Published in final edited form as:

Immunogenetics. 2017 May ; 69(5): 351–358. doi:10.1007/s00251-017-0978-6.

Peptide-binding motifs of two common equine class I MHC molecules in Thoroughbred horses

Tobias Bergmann¹, Mikaela Lindvall², Erin Moore², Eugene Moore², John Sidney², Donald Miller³, Rebecca L. Tallmadge⁴, Paisley T. Myers⁵, Stacy A. Malaker⁵, Jeffrey Shabanowitz⁵, Nikolaus Osterrieder¹, Bjoern Peters², Donald F. Hunt^{5,6}, Douglas F. Antczak³, and Alessandro Sette²

¹Institut für Virologie, Freie Universität Berlin, 14163 Berlin, Germany

²Department of Vaccine Discovery, La Jolla Institute for Allergy and Immunology, 9420 Athena Circle, La Jolla, CA 92037, USA

³Baker Institute for Animal Health, College of Veterinary Medicine, Cornell University, Ithaca, NY 14853, USA

⁴Department of Clinical Sciences, College of Veterinary Medicine, Cornell University, Ithaca, NY 14853, USA

⁵Department of Chemistry, University of Virginia, Charlottesville, VA 22904, USA

⁶Department of Pathology, University of Virginia, Charlottesville, VA 22908, USA

Abstract

Quantitative peptide-binding motifs of MHC class I alleles provide a valuable tool to efficiently identify putative T cell epitopes. Detailed information on equine MHC class I alleles is still very limited, and to date, only a single equine MHC class I allele, Eqca-1*00101 (ELA-A3 haplotype), has been characterized. The present study extends the number of characterized ELA class I specificities in two additional haplotypes found commonly in the Thoroughbred breed. Accordingly, we here report quantitative binding motifs for the ELA-A2 allele Eqca-16*00101 and the ELA-A9 allele Eqca-1*00201. Utilizing analyses of endogenously bound and eluted ligands and the screening of positional scanning combinatorial libraries, detailed and quantitative peptide-binding motifs were derived for both alleles. Eqca-16*00101 preferentially binds peptides with aliphatic/hydrophobic residues in position 2 and at the C-terminus, and Eqca-1*00201 has a preference for peptides with arginine in position 2 and hydrophobic/aliphatic residues at the C-terminus. Interestingly, the Eqca-16*00101 motif resembles that of the human HLA A02-supertype, while the Eqca-1*00201 motif resembles that of the HLA B27-supertype and two macaque class I alleles. It is expected that the identified motifs will facilitate the selection of candidate epitopes for the study of immune responses in horses.

Keywords

MHC class I; Peptide binding; MHC binding motif; T cell epitope; *Equus caballus* (Eqca)

Quantitative binding motifs of MHC class I alleles provide a valuable tool to efficiently identify putative T cell epitopes. However, detailed information on equine MHC class I alleles is still very limited. We recently determined a binding motif for the equine MHC class I allele Eqca-1*00101 of the ELA-A3 haplotype using sequencing of endogenous Eqca-1*00101 peptide binders. A more detailed and quantitative motif was further derived by testing the capacity of positional scanning combinatorial libraries (PSCL) to bind Eqca-1*00101 with in vitro assays based on the use of purified MHC molecules (Bergmann et al. 2015).

The MHC class I genes of the ELA A2, A3, and A9 haplotypes have been identified through full-length sequencing cDNA (Tallmadge et al. 2010). The high gene frequencies of the ELA A2 (0.168), A3 (0.168), and A9 (0.136) haplotypes in the Thoroughbred breed (Antczak et al. 1986; Tseng et al. 2010) make them attractive candidates for MHC class I molecule characterization. Here, we report information on quantitative binding motifs for two classical MHC class I genes, Eqca-16*00101 and Eqca-1*00201, of the ELA A2 and ELA A9 haplotypes. These specific genes were investigated because they are members of allelic series that are found in several horse haplotypes and share structural features with other classical horse MHC class I molecules; most of the other MHC class I genes on these haplotypes do not have these features (Antczak et al. 1986; Tseng et al. 2010).

In initial experiments, we sought to identify peptides that are endogenously bound by Eqca-16*00101 (ELA A2) and Eqca-1*00201 (ELA A9). As detailed previously (Bergmann et al. 2015; Sidney et al. 2013b), MHC molecules were purified from P815 cell clones transfected with either Eqca-16*00101 or Eqca-1*00201 genes (Azab et al. 2014) by immunoaffinity with the equine MHC-I-specific antibody (CZ3) (Rappocciolo et al. 2003). Peptides were then eluted from the respective MHC molecules as described previously (Slingluff et al. 1993) and analyzed by nanoflow high-performance liquid chromatography (HPLC)/microelectrospray ionization, coupled directly to a Thermo Fisher Scientific Orbitrap Fusion Tribrid mass spectrometer, equipped with a front-end electron transfer dissociation (FETD) source (Earley et al. 2013). Data were acquired as previously described (Udeshi et al. 2008) and searched against the SwissProt (Bairoch and Apweiler 2000) mouse database using the Open Mass Spectrometry Search Algorithm software (Geer et al. 2004), which enabled us to generate a list of candidate peptide sequences (source proteins of sequenced peptides is shown in Supplemental Table 1).

For both alleles, peptide sequences were determined for the 40 most abundant ligands. In the case of Eqca-16*00101, and as shown in Table 1, the preferred ligand size was nine residues, with 25 of the 40 ligands (63%) being nonamers. Of the remaining Eqca-16*00101 ligands, four were 10-mers (10%) and 11 (27.5%) were 11-mers. For Eqca-1*00201 (also see Table 1), 90% (36/40) of the ligands were 9-mers while only 10% (4/40) were 10-mers. Taken together, the MHC elution data has identified 9-mers as the preferred ligand size for both alleles, conforming to the canonical ligand size known for most human, mouse, and macaque MHC-I alleles (Rammensee et al. 1999), as well as the sole previously characterized equine MHC class I allele (Bergmann et al. 2015).

The panels of eluted ligands were next analyzed to derive corresponding main anchor binding motifs for each allele. For Eqca-16*00101, it was observed that the preponderance of residues found in position 2 were the aliphatic residues valine or isoleucine, with the majority (22/40; 55%) being valine (see Table 1 and Supplemental Table 2); other residues present were similarly small/polar (T, S, or Q) or aliphatic (M, A or L). At the C-terminus of the bound peptides, a preference for hydrophobic/aliphatic residues was noted, with leucine (19/40; 47.5%) being the most common, followed by methionine (7/40; 17.5%), isoleucine (6/40; 15%), and valine (5/40; 12.5%). Also noteworthy was position 1, where 80% (32/40) of the residues have polar characteristics and 47.5% (19/40) were positively charged.

For Eqca-1*00201, all peptides carried arginine in position 2, with only one exception (tyrosine) (see Table 1 and Supplemental Table 2). At the C-terminus, a preference for aliphatic/hydrophobic residues was apparent, as 65% (26/40) of peptides had leucine, and all but one of the remaining peptides had another small or hydrophobic residue (isoleucine, alanine, methionine or phenylalanine). Interestingly, position 1 also showed a narrow specificity, with 77.5% (31/40) of the eluted peptides carrying an asparagine residue. These results suggest that positions 1, 2, and the C-terminus play an important role in Eqca-1*00201 binding. Whether all three positions should be regarded as primary anchors was addressed in further experiments described below.

We took advantage of the eluted ligands identified above to develop quantitative, high-throughput binding assays. The strategy utilized has been successfully applied to several MHC class I and class II molecules expressed in humans, rodents, non-human primates, and horses (see, e.g., (Kubo et al. 1994; Sidney et al. 2013a)). Briefly, various eluted ligands were radiolabeled; in some cases, tyrosinated analogs of eluted ligands were synthesized to allow for radiolabeling. The radiolabeled probes were incubated with purified MHC molecules in direct binding assays and bound radioactivity determined as described previously (Sidney et al. 2013a). Ligands yielding optimal signal-to-noise ratios and specificity were selected for use as radiolabeled probes in competitive inhibition assays for the corresponding allele. For Eqca-16*00101, the ligand selected for subsequent assays was the *Mus musculus* geminin 140–148 peptide EVAEHVQYM (IC₅₀ 2.4 nM); for Eqca-1*00201, the *M. musculus* 60s ribosomal L4 protein 48–57 peptide NRQPYAVSEL (IC₅₀ 3.1 nM) was chosen. Quantitative in vitro competitive binding assays utilizing purified MHC-I and iodine¹²⁵-labeled standard probe peptides in a monoclonal antibody capture assay platform were performed essentially as described previously (Sidney et al. 2013a). Briefly, 0.1–1 nM of radiolabeled peptide was co-incubated at room temperature with 0.33 nM–33 μM inhibitory peptide and 1–10 nM of purified MHC-I in the presence of a protease inhibitor cocktail and 1 μM human β2-microglobulin (Scripps Laboratories) at pH 7 in PBS with 0.05% NP40. Following a 2-day incubation, MHC-I-bound radioactivity was determined by capturing MHC-I/peptide complexes on CZ3 (anti-ELA class I)-coated Lumitrac 600 plates (Greiner Bio-one, Frickenhausen, Germany) and bound radioactivity measured using the TopCount (Packard Instrument Co., Meriden, CT) microscintillation counter. The concentration of peptide yielding 50% inhibition of the binding of the radiolabeled peptide was calculated. Under the conditions utilized, where [label] < [MHC] and IC₅₀ [MHC], the measured IC₅₀ values were reasonable approximations of the true K_D values (Cheng and Prusoff 1973; Gulukota et al. 1997). Each competitor peptide was tested

at six different concentrations covering a 100,000-fold dose range and in three or more independent experiments. As a positive control, the unlabeled version of the radiolabeled probe was also tested in each experiment.

Utilizing this validated assay strategy, we found that 80% (32/40) of the Eqca-16*00101-eluted peptides bound Eqca-16*00101 with very high affinity ($IC_{50} < 50$ nM) and another 15% ($n = 6$) with an affinity in the 50–500-nM range (Table 1). In the case of Eqca-1*00201, 72.5% (29/40) of eluted peptides bound Eqca-1*00201 with very high affinity ($IC_{50} < 50$ nM) and 25% ($n = 10$) with an IC_{50} in the 50–500-nM range. For context, we note that the 500-nM threshold has previously been found to correspond with the majority of MHC-I epitopes (Assarsson et al. 2007; Sette et al. 1994). Thus, taken together, these data provide validation of the elution data and of the developed high-throughput binding assays.

To determine more detailed and quantitative binding motifs for both alleles, we utilized positional scanning combinatorial libraries (PSCL). Testing suitable PSCL for MHC binding allows for the evaluation of contributions to overall binding of any of the 20 natural amino acids in any position along the peptide backbone. PSCL have been successfully utilized to derive detailed quantitative motifs for more than a dozen different MHC molecules (see, e.g., (Bergmann et al. 2015; Sidney et al. 2008a; Sidney et al. 2007)). PSCL, synthesized as described previously (Pinilla et al. 1992), contain randomized 9-mer peptides with one fixed residue at a single position. With each of the 20 naturally occurring residues represented at each position along a 9-mer backbone, the entire library consisted of 180 peptide mixtures. A library completely randomized at each position was also tested.

PSCLs were tested in competition assays as above and data analyzed as described previously (Sidney et al. 2008b). Briefly, IC_{50} nM values for each mixture were standardized as a ratio to the geometric mean IC_{50} nM value of the entire set of 180 mixtures and then normalized at each position, so that the value associated with optimal binding at each position corresponded to 1. For each position, an average (geometric) relative binding affinity (ARB) was calculated, and then, the ratio of the ARB for the entire library to the ARB for each position was derived. This ratio, which describes the factor by which the normalized geometric mean binding affinity associated with all 20 residues at a specified position differs from that of the average affinity of the entire library, has been denominated as the specificity factor (SF). As calculated, positions with the highest specificity will have the highest SF value. As previously described (Sidney et al. 2008b), primary anchor positions may be defined as those with an $SF \geq 2.4$. This criterion generally identifies positions where the majority of residues are associated with significant decreases in binding capacity. Prominent secondary influences may arbitrarily be defined based on the number of 10-fold effects on binding capacity, relative to the optimal.

In the case of Eqca-16*00101 (Table 2), the PSCL analysis identified position 2 ($SF = 3.94$) and the C-terminus ($SF = 6.39$) as primary anchors. In position 2, valine was the most preferred, followed by glutamine, leucine, isoleucine, and the small residues alanine and threonine, representing a pattern in very good agreement with the elution data. At the C-terminus, phenylalanine was identified as the optimal residue, followed by valine, isoleucine, and leucine. The preference for phenylalanine, as characterized by the PSCL,

was surprising, as phenylalanine was not present at the C-terminus of any of the eluted ligands. However, it is similarly hydrophobic in nature to valine, isoleucine, and leucine, perhaps explaining its association with high-affinity binding. We speculate that the absence of phenylalanine in the sequenced ligands may reflect a processing bias of the mouse cell line used for transfection. In positions 1, 3, and 6, 10, 14, and 13 different residues, respectively, were associated with 10-fold reductions in affinity. Accordingly, these may be defined as strong secondary anchor positions. In position 1, the same preference for polar residues identified in the elution data was also apparent, while positions 3 and 6 were associated with a preference for hydrophobic/aliphatic residues.

For Eqca-1*00201, position 2 and the C-terminus were confirmed to be primary anchors, with SF values of 7.46 and 3.50, respectively (Table 3). In position 2, in agreement with the elution data, only arginine was tolerated, with all other residues being associated with at least 25-fold lower affinity. At the C-terminus, a preference for hydrophobic/aliphatic residues was apparent, with phenylalanine being the most preferred, followed by leucine, valine, and isoleucine. The preference for phenylalanine at the C-terminus was again surprising, as it was present in only 10% of the eluted ligands. At position 1, with an SF only slightly below the 2.4 threshold, 14 residues were associated with greater than 10-fold reduction in binding capacity, suggesting that position 1 is at least a very strong secondary anchor. The position 1 preference for asparagine identified in the elution set was confirmed by the PSCL, although other residues are also tolerated. Position 3, where five residues had 10-fold or greater reductions, appears to be a weak secondary anchor. Binding specificities for both alleles as characterized from the foregoing analyses are summarized in Fig. 1.

The present study presents experiments that utilized the elution of endogenously bound ligands and the screening of positional scanning combinatorial libraries to derive detailed and quantitative peptide-binding motifs for two common horse class I alleles. Eqca-16*00101 was found to preferentially bind peptides with aliphatic/hydrophobic residues in position 2 and the C-terminus, while Eqca-1*00201 demonstrated a preference for peptides with arginine in position 2 and hydrophobic/aliphatic residues at the C-terminus. Interestingly, the Eqca-16*00101 motif resembles that associated with the human HLA A02-supertype (Sidney et al. 2008c), while the Eqca-1*00201 motif resembles that associated with the HLA B27-supertype (Sidney et al. 2008c), and also two macaque MHC class I specificities (Loffredo et al. 2009). A particularly interesting finding in a previous study was that the equine class I allele Eqca-1*00101, the only other horse class I MHC allele characterized to date, was shown to have a narrow peptide repertoire, as compared to the repertoire breadth detected for a range of different HLA class I molecules. More specifically, Eqca-1*00101 only bound about 0.5% of all possible 9-mer peptides (Bergmann et al. 2015). By contrast, detailed analysis of the repertoires of a set of 27 common HLA class I alleles found an average repertoire of 3% (range 0.1–10.4%) (Paul et al. 2013). Future studies will take advantage of the present data to address whether the repertoires of Eqca-16*00101 and Eqca-1*00201 are also narrow, suggesting a commonality of most or all equine MHC class I alleles or more reflective of the repertoire breadth of alleles with the HLA A02- and B27-supertypes.

The combined genotype frequency of 0.47 for the ELA A2, A3, and A9 haplotypes in the Thoroughbred breed means that approximately 70% of Thoroughbreds will carry one of these haplotypes. The widespread use of Thoroughbred stallions as sires of sport horses increases the occurrence of these haplotypes in other horses around the world. It is expected that the present data will facilitate the selection of candidate epitopes for evaluation of T cell immunity in horses in experimental infection and vaccine trials.

Supplementary Material

Refer to Web version on PubMed Central for supplementary material.

Acknowledgments

The present work has been supported by funds provided through NIH NIAID contracts HHSN272200900042C and HHSN272201400045C, NIH U19AI090023 (Sette), NIH R01AI112566 (Sidney) and NIH AI 033993 (Hunt), the Harry M. Zweig Memorial Fund for Equine Research (Antczak), and Morris Animal Foundation grant D12EQ-026 (Antczak). DFA is an Investigator of the Dorothy Russell Havemeyer Foundation, Inc.

References

- Antczak DF, Bailey E, Barger B, Guerin G, Lazary S, McClure J, Mottironi VD, Symons R, Templeton J, Varewyck H. Joint report of the Third International Workshop on Lymphocyte Alloantigens of the Horse, Kennett Square, Pennsylvania, 25–27 April 1984. *Anim Genet.* 1986; 17:363–373. [PubMed: 3826760]
- Assarsson E, Sidney J, Oseroff C, Pasquetto V, Bui H-H, Frahm N, Brander C, Peters B, Grey H, Sette A. A quantitative analysis of the variables affecting the repertoire of T cell specificities recognized after vaccinia virus infection. *J Immunol.* 2007; 178:7890–7901. [PubMed: 17548627]
- Azab W, Harman R, Miller D, Tallmadge R, Frampton AR Jr, Antczak DF, Osterrieder N. Equid herpesvirus type 4 uses a restricted set of equine major histocompatibility complex class I proteins as entry receptors. *J Gen Virol.* 2014; 95:1554–1563. [PubMed: 24722677]
- Bairoch A, Apweiler R. The SWISS-PROT protein sequence database and its supplement TrEMBL in 2000. *Nucleic Acids Res.* 2000; 28:45–48. [PubMed: 10592178]
- Bergmann T, Moore C, Sidney J, Miller D, Tallmadge R, Harman RM, Oseroff C, Wriston A, Shabanowitz J, Hunt DF, Osterrieder N, Peters B, Antczak DF, Sette A. The common equine class I molecule Eqca-1*00101 (ELA-A3.1) is characterized by narrow peptide binding and T cell epitope repertoires. *Immunogenetics.* 2015
- Cheng Y, Prusoff WH. Relationship between the inhibition constant (K_i) and the concentration of inhibitor which causes 50 per cent inhibition (I₅₀) of an enzymatic reaction. *Biochem Pharmacol.* 1973; 22:3099–3108. [PubMed: 4202581]
- Earley L, Anderson LC, Bai DL, Mullen C, Syka JE, English AM, Dunyach JJ, Stafford GC Jr, Shabanowitz J, Hunt DF, Compton PD. Front-end electron transfer dissociation: a new ionization source. *Anal Chem.* 2013; 85:8385–8390. [PubMed: 23909443]
- Geer LY, Markey SP, Kowalak JA, Wagner L, Xu M, Maynard DM, Yang X, Shi W, Bryant SH. Open mass spectrometry search algorithm. *J Proteome Res.* 2004; 3:958–964. [PubMed: 15473683]
- Gulukota K, Sidney J, Sette A, DeLisi C. Two complementary methods for predicting peptides binding major histocompatibility complex molecules. *J Mol Biol.* 1997; 267:1258–1267. [PubMed: 9150410]
- Kubo RT, Sette A, Grey HM, Appella E, Sakaguchi K, Zhu NZ, Arnott D, Sherman N, Shabanowitz J, Michel H, et al. Definition of specific peptide motifs for four major HLA-A alleles. *J Immunol.* 1994; 152:3913–3924. [PubMed: 8144960]
- Loffredo JT, Sidney J, Bean AT, Beal DR, Bardet W, Wahl A, Hawkins OE, Piaskowski S, Wilson NA, Hildebrand WH, Watkins DI, Sette A. Two MHC class I molecules associated with elite control of

- immunodeficiency virus replication, Mamu-B*08 and HLA-B*2705, bind peptides with sequence similarity. *J Immunol.* 2009; 182:7763–7775. [PubMed: 19494300]
- Paul S, Weiskopf D, Angelo MA, Sidney J, Peters B, Sette A. HLA class I alleles are associated with peptide-binding repertoires of different size, affinity, and immunogenicity. *J Immunol.* 2013; 191:5831–5839. [PubMed: 24190657]
- Pinilla C, Appel JR, Blanc P, Houghten RA. Rapid identification of high affinity peptide ligands using positional scanning synthetic peptide combinatorial libraries. *BioTechniques.* 1992; 13:901–905. [PubMed: 1476743]
- Rammensee H, Bachmann J, Emmerich NP, Bachor OA, Stevanovic S. SYFPEITHI: database for MHC ligands and peptide motifs. *Immunogenetics.* 1999; 50:213–219. [PubMed: 10602881]
- Rappocciolo G, Birch J, Ellis SA. Down-regulation of MHC class I expression by equine herpesvirus-1. *J Gen Virol.* 2003; 84:293–300. [PubMed: 12560560]
- Sette A, Vitiello A, Rehman B, Fowler P, Nayarsina R, Kast WM, Melief CJ, Oseroff C, Yuan L, Ruppert J, Sidney J, del Guercio MF, Southwood S, Kubo RT, Chesnut RW, Grey HM, Chisari FV. The relationship between class I binding affinity and immunogenicity of potential cytotoxic T cell epitopes. *J Immunol.* 1994; 153:5586–5592. [PubMed: 7527444]
- Sidney J, Peters B, Moore C, Pencille TJ, Ngo S, Masterman KA, Asabe S, Pinilla C, Chisari FV, Sette A. Characterization of the peptide-binding specificity of the chimpanzee class I alleles A*0301 and A*0401 using a combinatorial peptide library. *Immunogenetics.* 2007
- Sidney J, Assarsson E, Moore C, Ngo S, Pinilla C, Sette A, Peters B. Quantitative peptide binding motifs for 19 human and mouse MHC class I molecules derived using positional scanning combinatorial peptide libraries. *Immunome research.* 2008a; 4:1. [PubMed: 18179690]
- Sidney J, Assarsson E, Moore C, Ngo S, Pinilla C, Sette A, Peters B. Quantitative peptide binding motifs for 19 human and mouse MHC class I molecules derived using positional scanning combinatorial peptide libraries. *Immunome Res.* 2008b; 4:2. [PubMed: 18221540]
- Sidney J, Peters B, Frahm N, Brander C, Sette A. HLA class I supertypes: a revised and updated classification. *BMC Immunol.* 2008c; 9:1. [PubMed: 18211710]
- Sidney, J., Southwood, S., Moore, C., Oseroff, C., Pinilla, C., Grey, HM., Sette, A. Measurement of MHC/peptide interactions by gel filtration or monoclonal antibody capture. In: Coligan, John E., et al., editors. *Current protocols in immunology.* 2013a. Chapter 18: Unit 18 3
- Sidney J, Southwood S, Moore C, Oseroff C, Pinilla C, Grey HM, Sette A. Measurement of MHC/peptide interactions by gel filtration or monoclonal antibody capture. *Curr Protoc Immunol.* 2013b chapter 18:unit 18 3.
- Slingluff CL Jr, Cox AL, Henderson RA, Hunt DF, Engelhard VH. Recognition of human melanoma cells by HLA-A2.1-restricted cytotoxic T lymphocytes is mediated by at least six shared peptide epitopes. *J Immunol.* 1993; 150:2955–2963. [PubMed: 7681084]
- Tallmadge RL, Campbell JA, Miller DC, Antczak DF. Analysis of MHC class I genes across horse MHC haplotypes. *Immunogenetics.* 2010; 62:159–172. [PubMed: 20099063]
- Tseng CT, Miller D, Cassano J, Bailey E, Antczak DF. Identification of equine major histocompatibility complex haplotypes using polymorphic microsatellites. *Anim Genet.* 2010; 41(Suppl 2):150–153.
- Udesi ND, Compton PD, Shabanowitz J, Hunt DF, Rose KL. Methods for analyzing peptides and proteins on a chromatographic timescale by electron-transfer dissociation mass spectrometry. *Nat Protoc.* 2008; 3:1709–1717. [PubMed: 18927556]

Fig. 1.

Summary motifs. The main features of the Eqca-16*00101 (**a**) and Eqca-1*00201 (**b**) motifs are summarized. As determined using eluted ligands and PSCL binding analyses, the motifs are summarized to show the main (*red fill*) and secondary (*blue fill*) anchor positions. The most preferred residues at the primary anchor positions are indicated by *enlarged font*. Tolerated anchor residues are also shown. The most preferred residues at secondary anchor positions are also shown

Table 1

Binding capacity of eluted ligands

Peptide	Len	IC50 nM
Ecqa-16*00101		
RLYPEGLAQL	10	0.20
SQYNGNLQDL	10	0.60
RVSDSVGFL	9	0.99
SVLKGLEKM	9	1.1
AVIGDVIRV	9	1.6
ETTGQLTKI	9	1.6
EQIRNQLTAM	10	1.7
EVAEHVQYM	9	2.4
EVVRQLQEI	9	2.5
EVVQKLTQM	9	3.0
GVIKVFNDM	9	3.1
DVKEQIYKL	9	3.2
EVTDHNHLEL	10	4.1
HVLPELNKM	9	6.7
EVEESINRL	9	7.4
GVYGDVQRV	9	8.3
AIVDKVPSV	9	8.6
DVIAQGVGKL	10	8.9
GTIDEIQKL	9	13
KIYEGQVEV	9	13
EVLEV MQHM	9	13
KVPDDIYKT	9	13
TMPKDIQL	9	14
SVVRDIQNT	9	14
SSLEKSYEL	9	16
EIKDFLLTA	9	18
TTIEAIHGL	9	27
EIKDIHNL	9	28
EVQNEIDRL	9	34
EMHNKIYSI	9	43
SVLGKVNEI	9	45
AVKDTILQL	9	50
DVANKIGII	9	52
EIQDKLSEV	9	86
EIPDDLKQL	9	203
IQKESTLHL	9	242
EAVKIINEL	9	277
AVHEQLAAL	9	321

Peptide	Len	IC50 nM
KVIDGLET	9	804
ETQKSIYYI	9	2552
Eqca-1*00201		
NRDKHSAVAYF	11	0.83
NRVAEFHTEL	10	1.3
NRMPQPPLSM	9	1.3
NRLNDGATAL	11	1.6
NRQPYAVSEL	10	3.0
MRVNDSTML	9	3.7
NRMADAIL	9	3.8
NRVEVGKDQEF	11	5.1
NRDKNSIYSG	11	5.2
NRTPTVTL	9	6.9
NRAPNTAEL	9	7.4
NRDRSWTEL	9	7.6
NRDLQADNQRL	11	9.5
MRPLDAVEL	9	10
NRVDPNGSRYL	11	10
NRMQAGEIGEM	11	10
NRDAAARLTI	10	10
MRANNPEQNRL	11	11
NRDKYIKDDF	10	13
NRLVQTAE	9	13
NRFSSYTRL	9	15
SRPLSAVEL	9	16
YRRKDGVFL	9	16
MRDSVLTHL	9	16
NRIANPDQL	9	20
NRTAPVSV	9	23
TRKPMTTKDLL	11	31
HRIPTNDEL	9	37
NRVTALTQM	9	49
NRILTDGL	9	53
NRIQHPSNVLH	11	67
NRRELTDKL	9	96
NRMITQTSI	9	107
NRDVAKSVL	9	145
NRVSQQNKINL	11	151
SRDPTVTMV	9	272
GYKAGMTHI	9	273
NRQVQHILA	9	356
NRDVPNETL	9	438

Peptide	Len	IC50 nM
NRGEVVTTF	9	950

Author Manuscript

Author Manuscript

Author Manuscript

Author Manuscript

Table 2

PSCL-derived quantitative peptide-binding motif for EpcA-16*00101

Position									
Residue	1	2	3	4	5	6	7	8	9
A	0.160	<i>0.161</i>	0.093	0.287	0.929	0.078	0.515	0.518	0.052
C	0.090	0.011	0.074	0.248	0.360	0.157	0.235	0.401	0.017
D	0.198	0.009	0.019	0.761	0.352	0.043	0.196	0.952	0.003
E	<i>0.697</i>	0.010	0.018	1.000	0.417	0.043	0.173	1.000	0.005
F	0.062	0.013	0.130	0.314	0.706	0.057	1.000	0.423	<i>1.000</i>
G	0.098	0.011	0.027	0.313	0.506	0.137	0.974	0.903	0.029
H	0.179	0.009	0.050	0.200	0.712	0.058	0.820	0.415	0.007
I	0.100	<i>0.144</i>	<i>0.769</i>	0.264	0.287	<i>1.000</i>	0.324	0.328	<i>0.469</i>
K	0.075	0.011	0.127	0.227	0.475	0.084	0.064	0.336	0.010
L	0.067	<i>0.166</i>	<i>1.000</i>	0.170	0.402	<i>0.515</i>	0.366	0.571	<i>0.402</i>
M	<i>0.355</i>	0.071	<i>0.702</i>	0.397	0.516	0.136	0.300	0.294	0.026
N	0.133	0.009	0.049	0.362	1.000	0.098	0.328	0.493	0.005
P	0.013	0.009	0.085	0.332	0.291	0.044	0.863	0.835	0.003
Q	<i>0.420</i>	<i>0.302</i>	0.033	0.274	0.359	0.068	0.350	0.250	0.003
R	<i>0.310</i>	0.015	0.014	0.325	0.539	0.085	0.383	0.182	0.002
S	<i>1.000</i>	0.039	0.018	0.856	0.672	0.088	0.217	0.429	0.004
T	0.137	<i>0.153</i>	0.017	0.581	0.219	0.064	0.093	0.479	0.039
V	0.087	<i>1.000</i>	0.110	0.287	0.216	0.151	0.197	0.318	<i>0.476</i>
W	0.017	0.024	0.026	0.309	0.310	0.039	0.200	0.260	0.024
Y	0.048	0.010	0.052	0.204	0.272	0.179	0.387	0.220	0.006
Geomean	0.12	0.03	0.07	0.34	0.43	0.10	0.31	0.43	0.02
SD	3.0	4.4	3.7	1.6	1.6	2.3	2.1	1.6	6.9
SF	1.05	3.94	1.97	0.38	0.30	1.31	0.42	0.31	6.39
10-fold	10	14	14	0	0	13	2	0	16

The PSCL was tested for binding, the data analyzed, and primary and secondary anchor positions defined, as described in the text. Values shown represent the average relative binding (ARB) of the corresponding library relative to other pools with the same fixed position. Values have been normalized to the optimal residue at the corresponding position. Geomean is the average ARB across all residues at the corresponding position. SD indicates the standard deviation between the ARB of pools at the same position. SF is the specificity factor, calculated as described in the text, representing the ratio of the average binding of the entire library to the average of pools at the indicated position. At the primary anchor positions (SF > 2.4), the most preferred residues, associated with an ARB > 0.1, are highlighted by

italicized font. 10-fold indicates the number of residues at each position associated with ARB 10-fold or more lower than the optimal residue at the same position. The library average binding for Eqca-16*00101 was 9844 nM

Author Manuscript

Author Manuscript

Author Manuscript

Author Manuscript

Table 3

PSCL-derived quantitative peptide-binding motif for Eqca-1*00201

Position									
Residue	1	2	3	4	5	6	7	8	9
A	0.039	0.018	0.172	0.348	0.450	0.391	0.554	0.570	0.083
C	0.154	0.026	0.265	0.298	0.304	0.230	0.418	0.331	0.077
D	0.029	0.018	0.339	0.369	0.368	0.166	0.335	0.696	0.036
E	0.021	0.020	0.299	0.361	0.230	0.367	0.229	0.838	0.024
F	0.117	0.018	0.236	0.344	1.000	0.459	0.257	0.540	1.000
G	0.029	0.018	0.066	0.330	0.561	0.519	0.174	0.602	0.033
H	0.220	0.018	0.061	0.227	0.380	0.228	0.250	0.218	0.016
I	0.160	0.018	0.443	0.395	0.586	0.467	0.701	0.420	0.294
K	0.033	0.018	0.090	0.321	0.191	0.185	0.142	0.553	0.012
L	0.086	0.018	0.450	0.416	0.545	0.331	1.000	0.649	0.835
M	0.886	0.021	1.000	0.582	0.582	0.209	0.308	1.000	0.041
N	1.000	0.018	0.167	0.230	0.537	0.599	0.538	0.568	0.013
P	0.035	0.018	0.243	1.000	0.742	0.278	0.647	0.524	0.019
Q	0.043	0.018	0.175	0.817	0.341	0.190	0.352	0.373	0.019
R	0.068	1.000	0.266	0.567	0.214	0.324	0.232	0.260	0.016
S	0.091	0.036	0.061	0.421	0.515	0.386	0.272	0.465	0.019
T	0.024	0.018	0.061	0.383	0.418	0.227	0.320	0.509	0.026
V	0.018	0.018	0.135	0.608	0.411	0.262	0.535	0.297	0.643
W	0.054	0.018	0.128	0.298	0.419	0.285	0.165	0.366	0.096
Y	0.079	0.018	0.281	0.390	0.715	1.000	0.359	0.521	0.012
Geomean	0.07	0.02	0.19	0.40	0.44	0.32	0.34	0.48	0.05
SD	3.1	2.5	2.1	1.5	1.5	1.6	1.7	1.5	4.3
SF	2.38	7.46	0.93	0.43	0.40	0.55	0.51	0.36	3.50
10-fold	14	19	5	0	0	0	0	0	16

See the notes to Table 2 for details. The library average binding for Eqca-1*00201 was 25,144 nM