

A probabilistic meta-predictor for the MHC class II binding peptides

Oleksiy Karpenko · Lei Huang · Yang Dai

Received: 10 May 2007 / Accepted: 21 November 2007 / Published online: 19 December 2007
© Springer-Verlag 2007

Abstract Several computational methods for the prediction of major histocompatibility complex (MHC) class II binding peptides embodying different strengths and weaknesses have been developed. To provide reliable prediction, it is important to design a system that enables the integration of outcomes from various predictors. The construction of a meta-predictor of this type based on a probabilistic approach is introduced in this paper. The design permits the easy incorporation of results obtained from any number of individual predictors. It is demonstrated that this integrated method outperforms six state-of-the-art individual predictors based on computational studies using MHC class II peptides from 13 HLA alleles and three mouse MHC alleles obtained from the Immune Epitope Database and Analysis Resource. It is concluded that this integrative approach provides a clearly enhanced reliability of prediction. Moreover, this computational framework can be directly extended to MHC class I binding predictions.

Keywords MHC class II binding · Epitope prediction · Meta-predictor · PM predictor

Introduction

The identification of antigen peptides that bind to major histocompatibility complex (MHC) molecules plays a crucial role in understanding the mechanisms of both humoral and adaptive immunity as well as in developing epitope-based vaccines. Two major types of MHC molecules are involved in the peptide-binding process. MHC class I molecules present endogenous antigens to CD8⁺ cytotoxic T cells. MHC class II molecules, on the other hand, present exogenously derived proteins through antigen-presenting cells to CD4⁺ helper T cells (Parham 2005). Generally, antigen peptides that bind to MHC class I molecules are approximately nine amino acid residues long (Bleek and Nathenson 1991). However, the peptide-binding groove of a MHC class II molecule is open at both ends, a property that makes it capable of accommodating longer peptides consisting of 10–30 residues (Castellino et al. 1997; Max et al. 1993; Sette et al. 1989).

Experimental determinations of the binding affinities of peptides to MHC molecules are time consuming and expensive. Therefore, considerable effort has been made on the development of computational tools for the identification of MHC-binding peptides (De Groot and Berzofsky 2004; Doytchinova et al. 2003; Flower 2004; Flower and Doytchinova 2002; Flower et al. 2002; Martin et al. 2003; Schirle et al. 2001). A host of computational methods has been developed for MHC class I prediction over the last two decades (De Groot et al. 2002; Flower et al. 2003; Martin et al. 2003; Nussbaum et al. 2003; Reche et al. 2002; Schirle et al. 2001). A comprehensive list of references can be found in a recent paper (Peters et al. 2006). The number of alleles

OK implemented the programs of the integrative system and improved the efficiency of the method. LH prepared data and obtained the results from the methods of LP and the Gibbs Sampler. YD initiated the work, designed the general framework, and supervised the project. All authors have read and approved the final manuscript.

O. Karpenko · L. Huang · Y. Dai (✉)
Department of Bioengineering (MC063),
University of Illinois at Chicago,
851 South Morgan Street,
Chicago, IL 60607, USA
e-mail: yangdai@uic.edu

O. Karpenko
e-mail: okarpe2@uic.edu

L. Huang
e-mail: lhuang7@uic.edu

covered by these methods is large, and the level of accuracy is relatively high.

Conversely, the situation for the prediction of MHC class II binding peptides is quite different. The variability in the peptide length complicates the prediction of peptide–MHC class II binding. The analyses of the binding motif and the structure of peptide–MHC class II complexes have suggested that a core of nine residues within a peptide is essential for peptide–MHC binding. Computational methods for MHC class II binding prediction include simple binding motifs (Borras-Cuesta et al. 2000; Rammensee et al. 1999; Singh and Raghava 2001), quantitative matrices (Bui et al. 2005; Peters and Sette 2005; Sturniolo et al. 1999), hidden Markov models (Kato et al. 2003; Noguchi et al. 2002), artificial neural networks (Brusic et al. 1998; Burden and Winkler 2005; Nielsen et al. 2003), iterative discriminant analysis (Mallios 1998, 2001), support vector machines/regression (Bhasin and Raghava 2004; Donnes and Elofsson 2002; Donnes and Kohlbacher 2006; Liu et al. 2006; Salomon and Flower 2006), the Gibbs sampler and its extension (Nielsen et al. 2007, 2004), partial least squares (Chang et al. 2006; Doytchinova and Flower 2003; Hattotuwigama et al. 2006), and other methods (Altiparmak et al. 2006; Chang et al. 2007; Cui et al. 2006, 2007; Doytchinova and Flower 2001; Hertz and Yanover 2006; Karpenko et al. 2005; Murugan and Dai 2005; Takahashi and Honda 2006; Tong et al. 2006; Wan et al. 2006). Because each method has its own strengths and weaknesses, it is hard for an immunologist to select a single method from the pool of existing predictors. Therefore, a system that produces reliable prediction through the integration of outcomes from major predictors is in clear need.

A consensus strategy for combining three human leukocyte antigen (HLA)-DR binding algorithms—SYFPEITHI (Rammensee et al. 1999), ProPred (Singh and Raghava 2001), and the iterative stepwise discriminant analysis meta-algorithm (Mallios 2001)—has been shown to be consistently best or second best (Mallios 2003) using sets of binding peptides in DRB1*0101 and DRB1*0401. In another integrative system, MULTIPRED (Zhang et al. 2005), the individual predictive engines implemented are hidden Markov models (HMMs) and artificial neural networks (ANNs). The system covers predictions of HLA protein binding peptides belonging to supertypes A2 and A3 (HLA class I) as well as DR (HLA class II). Users can choose either HMMs or ANNs as individual predictors. In addition, the system provides a mechanism that makes consensus prediction by combining the results from the two prediction methods. Significantly, recent work (Moise and De Groot 2006; Moutaftsi et al. 2006) has demonstrated the promise of an integrative approach for the computational identification of peptides with immunogenicity through the prediction of binding affinity to MHC class I molecules. Specifically, the computational prediction re-

duced the number of possible overlapping peptides by more than 85-fold, accelerating the discovery of 49 epitopes that account for ~95% of the immunome in a mouse model for vaccine development. The key step in their approach is a consensus prediction that combines four matrix-based epitope prediction algorithms (BIMAS: http://thr.cit.nih.gov/molbio/hla_bind/; Bui et al. 2005; Peters and Sette 2005; Udaka et al. 2000). Another recent integrative method for the MHC class I binding prediction uses the sum of the weighted votes from each individual predictor as a combined score for a peptide to make an improved prediction (Trost et al. 2007).

In the present work, a meta-predictive (called probabilistic meta-predictor or PM predictor thereafter) system based on a probabilistic approach is described. This method improves significantly our earlier work of a Naïve Bayesian meta-predictor (Huang et al. 2006, 2007) to achieve fast training and higher performance through a consensus score that combines predictive scores from each individual predictor. Like the previous integrative predictors (Huang et al. 2006, 2007; Trost et al. 2007), the framework presented in this work has the flexibility to incorporate an arbitrary number of predictors that provide predictions based on computed score correlated with the binding affinity.

To illustrate the basic framework of our PM predictor, the MHC class II binding prediction was taken as an example, although this approach can also be applied to the MHC class I. Six individual predictors of MHC class II binding predictions were selected based on their availability from the Internet. They are SVRMHC (Wan et al. 2006), ARB (Bui et al. 2005), RANKPEP (Reche et al. 2004), ProPred (Singh and Raghava 2001), Gibbs Sampler (Nielsen et al. 2004), and the LP model (Murugan and Dai 2005). The output from each of the first three methods for a peptide is a predictive score of binding affinity. Each of the latter three methods provides an allele-restricted position-specific scoring matrix (PSSM) that can be used to compute scores of the overlapping 9-mers of a peptide. The maximum score over these 9-mers is considered as the score of the peptide. Taking these scores of training peptides for a specific allele as inputs, we first estimate the probability distributions of the scores for both binding and nonbinding peptides in the training set for each individual predictor. Then, we combine these distributions a probabilistic model to obtain the integrative predictor. The effectiveness of our model is examined with the use of MHC class II peptides from 13 HLA alleles and three mouse MHC alleles obtained from the Immune Epitope Database and Analysis Resource (IEDB; Peters et al. 2005). The computational analysis shows that the PM predictor uniformly produces stable prediction and in general achieves statistically improved results in comparison with any individual predictor.

Materials and methods

Data set

The computational experiments were conducted using the data set available from the IEDB database (Peters et al. 2005). This data set comprises peptide data with IC₅₀ binding affinities for the 13 HLA (human MHC) and three mouse MHC class II alleles. This data set was also used in the recent study (Nielsen et al. 2007) for quantitative prediction of MHC class II peptide binding. The details of the data set are provided in Table 1.

Choosing individual predictors

Any predictor that is capable of assigning a predictive score to a peptide sequence can be employed as an individual predictor in our system. The six methods listed below were selected in this study. The coverage of their predictions is summarized in Table 2.

ARB predictions

The ARB predictions were obtained using a default parameter setting for the ARB web server (http://tools.immuneepitope.org/tools/matrix/iedb_input?matrixClass=II). Each peptide is assigned a predictive score.

Table 1 The data set used in this study. Peptide data for the 13 HLA-DR and 3 mouse H2-IA alleles are downloaded from <http://www.cbs.dtu.dk/suppl/immunology/NetMHCII/php>

Alleles	IEDB	Number of binders	Number of nonbinders
DRB1*0101	1,203	921	282
DRB1*0301	474	67	407
DRB1*0401	457	210	247
DRB1*0404	168	74	94
DRB1*0405	171	88	83
DRB1*0701	310	127	183
DRB1*0802	174	58	116
DRB1*0901	117	47	70
DRB1*1101	359	95	264
DRB1*1302	179	101	78
DRB1*1501	365	188	177
DRB4*0101	181	74	107
DRB5*0101	343	113	230
H2-IAb	76	43	33
H2-IAd	342	56	286
H2-IAs	126	35	91

These peptides are extracted from the IEDB database in November 2006. Binding peptides were identified with IC₅₀ binding threshold of 500 nM.

SVRMHC predictions

The SVRMHC (Wan et al. 2006) predictions were obtained using a default parameter setting for the SVRMHC web server (<http://svrmhc.umn.edu/SVRMHCdb>). This is a support vector machine regression-based method that makes predictions of the exact binding affinity of the peptide. The server returns pIC₅₀ prediction scores for each 9-mer within the query peptide, and the maximum score was assigned as the binding pIC₅₀ prediction value for the query peptide.

RANKPEP predictions

The RANKPEP (Reche et al. 2004) predictions were made by submitting peptides to the web server (<http://bio.dfci.harvard.edu/Tools/rankpep.html>) with default parameters. This method predicts binding peptides based on the scores calculated from a PSSM. The PSSMs are not available publicly, but the server returns a predictive score for each peptide.

Gibbs Sampler predictions

The PSSMs were obtained by submitting the binding peptides with default parameter settings to the web server (<http://www.cbs.dtu.dk/biotools/EasyGibbs/>). Gibbs Sampler (Nielsen et al. 2004) is an advanced motif sampler method based on the Gibbs sampling technique, which efficiently samples the possible alignment space of binder sequences. For each alignment, a log-odds weight matrix is calculated for the identified binding core subsequences. This matrix serves as the PSSM for the computation of a score for a 9-mer. It should be noted that Gibbs Sampler requires only binding peptides (with IC₅₀ < 500 nM) for the construction of PSSMs.

ProPred predictions

The PSSMs of ProPred (Singh and Raghava 2001) were obtained from its website (<http://www.imtech.res.in/raghava/propred/page4.html>). This predictor uses the quantitative matrices from 51 HLA-DR alleles for the prediction of MHC class II binding peptides. These matrices were generated from a pocket profile database previously described (Sturniolo et al. 1999) and covered the majority of human HLA-DR specificity. The matrices are the same as the ones in TEPITOPE (Sturniolo et al. 1999).

LP-top2 predictors

The LP method (Murugan and Dai 2005) was motivated by a text mining model. This LP-based iterative learning model enables the use of both binding and nonbinding peptides for the detection of the binding cores from a set of putative binding cores and for the construction of the

Table 2 The coverage of predictions for the six methods

Alleles	ARB	SVRMHC	RANKPEP	ProPred	Gibbs Sampler	LP-top2
DRB1*0101	*	*	*	*	*	*
DRB1*0301	*		*	*	*	*
DRB1*0401	*	*	*	*	*	*
DRB1*0404	*		*	*	*	*
DRB1*0405	*	*	*	*	*	*
DRB1*0701	*		*	*	*	*
DRB1*0802	*			*	*	*
DRB1*0901	*		*		*	*
DRB1*1101	*		*	*	*	*
DRB1*1302	*		*	*	*	*
DRB1*1501	*	*	*	*	*	*
DRB4*0101	*		*		*	*
DRB5*0101	*	*	*	*	*	*
H2-IAb	*		*		*	*
H2-IAd	*		*		*	*
H2-IAs					*	*

The empty entry means that the method cannot make prediction for the particular allele.

predictor simultaneously. The outcome of this predictor is a PSSM that can be used to score a 9-mer. The PSSMs were obtained by training the binding and non-binding peptides of each allele with the algorithm (Murugan and Dai 2005). Binding peptides were identified with IC50 binding threshold of 500 nM. The LP-top2 was selected for this study among several variants of the LP method because of its superior performance.

In summary, each of the former three predictors discussed above returns a predictive score that corresponds to the actual binding affinity for a peptide, while each of the three latter predictors returns a PSSM of size 20 by 9 for a set of training peptides of a specific allele. This PSSM will be used to calculate the score of each amino acid at each position of a 9-mer. The final score of a peptide is the maximum score over all overlapping 9-mers in the peptide. The scores derived from the latter three methods are not the actual binding affinity of a peptide; however, their magnitudes should correlate with the strength of the binding.

Several online predictors were not included in our study for various reasons. The current version of SVMHC (Donnes and Elofsson 2002; Donnes and Kohlbacher 2006) also makes MHC class II binding predictions. However, it uses the same matrices published in TEPITOPE (Sturniolo et al. 1999). Because those matrices were also used in ProPred, SVMHC was not selected. MHCpred (Doytchinova and Flower 2003) has an online prediction service; however, it covers only three alleles in our data set. Therefore, it was not included. The web server of MULTIPRED (Zhang et al. 2005) predicts eight HLA-DR variants. Because it only covers five alleles in our data set, it was not selected. For the other methods mentioned in the “Introduction,” the exclusion from this study was mainly due to lack of access to either the online predictors or the programs.

Recently, a new online MHC class II binding predictor was released (Nielsen et al. 2007). This method uses a novel stabilization matrix alignment method that allows for direct prediction of binding affinity. Comprehensive computational study has shown that it outperformed the other state-of-art MHC class II quantitative prediction methods. Ideally, it would be better to include this method as an individual predictor in our current study. However, because the goal of this work is the demonstration of the effectiveness of the integrative system and is not aimed at the best individual predictor, the exclusion of the above predictor does not affect the main results from this study.

The PM predictor

The PM predictor is based on a probabilistic model that combines prediction scores of a peptide from each predictor into a consensus score. The consensus score depends on the probability distribution of scores. A threshold has to be determined so that peptides with consensus scores above this threshold are predicted as binding and peptides with consensus scores below this threshold are correspondingly predicted as nonbinding. Figure 1 illustrates the framework of the PM predictor method. The details of this predictor are described as follows.

Calculation of the consensus score

Given a peptide, predictive scores are assigned by each of the m predictors. If there are n peptides in a test set, then the total number of scores from all individual predictors is nm . To consolidate the results from the m predictors, a consensus score is defined for a peptide. This consensus score provides the likelihood of a peptide to be classified into the binding class or the nonbinding class according to

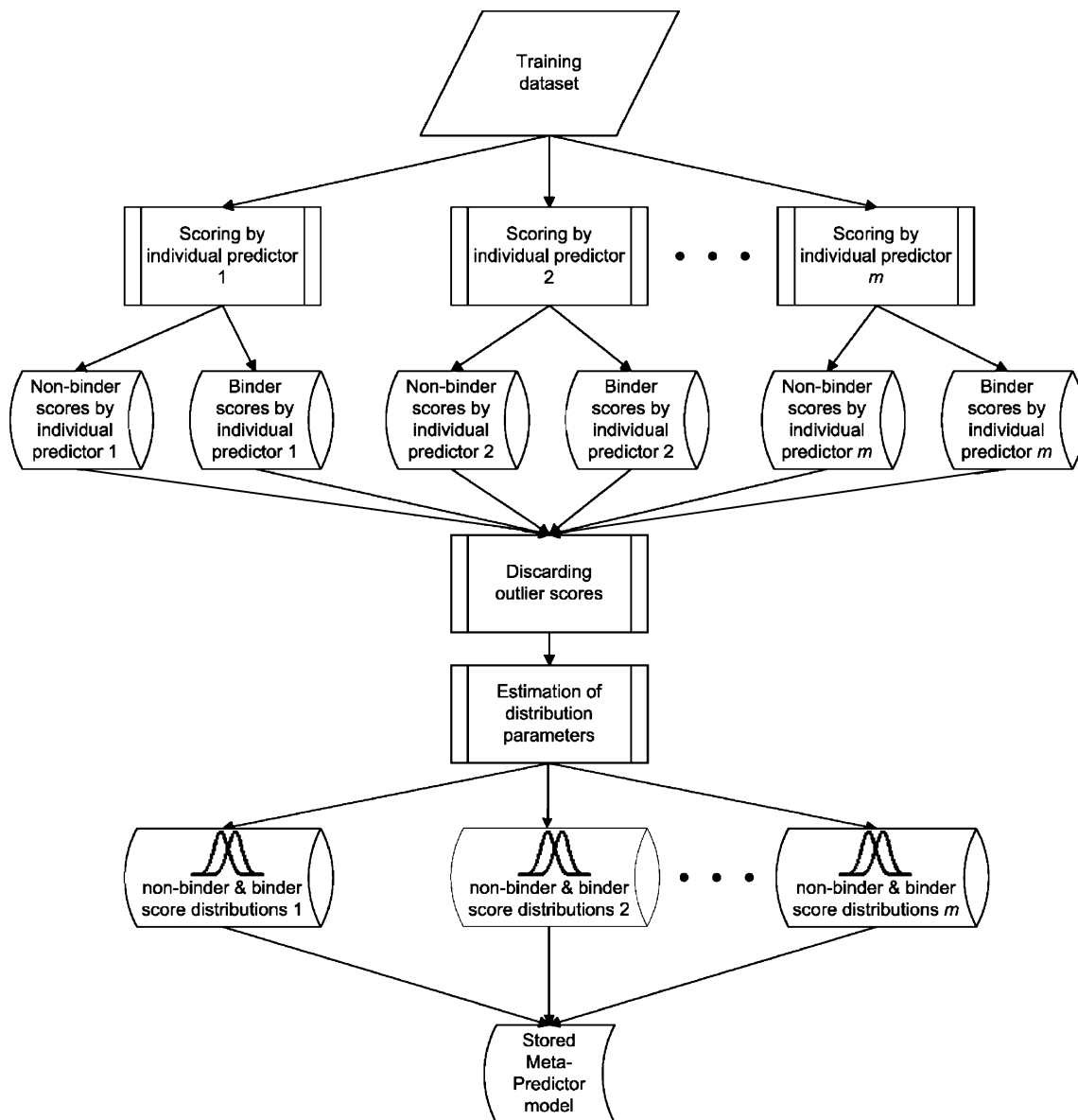


Fig. 1 Illustration of the framework for building the PM predictor

the information obtained from the probability distribution of scores for each individual predictor. The consensus score over all predictors is defined as the product of the likelihoods. If the estimations of the probability distributions of the scores for binding and nonbinding peptides are accurate, then the consensus scores of the binding peptides and nonbinding peptides should be grouped into two distinct intervals. This grouping would allow for a simple prediction by using a prescribed threshold of the consensus score. The details of the calculation of the consensus scores are presented in the [Appendix](#).

Estimating distributions of scores

In this study, an important assumption was made about the distributions of scores obtained from each predictor: Scores

from binding peptides and nonbinding peptides, respectively, follow normal distributions, respectively, with distinct means. The lowest and the highest 2.5% of the binder and nonbinder scores were dropped to exclude the influence of outliers on the estimation of the distribution parameters. Only the remaining 95% of the scores was used in the rest of the training procedure. With this assumption on the distributions, the estimate of the distribution parameters is straightforward. For each individual predictor, we calculated: (1) the mean and the standard deviation of scores for the binding peptides and (2) the mean and the standard deviation of scores for the nonbinding peptides. These parameters characterize the score distributions for each individual predictor.

The area under the receiver operating characteristic curve (Aroc; Swets 1988) over five different training and

test sets (Nielsen et al. 2007) was used for the evaluation of the PM predictor. More specifically, the probabilities were first determined from peptides in the training set, and then the consensus score was computed for each peptide in the test set. The Aroc value was subsequently calculated for the test set. This procedure was iterated through all five different training and test sets, and the average Aroc value was computed.

The majority voting algorithm

To compare the performance of different meta-predictors, the Majority Voting algorithm was also implemented. If a peptide is scored above a specified threshold σ_i by the i th predictor, the predictor casts a vote for that peptide to be binding; otherwise, the peptide is voted to be nonbinding. Once all m votes have been cast, the prediction is decided by the majority of the votes. Obviously, if m is even, a rule for breaking the tie is needed.

The Majority Voting algorithm utilized percentiles of peptides' scores in the training fold as thresholds for individual predictors for testing. For each of the m predictors, the scores of peptides in the four training folds were sorted. The percentiles of the sorted scores were then used as the thresholds. Each of the top z th percentile vectors, $\sigma = (\sigma_1, \dots, \sigma_m)$ yielded predictions for peptides in the test fold based on the majority votes. By varying z , an Aroc value can be calculated for the test fold. The average

Aroc value calculated from the different five training and test folds was reported.

The determinations of the best thresholds for δ in the PM-model and the z th percentile vectors $\sigma = (\sigma_1, \dots, \sigma_m)$ in the Majority Voting algorithm depend on the selected criterion. When an appropriate criterion is chosen, they can be optimized through a cross-validation procedure. For example, one may prefer a threshold that produces approximately equal sensitivity and specificity.

Results and discussion

The performance of the PM predictor was compared to those of the Majority Voting algorithm and the six individual predictors using the data set described above. We conducted 1,000 bootstrapping experiments for the Majority Voting algorithm and the PM predictor.

Table 3 summarizes the results obtained. In all cases, the performance was evaluated in terms of the Aroc value. The PM predictor demonstrated a higher accuracy than that of the Majority Voting algorithm for 14 out of 16 alleles used in the computational experiments. The only exceptions were the two mouse alleles H2-IAd and H2-IAs, for which the accuracies of the PM predictor are slightly lower than those of the Majority Voting algorithm. The average Aroc value (0.949) of the PM predictor over all tested alleles is slightly higher than that (0.936) given by the Majority

Table 3 Summary of the Aroc values for the six individual predictors and the two integrative methods

MHC allele	ARB	SVRMHC	RANKPEP	ProPred	Gibbs Sampler ^a	LP-top2 ^a	Majority Voting	PM predictor
DRB1*0101	0.660	0.613	0.620	0.648	0.657	0.781	0.824 (0.009)	<i>0.836 (0.007)</i>
DRB1*0301	0.787		0.755	0.723	0.825	0.893	0.947 (0.009)	<i>0.955 (0.007)</i>
DRB1*0401	0.740	0.739	0.723	0.754	0.834	0.778	0.924 (0.008)	<i>0.948 (0.006)</i>
DRB1*0404	0.786		0.640	0.827	0.898	0.994	0.976 (0.009)	<i>0.991 (0.005)</i>
DRB1*0405	0.724	0.685	0.661	0.788	0.803	0.935	0.940 (0.013)	<i>0.968 (0.008)</i>
DRB1*0701	0.742		0.672	0.761	0.826	0.862	0.923 (0.010)	<i>0.943 (0.007)</i>
DRB1*0802	0.802			0.768	0.922	0.986	0.977 (0.018)	<i>0.990 (0.008)</i>
DRB1*0901	0.710		0.690		0.921	0.941	0.947 (0.018)	<i>0.957 (0.008)</i>
DRB1*1101	0.728		0.650	0.708	0.885	0.878	0.932 (0.010)	<i>0.941 (0.007)</i>
DRB1*1302	0.916		0.577	0.720	0.851	0.962	0.961 (0.010)	<i>0.972 (0.007)</i>
DRB1*1501	0.792	0.735	0.641	0.724	0.810	0.868	0.921 (0.010)	<i>0.950 (0.006)</i>
DRB4*0101	0.803		0.523		0.911	0.975	0.945 (0.012)	<i>0.970 (0.007)</i>
DRB5*0101	0.681	0.620	0.639	0.660	0.805	0.913	0.882 (0.011)	<i>0.921 (0.008)</i>
H2-IAb	0.663		0.902		0.933	0.945	0.971 (0.019)	<i>0.976 (0.011)</i>
H2-IAd	0.818		0.618		0.846	0.942	0.943 (0.012)	0.910 (0.009)
H2-IAs					0.889	0.956	0.957 (0.012)	0.956 (0.007)
Average	0.757	0.678	0.665	0.735	0.851	0.913	0.936 (0.012)	0.949 (0.074)

The Aroc values for the Majority Voting algorithm and the PM predictor are the average of Aroc values computed from the 1,000 bootstrapping experiments. The standard deviations are provided in parentheses. The italicized numbers represent the better performance between the two methods: the Majority Voting and the PM predictor.

Aroc The area under the receiver operating characteristic curve, *PSSM* position-specific scoring matrix

^a The Arocs for the Gibbs Sampler and LP-top2 were calculated from the PSSMs trained by using the entire training set.

Voting algorithm (p value 0.002 for the one-tailed t test). The average standard deviation (0.007) of the Aroc values in the 1,000 bootstrapping experiments for the PM predictor, over all tested alleles, was lower than that (0.011) of the Majority Voting algorithm (p value 3.00×10^{-6} for the one-tailed t test), indicating a greater robustness in the prediction of the former.

The PM predictor outperformed every individual predictor for 12 out of 16 alleles. For the two alleles DRB1*0404 and DRB4*0101, the Aroc values of the PM predictor are comparable with those of LP-top2 (0.991 vs 0.994 and 0.970 vs 0.975, respectively); for H2-IAd, the Aroc value is lower than that of LP-top2 (0.910 vs 0.945), and for H2-IAs, the Aroc value is identical with that of the LP-top2.

The average Aroc value (0.949) of the PM predictor over all tested alleles is higher than those of the individual predictors: ARB (0.757, p value 1.75×10^{-9}), SVRMHC (0.678, p value 1.00×10^{-4}), RANKPEP (0.665, p value 1.57×10^{-9}), ProPred (0.735, p value 8.09×10^{-10}), Gibbs Sampler (0.851, p value 2.19×10^{-5}), and LP-top2 (0.913, p value 3.47×10^{-2}). All p values were derived from one-tailed t tests.

Because most of the online predictors do not allow for training of a new model based on training data submitted by the users, we did not obtain the Aroc values for individual predictors in a cross-validated fashion. The PSSMs of Gibbs Sampler and LP-top2 were obtained by training the entire data set only once. Therefore, the Aroc values of the Gibbs Sampler and the LP-top2 actually represent the training performance. Similarly, the ARB models (Bui et al. 2005)

were trained using the quantitative binding data contained within the IEDB database. For this reason, higher Aroc values for the Gibbs Sampler, LP-top2, and ARB predictions were obtained. On the other hand, SVRMHC was trained on relatively small sets of quantitative peptide binding data contained within the AntiJen database (Toseland et al. 2005), and the performance could probably be improved, if it were retrained on the data used here. The PSSMs for ProPred and the predictions of RANKPEP were obtained directly from the websites. The peptides used for training are not available. Therefore, the performance presented in the current study for the individual predictors should not be compared. The rigorous comparison for the state-of-art methods can be found elsewhere (Nielsen et al. 2007).

As the goal for our study is the construction of an integrative system that outperforms individual predictors, we accordingly compare first the outcome of an underperformed predictor to the integrative system. The MHC-BPS (Cui et al. 2006, 2007) was not included as an individual predictor in our system because of its low performance. However, we used this predictor to investigate how such a predictor would affect the performance of the integrative systems. The results were summarized in Table 4. The web server MHC-BPS (<http://bidd.cz3.nus.edu.sg/mhc/>) covers DRB1*0101, DRB1*0404, DRB1*0701, DRB1*0901, DRB1*1101, DRB1*1501, and DRB5*0101. The Aroc values are 0.470, 0.550, 0.555, 0.617, 0.562, 0.617, and 0.594, respectively, which yields an average Aroc value 0.566, a fairly low figure compared to other predictors

Table 4 Summary of the Aroc values for the seven individual predictors and the two integrative methods

MHC allele	ARB	SVRMHC	MHC-BPS	RANKPEP	ProPred	Gibbs Sampler ^a	LP-top2 ^a	Majority Voting	PM predictor
DRB1*0101	0.660	0.613	0.470	0.620	0.648	0.657	0.781	0.769 (0.009)	<i>0.808 (0.007)</i>
DRB1*0301	0.787			0.755	0.723	0.825	0.893	0.947 (0.009)	<i>0.955 (0.007)</i>
DRB1*0401	0.740	0.739		0.723	0.754	0.834	0.778	0.924 (0.008)	<i>0.948 (0.006)</i>
DRB1*0404	0.786		0.550	0.640	0.827	0.898	0.994	0.962 (0.011)	<i>0.990 (0.005)</i>
DRB1*0405	0.724	0.685		0.661	0.788	0.803	0.935	0.939 (0.013)	<i>0.968 (0.008)</i>
DRB1*0701	0.742		0.555	0.672	0.761	0.826	0.862	0.915 (0.010)	<i>0.930 (0.008)</i>
DRB1*0802	0.802				0.768	0.922	0.986	0.977 (0.018)	<i>0.990 (0.009)</i>
DRB1*0901	0.710		0.617	0.690		0.921	0.941	0.940 (0.018)	<i>0.940 (0.009)</i>
DRB1*1101	0.728		0.562	0.650	0.708	0.885	0.878	0.920 (0.011)	<i>0.916 (0.008)</i>
DRB1*1302	0.916			0.577	0.720	0.851	0.962	0.961 (0.010)	<i>0.972 (0.007)</i>
DRB1*1501	0.792	0.735	0.617	0.641	0.724	0.810	0.868	0.924 (0.009)	<i>0.949 (0.006)</i>
DRB4*0101	0.803			0.523		0.911	0.975	0.945 (0.012)	<i>0.970 (0.007)</i>
DRB5*0101	0.681	0.620	0.594	0.639	0.660	0.805	0.913	0.887 (0.011)	<i>0.914 (0.008)</i>
H2-IAb	0.663			0.902		0.933	0.945	0.971 (0.019)	<i>0.976 (0.011)</i>
H2-IAd	0.818			0.618		0.846	0.942	<i>0.943 (0.012)</i>	0.910 (0.009)
H2-IAs						0.889	0.956	<i>0.957 (0.012)</i>	0.956 (0.007)
Average	0.757	0.678	0.566	0.665	0.735	0.851	0.913	0.930 (0.012)	0.943 (0.008)

The Aroc values for the Majority Voting algorithm and the PM predictor are the average of Aroc values computed from the 1,000 bootstrapping experiments. The standard deviations are provided in parentheses. The italicized numbers represent the better performance between the two methods: the Majority Voting and the PM predictor. The Arocs for Gibbs Sampler and LP-top2 were calculated from the PSSMs trained by using the entire training set.

Aroc The area under the receiver operating characteristic curve, *PSSM* position-specific scoring matrix

Table 5 Summary of the Aroc values for the five individual predictors and the two integrative methods

MHC allele	ARB	SVRMHC	MHC-BPS	RANKPEP	ProPred	Majority Voting	PM predictor
DRB1*0101	0.660	0.613	0.470	0.620	0.648	0.676 (0.009)	<i>0.700</i> (0.007)
DRB1*0301	0.787			0.755	0.723	0.846 (0.013)	<i>0.850</i> (0.010)
DRB1*0401	0.740	0.739		0.723	0.754	0.849 (0.010)	<i>0.877</i> (0.008)
DRB1*0404	0.786		0.550	0.640	0.827	<i>0.828</i> (0.019)	0.815 (0.015)
DRB1*0405	0.724	0.685		0.661	0.788	0.821 (0.019)	<i>0.863</i> (0.014)
DRB1*0701	0.742		0.555	0.672	0.761	0.791 (0.013)	<i>0.809</i> (0.010)
DRB1*0802	0.802				0.768	0.819 (0.015)	<i>0.832</i> (0.011)
DRB1*0901	0.710		0.617	0.690		0.659 (0.029)	<i>0.782</i> (0.019)
DRB1*1101	0.728		0.562	0.650	0.708	<i>0.778</i> (0.015)	0.729 (0.010)
DRB1*1302	<i>0.916</i>			0.577	0.720	<i>0.841</i> (0.017)	0.784 (0.015)
DRB1*1501	0.792	0.735	0.617	0.641	0.724	0.836 (0.012)	<i>0.848</i> (0.010)
DRB4*0101	<i>0.803</i>			0.523		0.655 (0.015)	0.652 (0.012)
DRB5*0101	0.681	0.620	0.594	0.639	0.660	0.740 (0.014)	<i>0.762</i> (0.012)
H2-IAb	0.663			<i>0.902</i>		0.758 (0.030)	<i>0.900</i> (0.017)
H2-IAd	<i>0.818</i>			0.618		0.481 (0.011)	<i>0.712</i> (0.009)
Average	0.757	0.678	0.566	0.665	0.735	0.759	0.794

The Aroc values for the Majority Voting algorithm and the PM-predictor are the average of Aroc values computed from the 1,000 bootstrapping experiments. The standard deviations are provided in parentheses. The italicized numbers represent the better performance between the two methods: the Majority Voting and the PM predictor. The Arocs for Gibbs Sampler and LP-top2 were calculated from the PSSMs trained by using the entire training set.

Aroc The area under the receiver operating characteristic curve, *PSSM* position-specific scoring matrix

used in this work. The addition of MHC-BPS to the existing six predictors included in the integrative systems resulted in the Aroc values 0.930 and 0.943 for the Majority Voting algorithm and the PM predictor, respectively, which are moderately smaller than the corresponding Aroc values 0.936 and 0.949 for the same two integrative predictors without the use of MHC-BPS.

This result implies that the performance of the integrative systems may not be affected drastically by an individual predictor with a relatively low performance. Similar behavior is also observed in Tables 5 and 6, when the Gibbs Sampler and LP-top2 were removed from the system.

We now investigate whether the improved performances of the two integrative predictors were driven by the

Table 6 Summary of the Aroc values for the four individual predictors and the two integrative methods

MHC allele	ARB	SVRMHC	RANKPEP	ProPred	Majority Voting	PM predictor
DRB1*0101	0.660	0.613	0.620	0.648	<i>0.743</i> (0.010)	0.735 (0.007)
DRB1*0301	0.787		0.755	0.723	0.846 (0.014)	<i>0.851</i> (0.009)
DRB1*0401	0.740	0.739	0.723	0.754	0.849 (0.010)	<i>0.877</i> (0.008)
DRB1*0404	0.786		0.640	0.827	<i>0.851</i> (0.017)	0.841 (0.013)
DRB1*0405	0.724	0.685	0.661	0.788	0.820 (0.019)	<i>0.862</i> (0.014))
DRB1*0701	0.742		0.672	0.761	0.813 (0.013)	<i>0.833</i> (0.010)
DRB1*0802	0.802			0.768	0.820 (0.015)	<i>0.832</i> (0.011)
DRB1*0901	0.710		0.690		0.625 (0.022)	<i>0.747</i> (0.016)
DRB1*1101	0.728		0.650	0.708	<i>0.779</i> (0.013)	0.761 (0.009)
DRB1*1302	<i>0.916</i>		0.577	0.720	<i>0.841</i> (0.018)	0.784 (0.014)
DRB1*1501	0.792	0.735	0.641	0.724	0.815 (0.012)	<i>0.840</i> (0.010)
DRB4*0101	<i>0.803</i>		0.523		<i>0.657</i> (0.015)	0.652 (0.012)
DRB5*0101	0.681	0.620	0.639	0.660	0.723 (0.014)	<i>0.757</i> (0.011)
H2-IAb	0.663		<i>0.902</i>		0.757 (0.033)	<i>0.900</i> (0.017)
H2-IAd	<i>0.818</i>		0.618		0.480 (0.010)	<i>0.712</i> (0.009)
Average	0.757	0.678	0.665	0.735	0.761	0.799

The Aroc values for the Majority Voting algorithm and the PM predictor are the average of Aroc values computed from the 1,000 bootstrapping experiments. The standard deviations are provided in parentheses. The italicized numbers represent the better performance between the two methods: the Majority Voting and the PM predictor. The Arocs for Gibbs Sampler and LP-top2 were calculated from the PSSMs trained by using the entire training set.

Aroc The area under the receiver operating characteristic curve, *PSSM* position-specific scoring matrix

overfitted Gibbs Sampler and LP-top2. Accordingly, we excluded these two predictors and evaluate the performances of the Majority Voting algorithm and the PM predictor. Similarly, we considered two cases: with and without including MHC-BPS in the integrative systems. The results are summarized in Tables 5 and 6. The average Aroc value (0.799) of the PM predictor is slightly higher than that (0.761) of the Majority Voting algorithm with a p value of 0.128. The average Aroc value (0.799) of the PM predictor is higher than those of ARB (0.757 with p value 5.3×10^{-2}), SVRMHC (0.678 with p value 2.9×10^{-3}), RANKPEP (0.665 with p value 6.9×10^{-5}), and ProPed (0.735 with p value 6.93×10^{-3}). With the inclusion of MHC-BPS, the Aroc values for both of the integrative systems went very slightly down, from 0.799 to 0.794 for the PM predictor and from 0.761 to 0.759 for the Majority Voting algorithm. However, we still observed a higher average Aroc value for the PM predictor compared to those of the individual predictors. More precisely, the average Aroc value (0.794) of the PM predictor is higher than those of ARB (0.757 with p value 7.6×10^{-2}), SVRMHC (0.678 with p value 3.8×10^{-3}), MHC-BPS (0.566 with p value 1.2×10^{-7}), RANKPEP (0.665 with p value 1.2×10^{-4}), and ProPed (0.735 with p value 1.2×10^{-2}). Therefore, we conclude, from the results in Tables 3 to 6, that the PM predictor is reliable and consistently performs better than every individual predictor included in the system. Although the Majority Voting algorithm has a slightly lower performance than that of the PM predictor, a similar conclusion holds.

Comparison of the meta-prediction method proposed recently by our group (Huang et al. 2006, 2007) was not included in this study. The reason for that is the demanding training time required for the method when a relatively large number of individual methods are considered. However, a preliminary study using the three methods of ProPred, Gibbs Sampler, and LP-top2 has indicated an inferior performance compared to the PM predictor. Therefore, we conclude that the previous meta-predictor is not as competitive as the PM predictor when a large number of predictors are used in the system. As mentioned above, there are two algorithms (Mallios 2003; Zhang et al. 2005) making consensus predictions for MHC class II binding. However, the Maillios algorithm is not available, and the prediction provided by the website of Zhang et al. only covers six HLA-DR alleles. Accordingly, we did not compare our model with these two algorithms. In addition, we also did not include in this study the recently published method NetMHCII (Nielsen et al. 2007). It is not hard to speculate that this inclusion will improve the performance of both integrative predictors, as NetMHCII was shown to outperform the other state-of-the-art MHC class II prediction methods.

When building the PM predictor, we made the assumption of a normal distribution for the scores obtained from the individual predictors. This assumption greatly simplified the estimate of the parameters of the probability distributions of the predicted scores. However, this assumption may not be valid for the scores of some alleles. In such cases, the PM predictor may exhibit diminished performance. This limita-

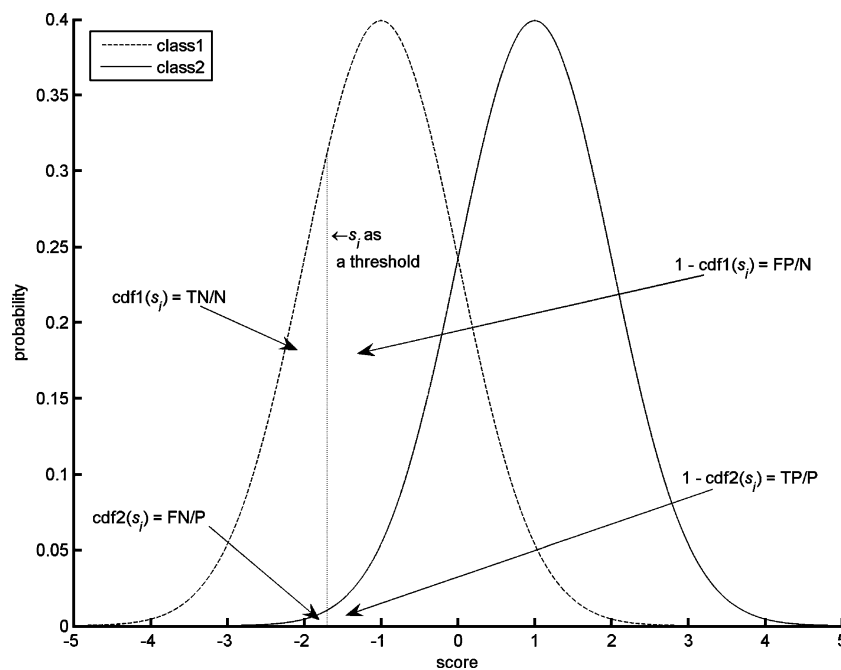


Fig. 2 Relationship between sample distributions and accuracy. The figure illustrates the distributions of scores of two classes: negative (*class1*) and positive (*class2*) samples; $cdf1(\cdot)$ and $cdf2(\cdot)$ are the corresponding cumulative distribution functions. The threshold s_i

separates true negatives (TN) from false positives (FP) in case of *class1* and false negatives (FN) from true positives (TP) in case of *class2*. N and P are the number of negatives and positives, respectively

tion can be overcome by incorporating appropriate distributions in the framework of the PM predictor.

Although the efficacy of the integrated framework has only been demonstrated through an application on the MHC class II binding predictions, this method can be readily extended to the MHC class I binding prediction. Similarly, the framework of the integrative system proposed recently for MHC class I binding (Troost et al. 2007) can be used for the MHC class II binding prediction. The performance comparison of these two systems on both MHC class I and class II binding predictions may lead to the construction of further improved prediction systems.

Conclusions

A new probabilistic meta-predictor (PM predictor) for MHC class II binding peptides has been developed that is based on the integration of predictions obtained from different methods. Using six state-of-the-art predictors, consistently enhanced performance of the PM predictor has been demonstrated in comparison with the individual methods using a data set including peptides from 13 HLA-DR and three mouse alleles from the IEDB database. The results also indicate that future improvement can be made through the incorporation of a more accurate probability estimate for the scores obtained from each individual predictor.

Acknowledgments This research is supported in part by the NIH under Grant 1 R03 AI069391-01.

Appendix

Calculation of the consensus scores

Suppose that a test peptide received a score s_i from predictor i , $i = 1, \dots, m$. We define two probability values:

$$x_i = P_{ib}(S > S_i), \quad 1 \leq i \leq m \quad (1)$$

$$y_i = P_{inb}(S \leq s_i), \quad 1 \leq i \leq m, \quad (2)$$

where S is the random variable representing scores obtained from predictor i . The probabilities $P_{ib}(S > S_i)$ and $P_{inb}(S \leq s_i)$ for binding peptides and nonbinding peptides can be easily calculated once the estimations of score distributions are made from the training peptides. That is, $P_{ib}(S > S_i) = 1 - P_{ib}(S \leq S_i) = 1 - \text{cdf}_{i,\text{binding}}(s_i, \Theta_{i,\text{binding}})$, and $P_{inb}(S \leq s_i) = \text{cdf}_{i,\text{nonbinding}}(s_i, \Theta_{i,\text{nonbinding}})$, where $\text{cdf}_{i,\text{binding}}(\cdot, \Theta_{i,\text{binding}})$ and $\text{cdf}_{i,\text{nonbinding}}(\cdot, \Theta_{i,\text{nonbinding}})$ are, respectively, the cumulative distribution functions for scores of

binding and nonbinding peptides; $\Theta_{i,\text{binding}}$ and $\Theta_{i,\text{nonbinding}}$ are, respectively, parameters of the distributions of scores made by the predictor i .

Next, we define the consensus score δ for a peptide by

$$\delta(x_i, y_i) = \prod_{i=1}^m \frac{1}{2} \left(\frac{1 - x_i}{x_i} + \frac{y_i}{1 - y_i} \right). \quad (3)$$

Note that this score depends on s_i when the probability distributions of the scores are given. One could consider that the goal of this function is to map the $2m$ probabilities values (x_i, y_i) associated with a peptide onto the one-dimensional space of the consensus scores. Such a mapping should efficiently separate the consensus scores of the peptides from the binding and nonbinding classes into two distinct one-dimensional clusters. It is interesting to observe the following relations:

$$P_{ib}(S \leq s_i) = 1 - \text{sensitivity}_i(s_i) = 1 - x_i, \quad (4)$$

$$P_{ib}(S > S_i) = \text{sensitivity}_i(s_i) = x_i, \quad (5)$$

$$P_{inb}(S \leq S_i) = \text{specificity}_i(s_i) = y_i, \quad (6)$$

$$P_{inb}(S > S_i) = 1 - \text{specificity}_i(s_i) = 1 - y_i. \quad (7)$$

In these relations, $\text{sensitivity}_i(s_i)$ and $\text{specificity}_i(s_i)$ are, respectively, sensitivity and specificity determined for the corresponding predictor i with a threshold value that is equal to s_i . Figure 2 explains such relationships between the distributions and classification accuracy.

This mapping possesses the following properties:

- It is defined and continuous for $\forall x_i, y_i \in (0, 1)$, $1 \leq i \leq m$
- For any $x_i \rightarrow 0$ or $y_i \rightarrow 1$, the consensus score is $\delta \rightarrow +\infty$
- For any $x_i \rightarrow 1$ or $y_i \rightarrow 0$, $\delta \rightarrow 0$
- For $\forall x_i \rightarrow 0.5$ and $\forall y_i \rightarrow 0.5$, the consensus score is $\delta \rightarrow 1$

Intuitively, if a higher specificity of the PM predictor is preferred, then the threshold of δ has to be greater than 1 for the binary prediction. Conversely, if a higher sensitivity of the PM predictor is preferred, then the threshold of δ has to be smaller than 1.

References

- Altıparmak F, Akalin A, Ferhatosmanoglu H (2006) Predicting the binding affinity of MHC class II peptides. In: Computational Systems Bioinformatics: Proceedings of the Conference CSB, pp 331–334

- Bhasin M, Raghava GP (2004) SVM based method for predicting HLA-DRB1*0401 binding peptides in an antigen sequence. *Bioinformatics* 20:421–423
- Bleek GMV, Nathenson SG (1991) The structure of the antigen-binding groove of major histocompatibility complex class I molecules determines specific selection of self-peptides. *PNAS* 88:11032–11036
- Borras-Cuesta F, Golvano J, Garcia-Granero M, Sarobe P, Riezu-Boj J, Huarte E, Lasarte J (2000) Specific and general HLA-DR binding motifs: comparison of algorithms. *Hum Immunol* 61:266–278
- Brusic V, Rudy G, Honeyman G, Hammer J, Harrison L (1998) Prediction of MHC class II-binding peptides using an evolutionary algorithm and artificial neural network. *Bioinformatics* 14:121–130
- Bui H-H, Sidney J, Peters B, Sathiamurthy M, Sinichi A, Purton K-A, Mothé BR, Chisari FV, Watkins DI, Sette A (2005) Automated generation and evaluation of specific MHC binding predictive tools: ARB matrix applications. *Immunogenetics* 57:304–314
- Burden FR, Winkler DA (2005) Predictive Bayesian neural network models of MHC class II peptide binding. *J Mol Graph Model* 23:481
- Castellino F, Zhong G, Germain RN (1997) Antigen presentation by MHC class II molecules: invariant chain function, protein trafficking, and the molecular basis of diverse determinant capture. *Hum Immunol* 54:159–169
- Chang ST, Ghosh D, Kirschner DE, Linderman JJ (2006) Peptide length-based prediction of peptide-MHC class II binding. *Bioinformatics* 22:2761–2767
- Chang KY, Suri A, Unanue ER (2007) Predicting peptides bound to I-Ag7 class II histocompatibility molecules using a novel expectation-maximization alignment algorithm. *Proteomics* 7:367–377
- Cui J, Han L, Lin H, Tang Z, Jiang L, Cao Z, Chen Y (2006) MHC-BPS: MHC-binder prediction server for identifying peptides of flexible lengths from sequence-derived physicochemical properties. *Immunogenetics* 58:607
- Cui J, Han LY, Lin HH, Zhang HL, Tang ZQ, Zheng CJ, Cao ZW, Chen YZ (2007) Prediction of MHC-binding peptides of flexible lengths from sequence-derived structural and physicochemical properties. *Mol Immunol* 44:866–877
- De Groot AS, Berzofsky JA (2004) From genome to vaccine—new immunoinformatics tools for vaccine design. *Methods* 34:425–428
- De Groot AS, Sbai H, Aubin CS, McMurry J, Martin W (2002) Immuno-informatics: mining genomes for vaccine components. *Immunol Cell Biol* 80:255–269
- Donnes P, Elofsson A (2002) Prediction of MHC class I binding peptides, using SVMHC. *BMC Bioinformatics* 3:25
- Donnes P, Kohlbacher O (2006) SVMHC: a server for prediction of MHC-binding peptides. *Nucleic Acids Res* 34:W194–W197
- Doytchinova IA, Flower DR (2001) Toward the quantitative prediction of T-cell epitopes: coMFA and coMSIA studies of peptides with affinity for the class I MHC molecule HLA-A*0201. *J Med Chem* 44:3572–3581
- Doytchinova IA, Flower DR (2003) Towards the in silico identification of class II restricted T-cell epitopes: a partial least squares iterative self-consistent algorithm for affinity prediction. *Bioinformatics* 19:2263–2270
- Doytchinova IA, Taylor P, Flower DR (2003) Proteomics in vaccinology and immunobiology: an informatics perspective of the immunone. *J Biomed Biotechnol* 2003:267–290
- Flower DR (2004) Vaccines in silico—the growth and power of immunoinformatics. *The Biochemist* 26:17–20
- Flower DR, Doytchinova IA (2002) Immunoinformatics and the prediction of immunogenicity. *Appl Bioinformatics* 1:167–176
- Flower DR, Doytchinova IA, Paine KPT, Blythe MJ, Lamponi D, Zygouri C, Guan P, McSparron H, Kirkbride H (2002) Computational vaccine design. In: Flower DR (ed) *Drug design: cutting edge approaches*. RSC, London, pp 136–180
- Flower DR, McSparron H, Blythe MJ, Zygouri C, Taylor D, Guan P, Wan S, Coveney PV, Walshe V, Borrow P, Doytchinova IA (2003) Computational vaccinology: quantitative approaches. *Novartis Found Symp* 254:102–120 discussion 120–125, 216–222, 250–252
- Hattotuwigama CK, Toseland CP, Guan P, Taylor DJ, Hemsley SL, Doytchinova IA, Flower DR (2006) Toward prediction of class II mouse major histocompatibility complex peptide binding Affinity: in silico bioinformatic evaluation using partial least squares, a robust multivariate statistical technique. *J Chem Inf Model* 46:1491–1502
- Hertz T, Yanover C (2006) PepDist: a new framework for protein-peptide binding prediction based on learning peptide distance functions. *BMC Bioinformatics* 7:S3
- Huang L, Karpenko O, Murugan N, Dai Y (2006) A meta-predictor for MHC class II binding peptides based on naive Bayesian approach. In: *Proceedings of the 28th International Conference of IEEE Engineering in Medicine and Biology Society (EMBS)*
- Huang L, Karpenko O, Murugan N, Dai Y (2007) Building a meta-predictor for MHC class II-binding peptides. In: Flower DR (ed) *Immunoinformatics: predicting immunogenicity in silico*. Humana, Totowa, NJ, pp 355–364
- Karpenko O, Shi J, Dai Y (2005) Prediction of MHC class II binders using the ant colony search strategy. *Artif Intell Med* 35:147–156
- Kato R, Noguchi H, Honda H, Kobayashi T (2003) Hidden Markov model-based approach as the first screening of binding peptides that interact with MHC class II molecules. *Enzyme Microb Technol* 33:472–481
- Liu W, Meng X, Xu Q, Flower D, Li T (2006) Quantitative prediction of mouse class I MHC peptide binding affinity using support vector machine regression (SVR) models. *BMC Bioinformatics* 7:182
- Mallios RR (1998) Iterative stepwise discriminant analysis: a meta-algorithm for detecting quantitative sequence motifs. *J Comput Biol* 5:703–711
- Mallios RR (2001) Predicting class II MHC/peptide multi-level binding with an iterative stepwise discriminant analysis meta-algorithm. *Bioinformatics* 17:942–948
- Mallios RR (2003) A consensus strategy for combining HLA-DR binding algorithms. *Hum Immunol* 64:852
- Martin W, Sbai H, De Groot AS (2003) Bioinformatics tools for identifying class I-restricted epitopes. *Methods* 29:289
- Max H, Halder T, Kropshofer H, Kalbus M, Muller CA, Kalbacher H (1993) Characterization of peptides bound to extracellular and intracellular HLA-DR1 molecules. *Hum Immunol* 38:193–200
- Moise L, De Groot AS (2006) Putting immunoinformatics to the test. *Nat Biotechnol* 24:791
- Moutaftsi M, Peters B, Pasquetto V, Tschärke DC, Sidney J, Bui H-H, Grey H, Sette A (2006) A consensus epitope prediction approach identifies the breadth of murine TCD8⁺-cell responses to vaccinia virus. *Nat Biotechnol* 24:817
- Murugan N, Dai Y (2005) Prediction of MHC class II binding peptides based on an iterative learning model. *Immunome Res* 1:6
- Nielsen M, Lundegaard C, Wornig P, Lauemøller SL, Lamberth K, Buus S, Brunak S, Lund O (2003) Reliable prediction of T-cell epitopes using neural networks with novel sequence representations. *Protein Sci* 12:1007–1017
- Nielsen M, Lundegaard C, Wornig P, Hvid CS, Lamberth K, Buus S, Brunak S, Lund O (2004) Improved prediction of MHC class I and class II epitopes using a novel Gibbs sampling approach. *Bioinformatics* 20:1388–1397
- Nielsen M, Lundegaard C, Lund O (2007) Prediction of MHC class II binding affinity using SMM-align, a novel stabilization matrix alignment method. *BMC Bioinformatics* 8:238
- Noguchi H, Kato R, Hanai T, Matsubara Y, Honda H, Brusic V, Kobayashi T (2002) Hidden Markov model-based prediction of antigenic peptides that interact with MHC Class II molecules. *J Biosci Bioeng* 94:264–270

- Nussbaum AK, Kuttler C, Tenzer S, Schild H (2003) Using the World Wide Web for predicting CTL epitopes. *Curr Opin Immunol* 15:69
- Parham P (2005) *The immune system*. Garland Science, New York, NY
- Peters B, Sette A (2005) Generating quantitative models describing the sequence specificity of biological processes with the stabilized matrix method. *BMC Bioinformatics* 6:132
- Peters B, Sidney J, Bourne P, Bui H-H, Buus S, Doh G, Fleri W, Kronenberg M, Kubo R, Lund O, Nemazee D, Ponomarenko JV, Sathiamurthy M, Schoenberger SP, Stewart S, Surko P, Way S, Wilson S, Sette A (2005) The design and implementation of the immune epitope database and analysis resource. *Immunogenetics* 57:326
- Peters B, Bui H-H, Frankild S, Nielsen M, Lundegaard C, Kostem E, Basch D, Lamberth K, Harndahl M, Fleri W, Wilson SS, Sidney J, Lund O, Buus S, Sette A (2006) A community resource benchmarking predictions of peptide binding to MHC-I molecules. *PLoS Comput Biol* 2:e65
- Rammensee H, Bachmann J, Emmerich NP, Bachor OA, Stevanovic S (1999) SYFPEITHI: database for MHC ligands and peptide motifs. *Immunogenetics* 50:213–219
- Reche PA, Glutting JP, Reinherz EL (2002) Prediction of MHC class I binding peptides using profile motifs. *Hum Immunol* 63:701–709
- Reche PA, Glutting JP, Zhang H, Reinherz EL (2004) Enhancement to the RANKPEP resource for the prediction of peptide binding to MHC molecules using profiles. *Immunogenetics* 56:405–419
- Salomon J, Flower DR (2006) Predicting Class II MHC-Peptide binding: a kernel based approach using similarity scores. *BMC Bioinformatics* 7:501
- Schirle M, Weinschenk T, Stevanovic S (2001) Combining computer algorithms with experimental approaches permits the rapid and accurate identification of T cell epitopes from defined antigens. *J Immunol Methods* 257:1–16
- Sette A, Buus S, Appella E, Smith JA, Chesnut R, Miles C, Colon SM, Grey HM (1989) Prediction of major histocompatibility complex binding regions of protein antigens by sequence pattern analysis. *Proc Natl Acad Sci USA* 86:3296–3300
- Singh H, Raghava GP (2001) ProPred: prediction of HLA-DR binding sites. *Bioinformatics* 17:1236–1237
- Sturniolo T, Bono E, Ding J, Raddrizzani L, Tuereci O, Sahin U, Braxenthaler M, Gallazzi F, Protti MP, Sinigaglia F, Hammer J (1999) Generation of tissue-specific and promiscuous HLA ligand databases using DNA microarrays and virtual HLA class II matrices. *Nat Biotechnol* 17:555–561
- Swets JA (1988) Measuring the accuracy of diagnostic systems. *Science* 240:1285–1293
- Takahashi H, Honda H (2006) Prediction of peptide binding to major histocompatibility complex class II molecules through use of boosted fuzzy classifier with SWEEP operator method. *J Biosci Bioeng* 101:137–141
- Tong JC, Zhang GL, Tan TW, August JT, Brusica V, Ranganathan S (2006) Prediction of HLA-DQ3.2{beta} ligands: evidence of multiple registers in class II binding peptides. *Bioinformatics* 22:1232–1238
- Toseland C, Clayton D, McSparron H, Hemsley S, Blythe M, Paine K, Doytchinova I, Guan P, Hattotuagama C, Flower D (2005) AntiJen: a quantitative immunology database integrating functional, thermodynamic, kinetic, biophysical, and cellular data. *Immunome Res* 1:4
- Trost B, Bickis M, Kusalik A (2007) Strength in numbers: achieving greater accuracy in MHC-I binding prediction by combining the results from multiple prediction tools. *Immunome Res* 3:5
- Udaka K, Wiesmuller KH, Kienle S, Jung G, Tamamura H, Yamagishi H, Okumura K, Walden P, Suto T, Kawasaki T (2000) An automated prediction of MHC class I-binding peptides based on positional scanning with peptide libraries. *Immunogenetics* 51:816–828
- Wan J, Liu W, Xu Q, Ren Y, Flower D, Li T (2006) SVRMHC prediction server for MHC-binding peptides. *BMC Bioinformatics* 7:463
- Zhang GL, Khan AM, Srinivasan KN, August JT, Brusica V (2005) MULTIPRED: a computational system for prediction of promiscuous HLA binding peptides. *Nucleic Acids Res* 33:W172–W179