

# A modern approach for epitope prediction: identification of foot-and-mouth disease virus peptides binding bovine leukocyte antigen (BoLA) class I molecules

Mital Pandya<sup>1</sup> · Michael Rasmussen<sup>2</sup> · Andreas Hansen<sup>2</sup> · Morten Nielsen<sup>3,4</sup> · Soren Buus<sup>2</sup> · William Golde<sup>5</sup> · John Barlow<sup>1</sup>

Received: 15 July 2015 / Accepted: 7 October 2015 / Published online: 24 October 2015  
© Springer-Verlag Berlin Heidelberg 2015

**Abstract** Major histocompatibility complex (MHC) class I molecules regulate adaptive immune responses through the presentation of antigenic peptides to CD8<sup>+</sup> T cells. Polymorphisms in the peptide binding region of class I molecules determine peptide binding affinity and stability during antigen presentation, and different antigen peptide motifs are associated with specific genetic sequences of class I molecules. Understanding bovine leukocyte antigen (BoLA), peptide-MHC class I binding specificities may facilitate development of vaccines or reagents for quantifying the adaptive immune response to intracellular pathogens, such as foot-and-mouth disease virus (FMDV). Six synthetic BoLA class I (BoLA-I) molecules were produced, and the peptide binding motif was generated for five of the six molecules using a combined approach of positional scanning combinatorial peptide libraries (PSCPLs) and neural network-based predictions (*NetMHCpan*). The updated *NetMHCpan* server was used to

predict BoLA-I binding peptides within the P1 structural polyprotein sequence of FMDV (strain A24 Cruzeiro) for BoLA-1\*01901, BoLA-2\*00801, BoLA-2\*01201, and BoLA-4\*02401. Peptide binding affinity and stability were determined for these BoLA-I molecules using the luminescent oxygen channeling immunoassay (LOCI) and scintillation proximity assay (SPA). The functional diversity of known BoLA alleles was predicted using the *MHCcluster* tool, and functional predictions for peptide motifs were compared to observed data from this and prior studies. The results of these analyses showed that BoLA alleles cluster into three distinct groups with the potential to define “BoLA supertypes.” This streamlined approach identifies potential T cell epitopes from pathogens, such as FMDV, and provides insight into T cell immunity following infection or vaccination.

**Keywords** MHC class I · Motif · Peptide · FMDV · Epitope · Immunoinformatics

**Electronic supplementary material** The online version of this article (doi:10.1007/s00251-015-0877-7) contains supplementary material, which is available to authorized users.

✉ John Barlow  
john.barlow@uvm.edu

- <sup>1</sup> Department of Animal Science, University of Vermont, 570 Main Street, 221 Terrill Building, Burlington, VT 05405, USA
- <sup>2</sup> Department of International Health, Immunology and Microbiology, University of Copenhagen, Copenhagen DK-2200, Denmark
- <sup>3</sup> Department of Systems Biology, Technical University of Denmark, Lyngby, Denmark
- <sup>4</sup> Instituto de Investigaciones Biotecnológicas, Universidad Nacional de San Martín, Buenos Aires, Argentina
- <sup>5</sup> Plum Island Animal Disease Center, Agricultural Research Service, USDA, Greenport, NY 11944, USA

## Introduction

Peptide binding to major histocompatibility complex (MHC) class I molecules is the most selective step in determining immunogenicity and immunodominance of antigens from intracellular pathogens such as viruses (Yewdell and Bennink 1999). The alpha-1 and alpha-2 domains of MHC class I molecules form a binding cleft that presents short polypeptides derived from cytosolic and nuclear proteins. Given the thousands of potential 8- to 11-amino acid polypeptides that can be generated from the proteins of intracellular pathogens, antigenic peptides that stably bind to major histocompatibility complex (MHC) class I molecules are relatively rare. The peptides that bind to class I molecules often share common features (i.e., motifs); this is a function of key amino acid

residues in the binding pockets of the MHC molecules. The antigenic peptides (epitopes) presented in the context of MHC class I molecules on the surface of virus-infected cells or activated antigen presenting dendritic cells are recognized by T cell receptors (TCRs) on CD8<sup>+</sup> T lymphocytes. Upon recognition of the peptide-MHC class I complexes, naïve CD8<sup>+</sup> T cells expand and differentiate into cytotoxic T lymphocytes (CTLs). These cells recognize and kill virus-infected cells through several mechanisms such as perforin, granzyme, and Fas ligand pathways (Groscurth and Filgueira 1998).

MHC class I molecules can be characterized based on their ability to bind specific peptide motifs (Stryhn et al. 1996). MHC class I molecules preferentially bind peptides of 8 to 11 amino acids, most commonly with dominant anchor residues located at positions 2 and 3 relative to the amino terminus and at the C-terminus. Residues at other positions (e.g., positions 1, 4, 5, 6, 7, 8) may also influence peptide binding and affect TCR binding to the MHC-peptide complex (Calis et al. 2013; Frankild et al. 2008). For example, a large fraction of human leukocyte antigen (HLA) alleles are associated with specificities for proline in position 2 (P2) and hydrophobic residues at the C-terminus (B7-supertype), or with hydrophobic residues in P2 and positive charges at the C-terminus (A3-supertype) (Sidney et al. 2008). Similar anchor residues are reported in other species such as cattle, pigs, birds, and chimpanzees (Follin et al. 2013; Hansen et al. 2014; Pedersen et al. 2012; Thomsen et al. 2013). Among cattle, there are currently 95 classical MHC class I alleles listed in the Immuno Polymorphism Database (IPD-MHC; <http://www.ebi.ac.uk/ipd/>) although additional alleles have been reported (Hansen et al. 2014). Peptide binding motifs or epitopes have been described for at least 13 bovine leukocyte antigen class I (BoLA-I) molecules (Bamford et al. 1995; De Groot et al. 2003; Gaddum et al. 1996a; Gaddum et al. 1996b; Graham et al. 2008; Guzman et al. 2010; Hansen et al. 2014; Hart et al. 2011; Hegde et al. 1999; Hegde et al. 1995; Hegde and Srikumaran 2000; Li et al. 2011; MacHugh et al. 2011; Momtaz et al. 2014; Nene et al. 2012; Sinnathamby et al. 2004; Svitek et al. 2014). Two of these studies focused on predicting antigenic peptides of foot-and-mouth disease virus (FMDV) (Guzman et al. 2010; Momtaz et al. 2014), while additional studies examined the induction of CTL response following FMDV infection or vaccination in cattle (Childerstone et al. 1999; Guzman et al. 2008) and swine (Patch et al. 2013; Patch et al. 2011). The potential role of CD8<sup>+</sup> T cell responses in FMDV infections or following vaccination remains controversial, as most FMDV vaccination research has focused on the role of virus-neutralizing antibodies (Grubman et al. 2010). Evidence of CD8<sup>+</sup> T cell responses from the few prior studies suggests additional research is justified, as understanding MHC class I FMDV epitopes and bovine CTL responses following infection or vaccination may identify opportunities to advance FMDV vaccine development.

FMDV is a picornavirus of the *Aphthovirus* genus with a linear (non-segmented) genome of single-stranded positive-sense RNA approximately 8.2 kilobases in length (Grubman et al. 2010). The genome is translated as a single polyprotein, which is cleaved by the activity of viral proteases. The P1-2A precursor portion of full-length polyprotein includes four structural capsid proteins: VP4, VP2, VP3, and VP1. FMDV causes an acute highly contagious febrile disease with vesicular lesions in cloven-hoofed animals, including domestic cattle and swine. Infection negatively impacts milk and meat production, and outbreaks in naïve populations are characterized by rapid spread and significant morbidity. Vaccination is an important component of FMDV control in endemic countries and will be increasingly essential in response to any future outbreaks in countries or regions currently free of FMDV (Grubman et al. 2010). An adenoviral vectored FMDV vaccine has recently received conditional approval for use in the USA. This vaccine construct contains nucleotide sequence encoding the FMDV P1-2A structural proteins and the 3C protease genes inserted into the E1 region of a replication-incompetent adenovirus 5 (Ad5) vector (Brake et al. 2012; Grubman et al. 2010). Multiple species have benefited from the use of adenovirus as a vaccine vector, including humans (human immunodeficiency virus), non-human primates (simian immunodeficiency virus), dogs (rabies), swine (porcine respiratory and reproductive syndrome virus), and cattle (bovine parainfluenza virus type 3, bovine herpes virus 1, and bovine viral diarrhea virus) (Tatsis and Ertl 2004). Human Ad5 vectors have been shown to specifically induce CD8<sup>+</sup> T cell responses in rodents, dogs, and non-human primates (Tatsis and Ertl 2004).

The bioinformatic tools, such as *NetMHCpan* and *NetMHC*, can be used to predict peptides that are likely bound by the BoLA molecules and to inform subsequent in vitro and ex vivo studies (Hansen et al. 2014; Nene et al. 2012; Svitek et al. 2014). These tools have been used to refine the peptide predictions obtained from previous studies (Nene et al. 2012) and have been applied to successfully identify immunogenic peptides of *Theileria parva* antigens driving CD8<sup>+</sup> T cell responses in cattle (Svitek et al. 2014). The reverse immunology approach has been used in bacterial and viral pathogens, such as meningococcus B, human immunodeficiency virus, yellow fever virus, West Nile virus, dengue virus, herpes virus, and hepatitis C virus (Bruno et al. 2015; Rappuoli 2001).

In this study, we applied bioinformatic prediction tools, *NetMHCpan-2.9* and *MHCcluster*, combined with high-throughput biochemical affinity and stability assays to identify potential FMDV epitopes bound by MHC molecules expressed by Holstein cattle in the research and teaching herd at the University of Vermont (UVM). We were able to identify several in vitro binding peptides and characterized the peptide motifs for five out of six BoLA molecules. This research provides insight into specific viral derived peptides that could

induce bovine CD8<sup>+</sup> T cells after infection or vaccination. This approach is a rapid and accurate method to identify T cell epitopes and sets the stage to examine T cell-specific immune responses in Holstein cattle, other domestic cattle breeds, and perhaps other bovinds.

## Methods

### Animals

Whole blood samples (8–10 ml in K3-EDTA) were taken by jugular or caudal tail venipuncture from 81 purebred Holstein cattle from the UVM Dairy Center of Excellence (DCE) research herd. Blood samples were held on ice, and within 2 h of sampling, peripheral blood mononuclear cells (PBMCs) were isolated by density gradient centrifugation using Histopaque 1083 (Sigma-Aldrich, St. Louis, MO). All animal use procedures were conducted in accordance with protocol #10-026 approved by the UVM Institutional Animal Care and Use Committee.

### RNA extraction and cDNA synthesis

RNA was purified from PBMCs using manufacturer's instructions for TRIzol Reagent in combination with RNeasy Mini Kit (Qiagen, Valencia, CA). Following RNA extraction, complementary DNA was generated using the ImProm-II Reverse Transcription System (Promega, Madison, WI) following manufacturer's recommendations.

### BoLA typing by polymerase chain reaction with sequence-specific primers (PCR-SSP)

PCR-SSP for a panel of 15 BoLA alleles was performed according to Ellis et al. (1998) with slight modifications. Primers to amplify a 16th allele (BoLA-2\*01201) were obtained from Codner (2010). PCR reaction mixtures consisted of 3.5 mM MgCl<sub>2</sub>, 0.2 mM dNTPs, 1 μM of each primer, and 1.25 units of Taq polymerase in 1X buffer (10 mM Tris-HCl, 50 mM KCl, 1.5 mM MgCl<sub>2</sub> (New England Biolabs, Ipswich, MA)) in a final reaction volume of 50 μl. Previously published annealing temperatures were adjusted for primer pairs 1, 5, 6, 7, 8, 10, 11, 13, 15, and 16, using the formula  $T_m(^{\circ}\text{C}) = [67.5 + 0.34(\%GC \times 100) - 395/\text{length of the primer}]$ . The cycling conditions were 95 °C for 60 s followed by 30 PCR cycles of 95 °C for 30 s, 55–61 °C for 30 s, and 72 °C for 30 s using a programmable C1000 Thermal Cycler (Bio-Rad, Hercules, California). Positive control complementary DNA (cDNA) for BoLA-3\*00201, BoLA-1\*02301, and BoLA-6\*01301 was kindly provided by Dr. Ivan Morrison (Roslin Institute, University of Edinburgh, Edinburgh, Scotland).

### Positional scanning combinatorial peptide libraries and FMDV peptides

Positional scanning combinatorial peptide library (PSCPL) peptides were synthesized using solid-phase 9-fluorenylmethyloxycarbonyl (Fmoc) chemistry on 2-chlorotrityl chloride resins and purchased from Schafer-N Copenhagen, Denmark ([www.schafer-n.com](http://www.schafer-n.com)). Briefly, an equimolar mixture of 19 of the common Fmoc amino acids (excluding cysteine) was prepared for each synthesis and used for coupling in eight positions, whereas a single Fmoc amino acid (including cysteine) was used in one position. This position was changed in each synthesis starting with the N-terminus and ending with the C-terminus. In one synthesis, the amino acid (AA) pool was used in all nine positions. Therefore, the library consisted of  $20 \times 9 + 1 = 181$  individual peptide libraries and includes:

- 20 PSCPL sublibraries describing position 1: AX<sub>8</sub>, CX<sub>8</sub>, DX<sub>8</sub>, ... YX<sub>8</sub>
- 20 PSCPL sublibraries describing position 2: XA X<sub>7</sub>, XC X<sub>7</sub>, XD X<sub>7</sub>, ... XY X<sub>7</sub>
- Etc.
- 20 PSCPL sublibraries describing position 9: X<sub>8</sub>A, X<sub>8</sub>C, X<sub>8</sub>D, ... X<sub>8</sub>Y
- A completely random nonameric peptide library pool: X<sub>9</sub>,

where X denotes the random incorporation of AA from the mixture and a single-letter AA abbreviation is used to denote the identity of a fixed AA (e.g., A, alanine). Following synthesis, the peptides were cleaved from the resin in trifluoroacetic acid/1,2-ethanedithiol/triisopropylsilane/water 95:2:1:3 v/v/v/v, precipitated in cold diethyl ether, and extracted with water before desalting on C18 columns, freeze drying, and weighing.

Peptides for the predicted FMDV CD8<sup>+</sup> T cell epitopes were synthesized by 9-fluorenylmethyloxycarbonyl (Fmoc) chemistry, purified by reverse-phase high-performance liquid chromatography (to at least >90 % purity), validated by mass spectrometry, and quantified by weight (GenScript, Piscataway, NJ).

### Synthesis of recombinant MHC-I heavy chain proteins

Recombinant molecules BoLA-1\*01901, BoLA-2\*00801, BoLA-2\*01601, BoLA-3\*01701, BoLA-4\*02401, and BoLA-6\*01302 were produced as previously described (Hansen et al. 2014). In brief, each BoLA molecule was generated in *Escherichia coli* including a biotinylation substrate peptide (BSP) at the N-terminal end, which was biotinylated in vivo using a co-induced BirA enzyme and the addition of biotin during the expression. *E. coli* were harvested as inclusion bodies, extracted into Tris-buffered 8 M urea and purified

using ion exchange, hydrophobic, and gel filtration chromatography. MHC-I heavy chain proteins were never exposed to reducing conditions, which allows for purification of highly active pre-oxidized BoLA molecules when diluted into an appropriate reaction buffer. The pre-oxidized, denatured proteins were stored at  $-20\text{ }^{\circ}\text{C}$  in Tris-buffered 8 M urea.

### Production of MHC-I light chain (beta-2-microglobulin) protein

Native, recombinant human and bovine beta-2-microglobulins ( $\beta_2\text{m}$ ) were expressed and purified as previously described (Hansen et al. 2014). Briefly, a histidine affinity tag (HAT) followed by an FXa restriction enzyme site was inserted N-terminally of a synthetic gene encoding the native, mature human or bovine  $\beta_2\text{m}$ . The recombinant gene was expressed in the *E. coli* expression host, BL21 (DE3), harvested as inclusion bodies, extracted into a urea buffer, and purified. The tagged  $\beta_2\text{m}$  protein was digested for 48 h at room temperature with the FXa protease releasing intact natively folded  $\beta_2\text{m}$ . The folded  $\beta_2\text{m}$  was purified as previously described, and fractions containing  $\beta_2\text{m}$  were identified by  $A_{280}$  UV absorbance and SDS-PAGE and then pooled. Protein concentrations were determined by BCA assay. The native  $\beta_2\text{m}$  proteins were stored at  $-20\text{ }^{\circ}\text{C}$ .

### Determining peptide-MHC-I binding motifs by a scintillation proximity assay

Nonameric peptide binding motifs were determined for BoLA-1\*01901, BoLA-2\*00801, BoLA-3\*01701, BoLA-4\*02401, and BoLA-6\*01302 using the PSCPL as previously described (Hansen et al. 2014; Harndahl et al. 2011; Rasmussen et al. 2014). Recombinant, biotinylated MHC-I heavy chain molecules in 8 M urea were diluted at least 100-fold into PBS buffer containing bovine  $\beta_2\text{m}$  and peptide to initiate pMHC-I complex formation. The final concentration of MHC-I heavy chain was between 10 and 100 nM, depending on the specific activity of the MHC-I heavy chain. The reactions were carried out in the wells of streptavidin-coated scintillation 384-well FlashPlate® PLUS microplates (Perkin Elmer, Waltham, MA). Recombinant bovine  $\beta_2\text{m}$  was expressed, purified, and radiolabeled with iodine ( $^{125}\text{I}$ ) as previously described (Harndahl et al. 2011).  $^{125}\text{I}$ -labeled  $\beta_2\text{m}$  (approximately 1 nM, corresponding to approximately 25,000 cpm/well) and saturating concentrations (10  $\mu\text{M}$ ) of peptide were allowed to reach steady state by overnight incubation at  $18\text{ }^{\circ}\text{C}$ . After overnight incubation, excess unlabeled bovine  $\beta_2\text{m}$  was added to a final concentration of 1  $\mu\text{M}$  and the temperature was raised to  $37\text{ }^{\circ}\text{C}$  to initiate dissociation. pMHC-I dissociation was monitored for 24 h by consecutive measurement of the scintillation microplate on a

scintillation TopCount NXT multiplate counter (Perkin Elmer, Waltham, MA). PSCPL dissociation data were analyzed as described (Rasmussen et al. 2014). Briefly, following background correction, the area under the dissociation curve (AUC) was calculated for each sublibrary by summing the counts from 0 to 24 h. The relative contribution of each residue in each position (i.e., the relative binding (RB)) was calculated as  $\text{RB} = (\text{AUC}_{\text{sublibrary}} / \text{AUC}_{\text{X9}})$ . The RB values were normalized for each peptide position to give a sum of 20 for each residue. The anchor position (AP) value for each peptide position was calculated as:

$$\text{AP} = \sum_{i=1}^{20} (1 - \text{RB}_{\text{sublibrary}})^2$$

Data generated by the PSCPL assay for the five BoLA molecules were used to select a library of peptides to evaluate in vitro binding affinity of these molecules.

### Measuring peptide-MHC-I affinity interactions by luminescent oxygen channeling immunoassay AlphaScreen™ assay

A set of up to 192 peptides was selected from our peptide repository, which contains over 9000 nonamer peptides. Based on availability in the library, up to 96 high-scoring peptides (rank score  $\leq 0.5$ ) according to the PSCPL matrices were selected, and as many as possible up to 96 additional peptides were selected based on *NetMHCpan-2.8* predictions, but with low affinity according to PSCPL. This combined method has been demonstrated to successfully identify high-affinity MHC binding peptides, and inclusion of such data has been shown to improve the predictive performance of the *NetMHCpan* algorithm (Hansen et al. 2014; Hoof et al. 2009; Nielsen et al. 2008; Pedersen et al. 2012; Rasmussen et al. 2014).

Affinity measurements of peptide-MHC-I interactions with MHC-I molecules were conducted using the AlphaScreen™ technology, as previously described (Harndahl et al. 2009). The donor and acceptor beads of the luminescent oxygen channeling immunoassay (LOCI) were purchased from PerkinElmer (Waltham, MA). Donor beads were obtained pre-conjugated to streptavidin; acceptor beads were in-house conjugated to the monoclonal antibody W6/32 (gift from Dr. Jonathan Boyson), to a final concentration of 1 mg/ml using standard procedures as described by the manufacturer. Recombinant, biotinylated MHC-I heavy chains, human  $\beta_2\text{m}$  and peptide were titrated in phosphate-buffered saline (PBS) with 0.1 % Lutrol F-68 as surfactant. They were allowed to fold for 48 h at  $18\text{ }^{\circ}\text{C}$ . The folding mixture was transferred to a 384-well OptiPlate (PerkinElmer, Waltham, MA). The



donor and acceptor beads were each added to a final concentration of 5 µg/ml and incubated overnight at 18 °C. To ensure temperature equilibrium during the reading time, plates were placed at room temperature 1 h prior to reading on BioTek Synergy H4 (BioTek, Winooski, VT) or on an EnVision® multilabel reader (PerkinElmer, Waltham, MA). All handling of LOCI reagents was done in the dark or in green light.

LOCI signals were converted to peptide-MHC-I complex concentration via a pre-folded peptide-MHC-I complex standard. The peptide-MHC-I binding affinity was calculated by fitting peptide dose-response data to a one-site binding model ( $Y = B_{\max} * X / (K_d + X)$ ) by non-linear regression where  $Y$  is the concentration of pMHC-I complexes formed and  $X$  is the concentration of ligand (i.e., peptide) (Harndahl et al. 2009). The  $EC_{50}$  approximates the equilibrium dissociation constant ( $K_d$ ) as long as the receptor concentration used is less than the  $K_d$ , thus avoiding ligand depletion. Under conditions of limited receptor concentration ( $[MHC-I \text{ HC}] \leq K_d$ ), the  $EC_{50}$  is a reasonable approximation of the  $K_d$ .  $K_d$  calculations were done in GraphPad Prism® 6 Software (GraphPad Software, Inc., San Diego, CA).

### Bioinformatic analysis of the FMDV peptide binding specificity of BoLA molecules

The updated pan-specific algorithm *NetMHCpan-2.9* (Hoof et al. 2009) was used to predict peptide-MHC-I interactions of BoLA-1\*01901, BoLA-2\*00801, BoLA-2\*01201, and BoLA-4\*02401 with nonameric and decameric peptides of the P1 region of FMDV serotype A24 Cruzeiro. The FMDV P1 sequence used in this analysis was obtained from Dr. Marvin Grubman (Personal Communication, Supplemental Fig. 1) and is the sequence used in construction of the new adenovirus-vectored FMDV vaccine (Brake et al. 2012).

### Measuring specific FMDV-peptide-MHC-I interactions by LOCI and SPA

The top strong binding FMDV peptides, i.e., a *NetMHCpan* rank score of  $\leq 0.5$  %, were synthesized as described above for use in subsequent binding affinity (LOCI) and stability (scintillation proximity assay (SPA)) assays with synthetic BoLA-1\*01901, BoLA-2\*00801, BoLA-2\*01201, and BoLA-4\*02401. We also tested two weak binding peptides, i.e., those with a *NetMHCpan* rank score of 0.51–2 %; thus, a total of 30 synthetic peptides were selected. Peptide-MHC stability in the dissociation SPA was expressed as a half-life in hours, and based on a previous study, we used a threshold of 1 h to identify potential immunogenic epitopes (Harndahl et al., 2012).

### Functional predictions of BoLA molecules using *MHCcluster*, visualized with sequence logos

*MHCcluster* server 2.1 was used to generate an unrooted tree and heat-map of the functional relationship between the BoLA molecules using full-length BoLA class I protein sequences, as previously described in Thomsen et al. (2013). Ovar-1-N\*00201 MHC-I sequence (GI:108792434) was also included. As part of the *MHCcluster* output, a sequence logo for each allele was generated using the *Seq2Logo* server (Thomsen and Nielsen 2012). The logos are visual representations of predicted nonameric amino acid binding motifs of individual BoLA molecules. Briefly, logos were created from the strongest binders (predicted top 1 %) from a set of 100,000 random 9-mer peptides and were clustered using the Hobohm 1 algorithm with the similarity threshold of 63 % to remove redundancy, and pseudocounts were applied with a weight on prior of 200 to account for a low number of observations. The accuracy of the predicted sequence motif is estimated from the distance to the “nearest neighbor” MHC molecule included in the training of the peptide binding prediction method; a value  $\geq 0.70$  is considered an accurate prediction and  $\geq 0.90$  is considered highly accurate (Hoof et al. 2009; Thomsen et al. 2013).

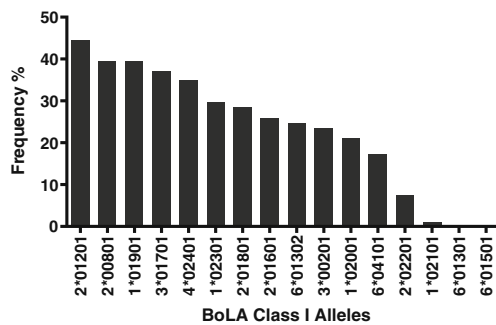
## Results

### BoLA class I typing of individual animals

Frequency of known BoLA gene expression was determined by PCR-SSP (Ellis et al. 1998) for 81 Holstein cattle from the UVM teaching and research herd (Fig. 1). We used the observed frequencies in our cattle population and availability of other recombinant BoLA molecules (e.g., BoLA-2\*01201 (Hansen et al. 2014)) to inform synthesis of recombinant BoLA molecules for this study. We chose to synthesize BoLA-1\*01901, BoLA-2\*00801, BoLA-2\*01601, BoLA-3\*01701, BoLA-4\*02401, and BoLA-6\*01302.

### Determining the peptide binding specificity of BoLA-1\*01901, BoLA-2\*00801, BoLA-3\*01701, BoLA-4\*02401, and BoLA-6\*01302 by dissociation-driven PSCPL analysis

We applied a positional scanning combinatorial peptide libraries (PSCPLs) approach combining SPA-based peptide-MHC-I dissociation assay with nonameric PSCPLs (Rasmussen et al. 2014) to analyze the peptide binding specificity of six BoLA-I molecules: BoLA-1\*01901, BoLA-2\*00801, BoLA-2\*01601, BoLA-3\*01701, BoLA-4\*02401, and BoLA-6\*01302. Peptide binding motifs (Fig. 2) and PSCPL matrices (Supplemental Fig. 2) were derived for five BoLA MHC class



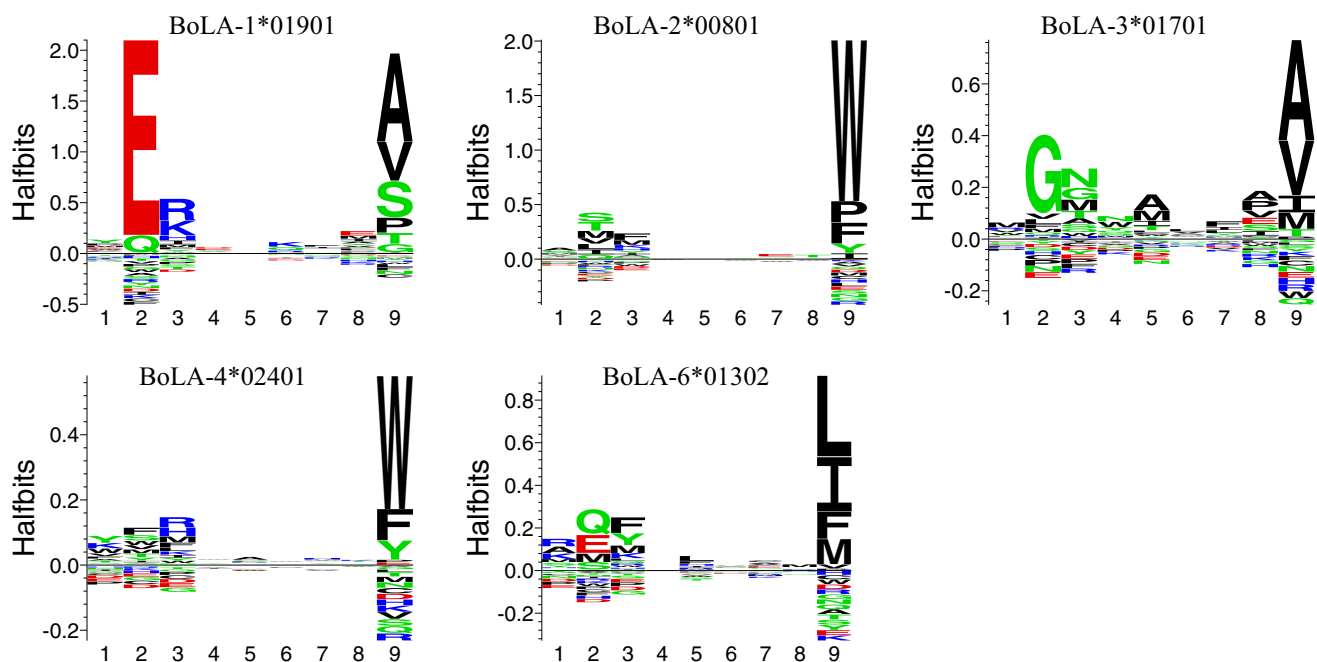
**Fig. 1** Frequency of BoLA class I alleles in 81 Holstein cattle from the UVM Dairy Center of Excellence research herd assayed by PCR-SSP

I heavy chains, BoLA-1\*01901, BoLA-2\*00801, BoLA-3\*01701, BoLA-4\*02401, and BoLA-6\*01302. Recombinant BoLA-2\*01601 was also synthesized, but we were unable to derive a peptide binding motif for this molecule due to low signal-to-background ratio in the SPA-based pMHC-I dissociation assay (data not shown). Overall, the shape of the motifs of these BoLA molecules appeared to be very similar to previously reported BoLA motifs with anchor positions at the N- and C-terminal parts of the peptide (Hansen et al. 2014). All molecules had prominent anchors in C-terminal position 9 (P9), and a few motifs had a less defined primary anchor position at position 2 (P2) or position 3 (P3). The

BoLA-1\*01901 molecule had a pronounced anchor position in P2, P3, and P9 with strong preference for glutamic acid in P2; preference for arginine or lysine in P3; and alanine, valine, or serine in P9. BoLA-2\*00801 and BoLA-4\*02401 have very similar peptide binding motifs; both contain less pronounced anchors in P2, but had strongly pronounced preference for aromatic amino acids in the C-terminal (P9) peptide position (Fig. 2). The motif for BoLA-3\*01701 was rather unique with potential anchor residues in P2, P3, P5, P8, and P9. BoLA-3\*01701 had a strong pronounced preference for glycine in P2, and preferences for amino acids with more neutral side chains in P3. BoLA-3\*01701 also appeared to favor asparagine or methionine in P5 and asparagine, proline, or valine in P8. Similar to BoLA-1\*01901, BoLA-3\*01701 strongly favored alanine and valine in P9. The motif for BoLA-6\*01302 had a distinct preference for hydrophobic amino acids, such as leucine, phenylalanine, isoleucine, and methionine in P9 and a less pronounced P2 anchor with preference for glutamic acid and glutamine.

#### BoLA binding affinity for predicted nonameric peptides and retraining NetMHCpan for specific BoLA molecules

Peptide-MHC-I affinity interactions were measured by LOCI, and these data were used to update the existing bioinformatics



**Fig. 2** Sequence logo representations of BoLA-1\*01901, BoLA-2\*00801, BoLA-4\*02401, BoLA-6\*01302, and BoLA-3\*01701 binding motifs. The logo for BoLA-2\*01201 has been previously described (Hansen et al. 2014). The logos were calculated from the top 1 % highest-scoring peptides selected from a pool of 9-mer peptides using the positional scanning combinatorial peptide library matrix. In the

sequence logo, each peptide position is represented by a stack of letters indicating its significance for binding and the height of each amino acid is proportional to its relative frequency (Thomsen and Nielsen 2012). Acidic residues are displayed in red, basic in blue, neutral in green, and hydrophobic in black

predictor *NetMHCpan v2.8*. In this approach, the PSCPL matrices are used to select predicted BoLA-I binding peptides from an in-house repository of >9000 nonameric peptides. To complement this selection method, *NetMHCpan v2.8* was used to select additional strong binding peptides from the repository that were not predicted to bind according to the PSCPL-derived binding matrix. We have previously demonstrated that using either the PSCPL or the NetMHCpan method alone is suboptimal to improve the predictive algorithm (Hansen et al. 2014; Rasmussen et al. 2014). These previous studies concluded that PSCPL matrix-guided selection will be able to identify a large proportion of binders, although in many cases this will not cover the complete binding space of a particular MHC molecule. The number of predicted binders available from the nonameric peptide repository included 170, 170, 145, 170, and 157 peptides for BoLA-1\*01901, BoLA-2\*00801, BoLA-3\*01701, BoLA-4\*02401, and BoLA-6\*01302 molecules, respectively (Supplemental Table 1). For each BoLA molecule-peptide combination, the binding affinity was subsequently determined using the LOCI assay (Hansen et al. 2014; Harndahl et al. 2009). Binding affinity data was already available for BoLA-2\*01201 (Hansen et al. 2014), the molecule found at the highest frequency in our herd. For BoLA-2\*00801 and 6\*01302, >90 % of the peptides from the pool of PSCPL matrix predicted peptides were found to bind with an affinity <500 nM. Greater than 60 % of the predicted peptides identified by the complementary *NetMHCpan v2.8* prediction bound at the same affinity (Supplemental Table 1). Similarly, the PSCPL matrix-based prediction for BoLA-1\*01901 and BoLA-4\*02401 found 31 and 40 % of the predicted peptides for each molecule to bind with an affinity <500 nM. The *NetMHCpan v2.8* prediction identified additional strong binding peptides (4 and 10 % of peptides predicted by *NetMHCpan v2.8* were found to bind to BoLA-1\*01901 and BoLA-4\*02401, respectively, Supplemental Table 1). The complementary method of PSCPL matrix plus *NetMHCpan* predictions identified a large number of MHC-I binding peptides that can in turn be used to update the prediction algorithm (Hansen et al. 2014; Rasmussen et al. 2014). One exception was BoLA-3\*01701 where only 4 % of the peptides predicted to bind according to the PSCPL matrix-based-prediction bound with an affinity <500 nM, and the *NetMHCpan v2.8* prediction identified no additional binders (Supplemental Table 1). The small number of binders found for BoLA-3\*01701 is potentially due to the bias in the peptide repository. The prevalence of peptides matching BoLA-3\*01701 motif is much lower compared to the other molecules, and in particular, there are no peptides in our repository that match both preferred anchor amino acids, glycine in P2, and alanine or valine in P9. Similar results have been seen for HLA-C\*0401 (Rasmussen et al. 2014).

The data generated from the LOCI binding assay were used to retrain and update the *NetMHCpan* algorithm for the

BoLA-1\*01901, BoLA-2\*00801, BoLA-2\*01201, BoLA-4\*02401, and BoLA-6\*01302. The estimated prediction accuracies in *NetMHCpan v2.8* for these five molecules were 0.489, 0.437, 0.853, 0.398, and 0.809, respectively, and following retraining, the accuracies improved to 0.853 for each molecule in *NetMHCpan v2.9*. The accuracy predictions were extracted from *MHCcluster* output and were calculated from the distance to the nearest neighbor in the training data of *NetMHCpan v2.9*.

#### Identification of peptides derived from FMDV and bound by BoLA-1\*01901, BoLA-2\*00801, BoLA-2\*01201, and BoLA-4\*02401

The updated version of *NetMHCpan v2.9* was used to make FMDV peptide predictions and is available at (<http://www.cbs.dtu.dk/services/NetMHCpan/>). The 736-amino acid fragment of FMDV P1 capsid proteins used in this study generates a total of 1455 peptides (728 nonameric and 727 decameric peptides). Screening of these potential epitopes using a *NetMHCpan v2.9* rank score threshold of  $\leq 0.5$  % identified 28 FMDV candidate peptides binding to four BoLA molecules found frequently in our herd: 1\*01901, 2\*00801, 2\*01201, and 4\*02401 (Table 1 and Fig. 2). We also included two peptides with rank score of 0.51–2 %. Although functional synthetic molecules are available for BoLA-6\*01302, our goal was to focus on molecules most frequently observed in our herd and FMDV peptide binding assays was not performed for BoLA-6\*01302 at this time.

Among the 30 peptides, 25 (13 nonamers and 12 decamers) were predicted to bind with strong or weak affinity to a single BoLA molecule, while five peptides, <sup>240</sup>MTAHITVPY<sub>248</sub>, <sup>444</sup>A A H C I H A E W<sub>452</sub>, <sup>207</sup>L L V A M V P E W<sub>215</sub>, <sup>443</sup>RAAHCIHAEW<sub>452</sub>, and <sup>604</sup>VVRHEGNLTW<sub>613</sub>, were common to more than one BoLA molecule. Five nonamer peptides were nested within decamer peptides that were predicted to bind the same molecule. Only two 10-mer peptides were predicted to be strong binders (rank score  $\leq 0.5$ ) to BoLA-1\*01901 allele, while 13 peptides were predicted to bind to BoLA-2\*00801 (six 9-mers and seven 10-mers) and nine peptides were predicted as strong binders to BoLA-2\*01201 (six 9-mers and three 10-mers). Two peptides, <sup>386</sup>AAKHMSNTY<sub>394</sub> and <sup>240</sup>MTAHITVPY<sub>248</sub>, were predicted to be weak binders (rank score >0.5) to BoLA-2\*00801 and BoLA-2\*01201, respectively.

The 30 peptides were tested for in vitro binding affinity and stability using the LOCI and SPA assays (Table 1). In this study, we included two peptides (<sup>386</sup>AAKHMSNTY<sub>394</sub> and <sup>240</sup>MTAHITVPY<sub>248</sub>), which were predicted by NetMHCpan to be weak binders; these peptides were found to be either weak- or non-binding in the LOCI assay and had short binding half-lives ( $\leq 0.3$  h.) in the SPA. In comparison, among the 28 peptides predicted to be strong binders by *NetMHCpan v2.9*,

**Table 1** FMDV (strain A24) P1 region peptides predicted as strong ( $n=28$ ) or weak ( $n=2$ ) binders using NetMHCpan v2.9

BoLA molecule	Peptide sequence <sup>a</sup>	NetMHCpan predicted % rank score <sup>b</sup>	$K_d$ (nM) <sup>c</sup>	Binding status category <sup>c</sup>	pMHC-I complex stability <sup>d</sup>	
1*01901	<sup>200</sup> NQFNGGCLLV <sub>209</sub>	0.5	1086	Weak	0.3	
	<sup>499</sup> AENDTLVVSV <sub>508</sub>	0.12	7817	Non-binder	0.5	
2*00801	<sup>60</sup> TTNTQNNDW <sub>68</sub>	0.4	9	Strong	2.3	
	<sup>653</sup> TVYNGTSKY <sub>661</sub>	0.4	5	Strong	1.5	
	<sup>114</sup> SSVGVTHGY <sub>122</sub>	0.15	6	Strong	0.8	
	<sup>444</sup> AAHCIHAEW <sub>452</sub>	0.25	21	Strong	17.8	
	<sup>207</sup> LLVAMVPEW <sub>215</sub>	0.3	2	Strong	6.1	
	<sup>443</sup> RAAHCIIHAEW <sub>452</sub>	0.08	1.7	Strong	5.4	
	<sup>604</sup> VVRHEGNLTW <sub>613</sub>	0.15	11	Strong	1.1	
	<sup>59</sup> HTTNTQNNDW <sub>68</sub>	0.3	40	Strong	0.2	
	<sup>182</sup> DSYAYMRNGW <sub>191</sub>	0.25	2	Strong	2.7	
	<sup>240</sup> MTAHITVPY <sub>248</sub>	0.25	51	Intermediate	0.8	
	<sup>332</sup> KVYNPPRTNY <sub>341</sub>	0.5	393	Intermediate	0.6	
	<sup>386</sup> AAKHSMSNTY <sub>394</sub>	1.5	1299	Weak	0.2	
	<sup>385</sup> LAAKHSMSNTY <sub>394</sub>	0.5	1380	Weak	0.2	
	<sup>284</sup> KVYANIPTY <sub>293</sub>	0.4	>10,000	Non-binder	0.2	
	2*01201	<sup>685</sup> ASFNYGAIK <sub>693</sub>	0.01	67	Intermediate	4.6
		<sup>626</sup> TSNPTAYNK <sub>634</sub>	0.25	226	Intermediate	2.2
<sup>324</sup> KTADPAYGK <sub>332</sub>		0.25	443	Intermediate	6.0	
<sup>490</sup> CIYQITHGK <sub>498</sub>		0.5	1035	Weak	6.8	
<sup>371</sup> RTDDTRLAK <sub>380</sub>		0.4	4176	Weak	0.3	
<sup>140</sup> RVVQAERFYK <sub>149</sub>		0.4	1353	Weak	0.6	
<sup>141</sup> VVQAERFYK <sub>149</sub>		0.25	>20,000	Non-binder	0.2	
<sup>250</sup> GVNRYDQYK <sub>258</sub>		0.5	8693	Non-binder	0.5	
<sup>240</sup> MTAHITVPY <sub>248</sub>		1.5	5547	Non-binder	0.3	
<sup>630</sup> TAYNKAPFTR <sub>639</sub>		0.4	>20,000	Non-binder	0.3	
4*02401	<sup>444</sup> AAHCIHAEW <sub>452</sub>	0.2	572	Weak	2.7	
	<sup>207</sup> LLVAMVPEW <sub>215</sub>	0.15	588	Weak	0.5	
	<sup>443</sup> RAAHCIIHAEW <sub>452</sub>	0.08	1792	Weak	1.7	
	<sup>604</sup> VVRHEGNLTW <sub>613</sub>	0.4	1001	Weak	0.4	

<sup>a</sup> Peptide positions are relative to the FMDV P1 capsid polyprotein amino acid sequence shown in Supplemental Fig. 1

<sup>b</sup> 9-mer or 10-mer peptides predicted as strong (rank score  $\leq 0.5$ ) or weak (rank score  $> 0.5$ ) binders using NetMHCpan v2.9

<sup>c</sup> Observed binding affinity in the LOCI assay and associated binding status category based on  $K_d$  threshold values: strong binder ( $K_d < 50$  nM), intermediate binder ( $K_d = 50$ – $499$  nM), weak binder ( $K_d = 500$ – $5000$  nM), and non-binder ( $K_d > 5000$  nM)

<sup>d</sup> Peptide binding stability half-life in hours

the LOCI assay identified nine strong ( $K_d < 50$  nM), five intermediate ( $K_d = 50$ – $499$  nM), and 10 weak ( $K_d = 500$ – $5000$  nM) binders (Table 1), and 13 molecules had binding stabilities  $> 1$  h in the SPA. LOCI binding affinity thresholds suggest at least one weak binder ( $K_d \leq 5000$  nM) was identified for each BoLA molecule. Only two molecules, BoLA-2\*00801 and BoLA-2\*01201, had peptides that were predicted strong or intermediate binders, whereas only weak binding or non-binding peptides were found for BoLA-1\*01901 and BoLA-4\*02401. We identified nine strong binders ( $K_d < 50$  nM) for

BoLA-2\*00801. A SPA binding stability half-life greater than 1 h was identified for 7, 4, and 2 peptides for BoLA-2\*00801, BoLA-2\*01201, and BoLA-4\*02401 molecules, respectively, suggesting there may be immunogenic peptides among the predictions for these three molecules.

### Functional predictions of BoLA class I allele groups

The *MHCcluster* tree was generated to predict the functional relationship of the BoLA molecules. Three large clusters were





identified as epitopes from experimental studies are frequently identified as potential epitopes by the *NetMHCpan* predictor. Second, the predictor will often identify alternative peptides that are predicted to have higher binding affinity compared to CD8<sup>+</sup> T cell epitopes previously identified in experimental studies, and these alternatives are frequently shorter peptides nested within the previously identified CD8<sup>+</sup> T cell epitope (Nene et al. 2012). Third, the process of using epitope predictions to guide in vitro and ex vivo experiments generates new and sometimes unanticipated data that can be used to iteratively inform improvements in the *NetMHCpan* and subsequent live animal immune response studies.

Only a fraction of peptides that bind to MHC molecules are CD8<sup>+</sup> T cell epitopes, with potential epitopes first narrowed by MHC class I determinants and further by TCR repertoire restriction and antigen processing (Yewdell and Bennink 1999). In order to be considered a CD8<sup>+</sup> T cell epitope, the pMHC complex must trigger a specific T cell immune response. Peptide affinity and stability are two factors that play an important role in determining the immunogenicity of a peptide (Busch and Pamer 1998; Harndahl et al. 2012). Positional scanning combinatorial peptide libraries (PSCPLs) have successfully been applied to address peptide-MHC-I specificity of humans (Rasmussen et al. 2014; Stryhn et al. 1996), swine (Pedersen et al. 2012), and cattle (Hansen et al. 2014). In cattle, epitopes or peptide binding motifs have been characterized for at least 13 BoLA molecules (Gaddum et al. 1996a; Gaddum et al. 1996b; Graham et al. 2007; Graham et al. 2008; Hansen et al. 2014; Hegde et al. 1999; Hegde et al. 1995; Li Pira et al. 2010; Macdonald et al. 2010; MacHugh et al. 2011; Momtaz et al. 2014; Nene et al. 2012; Svitek et al. 2014), and we have expanded characterization of peptide binding motifs for five additional molecules. We also produced recombinant BoLA-2\*01601, but this molecule had a poor  $\beta_2m$  complex-specific signal in the SPA using the PSCPL and predicted FMDV-specific peptides were not tested in the LOCI or SPA assays. Possible reasons for this non-functional molecule could be poor quality of the protein preparation or intrinsic biochemical properties of the molecule that result in incompatibility with the assay. For example, it is possible this particular molecule has a preference for 8-mers, 10-mers, or 11-mers instead of 9-mers used in the PSCPL, which would explain why a motif could not be derived (Stevens et al. 1998). Alternative approaches, such as peptide elution experiments, might be considered for molecules that appear to be incompatible with in vitro binding assays. While we were able to generate data for BoLA-3\*01701 using the PSCPL, subsequent testing in the SPA was unsuccessful, possibly due to loss of protein function, and specific FMDV peptide binding data was not generated in the LOCI or SPA. Another potential reason for the difficulty in characterizing the peptide binding affinity of BoLA-3\*01701 could be due to a relatively limited peptide binding repertoire outside of the

PSCPL. Further work to identify peptide motifs and epitopes for these two molecules is warranted as both appear to occupy regions of the *MHCcluster* diagram that lack in vivo or in vitro data supporting motif predictions (Fig 3). For example, BoLA-3\*01701 estimated prediction accuracy is 0.47 using nearest neighbor Rhesus macaque Mamu-B3901 and this molecule has been previously described as potentially not conforming to more commonly found peptide motifs (De Groot et al. 2003). Again, peptide elution experiments might be the most effective method to characterize the peptide motif for BoLA-3\*01701. In comparison, for BoLA-6\*01301 where in vitro and in vivo data are available from eight different studies (Gaddum et al. 1996b; Graham et al. 2006; Graham et al. 2008; Guzman et al. 2008; Hansen et al. 2014; Macdonald et al. 2010; Nene et al. 2012; Svitek et al. 2014), the estimated prediction accuracy is 0.853. While we did not identify BoLA-6\*01301 in our herd, we did identify BoLA-6\*01302 in 25 % of UVM-DCE Holstein cattle (Fig. 1). A higher frequency of 6\*01302 relative to 6\*01301 has also recently been reported in other cattle populations (Svitek et al., 2015). A PSCPL matrix was derived for 6\*01302; however, at this time, we elected to complete in vitro binding assays for the four other molecules more frequently identified in our herd. BoLA-2\*01201 was previously synthesized and a motif was generated (Hansen et al. 2014), and we decided to test FMDV-predicted peptides for this molecule because it was found in relatively high frequency in our herd.

*NetMHCpan-2.9* predicted 28 peptides derived from the FMDV structural proteins that would most likely bind four common BoLA molecules (BoLA-1\*01901, BoLA-2\*00801, BoLA-2\*01201, and BoLA-4\*02401) identified among animals from our research herd. *NetMHCpan* is trained on human, mice, swine, and a handful of BoLA peptide-MHC binding data. It is capable of making accurate predictions where no binding data are available (Nielsen et al. 2008). In our current study, 50 % (14/28) of the peptides predicted by the *NetMHCpan-2.9* algorithm with a rank score  $\leq 0.5$  % were demonstrated to have high or intermediate binding affinity in the LOCI assay. The number of predicted peptide binders for a particular molecule is affected by the rank cutoff value selected. For example, the algorithm predicted only two strong binding peptides for BoLA-1\*01901, and by increasing the rank threshold cutoff to greater than 2 %, we could potentially identify more binders. However, previous benchmark studies determined the vast majority of epitopes have binding with a rank score of 2 % or less (Erup Larsen et al. 2011). Therefore, applying a *NetMHCpan* threshold of 2 % allowed us to reduce the number of FMDV capsid protein 9- and 10-mer peptides from over 700 to 10 and 14 for alleles BoLA-2\*01201 and BoLA-2\*00801, respectively. The results of the binding assays allow us to further narrow the pool of peptides that might be evaluated as potential epitopes in T cell response assays for vaccinated or virus-challenged cattle, especially cattle

expressing BoLA-2\*00801 where we observed seven peptides with binding stability half-lives >1 h. The limited number of potential epitopes identified for the remaining molecules could either be a function of a limited number of epitopes in FMDV structural proteins or current limitations in the predictive models.

BoLA-2\*00801 and BoLA-4\*02401 share 92 % sequence identity in the peptide binding groove and were predicted to have similar peptide binding specificities. The algorithm predicted four FMDV peptides (<sup>444</sup>AAHCIHAEW<sub>452</sub>, <sup>207</sup>LLVAMVPEW<sub>215</sub>, <sup>443</sup>RAAHCIHAEW<sub>452</sub>, <sup>604</sup>VVRHEGNLTW<sub>613</sub>) as potential binders to both BoLA-2\*00801 and BoLA-4\*02401. However, our in vitro results indicate that these peptides had strong affinity for BoLA-2\*00801 and weak affinity for BoLA-4\*02401. A potential reason for the differences in binding could be due to the subtle differences in the key amino acids in the anchor positions. Both molecules have common amino acids, tryptophan, and phenylalanine in P9; however, the P2 anchor appears less important for BoLA-4\*02401. The amino acids in P2 for BoLA-4\*02401 have aromatic hydrophobic side chains, as opposed to BoLA-2\*00801, which has amino acids with polar, neutral side chains in this position (Fig. 2). This subtle difference could be the reason for weak binding by BoLA-4\*02401. As more BoLA binding data are generated, the algorithm can be iteratively retrained to capture these differences.

Perhaps more importantly, these data can be used to inform ex vivo CTL response or CD8+ tetramer assays using cells from cattle expressing these BoLA molecules. For example, Svitek et al. (2014) recently used tetramer assays to identify a *T. parva* epitope (Tp5<sub>87-95</sub>) that binds to two BoLA molecules (BoLA-1\*02301 and BoLA-T5). The role of cattle CD8+ T cell responses to FMDV infection or vaccination remains controversial, and a limited number of studies have evaluated class I-restricted CD8+ T cell responses following FMDV infection or vaccination (Childerstone et al. 1999; Guzman et al. 2010; Guzman et al. 2008; Patch et al. 2013; Patch et al. 2011). An important next step is to test whether the peptides identified in this study are MHC-I epitopes in cattle expressing the BoLA molecules evaluated in this study. For example, the CD8+ response to peptides identified in this study could be evaluated among cattle vaccinated with the AD5-FMDV vaccine expressing A24 P1 capsid proteins.

Thus far, it appears that leucine, phenylalanine, isoleucine, and methionine are important anchor residues in P9 of the BoLA class I binding motif. Position 2 is more variable, with glutamic acid, glutamine, leucine, arginine, threonine, and serine being the most commonly favored amino acids in this anchor position. In one specific cluster, tryptophan was identified as a strong anchor residue in P9. In another cluster, lysine and arginine were identified as favored anchor residues in P9. Some motif predictions have been validated in animals. For example, the CD8+ T cell antigen Tp2 and epitope

<sup>49</sup>KSSHGMGKVGK<sub>59</sub> from the Muguga strain of *T. parva* matched the binding specificity for BoLA-2\*01201 predicted by the algorithms (Nene et al. 2012). In contrast, MHC cluster did not capture the P5 anchor of the BoLA-6\*01301 molecule previously identified to bind an immunogenic peptide (Macdonald et al. 2010). This is not surprising as the P5 anchor appears to be relatively rare, and these rare events are less well recognized by the predictor tool. We also find it worth noting that peptides without the P5 anchor appear to have strong binding affinities for BoLA-6\*01301 and BoLA-6\*01302, and these peptides are immunogenic (Hansen et al. 2014; Svitek et al. 2015).

The functional relationship between all known BoLA molecules was predicted using the *MHCcluster* algorithm. The characterization of the BoLA binding specificities may also make it possible to cluster molecules that share largely overlapping peptide binding specificities (i.e., BoLA supertypes) (Lund et al. 2004; Sidney et al. 2008). A MHC cluster tree generated following inclusion of our data shows there are currently two clusters lacking binding data; representative molecules from each of these clusters are highlighted in blue (Fig. 3). The predictions of the *NetMHCpan v2.9* and *MHCcluster* algorithms can be improved by characterizing the binding specificity of additional BoLA molecules, perhaps by first addressing regions of the tree currently lacking coverage.

In conclusion, we have translated technologies that have been applied to the study of human T cell responses to the study of cattle immunity. These technologies include bioinformatics, peptide synthesis, and proteomics. Using FMDV or other important pathogens, extending this research should advance our understanding of how viral antigens are recognized by the bovine immune system and contribute to improved vaccine development against bovine pathogens.

**Acknowledgments** This work was funded by a cooperative agreement (58-1940-9-953) between UVM and the Agricultural Research Service, US Department of Agriculture, which was funded in part by an interagency agreement (HSHQDC-09-X-00373) and (HSHQPM-13-X-00110) between the Department of Homeland Security and US Department of Agriculture. This work was in part funded by a grant (0965346) from the USA National Science Foundation under the BREAD program. We also acknowledge support from USDA-Agriculture and Food Research Initiative, National Institute of Food and Agriculture pre-doctoral fellowship (28832), as well as matching funds and research support provided by the UVM College of Agriculture and Life Sciences. Matthew Sammons provided technical assistance with PCR-SSP typing of the cattle. Tim Hunter and Jessica Hoffman at the DNA Analysis facility at the UVM Cancer Center provided technical assistance with the BioTek plate reader used for LOCI assays.

#### Compliance with ethical standards

**Conflict of interest** The authors declare that they have no competing interests.

## References

- Bamford AI, Douglas A, Friede T, Stevanovic S, Rammensee HG, Adair BM (1995) Peptide motif of a cattle MHC class I molecule. *Immunol Lett* 45:129–36
- Brake DA, McIlhane M, Miller T, Christianson K, Keene A, Lohns G, Purcell C, Neilan J, Schutta C, Barrera J, Burrage T, Brough DE, Butman BT (2012) Human adenovirus-vectored foot-and-mouth disease vaccines: establishment of a vaccine product profile through *in vitro* testing. *Dev Biol (Basel)* 134:123–33
- Bruno L, Cortese M, Rappuoli R, Merola M (2015) Lessons from Reverse Vaccinology for viral vaccine design. *Curr Opin Virol* 11: 89–97
- Busch DH, Pamer EG (1998) MHC class I/peptide stability: implications for immunodominance, *in vitro* proliferation, and diversity of responding CTL. *J Immunol* 160:4441–8
- Calis JJ, Maybeno M, Greenbaum JA, Weiskopf D, De Silva AD, Sette A, Kesmir C, Peters B (2013) Properties of MHC class I presented peptides that enhance immunogenicity. *PLoS Comput Biol* 9, e1003266
- Childerstone AJ, Cedillo-Baron L, Foster-Cuevas M, Parkhouse RM (1999) Demonstration of bovine CD8+ T-cell responses to foot-and-mouth disease virus. *J Gen Virol* 80(Pt 3):663–9
- Codner G (2010) Assessing the MHC class I diversity in dairy populations. University of Glasgow
- De Groot AS, Nene V, Hegde NR, Srikumaran S, Rayner J, Martin W (2003) T cell epitope identification for bovine vaccines: an epitope mapping method for BoLA A-11. *Int J Parasitol* 33:641–53
- Ellis SA, Staines KA, Stear MJ, Hensen EJ, Morrison WI (1998) DNA typing for BoLA class I using sequence-specific primers (PCR-SSP). *Eur J Immunogenet* 25:365–370
- Erup Larsen M, Kloverpris H, Stryhn A, Koofhethile CK, Sims S, Ndung'u T, Goulder P, Buus S, Nielsen M (2011) HLA-restrictor—a tool for patient-specific predictions of HLA restriction elements and optimal epitopes within peptides. *Immunogenetics* 63:43–55
- Follin E, Karlsson M, Lundegaard C, Nielsen M, Wallin S, Paulsson K, Westerdahl H (2013) *In silico* peptide-binding predictions of passerine MHC class I reveal similarities across distantly related species, suggesting convergence on the level of protein function. *Immunogenetics* 65:299–311
- Frankild S, de Boer RJ, Lund O, Nielsen M, Kesmir C (2008) Amino acid similarity accounts for T cell cross-reactivity and for “holes” in the T cell repertoire. *PLoS One* 3:e1831
- Gaddum RM, Ellis SA, Willis AC, Cook RS, Staines KA, Thomas LH, Taylor G (1996a) Identification of potential CTL epitopes of bovine RSV using allele-specific peptide motifs from bovine MHC class I molecules. *Vet Immunol Immunopathol* 54:211–9
- Gaddum RM, Willis AC, Ellis SA (1996b) Peptide motifs from three cattle MHC (BoLA) class I antigens. *Immunogenetics* 43:238–9
- Graham SP, Honda Y, Pelle R, Mwangi DM, Glew EJ, de Villiers EP, Shah T, Bishop R, van der Bruggen P, Nene V, Taracha EL (2007) A novel strategy for the identification of antigens that are recognised by bovine MHC class I restricted cytotoxic T cells in a protozoan infection using reverse vaccinology. *Immunome Res* 3:2
- Graham SP, Pelle R, Honda Y, Mwangi DM, Tonukari NJ, Yamage M, Glew EJ, de Villiers EP, Shah T, Bishop R, Abuya E, Awino E, Gachanja J, Luyai AE, Mbwika F, Muthiani AM, Ndegwa DM, Njehira M, Nyanjui JK, Onono FO, Osaso J, Saya RM, Wildmann C, Fraser CM, Maudlin I, Gardner MJ, Morzaria SP, Loosmore S, Gilbert SC, Audonnet JC, van der Bruggen P, Nene V, Taracha EL (2006) *Theileria parva* candidate vaccine antigens recognized by immune bovine cytotoxic T lymphocytes. *Proc Natl Acad Sci U S A* 103:3286–91
- Graham SP, Pelle R, Yamage M, Mwangi DM, Honda Y, Mwakubambanya RS, de Villiers EP, Abuya E, Awino E, Gachanja J, Mbwika F, Muthiani AM, Muriuki C, Nyanjui JK, Onono FO, Osaso J, Riitho V, Saya RM, Ellis SA, McKeever DJ, MacHugh ND, Gilbert SC, Audonnet JC, Morrison WI, van der Bruggen P, Taracha EL (2008) Characterization of the fine specificity of bovine CD8 T-cell responses to defined antigens from the protozoan parasite *Theileria parva*. *Infect Immun* 76:685–94
- Groscurth P, Filgueira L (1998) Killing mechanisms of cytotoxic T Lymphocytes. *News Physiol Sci* 13:17–21
- Grubman MJ, Moraes MP, Schutta C, Barrera J, Neilan J, Eddyreddy D, Butman BT, Brough DE, Brake DA (2010) Adenovirus serotype 5-vectored foot-and-mouth disease subunit vaccines: the first decade. *Futur Virol* 5:51–64
- Guzman E, Taylor G, Charleston B, Ellis SA (2010) Induction of a cross-reactive CD8+ T-cell response following foot-and-mouth disease virus vaccination. *J Virol* 84:12375–84
- Guzman E, Taylor G, Charleston B, Skinner MA, Ellis SA (2008) An MHC-restricted CD8+ T-cell response is induced in cattle by foot-and-mouth disease virus (FMDV) infection and also following vaccination with inactivated FMDV. *J Gen Virol* 89:667–75
- 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* 66:705–18
- 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* 14:173–80
- 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* 42:1405–16
- Hart J, MacHugh ND, Morrison WI (2011) *Theileria annulata*-transformed cell lines are efficient antigen-presenting cells for *in vitro* analysis of CD8 T cell responses to bovine herpesvirus-1. *Vet Res* 42:119
- Hegde NR, Deshpande MS, Godson DL, Babiuk LA, Srikumaran S (1999) Bovine lymphocyte antigen-A11-specific peptide motif as a means to identify cytotoxic T-lymphocyte epitopes of bovine herpesvirus 1. *Viral Immunol* 12:149–61
- Hegde NR, Ellis SA, Gaddum RM, Tregaskes CA, Sarath G, Srikumaran S (1995) Peptide motif of the cattle MHC class I antigen BoLA-A11. *Immunogenetics* 42:302–3
- Hegde NR, Srikumaran S (2000) Reverse immunogenetic and polyepitopic approaches for the induction of cell-mediated immunity against bovine viral pathogens. *Anim Health Res Rev* 1:103–18
- 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
- Li Pira G, Ivaldi F, Moretti P, Manca F (2010) High throughput T epitope mapping and vaccine development. *J Biomed Biotechnol* 2010:1–36
- Li X, Liu J, Qi J, Gao F, Li Q, Li X, Zhang N, Xia C, Gao GF (2011) Two distinct conformations of a rinderpest virus epitope presented by bovine major histocompatibility complex class I N\*01801: a host strategy to present featured peptides. *J Virol* 85:6038–48
- Lund O, Nielsen M, Kesmir C, Petersen AG, Lundegaard C, Worning P, Sylvester-Hvid C, Lamberth K, Roder G, Justesen S, Buus S, Brnak S (2004) Definition of supertypes for HLA molecules using clustering of specificity matrices. *Immunogenetics* 55:797–810
- Macdonald IK, Harkiolaki M, Hunt L, Connelley T, Carroll AV, MacHugh ND, Graham SP, Jones EY, Morrison WI, Flower DR,



- Ellis SA (2010) MHC class I bound to an immunodominant *Theileria parva* epitope demonstrates unconventional presentation to T cell receptors. *PLoS Pathog* 6:e1001149
- MacHugh ND, Weir W, Burrells A, Lizundia R, Graham SP, Taracha EL, Shields BR, Langsley G, Morrison WI (2011) Extensive polymorphism and evidence of immune selection in a highly dominant antigen recognized by bovine CD8 T cells specific for *Theileria annulata*. *Infect Immun* 79:2059–69
- Momtaz S, Rahman A, Sultana M, Hossain MA (2014) Evolutionary analysis and prediction of peptide vaccine candidates for foot-and-mouth-disease virus types A and O in Bangladesh. *Evol Bioinform Online* 10:187–96
- Nene V, Svitek N, Toye P, Golde WT, Barlow J, Harndahl M, Buus S, Nielsen M (2012) Designing bovine T cell vaccines via reverse immunology. *Ticks Tick Borne Dis* 188–192
- 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* 4:100–107
- Patch JR, Kenney M, Pacheco JM, Grubman MJ, Golde WT (2013) Characterization of cytotoxic T lymphocyte function after foot-and-mouth disease virus infection and vaccination. *Viral Immunol* 26:239–49
- Patch JR, Pedersen LE, Toka FN, Moraes M, Grubman MJ, Nielsen M, Jungersen G, Buus S, Golde WT (2011) Induction of foot-and-mouth disease virus-specific cytotoxic T cell killing by vaccination. *Clin Vaccine Immunol* 18:280–8
- 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* 44:251–8
- Rappuoli R (2001) Reverse vaccinology, a genome-based approach to vaccine development. *Vaccine* 19:2688–91
- 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* 193:4790–802
- Sidney J, Peters B, Frahm N, Brander C, Sette A (2008) HLA class I supertypes: a revised and updated classification. *BMC Immunol* 9:1
- Sinnathamby G, Seth S, Nayak R, Shaila MS (2004) Cytotoxic T cell epitope in cattle from the attachment glycoproteins of rinderpest and peste des petits ruminants viruses. *Viral Immunol* 17:401–10
- Stevens J, Wiesmuller KH, Walden P, Joly E (1998) Peptide length preferences for rat and mouse MHC class I molecules using random peptide libraries. *Eur J Immunol* 28:1272–1279
- 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
- Svitek N, Awino E, Nene V, Steinaa L (2015) BoLA-6\*01301 and BoLA-6\*01302, two allelic variants of the A18 haplotype, present the same epitope from the Tp1 antigen of *Theileria parva*. *Vet Immunol Immunopathol* 167:80–5
- Svitek N, Hansen AM, Steinaa L, Saya R, Awino E, Nielsen M, Buus S, Nene V (2014) Use of “one-pot, mix-and-read” peptide-MHC class I tetramers and predictive algorithms to improve detection of cytotoxic T lymphocyte responses in cattle. *Vet Res* 45:50
- Tatsis N, Ertl HCJ (2004) Adenoviruses as vaccine vectors. *Mol Ther* 10:616–629
- Thomsen M, Lundegaard C, Buus S, Lund O, Nielsen M (2013) *MHCcluster*, a method for functional clustering of MHC molecules. *Immunogenetics* 65:655–665
- 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* 40:281–287
- Yewdell JW, Bennink JR (1999) Immunodominance in major histocompatibility complex class I-restricted T lymphocyte responses. *Annu Rev Immunol* 17:51–88