcine MHC class I molecules SLA-1*04:01, SLA-2*04:01, and SLA-3*04:01

**Lasse Eggers Pedersen**[1] · **Michael Rasmussen**[2] · **Mikkel Harndahl**[2] · **Morten Nielsen**[3,4] · **Søren Buus**[2] · **Gregers Jungersen**[1]

Received: 8 September 2015 / Accepted: 9 November 2015 / Published online: 14 November 2015


**Abstract** Affinity and stability of peptides bound by major histocompatibility complex (MHC) class I molecules are important factors in presentation of peptides to cytotoxic T lymphocytes (CTLs). In silico prediction methods of peptide-MHC binding followed by experimental analysis of peptide-MHC interactions constitute an attractive protocol to select target peptides from the vast pool of viral proteome peptides. We have earlier reported the peptide binding motif of the porcine MHC-I molecules SLA-1*04:01 and SLA-2*04:01, identified by an ELISA affinity-based positional scanning combinatorial peptide library (PSCPL) approach. Here, we report the peptide binding motif of SLA-3*04:01 and combine two prediction methods and analysis of both peptide binding affinity and stability of peptide-MHC complexes to improve rational peptide selection. Using a peptide prediction strategy combining PSCPL binding matrices and in silico prediction algorithms (NetMHCpan), peptide ligands from a repository of 8900 peptides were predicted for binding to SLA-1*04:01, SLA-2*04:01, and SLA-3*04:01 and validated by affinity and stability assays. From the pool of predicted peptides for SLA-1*04:01, SLA-2*04:01, and SLA-3*04:01, a total of 71, 28, and 38 % were binders with affinities below 500 nM, respectively. Comparison of peptide-SLA binding affinity and complex stability showed that peptides of high affinity generally, but not always, produce complexes of high stability. In conclusion, we demonstrate how state-of-the-art prediction and in vitro immunology tools in combination can be used for accurate selection of peptides for MHC class I binding, hence providing an expansion of the field of peptide-MHC analysis also to include pigs as a livestock experimental model.

**Keywords** MHC · Swine · Peptide binding prediction · Affinity · Stability

**Electronic supplementary material** The online version of this article (doi:10.1007/s00251-015-0883-9) contains supplementary material, which is available to authorized users.

✉ Gregers Jungersen
grju@vet.dtu.dk

[1] National Veterinary Institute, Technical University of Denmark, Bülowsvej 27, 1870 Frederiksberg C, Denmark

[2] Laboratory of Experimental Immunology, Department of International Health, Immunology and Microbiology, Faculty of Health and Medical Sciences, University of Copenhagen, København, Denmark

[3] Center for Biological Sequence Analysis, Technical University of Denmark, Lyngby, Denmark

[4] Instituto de Investigaciones Biotecnológicas, Universidad Nacional de San Martín, Buenos Aires, Argentina

## Introduction

MHC class I molecules can be seen as telltales of the immune system displaying the intracellular state of all nucleated cells to the surroundings by binding endogenously derived protein fragments from degraded self- as well as non-self proteins. Recognition of peptide-MHC complexes (pMHC) by cytotoxic T lymphocytes (CTLs) can lead to specific immune activation depending on the origin of the peptide. Cells infected by an intracellular pathogen will display non-self peptides on their cell surface hereby signaling a state of alert to circulating CTLs leading to their activation and subsequent elimination of the transformed or infected cells (Bevan 1995; Doherty and Zinkernagel 1975; Harty et al. 2000).

Peptide expression and processing is the prime prerequisite for epitope binding, and such MHC peptide binding is the most selective event in antigen presentation. Once bound, the stability of the pMHC influences the timeframe in which

peptides are displayed at the cell surface. Hence, the stability of individual pMHCs has a proposed impact on the immunogenicity of peptides, identifying stably bound peptides as being the more immunogenic (Busch and Pamer 1998; Harndahl et al. 2012; van der Burg et al. 1996).

Knowledge of the MHC-I binding specificity is paramount in epitope identification studies to limit the pool of peptides available for analysis. The peptide binding specificities of human leukocyte antigen (HLA) molecules have been thoroughly studied and several high performance prediction tools, such as *NetMHCpan*, exist to predict peptide-HLA binding (Lundegaard et al. 2011; Nielsen et al. 2007; Nielsen et al. 2008; Zhang et al. 2009). However, with exception of a few non-human primates (Allen et al. 1998; de Groot et al. 2013; Dzuris et al. 2000; McKinney et al. 2000; Sidney et al. 2007) and cattle species (Hansen et al. 2014), knowledge of the peptide binding motifs of MHC-I molecules from other species is limited, complicating epitope identification.

Mapping the MHC specificity in swine is important for epitope studies in infectious diseases of swine and rational development of novel porcine vaccines and for detailed studies of CTL responses when pigs are used as large animal models of human immunology and vaccine development. Recently, we reported the peptide binding motif of the porcine MHC-I molecules SLA-1*04:01 and SLA-2*04:01, identified by an ELISA affinity-based positional scanning combinatorial peptide library (PSCPL) approach (Pedersen et al. 2011; Pedersen et al. 2012). Here, we describe the PSCPL-derived peptide binding motif of the SLA-3*04:01 molecule by peptide affinity and a newly developed pMHC-I complex stability-based methodology (Rasmussen et al. 2014) which is also used to determine peptide stability with SLA-1*04:01 and SLA-2*04:01. Following the approach described previously for human and cattle regarding peptide selection and binding specificity characterization of MHCs (Hansen et al. 2014; Rasmussen et al. 2014), an extension was here made to include swine using a combination of the individual binding matrices for SLA-1*04:01, SLA-2*04:01, and SLA-3*04:01 with in silico peptide binding predictions using *NetMHCpan* (Hoof et al. 2009; Lundegaard et al. 2011). Similar to the approach taken earlier for human and cattle, we predicted possible peptide binders from a set of some 8900 nonameric peptides to investigate if combined predictions were more accurate than using either of the prediction tools alone. Predicted peptides were screened for binding, and high-affinity binders were identified for SLA-1*04:01, SLA-2*04:01, and SLA-3*04:01. In addition, the stability of peptide-SLA complexes (pSLA) was measured for peptides identified to be bound with high affinity, showing that our recently published peptide-HLA stability assay can easily be modified to measure pMHC stability in other species (Harndahl et al. 2011). By measuring pSLA binding affinity and complex stability peptide ligands of both high affinity and high stability were identified for all three SLA proteins and revealed a strong correlation between high-affinity binding and stability.

All together, we here provide a set of state-of-the-art tools for the prediction, analysis, and selection of peptides for MHC class I binding, hereby paving the way to identification of strongly, and stably, bound peptides harboring the potential to be important CTL peptide epitopes. These experiments introduce an expansion of the field of pMHC analysis also to include pigs as a livestock experimental model.

## Materials and methods

### Recombinant construct encoding the SLA-3*04:01 protein

Recombinant SLA-3*04:01 protein was produced as described elsewhere (Pedersen et al. 2011). Briefly, DNA encoding residues 1 to 276 of the SLA-3*04:01 alpha chain (purchased from Genscript) was cloned into a pET28a+ vector (Novagen) upstream of a polytag consisting of a FXa cleavage site, a biotinylation signal peptide (BSP), and a histidine affinity tag (HAT). Correct constructs were verified by sequencing (ABI Prism 3100Avant, Applied Biosystems) (Hansen et al. 2001). The validated construct of interest was transformed into *Escherichia coli* BL21(DE3), containing the pACYC184 expression plasmid (Avidity) encoding an IPTG inducible *BirA* gene to express biotin-ligase. This leads to almost complete in vivo biotinylation (85–100 %) of the desired product (Leisner et al. 2008).

### Expression and purification of recombinant SLA-3*04:01, human, and porcine beta-2 microglobulin ($\beta_2$m)

Recombinant SLA-I heavy chain proteins were produced by overexpression in *E. coli* and purified as described previously (Pedersen et al. 2011).

In brief, SLA-I proteins were produced in *E. coli* BL21(DE3) transformed with the pET28a+ vector encoding SLA-I residues 1-275 fused C-terminally to a biotinylation sequence peptide (BSP) for in vivo biotinylation and a histidine affinity tag (HAT) for purification purposes. Following cell lysis by mechanical cell disruption (Constant Systems cell disruptor), inclusion bodies containing SLA-I proteins were isolated by centrifugation. Inclusion bodies were washed to remove cell debris and the protein extracted into Tris-buffered urea (8 M urea, 25 mM Tris, pH 8.0). SLA-I proteins were purified under denaturing conditions by successive immobilized metal adsorption, hydrophobic interaction, and size-exclusion chromatography.

Recombinant human and porcine $\beta_2$m were expressed and purified as described elsewhere (Ferre et al. 2003; Ostergaard et al. 2001; Pedersen et al. 2011).

### Peptides

All peptides were purchased from Schafer-N, Denmark (www.schafer-n.com). Briefly, they were synthesized by standard 9-fluorenylmethyloxycarbonyl (Fmoc)-chemistry, purified by reverse-phase HPLC (to at least >80 % purity, frequently 95–99 % purity), validated by mass spectrometry, and quantified by weight.

### ELISA-based peptide-SLA binding assay

Peptide binding analysis of peptide interactions with the SLA-1*04:01 was performed using a previously described ELISA-based pMHC-I binding assay (Pedersen et al. 2012).

### Peptide-SLA binding affinity assay using luminescent oxygen channeling assay (LOCI) detection

Peptide-SLA binding affinity to SLA-3*04:01 and SLA-2*04:01 proteins was measured using a conformation-dependent anti-SLA monoclonal antibody-based LOCI technology as previously described for HLA (Harndahl et al. 2009). In brief, recombinant, biotinylated SLA heavy chains were diluted in PBS/0.1 % Lutrol F68 (BASF, Art. No. S30101) to a concentration of 2 nM followed by the addition of 15 times excess of porcine $\beta_2$m and increasing concentrations of peptide (spanning from 0.01 nM to 66.6 μM), and incubated at 18 °C for 48 h. pSLA complexes were detected using Alphascreen® streptavidin donor beads (PerkinElmer 6760002) and anti-SLA monoclonal antibody (clone PT85A, VMRD) conjugated Alphascreen® acceptor beads (PerkinElmer 6762002). Alphascreen® beads and pSLA complexes were incubated at 18 °C overnight at a final bead concentration of 5 μg/ml of each and equilibrated to reader temperature 1 h prior to detection by an EnVision® Multilabel Reader (PerkinElmer).

### Peptide-SLA binding stability assay using scintillation proximity assay (SPA)-based pMHC-I dissociation

The measurement of pSLA complex stability was done as previously described for HLA (Harndahl et al. 2011). Briefly, recombinant, biotinylated SLA-I heavy chain was diluted into renaturing buffer containing $^{125}$I-radiolabeled porcine $\beta_2$m and a peptide of interest in streptavidin-coated scintillation microplates (SA FlashPlate® PLUS, PerkinElmer, Boston, USA). Steady state of the folding reaction was reached by overnight incubation at 18 °C, and dissociation was initiated by addition of excess of unlabeled $\beta_2$m and by elevating the temperature to 37 °C. Dissociation was monitored continuously in a scintillation microplate reader (Topcount NXT, PerkinElmer).

Peptide-SLA dissociation data were fitted to a monophasic dissociation model by nonlinear regression. PSCPL data were analyzed as described previously (Rasmussen et al. 2014), and the obtained binding matrix was visualized by the Seq2Logo server (Thomsen and Nielsen 2012) normalizing the values at each position to 1, using P-weighted Kullback-Leibler logo types and an even amino acid background distribution (0.05 for all amino acids). This assay was also used for mapping of the SLA-3*04:01 peptide binding motif by the PSCPL approach. The experimental strategy of PSCPL has previously been described for both HLA (Stryhn et al. 1996) and SLA (Pedersen et al. 2011) proteins.

## Results

### Mapping of the SLA-3*04:01 peptide binding motif using positional scanning combinatorial peptide libraries

We have recently reported the peptide binding specificities of swine leukocyte antigens SLA-1*04:01 and SLA-2*04:01 (Pedersen et al. 2011; Pedersen et al. 2012). These peptide binding specificities were determined in an unbiased way by the use of a nonameric positional scanning combinatorial peptide library (PSCPL) analysis using an ELISA-based peptide-SLA-I binding assay. Here, we have applied a recently developed method to determine human MHC-I binding specificities (Hansen et al. 2014; Rasmussen et al. 2014), based on a pMHC-I dissociation assay (Rasmussen et al. 2014). This method is based on the dissociation of $^{125}$I-radiolabeled $\beta_2$m from in situ formed pHLA complexes. By $^{125}$I radiolabeling of porcine $\beta_2$m, we adopted this methodology to allow determination of the SLA-3*04:01 binding motif, illustrated in Fig. 1 and Table 1.

PSCPL analysis of the SLA-3*04:01 molecule revealed a binding motif with a primary anchor in the position 2 (P2) of the peptide with an anchor position (AP) value of 27 and showing preferences for asparagine (N) and arginine (R) with relative binding (RB) values of 5.6 and 2.3, respectively (Fig. 1 and Table 1). This suggests a polar, acidic P2-pocket within the binding groove. Auxiliary anchors were observed in P1, P3, P7, and P9 with AP values of 9, 11, 11, and 9, respectively. For P9, it was found that the hydrophobic amino acids methionine (M), phenylalanine (F), and tyrosine (Y) were preferred with RB values of 2.4, 2.3, and 2.0, respectively (Table 1). Furthermore, it was observed that arginine (R), glutamine (Q), lysine (K), and proline (P) were all disfavored at P9 (RB values of 0.2, 0.3, 0.3, and 0.4, respectively). Altogether, these results indicate the presence of a hydrophobic pocket accommodating the C-terminal peptide residue within the binding groove of SLA-3*04:01. An auxiliary anchor expressing a preference for the tryptophan (W), leucine (L), and phenylalanine (F) was observed in P7. These results

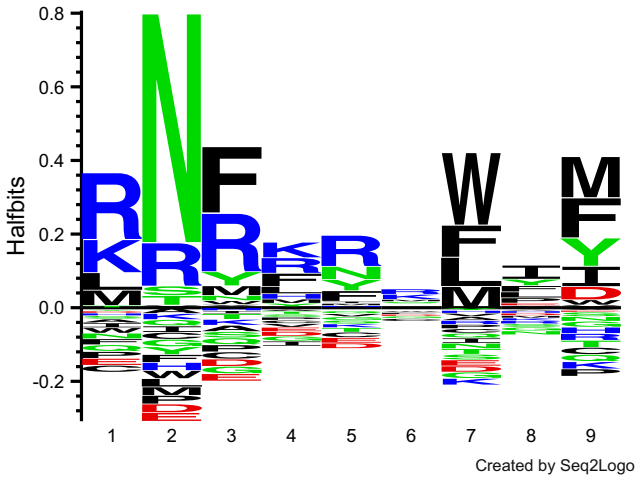


**Fig. 1** SLA binding motif. Peptide binding motif of the SLA-3*04:01 molecule determined by pMHC dissociation-driven PSCPL analysis and shown as a sequence logo generated from the PSCPL-derived peptide binding matrix using the Seq2Logo server (Thomsen and Nielsen 2012) normalizing scores for each peptide position to 1, using P-weighted Kullback-Leibler logo types with an even amino acid background distribution (0.05 for all amino acids). The height of each stack represents the relative importance of the position to peptide binding. The height of each individual amino acid within a given position relates the preference of the given amino acid residue in that position. *Positive values* indicate preferred residues, *negative values* are disfavored. (Note: The color version of this figure only appears in the online version)

indicate that P2 is the most important position with respect to peptide binding for SLA-3*04:01.

### Analysis of peptide-SLA binding affinity

In order to validate the peptide library-derived binding SLA-I motif obtained here and those described previously (Pedersen et al. 2011; Pedersen et al. 2012), we employed a combined peptide binding prediction strategy. Some 8900 nonamer peptides from an in-house repository were assigned scores using both the *NetMHCpan* prediction method (version 2.4) (Nielsen et al. 2007; Zhang et al. 2009) and the position-specific binding matrices for each of the three SLA-1 molecules (SLA-1*04:01, SLA-2*04:01, and SLA-3*04:01). Next, the two scores were combined taking the harmonic mean of the two percentile rank scores, obtained by comparing the prediction values to a distribution obtained from a set of 200,000 random natural peptides. Applying this prediction strategy and setting a threshold of 0.5 % percentile rank score, a total of 156, 72, and 76 of the in-house peptides were predicted to be bound by SLA-1*04:01, SLA-2*04:01, and SLA-3*04:01, respectively (Table 2 and Supporting Table 1 A–C). The peptide-SLA-I binding affinity was then measured for the predicted binding peptides using direct binding assays (Pedersen et al. 2011), and peptides were grouped as high, intermediate, weak, or non-binders (Table 2). Individual peptide sequences, origin, prediction rank scores, and $K_D$ affinity values are shown in Supporting Table 1 A–C.

**Table 1** Nonameric peptide binding motif

| SLA-3*04:01 | Amino acid position in peptide | | | | | | | | |
|---|---|---|---|---|---|---|---|---|---|
| | 1 | 2 | 3 | 4 | 5 | 6 | 7 | 8 | 9 |
| A | 0.8 | 0.7 | 0.6 | 0.8 | 0.9 | 1.0 | 0.7 | 0.9 | 0.9 |
| C | *0.4* | 0.6 | *0.5* | 0.7 | *0.5* | 0.5 | 0.7 | 0.9 | *0.5* |
| D | 0.9 | *0.4* | *0.3* | 0.5 | *0.4* | 0.7 | 0.5 | 0.8 | 1.5 |
| E | *0.4* | *0.4* | *0.4* | 0.6 | *0.4* | 0.9 | 0.5 | 1.2 | 0.9 |
| F | 0.6 | 0.5 | **3.0** | 1.7 | 1.6 | 0.8 | **2.2** | 1.5 | **2.3** |
| G | 0.9 | 0.6 | *0.4* | *0.3* | 0.7 | 0.6 | *0.5* | 0.5 | 0.6 |
| H | 0.8 | 0.5 | 0.8 | 1.3 | 1.1 | 1.0 | 0.8 | 0.8 | 0.6 |
| I | 0.7 | 0.7 | 0.8 | *0.5* | 1.1 | 1.0 | 0.6 | 1.9 | 1.8 |
| K | **2.2** | 0.7 | 0.6 | 1.8 | 0.7 | 1.5 | *0.4* | 0.7 | *0.3* |
| L | 1.7 | *0.5* | 1.1 | 1.4 | 0.9 | 1.1 | **2.1** | 1.4 | 0.9 |
| M | 1.6 | *0.5* | 1.4 | 1.3 | 1.1 | 1.2 | 1.9 | 0.7 | **2.4** |
| N | 0.6 | **5.6** | 1.2 | 0.9 | 1.7 | 1.1 | 0.6 | *0.5* | 0.7 |
| P | *0.3* | *0.4* | *0.5* | 0.7 | 0.9 | 0.9 | 0.7 | 1.4 | *0.4* |
| Q | *0.5* | 0.7 | 0.5 | 1.1 | 0.8 | 1.1 | 0.7 | 0.9 | *0.3* |
| R | **3.1** | **2.3** | **2.8** | 1.8 | **2.4** | 1.5 | 0.7 | 0.7 | *0.2* |
| S | 0.8 | 1.3 | 0.6 | 0.8 | 0.7 | 1.2 | 0.6 | 0.6 | 0.7 |
| T | 1.1 | 1.3 | 0.7 | 0.8 | 0.6 | 1.1 | 0.6 | 1.0 | *0.5* |
| V | 1.0 | 1.1 | 1.1 | 0.7 | 0.8 | 0.9 | 0.8 | 1.2 | 1.1 |
| W | 0.6 | 0.5 | 1.2 | 1.0 | 1.0 | 0.8 | **3.4** | 0.9 | 1.2 |
| Y | 0.9 | 0.6 | 1.6 | 1.2 | 1.6 | 1.0 | 0.8 | 1.5 | **2.0** |
| Sum | 20 | 20 | 20 | 20 | 20 | 20 | 20 | 20 | 20 |
| AP value | 9 | 27 | 11 | 4 | 5 | 1 | 11 | 3 | 9 |

Nonameric peptide binding motif of SLA-3*04:01 derived from pMHC-I dissociation assay-driven PSCPL analysis. Relative binding (RB) values greater than 2 are defined as favored AAs for peptide binding (marked in **bold**), whereas values below 0.5 are defined as destabilizing (marked in *italic*), (Lauemoller et al. 2001; Stryhn et al. 1996). As previously defined, AP values above 15 are considered as dominating anchors of importance illustrated by underlining (Lamberth et al. 2008)

### Evaluation of the combined prediction strategy

To quantify the complementarities of the PSCPL matrix- and *NetMHCpan*-based prediction methods, the peptides predicted to bind were split into subsets predicted (a) by *either* the matrix method *or* the *NetMHCpan* method, (b) by *both* the matrix and

**Table 2** Peptide distribution

| Peptide affinity ($K_D$, nM) | | | | |
|---|---|---|---|---|
| Molecule | ≤50 | 50–500 | 500–5000 | ≥5000 |
| SLA-1*04:01 | 66 | 45 | 20 | 25 |
| SLA-2*04:01 | 6 | 14 | 20 | 32 |
| SLA-3*04:01 | 20 | 9 | 9 | 38 |

Distribution of peptides predicted for SLA binding by the combined *NetMHCpan*-PSCPL matrix prediction method by measured peptide affinity values. The total numbers of identified strong ($K_D$≤50 nM), intermediate ($K_D$ 50–500 nM), and weak ($K_D$ 500–5000) as well as non-binding ($K_D$≥5000 nM) peptides are indicated for each SLA molecule

*NetMHCpan* methods (combined), or (c) *exclusively* by one of the two (Table 3). From this comparison, it is seen that the two methods are highly complementary. Depending on the SLA molecule, between 45 and 59 % of the peptides validated to be bound were predicted by both methods (Table 3, combined prediction—sensitivity). Also, it is apparent that one method is not consistently more accurate in identifying SLA peptide binders than the other. For SLA-2*04:01, only 2 of the 20 peptides (10 %) predicted exclusively by the matrix-based method turned out to be actual binders, corresponding to a gain in sensitivity of 10 % (2 of 20). For the *NetMHCpan* method, the same numbers were 36 % (9 of 25) and 45 % (9 of 20) gain in sensitivity. In contrast to this, the exclusive matrix-based method performed predictions for the SLA-3*04:01 molecule predicted 47 % actual binding peptides (8 of 17) with a corresponding gain in sensitivity of almost 28 % (8 of 29), whereas the *NetMHCpan* exclusive predictions for this SLA molecule only identified 11 % binding peptides (4 of 35) corresponding to a gain in sensitivity of 14 % (4 of 29). The predictive positive value for the two methods on the two SLA molecules (SLA-2*04:01 and SLA-3*04:01) is statistically significant ($p<0.05$ in both cases, comparison of ratio), emphasizing the complementarities of the two individual methods and underlining the importance of including both to ensure optimal capture of the binding repertoire of SLA molecules.

## pSLA stability analysis

pMHC-I complex stability has been proposed to be a better determinant of immunogenicity compared to pMHC-I binding affinity (Busch and Pamer 1998; Harndahl et al. 2012; van der Burg et al. 1996). In order to investigate the stability of pSLA complexes, we modified an existing scintillation proximity assay (SPA)-based pMHC dissociation assay (Harndahl et al. 2011) enabling us to measure pSLA complex stability. As peptide binding is a prerequisite for complex stability, only peptides binding with an affinity below 500 nM to SLA-1*04:01, SLA-2*04:01, or SLA-3*04:01 were screened for pSLA complex stability. It was found that the pSLA stability ranged from unstable ($t_{½}$ of 0.18 h) to extremely stable ($t_{½}$ of 158 h). The relationship between pSLA affinity and stability is shown for all three SLA molecules in Fig. 2a–c. From these plots, it is apparent that in most cases, peptides of high affinity also produce highly stable peptide-SLA complexes (Spearman's rank correlation between affinity and stability is highly significant—$p<0.01$—in all three plots), and in general, it was observed that peptides found to bind with high affinity also formed highly stable pSLA complexes with half-lives ≥1 h (Fig. 2a–b). However, several examples are also displayed in the figure where high-affinity peptides have low stability, confirming the earlier finding that stability carries additional information not captured by the affinity measurement (Harndahl et al. 2012; Jorgensen et al. 2014). It was observed that SLA-3*04:01 complexes were more unstable than the equivalent pSLA-1*04:01 and pSLA-2*04:01, having only six peptides with stability $t_{½}$ periods ranging between 1.1 and 6.7 h (Fig. 2c + Supporting Table 1 C). In comparison, the SLA-1*04:01 had 65 peptides with stabilities ranging from $t_{½}$ of 1 h to more than 158 h, and several of those had $t_{½}$ of more than 10 h (Fig. 2a + Supporting Table 1 A). For

**Table 3** Comparison of prediction methods

| PSCPL matrix *or* NetMHCpan prediction | Peptides predicted by *either* PSCPL *or* NetMHCpan | Peptides bound§ | Sensitivity |
|---|---|---|---|
| SLA-1*04:01 | 156 | 111 (71 %) | NA |
| SLA-2*04:01 | 72 | 20 (28 %) | NA |
| SLA-3*04:01 | 76 | 29 (38 %) | NA |
| Combined prediction | Peptides predicted by PSCPL *and* NetMHCpan | Peptides bound§ | Sensitivity |
| SLA-1*04:01 | 61 | 53 (87 %) | 53/111 (48 %) |
| SLA-2*04:01 | 27 | 9 (33 %) | 9/20 (45 %) |
| SLA-3*04:01 | 24 | 17 (71 %) | 17/29 (59 %) |
| NetMHCpan prediction | Peptides predicted by NetMHCpan *exclusively* | Peptides bound§ | Sensitivity |
| SLA-1*04:01 | 74 | 43 (58 %) | 43/111 (39 %) |
| SLA-2*04:01 | 25 | 9 (36 %) | 9/20 (45 %) |
| SLA-3*04:01 | 35 | 4 (11 %) | 4/29 (14 %) |
| PSCPL matrix prediction | Peptides predicted by PSCPL *exclusively* | Peptides bound§ | Sensitivity |
| SLA-1*04:01 | 20 | 15 (75 %) | 15/111 (14 %) |
| SLA-2*04:01 | 20 | 2 (10 %) | 2/20 (10 %) |
| SLA-3*04:01 | 17 | 8 (47 %) | 8/29 (28 %) |

Comparison of the PSCPL matrix- and NetMHCpan-based prediction methods to identify binding peptides for SLA-1*04:01, SLA-2*04:01, and SLA-3*04:01. The sensitivity of each prediction scenario is given as number of validated peptide binders (§) within the pool of peptides predicted by the respective criteria relative to all identified peptide binders from the pool of PSCPL matrix- or NetMHCpan-predicted candidates

§ Peptides bound are defined as peptides bound with affinities<500 nM

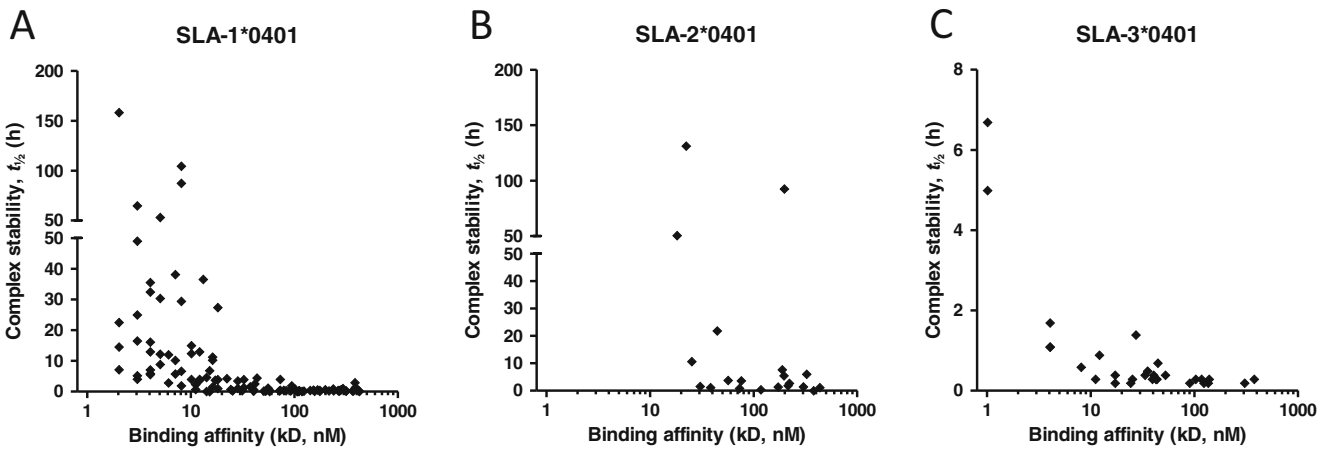


**Fig. 2** High-affinity SLA-I peptide ligands form stable complexes. Peptides with binding affinities below 500 nM were screened for complex stability using a pMHC-I dissociation assay. The measured pSLA-I half-lives are shown as a function of the measured binding affinity for **a** SLA-1*04:01 ($n$=111), **b** SLA-2*04:01 ($n$=20), and **c** SLA-3*04:01 ($n$=29)

SLA-2*04:01, 18 of the 20 peptides with measured $K_D$ affinity <500 nM were shown to be stably bound with $t_{½}$ >1 h (Supporting Table 1 B). Note that the analysis here for SLA-2*04:01 and SLA-3*04:01 is limited to the relatively few peptides identified with $K_D$ values below 500. In addition, our in-house peptide repository could lack peptide sequences that match the SLA-3*04:01 molecule compared to SLA-1*04:01, and to some extent also SLA-2*04:01. This is based on the fact that the majority of the peptides in the library pool has been synthesized to match HLA-I specificities (more closely resembling, although not equal to, SLA-1 and SLA-2 specificities). Additional analyses including a larger number of peptides would be necessary to draw a more general conclusion in regard to those two molecules and their pSLA stabilities.

## Discussion

Characterization of the peptide binding properties of swine MHC class I molecules can contribute tremendously to novel epitope discovery as well as to the overall validation and analysis of currently available or newly developed livestock vaccines including a CTL component. Experimental identification of epitope candidates continues, however, to be a highly cost-intensive task given the large number of different peptides found within a viral proteome and the functional diversity of MHC molecules.

In this light, and due to the fact that haplotype Lr-4.0 expressing the SLA-1*04:01, SLA-2*04:01, and SLA-3*04:01 allele combination is one of the most common haplotypes among swine populations in Denmark as well as worldwide (Ho et al. 2009; Pedersen et al. 2014), we wanted to analyze the peptide binding specificity of SLA-3*04:01 and validate the specificity of SLA-3*04:01 and previously published binding motifs of SLA-1*04:01 and SLA-2*04:01 (Pedersen et al. 2011; Pedersen et al. 2012) with discrete binding data. We employed a recently developed method to determine pHLA-I peptide binding specificities using peptide libraries (Rasmussen et al. 2014) that have been shown to be easily adaptable to determine peptide binding specificities of MHC-I molecules from other species (Hansen et al. 2014). Applying a nonameric peptide library to analyze the SLA-3*04:01, peptide binding specificity revealed a single dominating anchor in peptide position P2 (Table 1) preferring arginine and glutamine. No dominating anchor in the C-terminal peptide position was observed, a position commonly harboring a primary anchor in other SLA-I and HLA-I proteins (Lamberth et al. 2008; Pedersen et al. 2011; Pedersen et al. 2012; Rammensee et al. 1999; Rapin et al. 2008; Sette and Sidney 1999). Auxiliary anchors positions were observed in peptide positions 1, 3, 7, and 9.

To validate the peptide library-derived binding motifs with discrete binding data, we selected peptides for binding analysis using a prediction strategy combining the obtained binding matrices with the state-of-the-art predictor of pMHC-I binding *NetMHCpan* (Hoof et al. 2009; Nielsen et al. 2007) that have previously proved effective for identification of high-affinity MHC-I ligands (Hansen et al. 2014; Rasmussen et al. 2014). Using this approach, we identified a number of high-affinity ligands for SLA-1*04:01, SLA-2*04:01, and SLA-3*04:01 showing that combining prediction methods results in a larger number of actual binding peptides.

High-affinity binding is a prerequisite for a peptide to be immunogenic, but high-affinity HLA-I ligands can form complexes of varying stability and pMHC-I complex stability has been shown to correlate with immunogenicity (Busch and Pamer 1998; Harndahl et al. 2012; van der Burg et al. 1996). In order to test if the high-affinity SLA-I binding peptides found could form stable complexes, peptides found to bind with an affinity better than 500 nM were screened for pSLA-I

complex stability. Indeed, stable pSLA-I complexes were observed for each of the three SLA-I allotypes. For both SLA-1*04:01 and SLA-2*04:01, peptides forming complexes with half-lives of >100 h were found, while the most stable complex observed for SLA-3*04:01 had a $t_{½}$ of 6.7 h (Fig. 2a–c). Why SLA-3*04:01 forms less stable complexes compared to SLA-1*04:01 and SLA-2*04:01 remains to be addressed in detail. Factors such as the low number of peptides tested, limitations in regard to the diversity of peptides in the in-house peptide repository, or the fact that SLA-3*04:01 seems to lack the otherwise commonly expressed C-terminal anchor could all be causative for the relatively low stability of the pSLA-3*04:01 complexes analyzed. Interaction between the $\beta_2$-microglobulin subunit and the MHC class I molecule exists by contact with the much more conserved $\alpha_3$-domain of the MHC molecule, as compared to the more diverse peptide-binding $\alpha_1$ and $\alpha_2$ regions. Hence, the affinity of this interaction is not expected to be accounted for with regard to the number of binding peptides that are found for the individual MHC molecules. As seen in Supporting Tables 1 A + B, a trend for peptides satisfying *both* the P2 and P9 anchors of SLA-1*04:01 and SLA-2*04:01, as well as the P3 anchor of SLA-1*04:01, is that they are bound with higher stability, as compared to peptides that satisfy only *one* of the two anchors (or only *two* of the *three* anchors with regard to SLA-1*04:01) which tend to be bound with lower stability. It has been proposed that optimal residues at both the N- and C-terminus of the peptide are required to form a high stability interaction (Harndahl et al. 2012; Jorgensen et al. 2014). Hence, for each anchor satisfied, the stability of the bound peptide seems to increase, which could explain why the "one-anchor-only" molecule, SLA-3*04:01, produces mostly complexes of low stability. Reflecting on the biological relevance, expression of a single major anchor for P2 could be a phenotypic preference in regard to SLA-3*04:01 leading to faster turnover rates in the sense of peptide association and dissociation, which again would claim pSLA-3*04:01 of lower stabilities as compared to other molecules showing slow peptide turnovers but higher stability. These hypothetically functional differences between SLA-3 alleles compared to SLA-1 and SLA-2 alleles might also be related to their expression levels as it has previously been shown that SLA-3 alleles are generally expressed lower than SLA-2 and SLA-1 alleles (Crew et al. 2004; Kita et al. 2012). Maybe, the high turnover needs to be limited compared to the slower turnover rate provided by the SLA-1 and SLA-2 alleles to avoid excessive competition between the too many different T cell clones activated. In conclusion, we here present state-of-the-art tools for peptide prediction, selection, and binding analysis—both in terms of affinity and stability for porcine MHC class I molecules. These analyses were based on combining predictions based on individual SLA peptide binding matrix with those of trained neural networks in *NetMHCpan* and measured peptide-MHC binding affinity and pMHC complex stability using homogenous high throughput assays. The combination of the two individual prediction methods resulted in an enhancement of the overall peptide selection in the range of 10 to 45 % increase in the number of verified peptide binders (Supporting Table 1 A–C). In Supporting Table 1 A–C, there are several examples such as peptides TSDGFINGW and ITDYIVGYY. Both of these peptides bound to SLA-1*04:01 with high affinity and formed highly stable pSLA complexes, but TSDGFINGW was only predicted as a candidate ligand for SLA-1*04:01 by the matrix-based predictions (matrix score of 195, *NetMHCpan* rank of 15.00), while the ITDYIVGYY peptide was predicted exclusively by the *NetMHCpan* (matrix score of 0, NetMHCpan rank of 0.03).

Having determined the SLA-3*04:01 binding matrix, and mastering the approach of specifically mapping such matrices for additional SLA molecules, can be of great value when screening for peptide candidates during binding and pSLA studies. Not only does it provide an overview and insight of specific biochemical features within the binding groove, it also contributes to the overall predictions of possible peptide candidates for binding. The combined prediction strategy applied here has again proved extremely effective in the identification of high affinity and stable MHC-I ligands as it has been shown for human and bovine allotypes (Hansen et al. 2014; Rasmussen et al. 2014).

We expect that the biochemical data presented here on SLA-I peptide binding preferences should aid rational epitope discovery of porcine CTL epitopes. Furthermore, implementation of the binding data reported here in state-of-the-art prediction tools such as the *NetMHCpan*, which only include a very limited number of data for SLA-1*04:01 and SLA-2*04:01 (15 and 1, respectively) and none for the SLA-3*04:01 so far, could also benefit the process of identifying SLA-I restricted T cell epitopes.

Finally, we here present several virally derived peptides bound with both high affinity and stability by three of the most commonly expressed SLA molecules worldwide (Supporting Table 1 A–C). Such peptides, required their pathogenic origin is infectious to swine, can be categorized as potential CTL epitopes expressing the ability to generate stable pSLA complexes, in that they are likely to be targets for circulating CTLs, eventually leading to cell activation and elimination of pathogenic intruders.

## References

Allen TM, Sidney J, del Guercio MF, Glickman RL, Lensmeyer GL, Wiebe DA, DeMars R, Pauza CD, Johnson RP, Sette A, Watkins DI (1998) Characterization of the peptide binding motif of a rhesus MHC class I molecule (Mamu-A*01) that binds an immunodominant CTL epitope from simian immunodeficiency virus. J Immunol 160:6062–6071

Bevan MJ (1995) Antigen presentation to cytotoxic T lymphocytes in vivo. J Exp Med 182:639–641

Busch DH, Pamer EG (1998) MHC class I/peptide stability: implications for immunodominance, in vitro proliferation, and diversity of responding CTL. J Immunol 160:4441–4448

Crew MD, Phanavanh B, Garcia-Borges CN (2004) Sequence and mRNA expression of nonclassical SLA class I genes SLA-7 and SLA-8. Immunogenetics 56:111–114

de Groot NG, Heijmans CM, de Ru AH, Hassan C, Otting N, Doxiadis GG, Koning F, van Veelen PA, Bontrop RE (2013) Unique peptide-binding motif for Mamu-B*037:01: an MHC class I allele common to Indian and Chinese rhesus macaques. Immunogenetics 65:897–900

Doherty PC, Zinkernagel RM (1975) H-2 compatibility is required for T-cell-mediated lysis of target cells infected with lymphocytic choriomeningitis virus. J Exp Med 141:502–507

Dzuris JL, Sidney J, Appella E, Chesnut RW, Watkins DI, Sette A (2000) Conserved MHC class I peptide binding motif between humans and rhesus macaques. J Immunol 164:283–291

Ferre H, Ruffet E, Blicher T, Sylvester-Hvid C, Nielsen LL, Hobley TJ, Thomas OR, Buus S (2003) Purification of correctly oxidized MHC class I heavy-chain molecules under denaturing conditions: a novel strategy exploiting disulfide assisted protein folding. Protein Sci 12: 551–559

Hansen NJ, Pedersen LO, Stryhn A, Buus S (2001) High-throughput polymerase chain reaction cleanup in microtiter format. Anal Biochem 296:149–151

Hansen AM, Rasmussen M, Svitek N, Harndahl M, Golde WT, Barlow J, Nene V, Buus S, Nielsen M (2014) Characterization of binding specificities of bovine leucocyte class I molecules: impacts for rational epitope discovery. Immunogenetics 66:705–718

Harndahl M, Justesen S, Lamberth K, Roder G, Nielsen M, Buus S (2009) Peptide binding to HLA class I molecules: homogenous, high-throughput screening, and affinity assays. J Biomol Screen 14:173–180

Harndahl M, Rasmussen M, Roder G, Buus S (2011) Real-time, high-throughput measurements of peptide-MHC-I dissociation using a scintillation proximity assay. J Immunol Methods 374:5–12

Harndahl M, Rasmussen M, Roder G, Dalgaard Pedersen I, Sorensen M, Nielsen M, Buus S (2012) Peptide-MHC class I stability is a better predictor than peptide affinity of CTL immunogenicity. Eur J Immunol 42:1405–1416

Harty JT, Tvinnereim AR, White DW (2000) CD8+ T cell effector mechanisms in resistance to infection. Annu Rev Immunol 18:275–308

Ho CS, Franzo-Romain MH, Lee YJ, Lee JH, Smith DM (2009) Sequence-based characterization of swine leucocyte antigen alleles in commercially available porcine cell lines. Int J Immunogenet 36: 231–234

Hoof I, Peters B, Sidney J, Pedersen LE, Sette A, Lund O, Buus S, Nielsen M (2009) NetMHCpan, a method for MHC class I binding prediction beyond humans. Immunogenetics 61:1–13

Jorgensen KW, Rasmussen M, Buus S, Nielsen M (2014) NetMHCstab—predicting stability of peptide-MHC-I complexes; impacts for cytotoxic T lymphocyte epitope discovery. Immunology 141:18–26

Kita YF, Ando A, Tanaka K, Suzuki S, Ozaki Y, Uenishi H, Inoko H, Kulski JK, Shiina T (2012) Application of high-resolution, massively parallel pyrosequencing for estimation of haplotypes and gene expression levels of swine leukocyte antigen (SLA) class I genes. Immunogenetics 64:187–199

Lamberth K, Roder G, Harndahl M, Nielsen M, Lundegaard C, Schafer-Nielsen C, Lund O, Buus S (2008) The peptide-binding specificity of HLA-A*3001 demonstrates membership of the HLA-A3 supertype. Immunogenetics 60:633–643

Lauemoller SL, Holm A, Hilden J, Brunak S, Holst NM, Stryhn A, Ostergaard PL, Buus S (2001) Quantitative predictions of peptide binding to MHC class I molecules using specificity matrices and anchor-stratified calibrations. Tissue Antigens 57:405–414

Leisner C, Loeth N, Lamberth K, Justesen S, Sylvester-Hvid C, Schmidt EG, Claesson M, Buus S, Stryhn A (2008) One-pot, mix-and-read peptide-MHC tetramers. PLoS One 3, e1678

Lundegaard C, Lund O, Nielsen M (2011) Prediction of epitopes using neural network based methods. J Immunol Methods 374:26–34

McKinney DM, Erickson AL, Walker CM, Thimme R, Chisari FV, Sidney J, Sette A (2000) Identification of five different Patr class I molecules that bind HLA supertype peptides and definition of their peptide binding motifs. J Immunol 165:4414–4422

Nielsen M, Lundegaard C, Blicher T, Lamberth K, Harndahl M, Justesen S, Roder G, Peters B, Sette A, Lund O, Buus S (2007) NetMHCpan, a method for quantitative predictions of peptide binding to any HLA-A and -B locus protein of known sequence. PLoS One 2, e796

Nielsen M, Lundegaard C, Blicher T, Peters B, Sette A, Justesen S, Buus S, Lund O (2008) Quantitative predictions of peptide binding to any HLA-DR molecule of known sequence: NetMHCIIpan. PLoS Comput Biol 4, e1000107

Ostergaard PL, Nissen MH, Hansen NJ, Nielsen LL, Lauenmoller SL, Blicher T, Nansen A, Sylvester-Hvid C, Thromsen AR, Buus S (2001) Efficient assembly of recombinant major histocompatibility complex class I molecules with preformed disulfide bonds. Eur J Immunol 31:2986–2996

Pedersen LE, Harndahl M, Rasmussen M, Lamberth K, Golde WT, Lund O, Nielsen M, Buus S (2011) Porcine major histocompatibility complex (MHC) class I molecules and analysis of their peptide-binding specificities. Immunogenetics 63:821–834

Pedersen LE, Harndahl M, Nielsen M, Patch JR, Jungersen G, Buus S, Golde WT (2012) Identification of peptides from foot-and-mouth disease virus structural proteins bound by class I swine leukocyte antigen (SLA) alleles, SLA-1*0401 and SLA-2*0401. Anim Genet

Pedersen LE, Jungersen G, Sorensen MR, Ho CS, Vadekaer DF (2014) Swine leukocyte antigen (SLA) class I allele typing of Danish swine herds and identification of commonly occurring haplotypes using sequence specific low and high resolution primers. Vet Immunol Immunopathol 162:108–116

Rammensee H, Bachmann J, Emmerich NP, Bachor OA, Stevanovic S (1999) SYFPEITHI: database for MHC ligands and peptide motifs. Immunogenetics 50:213–219

Rapin N, Hoof I, Lund O, Nielsen M (2008) MHC motif viewer. Immunogenetics 60:759–765

Rasmussen M, Harndahl M, Stryhn A, Boucherma R, Nielsen LL, Lemonnier FA, Nielsen M, Buus S (2014) Uncovering the peptide-binding specificities of HLA-C: a general strategy to determine the specificity of any MHC class I molecule. J Immunol 193: 4790–4802

Sette A, Sidney J (1999) Nine major HLA class I supertypes account for the vast preponderance of HLA-A and -B polymorphism. Immunogenetics 50:201–212

Sidney J, Peters B, Moore C, Pencille TJ, Ngo S, Masterman KA, Asabe S, Pinilla C, Chisari FV, Sette A (2007) Characterization of the peptide-binding specificity of the chimpanzee class I alleles A 0301 and A 0401 using a combinatorial peptide library. Immunogenetics 59:745–751

Stryhn A, Pedersen LO, Romme T, Holm CB, Holm A, Buus S (1996) Peptide binding specificity of major histocompatibility complex class I resolved into an array of apparently independent subspecificities: quantitation by peptide libraries and improved prediction of binding. Eur J Immunol 26:1911–1918

Thomsen MC, Nielsen M (2012) Seq2Logo: a method for construction and visualization of amino acid binding motifs and sequence profiles including sequence weighting, pseudo counts and two-sided representation of amino acid enrichment and depletion. Nucleic Acids Res 40:W281–W287

van der Burg SH, Visseren MJ, Brandt RM, Kast WM, Melief CJ (1996) Immunogenicity of peptides bound to MHC class I molecules depends on the MHC-peptide complex stability. J Immunol 156:3308–3314

Zhang H, Lundegaard C, Nielsen M (2009) Pan-specific MHC class I predictors: a benchmark of HLA class I pan-specific prediction methods. Bioinformatics 25:83–89