The Journal of Immunology

Empirical Evaluation of a Dynamic Experiment Design Method for Prediction of MHC Class I-Binding Peptides

Keiko Udaka, Hiroshi Mamitsuka, Yukinobu Nakaseko and Naoki Abe

J Immunol 2002; 169:5744-5753; ;
doi: 10.4049/jimmunol.169.10.5744
http://www.jimmunol.org/content/169/10/5744

The Journal of Immunology is published twice each month by
The American Association of Immunologists, Inc.,
1451 Rockville Pike, Suite 650, Rockville, MD 20852

Print ISSN: 0022-1767 Online ISSN: 1550-6606.

Empirical Evaluation of a Dynamic Experiment Design Method for Prediction of MHC Class I-Binding Peptides[1]

Keiko Udaka,[2]* Hiroshi Mamitsuka,[3†] Yukinobu Nakaseko,* and Naoki Abe[†‡]

The ability to predict MHC-binding peptides remains limited despite ever expanding demands for specific immunotherapy against cancers, infectious diseases, and autoimmune disorders. Previous analyses revealed position-specific preference of amino acids but failed to detect sequence patterns. Efforts to use computational analysis to identify sequence patterns have been hampered by the insufficiency of the number/quality of the peptide binding data. We propose here a dynamic experiment design to search for sequence patterns that are common to the MHC class I-binding peptides. The method is based on a committee-based framework of query learning using hidden Markov models as its component algorithm. It enables a comprehensive search of a large variety (20^9) of peptides with a small number of experiments. The learning was conducted in seven rounds of feedback loops, in which our computational method was used to determine the next set of peptides to be analyzed based on the results of the earlier iterations. After these training cycles, the algorithm enabled a real number prediction of MHC binding peptides with an accuracy surpassing that of the hitherto best performing positional scanning method. *The Journal of Immunology,* **2002, 169: 5744–5753.**

Major histocompatibility complex class I molecules present peptides to CTL for immune surveillance. After the major anchor amino acids were identified by pool sequencing (1, 2), research efforts have focused on further clarifying the detailed specificity of the MHC molecules. These efforts include identification of secondary anchors (3), side chain scanning (4, 5), positional scanning with peptide libraries (hereafter called the library method) (6–8), and computational analyses (9–11). The latter body of work has resulted in two predictive algorithms: BIMAS (4) and SYFPEITHI (12). In these methods, the position-specific information obtained by the pool sequencing or side chain scanning is used with the assumption of independence between multiple amino acid positions. The structure of proteins, however, is generally flexible, and the fitness of an amino acid depends on the sequential context within the protein or the peptide. It is also hard to translate the sequencing data into quantitative terms. These problems have limited the performance of prediction (13). Peptide library scanning, developed by us and others (6–8, 14), reveals detailed profiles of amino acid preferences at every position of the peptide in quantitative values and thus enables automated prediction (14). However, it also ignores the relationship between multiple positions; thus, the prediction can be inaccurate when amino acids behave interactively, e.g., by inducing conformational adaptations in the MHC molecules (8, 14, 15).

Computational approaches have been introduced to extract not only position-specific information but also to capture some sequence patterns in peptides (16, 17). To attain satisfactory prediction accuracy, usually a large amount of experimental data are needed. To date, the source of MHC-binding peptides has been databases derived from the literature, in which peptide binding was determined using a variety of methods and criteria (2, 18). The lack of coherency of the reported binding values has limited the use of data to yes/no binary outputs. This, together with the limited number of data, has hampered the full exploitation of the existing databases.

A supervised learning algorithm based on hidden Markov models (HMM)[4] was developed by Mamitsuka and has been successfully applied to the analysis of the HLA-A*0201-binding peptides from MHCPEP (18) with a binary output (16). The effectiveness of this approach prompted us to extend the method to predict the degree of peptide binding with real number values. However, training such a learning algorithm in an ordinary way would require a large number of binding data. For example, for the 9-mer peptides that are most common among MHC class I ligands in humans and mice, sequence variation reaches $20^9 = {\sim}5 \times 10^{11}$. Given that only 1 or 2 peptides in every 200 random peptides bind with high affinity to a given MHC class I molecule (7, 8), it is essential to develop a disciplined way of selecting a small number of informative peptides for experimentation, from which the specificity of a given MHC molecule can be extracted effectively. In the present study, we propose a novel method for dynamic experimental design based on a query learning technique to address this issue. The proposed method works by iterating a feedback loop consisting of computational analysis and experimental measurement.

Materials and Methods

Supervised learning algorithm for HMM

The general scheme of the supervised learning algorithm for HMM has been described elsewhere (19). The algorithm is devised to make HMM,

*Department of Biophysics, Kyoto University, Kyoto, Japan; †Theory NEC Laboratory, Real World Computing Partnership, Kawasaki, Japan; and ‡IBM Thomas J. Watson Research Center, Yorktown, NY 10598

Received for publication January 3, 2002. Accepted for publication September 24, 2002.

The costs of publication of this article were defrayed in part by the payment of page charges. This article must therefore be hereby marked *advertisement* in accordance with 18 U.S.C. Section 1734 solely to indicate this fact.

[1] This work was supported by the Real World Computing Partnership, sponsored by the Ministry of Economy, Trade and Industry. Support from Hayashi Memorial Foundation for Female Natural Scientists (to K.U.) is also acknowledged. Most of the work presented herein was performed while N.A. was with Real World Computing Partnership and NEC.

[2] Address correspondence and reprint requests to Dr. Keiko Udaka, Department of Biophysics, Kyoto University, Kyoto, 606-8502, Japan. E-mail address: keiko@molbio.biophys.kyoto-u.ac.jp

[3] Current address: Institute for Chemical Research, Kyoto University, Uji, Kyoto, 611-0011, Japan.

[4] Abbreviations used in this paper: HMM, hidden Markov model; Qbag, query by bagging method; ANN, artificial neural network; Lib, library method; P-x, position x; P-C curve, precision-coverage curve.

0022-1767/02/$02.00

which is normally used to analyze sequence patterns, applicable for sequence data that are labeled with another variable. A peptide sequence with its log K_d value is a typical example for such an application. The purpose of the algorithm is to relate given sequences to the values of the corresponding label (log K_d).

Let H be an HMM and O be a training peptide. As shown in Fig. 1*a*, an HMM consists of a number of states, which are represented by circles. Each state bears the probability for every amino acid that varies from 0 to 1, and their sum over the 20 aa is 1 for each state. When a peptide sequence is given, say KLFGINMPL, the probabilities for K in the 1st state, L at the 2nd state, . . . and L in the last state are multiplied. For the states in the six independent paths, this product is further multiplied by the probability that each path is taken and the sum total for all the six paths is obtained. The logarithm of this final probability value is defined as $L_{O|H}$ for a peptide O in a given H. The $L_{O|H}$ value depends on the size of a given H and O. Generally, the longer O is, the smaller is the $L_{O|H}$ value. The size of O is fixed in the present problem as 9-mer. Therefore, to relate $L_{O|H}$ to the log K_d value, a constant, *C*, is defined as follows: $C = L_{O|H'} - L_{K_a}$, where L_{K_a} is the average log K_a of all the input peptides and H′ is the reference HMM having uniform probability distribution. The limit of the binding measurement was set for −3 in log K_d. This was mainly due to solubility of most peptides. For training purposes, peptides that exhibit no binding at this concentration were all labeled as −3.

1. Training of HMM. Let $L^*_{O|H}$ be the tentative $L_{O|H}$ of O in a given H. $L^*_{O|H}$ is defined because $L_{O|H}$ changes, depending on the probability values in the states (Fig. 1*a*, ○, ●) in a given H. The $L^*_{O|H}$ is related to log K_d as follows: $L^*_{O|H} = C + L^{O}_{K_a}$, where $L^{O}_{K_a}$ is the log K_a value of the peptide O.

A function $g_{O|H}$ is defined as follows:

$$g_{O|H} = (D_{max} - D_{O|H})/D_{max}$$

where $D_{O|H} = |L^*_{O|H} - L_{O|H}|^2$, the square of the distance between $L^*_{O|H}$ and $L_{O|H}$. D_{max} is a constant that satisfies $D_{max} > \max_O D_{O|H}$.

Function $g_{O|H}$ is introduced to find an optimal set of probability parameters for the states that gives the least discrepancy between $L_{O|H}$ and $L^*_{O|H}$. Note that when $D_{O|H}$ is close to zero, $g_{O|H}$ approaches 1. Thus, the goal of the algorithm is to seek the parameter set that gives the maximum $g_{O|H}$.

Finally, probability parameters of H are trained so as to minimize the energy function, *E*, which is based on a gradient descent algorithm, for all the input peptides.

$$E = \sum_O - \log g_{O|H}$$

2. Prediction of log $\mathbf{K_a}$***.*** When the HMM is optimally trained and settled with a set of probability parameters, the training can be stopped by fixing the parameters, and then the HMM can be used to predict the log K_d for unknown peptides. The log K_a of a peptide O is estimated by $L_{O|H} - C$.

Peptides

Peptides were synthesized by fluorenylmethoxycarbonyl chemistry and purified by HPLC to the purity of > 95%. Peptides were analyzed using a MALDI-TOF mass spectrometer (Voyager DE-RP; Applied Biosystems, Foster City, CA). The concentration of peptide was determined by the MicroBCA assay (Pierce, Rockford, IL) using BSA as standard. For rare peptides that react poorly with the bicinchoninic acid reagent, the quantitative ninhydrin reaction was used instead.

Peptide binding

Peptide binding was measured by a stabilization assay using TAP-deficient RMAS cells (8). Briefly, 1×10^5 RMAS cells incubated overnight at 26°C were mixed with graded concentrations of peptide in 0.25% BSA-containing DMEM, incubated for 30 min at room temperature, and then exposed to 37°C for 70 min. The remaining D^b molecules on the cell surface were stained with FITC-labeled anti-D^b mAb (B22.249) and analyzed by flow cytometry using a FACScan (BD Biosciences, Tokyo, Japan). Differences between experiments were normalized by including the D^b binding reference peptides in every assay (14).

Results

A dynamic experiment design using a query learning algorithm with HMM

We use a query learning algorithm called query by bagging (Qbag), the performance of which has been demonstrated to be superior in several applications (20, 21). Qbag combines the advantage of the query by committee method, which helps select the most informative samples, and the averaging by bagging (Bootstrapping by Aggregation) method, which helps reduce the part of the prediction error that is attributable to variability in the sample. Qbag can be thought of as a general scheme for obtaining a query learning method, using an arbitrary supervised learning algorithm as a component. Here, a supervised learning algorithm is an algorithm that estimates a function $F: X \rightarrow Y$, given labeled examples, or pairs of the form $(x, F(x))$. A query learning algorithm is one that estimates a function *F*, based on labeled examples $(x, F(x))$ for *x* values of its choice. Qbag enjoys the following advantages: 1) the number of data (queries) required to attain a given predictive performance has been shown to be minimal; 2) it can deal with data having real number values (labels); 3) it places no restriction on the type of learning algorithm as its component.

We chose as the component algorithm a supervised learning algorithm for HMMs. The HMM has been applied to multiple alignment, because it can capture a sequence pattern common to a given set of multiple sequences even if the pattern appears at different places in the sequences (22, 23). Other applications of HMMs include the prediction of protein structures (24) and open reading frame finding in the genome (25). HMM is chosen for the following reasons. 1) It can help visualize amino acid preferences and sequence patterns hidden in a given set of training peptides. This is not normally possible with an artificial neural network (ANN), for example. 2) The learning and predictive abilities of HMM are at least comparable with those of other learning models such as ANN, if not better. Because general unsupervised learning algorithms for HMMs can deal only with unlabeled data, we used the supervised learning algorithm developed by Mamitsuka (16) for the analysis of peptide sequences labeled with their binding values.

The HMM used in this study was designed as shown in Fig. 1*a*. The HMM consists of states, represented by open circles that correspond roughly to amino acid positions in this case. Each state bears symbol generation probabilities for all the amino acids, specifying the probability that each amino acid assumes that position. A parallel circular model was designed that shares the variable state at the top. This design was aimed at accommodating lateral shifts of subsequences by 1 aa. The variable state can be occupied by the first or last amino acid. The core states of 8 aa are connected by one-directional transitions. The first and eighth states have alternative state transitions to and from the variable state at the top. Six independent paths were designed in parallel without interconnection between them. This simple model exploits the fact that, for MHC class I-binding peptides, orientation is fixed, and size variation is limited. When an HMM is trained with a set of peptide-binding data, symbol generation probabilities and state transition probabilities are optimized so that the log likelihood value (predictive score) of a given peptide approximates its log K_a as described in *Materials and Methods*. We chose a mouse MHC class I molecule, D^b, as a model MHC molecule in this study. D^b predominantly binds 9-mer peptides.

Learning process

The Qbag-designed experiments proceed through the following six steps (Fig. 2).

Steps 1 and 2: training of HMMs with subsets of the binding data. HMMs were first trained with 186 peptide-binding data. Because Qbag tries to maximize information gain by choosing peptides on which it cannot make predictions reliably, at the initial stage the choice of training data might as well be random. To take advantage of prior information, however, we included 80 peptides with known binding abilities that were available in our laboratory, in addition to 106 random peptides that we synthesized anew. To

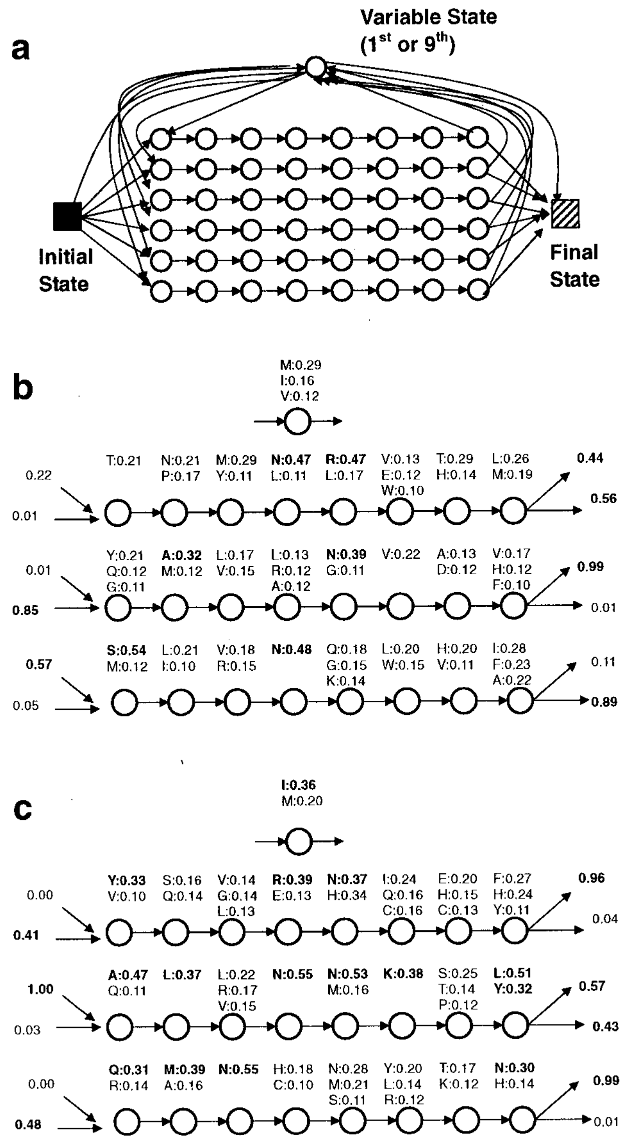
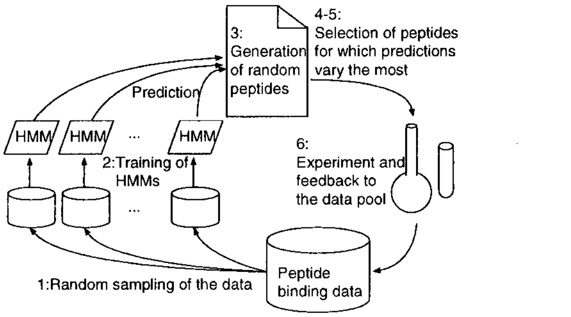


FIGURE 1. HMMs used for analysis of the MHC class I-binding peptides. *a*, Basic structure of HMM designed for this study. The model is explained in the text. *b* and *c*, Two examples of probability parameters found in the trained HMMs. The three most frequently used paths are shown. Only the symbol generation probabilities and state transition probabilities that exceed 0.1 (in bold for >0.3) are presented.



```
Procedure: Query-by-Bagging (Qbag)
Input Parameters: $M, A, T, T', R, I, D$
0. Obtain initial sample $S_1 = \langle (x_1, y_1), \cdots, (x_I, y_I) \rangle$.
For $i = 1, \ldots, M$
  For $j = 1, \ldots, D$
    1. By re-sampling from $S_i$ with uniform distribution,
       obtain sub-samples $S'_1, \ldots, S'_T$ (of same size as $S_i$).
    2. Run component learning algorithm $A$ on $S'_1, \ldots, S'_T$
       to obtain hypotheses $h_1, \ldots, h_T$.
    3. Randomly generate $R$ examples (peptides) $C_i$.
    4. For all examples (peptides) $x$ in $C_i$, compute the variance $v(x)$ by:
       $v(x) = \frac{\sum_{t=1,\ldots,T} |h_t(x) - h_{ave}(x)|^2}{T}$ where $h_{ave}(x) = \frac{\sum_{t=1,\ldots,T} h_t(x)}{T}$
    5. Select an example (peptide) $x^*_j$ from $C_i$ having
       the largest value of $v(x)$.
   6. Empirically determine the binding values of $E_i (= x^*_1, \cdots, x^*_D)$, $y^*_1, \cdots, y^*_D$
      and update the training data by
   $S_{i+1} = append(S_i, \langle (x^*_1, y^*_1), \cdots, (x^*_D, y^*_D) \rangle)$
End For
 7. By re-sampling from $S_{M+1}$ with uniform distribution,
    obtain sub-samples $S^f_1, \ldots, S^f_{T'}$ (of same size as $S_{M+1}$).
 8. Run component learning algorithm $A$ on $S^f_1, \ldots, S^f_{T'}$
    to obtain hypotheses $h^f_1, \ldots, h^f_{T'}$.
 9. Output final hypothesis by taking average over the hypotheses $h^f_1, \ldots, h^f_{T'}$.
```

In our experiments, the input parameters were set as follows:

$M = 7$.
A = Supervised HMM algorithm.
$T = 50$.
$T' = 100$.
$R = 100,000$.
$I = 148$.
$D = 50$ (but actual binding experiments are done for only 24 to 36 peptides).

FIGURE 2. Outline of the dynamic experiment design by Qbag. The algorithm is described in pseudocode. Parameters are as follows: *M*, the round of iterative learning cycle; *T*, the number of the committee HMMs used for selection of the test peptides; *T′*, the number of trained HMMs used for prediction of log K_d for peptides; *R*, the number of random peptides from which test peptides are selected; *I*, the number of peptide binding data used for the initial training; *D*, the number of test peptides selected for experiments.

train a multiple number of HMMs, which will collectively serve as a committee for prediction, the same number (186) of peptides were sampled at random with replacement from the above-mentioned data pool of size 186. Because replacement was allowed during sampling, the sampled data pool consisted of a slightly different subset. Fifty subsets were independently sampled altogether, and one HMM was trained with each of these subsets. Fifty independent HMMs were thus trained, and each settled with slightly different parameters. From the second round on, the training data obtained in step 6 were added to the data pool for Step 1, and the same number of data as the expanded data pool was sampled with replacement for training. This gradually increased the work load of computation, but it remained within the capacity of two workstations arranged in tandem.

Step 3: generation of test peptides. To search the space of size 20^9 for D^b-binding peptides, 100,000 nonamer peptides were randomly generated for each cycle of training.

Steps 4 and 5: judgment by the committee of HMMs. The 50 independently trained HMMs from step 2 were each given the task of predicting the binding ability of 100,000 random peptides. A peptide for which the predictions made by the committee HMMs were most spread, i.e., bore the largest variance, was selected. By excavating the most unpredictable peptides, the most wanted information can be sought efficiently. In effect, this method preferentially explores the unexamined parts of the search space. By repeating 50 times steps 3–5, 50 such peptides were selected. Through these steps, ~5 million peptides (1 of 100,000 of the whole search space) were, in effect, screened.

Step 6: measurement of D^b binding and feedback to the database. From 24 to 36 peptides were arbitrarily chosen from the 50 selected peptides, synthesized in dozens, and subjected to measurement of D^b binding. Peptide binding data are submitted as the supplemental material. The number of peptides examined in each training cycle varied for practical reasons, such as the inability to synthesize some peptides or the time allowed between experiments. A subset of peptides was chosen of 50 to omit peptides that would be hard to synthesize or would not be likely to be soluble during the cell binding assay, e.g., FWLLLLLLL. The proportion of such peptides was 10% or less. Thus, selecting from the 40 candidate peptides should have been sufficient. One or two peptides of 24–36, on average, could not be synthesized. Finally, the results of peptide binding were added to the data pool used in step 1.

Our past research demonstrated that the Qbag procedure, basically as used in the present study, substantially improved the prediction accuracy in various application domains (20, 21). This performance improvement is a combined effect of the query by committee aspect, which helps select the most informative samples, and the averaging by bagging aspect, which helps reduce the component of prediction error attributable to variance of a particular data set.

Also, the choice of the various parameters in the Qbag procedure (M, T, T', R, I, D) is mostly a matter of pragmatic decision, usually determined by the amount of computational and human resources available. For example, the parameters T and T' in Fig. 2 determine the number of subsamples of the binding data; thus, the number of HMMs to obtain in each iteration was determined based on the past experimental evidence, suggesting that the performance tends to saturate at the order of ~100 (26) and the fact that the computational time required to obtain each hypothesis is ~10 min.

The MHC stabilization assay used for measurement of peptide binding was conducted as described before (14). Typical binding curves are shown in Fig. 3*a*. To obtain the log K_d values that correspond to the peptide concentrations at half-maximal binding,

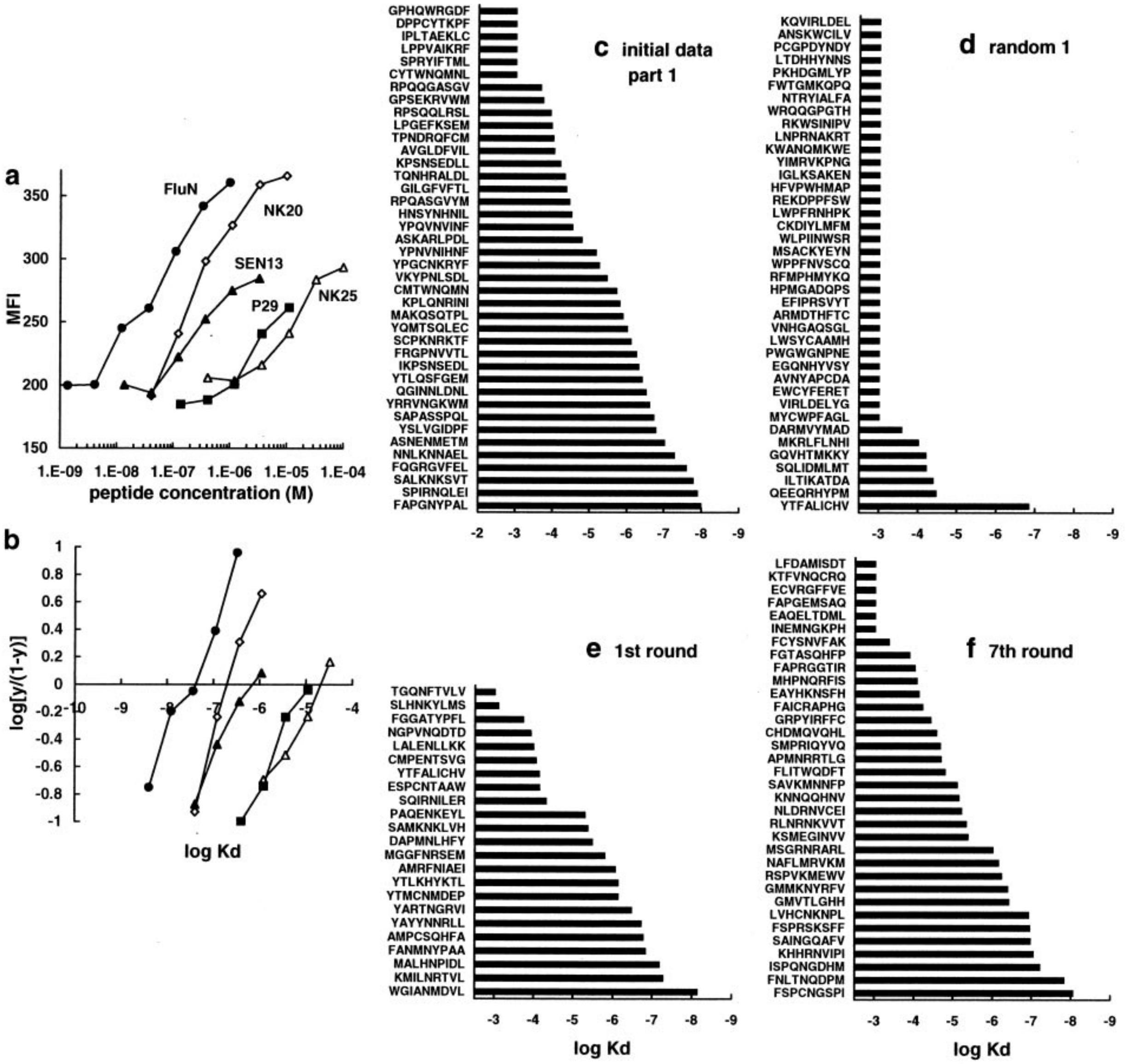


FIGURE 3. Progress of Qbag learning. *a,* Peptide binding curves measured by MHC stabilization assay. Mean fluorescence intensity (MFI) of the peptide-pulsed RMAS cells was measured after staining with an FITC-labeled anti-D^b mAb. Three reference peptides with high, medium and low affinities (FluNP, SEN135, P29) were included for normalization of the binding data in every assay. *b,* Raw peptide binding curves were linearized as described (14). The symbols are the same as in a. The *x*-intercepts, which correspond to the log K_d values, were deduced by the least squares method. *c,* MHC binding data of one-half of the initial peptides; *d,* the fraction of the random peptides used for initial training of the HMMs; *e,* peptides selected for the experiment after the first round of the Qbag learning; *f,* peptides selected for the experiment after the seventh round of the Qbag learning.

the binding curves were linearized as described in *Materials and Methods*. *X* intercepts corresponding to the log K_d values in Fig. 3*b* were deduced by the least squares method. The relative, but not absolute, binding abilities among peptides were consistent between experiments. Therefore, three reference peptides of high, medium, and low binding affinity were always included to normalize the raw binding data for experimental variation.

Progress of the training can be seen in Fig. 3, *c–f*. The initial set included peptides with various binding properties (Fig. 3*c*).[5] Among random peptides, few binders had K_d values 10^{-6} (M) or below (Fig. 3*d*). After the first iteration of learning, however, most of the selected peptides bound more or less to D^b (Fig. 3*e*). The Qbag selects the peptide with the prediction by the committee HMMs that varies most. As a consequence, obvious binders/nonbinders were eliminated, and the peptides that may or may not bind were preferentially selected. Peptides that are likely to bind, but with degrees of binding that are in controversy, were also selected. The fact that peptides with various degrees of binding were evenly selected has a crucial impact on information gathering. Such peptides can often reveal fine details of the structure because even the smallest impact on binding is noticeable by the K_d values. We repeated the learning cycle for seven rounds (Fig. 3*f*). In all, 181 peptides were synthesized and tested during the 7 rounds. We stopped after seven rounds partly because if the method required more than a couple of hundred peptides for training it would not be practical given a large number of allelic MHC molecules.

Sequence patterns seen in the HMMs after training

Unlike ANN, HMM has the advantage that its state transitions generating sequence patterns can be examined after the training. As two examples are seen in Fig. 1, *b* and *c*, the patterns are in fact quite interesting. D^b has the major anchor amino acids of N at position 5 (P5) and M or I at P9 (1). Preferences for these amino acids can more or less be seen in the major paths (Fig. 1, *b* and *c*). We tentatively designed the HMM with six parallel paths. In the HMMs obtained as a result of training, there were typically two or three major paths that peptides would take frequently (Fig. 1, *b* and *c*). Thus, restricting the number of paths to six in the model seemed to achieve computational efficiency with no significant loss in representational power. The probability parameters in the HMM itself represent the dynamic specificity of the MHC molecule. It would be interesting to refer to this information to aid computer modeling in the future. For instance, better binding variants may be designed by replacing the MHC-binding residues with a pattern of sequence found in the major paths. Interestingly, some of the patterns exhibit sequential properties; e.g., the combination of A or S at P2 and N at P5 is favored (Fig. 1, *b* and *c*), but Q is rather favored for P2 only if P3 and P4 are MN (Fig. 1*c*). This type of specificity can be better captured/visualized in HMMs than positional scanning or ANN. It will be interesting to examine the structural basis for such sequence-specific patterns by structural analysis.

We designed the HMM as a cyclic model. This is effective for accommodating some lateral slide of subsequences. However, considering the critical role of the C-terminal amino acid (P9) (2) and the substantial preference of the amino acid at P1 (7), a noncyclic model with entirely independent paths may have been more appropriate. In retrospect, however, this did not appear to have greatly affected the outcome, because prominent patterns at the C terminus often ended at the end of the core parallel paths and had a limited influence on P1 selection in the variable state at the top (Fig. 1, *b* and *c*).

Comparison of the previous prediction methods

The 181 peptide binding data obtained during Qbag learning are of random source, and they exhibit a wide range of binding abilities. Thus, this data set provides an excellent test bed for comparing the performance of existing methods of automated prediction. Although it is not the focus of this study, such a direct comparison will be useful. Fig. 4 shows the correlation between predicted and actual binding for the library method (Lib) (Ref. 14 and Fig. 4*a*), SYFPEITHI (Ref. 12 and Fig. 4*b*) and BIMAS (Ref. 4 and Fig. 4*c*). All these methods are accessible online (http://www.ddbj.nig.ac.jp/analysesp-e.html, http://www.uni-tuebingen.de/uni/kxi/, http://bimas.dcrt.nih.gov/molbio/hla_bind/index.html). These three methods differ with respect to the source of peptides used to probe the MHC molecules (synthetic peptide libraries for Lib, natural peptide libraries for SYFPEITHI, and the amino acid-substituted variants of binder peptides for BIMAS). Linear correlation can be seen for all three methods. Linear correlation has been demonstrated previously for Lib (6, 14). This is not surprising because these methods are based on position-specific information and they all assume the independence of amino acid positions. The expected log K_d scales are deduced by

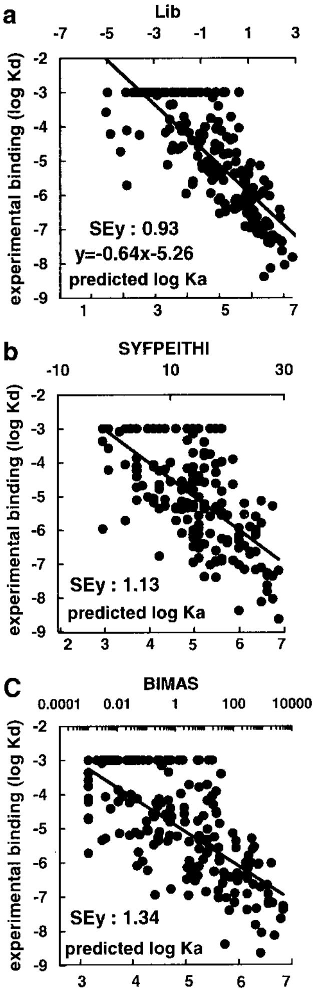


FIGURE 4. Performance comparison of three publicly accessible methods of prediction. MHC binding data of the peptides that were used for Qbag learning were used to evaluate the predictive performance of Lib (14) (*a*), SYFPEITHI (12) (*b*), and BIMAS (4) (*c*). *Abscissa,* Predictive scores of the respective method. *Inset,* SE along the *y*-axis (SE*y*). The scale of the second *x*-axis shown at the bottom was determined by linear regression of the input data.

[5] The on-line version of this article contains supplemental material.

linear regression, and they are added as the second *x*-axes at the bottom of Fig. 4. There are, however, substantial differences in the accuracy of the prediction. The *x*-axes use different values in the respective figures; therefore, standard error along the *y*-axis (SE*y*) is used as a measure of assessment. SE*y* becomes smaller in the order BIMAS, SYFPEITHI, and Lib. The fitness of amino acids used for BIMAS is measured using a series of variant peptides. Thus, the fitness values are likely to be influenced by the sequence context of peptides, as reported previously (8). These sequence-dependent properties are averaged out by the use of the libraries in the cases of SYFPEITHI and Lib. Now SYFPEITHI scores are given in integers, whereas Lib uses real number values from the binding measurement. Unlike the natural library of SYFPEITHI, equimolar representations of amino acids in the synthetic libraries used in Lib, and its quantitatively more accurate measurement, appear to be accountable for improved prediction over SYFPEITHI. Next, we compared the predictive performance of Qbag with Lib, the best performing method among the three.

Assessment of the predictive power of trained HMMs

To assess the effectiveness of the Qbag learning, a program for prediction was constructed using 367 binding data described above (186 before learning, 181 during learning). First, 100 HMMs were loaded with different initial parameters, and each of them was trained with a subset of 367 binding data. Subsets of the binding data were obtained by sampling the same number of peptides (367) as the size of the entire data pool with replacement as described in steps 1 and 2. Increasing the number of HMMs from 50 to 100 this time was due to upgrading of the computational capacity. It took ~10 min to train 1 HMM with 367 peptides. Test peptides were scored by 100 HMMs, and the output score was given by averaging the 100 predictions. The average value was used as the final prediction, according to the theoretical and empirical bases given by Breiman (26). We chose the number 100, because past work has indicated that the prediction accuracy of learning saturates on the order of ~100 (26).

Predictive power of Qbag for random peptides

To make a fair judgment on an unbiased population, we compared the predictive power of Qbag with that of Lib for random peptides. In the past publications describing different computational methods for predicting MHC-binding peptides, four measures, true/false positives and true/false negatives, were often used as criteria for performance evaluation. However, the same technique is not feasible for our problem. We set out to explore the search space of 20^9 peptides where the chances to encounter binders were rare (<1%). To select a feasible number of test peptides with a variety of binding abilities, we must use some predictive method. This selection process inevitably introduces a bias that depends on the selection method used. Therefore, instead of the above mentioned four-measure approach, we used the so-called precision-recall curve analysis (for ease of understanding, in this paper we refer to it as the precision-coverage curve (P-C curve) for assessment (Fig. 5*g*). The P-C curve is a standard measure of assessment in the fields of information retrieval, machine learning, and statistics, where a relatively small number of relevant items (binder peptides, here) must be predicted from a huge population. It essentially examines whether the ranking of the predictions is in the correct order. Here, precision is the fraction of actual binders among peptides that are predicted to bind. Coverage (recall) is the fraction of peptides that are predicted to bind, among all the binders. A low precision indicates many false positives, whereas a low coverage indicates many false negatives. The two measures are in general in conflict, given that higher precision can be attained by sacrificing coverage and vice versa. (For example, by predicting just the very few top ranked peptides as binders, high precision can be achieved but coverage will be low.) Generally, a prediction method with a P-C curve that lies above another has a better predictive performance.

To calculate precision and coverage, we first had both methods predict the binding ability of a large number (1 million) of randomly generated peptides and rank them according to their predicted scores. We then randomly selected a feasibly small number (41) of peptides from the top 1% of the respective rankings and measured their binding. These random test peptides were synthesized anew, and none of these peptides was overlapping with the peptides used for training. Previously, we found that most T cell epitope peptides have log K_d values of −6 or below (14). In that paper, this cutoff threshold of −6 approximately corresponded to the Lib score that was 2 SD away from the mean score. If the Gaussian distribution is assumed, the 2-SD threshold by the Lib score would indicate the top ~2.5% ranked peptides. This is a higher percentage than the ~1% estimated within the random peptide library (7, 8, 14), suggesting the presence of false positives. As shown in Fig. 5, *a* and *b* and Table I, the Qbag-scored peptides exhibit higher binding (i.e., lower log K_d values). If the cutoff threshold for biological activity is set at −6 in log K_d, 61% of the Qbag-scored peptides are binders. This is significantly higher than the 39% obtained with Lib (Table I). Although Qbag alone achieves 61% precision, combining it with Lib and selecting the peptides that are predicted to be high binders by both methods improve the precision to 84%.

We next compared the P-C curves between the peptides predicted by the respective methods. For calculation of precision and coverage, we assumed here that the top 1% of ranked peptides are binders, because it was estimated in our past publications that peptides with comparable binding affinity with the natural CTL epitopes exist at a frequency of 1 in every 100–200 random peptides for a given MHC class I molecule (7, 8, 14). To calculate precision and coverage, peptides were first ordered in descending order with respect to their scores. Precision was calculated as the fraction of peptides the log K_d values of which were <−6. For instance, if 8 peptides are actual binders among peptides ranked 10th or higher, precision at the 10th peptide is 0.8. Coverage was calculated as the fraction of binders that are successfully predicted among all the binders. Because all 41 peptides belong to the top 1%, they were assumed here to be binders. If 4 binders of a total of 41 would-be binders are predicted within the top 10 ranked peptides, coverage at the 10th peptide is 4/41 = 0.098. As shown in Fig. 5*g*, Qbag exhibits better performance, indicating that predictions are in better order and exhibit broader coverage.

The relationship between the Lib and Qbag (Que) scores for the $\mathrm{Lib}^{\mathrm{high}}$ and $\mathrm{Que}^{\mathrm{high}}$ peptides is shown in Fig. 5, *c* and *d*. The actual binding of these peptides indicates that the two methods are complementary. The $\mathrm{Lib}^{\mathrm{high}}$ peptides include many peptides the Que scores of which predict lower binding (Fig. 5*c*). These peptides tend to exhibit lower binding (Fig. 5, *c* and *e*, ○ and ▲). The $\mathrm{Que}^{\mathrm{high}}$ peptides, in contrast, include peptides that are predicted to be higher and lower binders by Lib. Peptides with Lib scores that predict higher binding tend to bind better (Fig. 5, *d* and *f*, □) whereas those with lower Lib scores bind less (Fig. 5, *d* and *f*, ○). Taken together, Lib tends to overestimate the binding for a number of peptides. This is consistent with the lower precision of Lib demonstrated by the P-C analysis, which indicates a high rate of false positives.

This time, active training of HMMs was halted at 181 peptides but the prediction should improve further if the training were continued. This is an encouraging performance for a sequence analysis

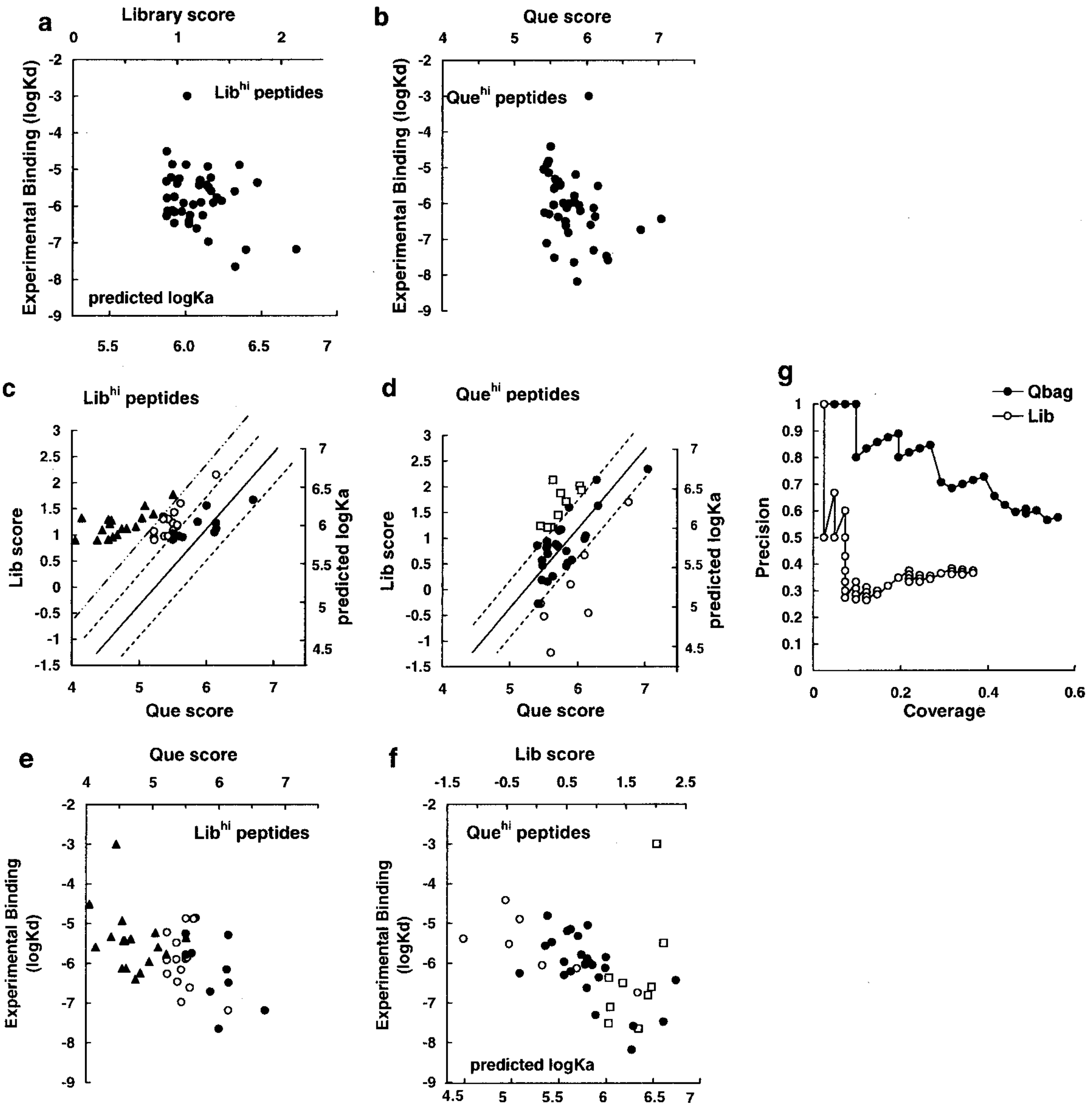


FIGURE 5. Predictive performance of Qbag tested on random peptides. Forty-one peptides were randomly selected from the peptides scored within the top 1% by Lib (*a*) or Qbag (*b*), respectively, and their MHC binding was measured. The Lib and Que scores for the Lib^{high} (*c*) and Que^{high} (*d*) peptides were plotted, and their relationship was investigated. Arbitrary thresholds were set at 0.3 (– – – –) and 0.6 (••—••) away from the linear correlation between the Que score and log K_a predicted by Lib. Compared with the peptides with predictions that were consistent by the two methods (●) peptides with Que scores that suggested weaker binding (○, ▲ in *c* and *e*) actually bound less (*e*). The Que^{high} peptides (*d*) were arbitrarily grouped into three categories. Peptides with predictions by Lib that suggest stronger binding than expected from the Que scores (□ in *d* and *f*) actually bound better (*f*), whereas peptides with Lib scores that predicted weaker binding bound less (○ in *d* and *f*). The expected log K_d scale that was deduced from the Lib analysis (Fig. 4*a*) was added as the second abscissa on the right (in *c* and *d*) or at the bottom (in *f*). *g*, P-C curves. The method and rationale of the P-C curve analysis are described in the text. Briefly, peptides selected by the respective method were first arranged in the order of decreasing scores. Starting from the top scored peptides, precision and coverage were calculated. Precision designates the fraction of peptides that actually bind among peptides that are predicted to bind. Coverage calculates the inclusion of binder peptides within the predicted peptides. These parameters given for the *i*th ranked peptide were calculated for the peptides that were ranked *i*th or above.

that explores the entire random space with real number representation. In a previous report of passive learning using published binding data, at least 1000 data or more can be estimated to be necessary for a prediction with binary output to reach an error rate of 15% and a confidence level of 95% (9, 27). In another report, a simulation using binding data that contained many binder peptides (59 of 329 for A24, 221 of 404 for B27, 72 of 285 for B35 with all A24 and B35 peptides bearing major anchors) suggested that 350–500 binding data would be necessary (Fig. 2 in Ref. 9) to derive reasonable matrix models with binary output. Also, 300–400 data containing many binders (61 of 317 peptides with $IC_{50} < 500$ nM; 273 of 463 with both major anchors) were used for ANN analysis by Gulukota et al. (10). Because the above studies used part of the binding data for training and the rest for assessment of the predictive power, it is not clear how well these methods would predict random peptides. Although direct comparison is difficult, it is encouraging that Qbag can achieve the reported level of predictive performance in real number prediction with a comparable number of training data. If we had started from random peptides without the guidance of Qbag (such as in Fig. 3*d*), it would have taken much longer to obtain a sufficient number of informative data. It is rare to find binding data containing >100 peptides for an MHC

Table I. *Prediction of D^b binding peptides for random peptides*[a]

Peptides	Mean Log K_d (M)	No. of Peptides with Log $K_d < -6$	
Que^{high} sampled peptides			(%)
Total	−6.08	25/41	61
Que^{high} and Lib^{high}	−6.59	16/19	84
Que^{high} but Lib^{low}	−5.64	9/22	41
Lib^{high} sampled peptides			
Total	−5.78	16/41	39
Lib^{high} and Que^{high}	−6.03	9/19	47
Lib^{high} but Que^{low}	−5.56	7/22	32

[a] Forty-one peptides ranked within the top 1% by the Qbag or library method were examined for their MHC-binding abilities. No peptide was overlapping between the two groups. Previously, the Lib score at 2 SD from the mean score (i.e., 0.92) corresponded approximately to −6 in log K_d, and that often coincided with the biological activities (14). The Qbag scores >5.4 rank approximately the same number of peptides as Lib > 0.92. Therefore, these values were used as thresholds for Que$^{high/low}$ and Lib$^{high/low}$.

molecule in the current databases, and passive analysis of published data alone would not meet future demands.

*Prediction of D*b*-binding peptides from proteins in the database*

To test the predictive performance on peptides existing in the real world, we next chose protein sequences from GenBank that had been reported to be up-regulated in tumor cells. From 21 proteins, 14,071 overlapping 9-mer peptides were obtained and subjected to prediction by Qbag or Lib. No peptide was overlapping with the peptides used for training of the Qbag algorithm.

A positive overall correlation can be seen between the Lib and Qbag scores (Fig. 6*a*). This suggests that a large part of the binding energy is, in fact, supplied additively by the independent binding of amino acids, which is captured well by Lib. Note, however, that the points form an upward-pointing triangular shape, indicating that Lib may tend to overestimate the binding ability, as has been suggested for random peptides. In contrast, Qbag can capture both position-specific preferences of amino acids and sequence patterns and is therefore able to differentiate false positives that are hard to distinguish by positional scanning type methods.

We chose for binding assays mostly peptides from the top ranked peptides (Fig. 6*a*, ○) by the Lib and/or Qbag method because binders, if identified, may have a clinical importance. For this reason, selection of peptides was arbitrary, and it so happened that the selected peptides were among the top 0.6% of ranked peptides according to the Lib score, but 1.3% according to the Que score. Thus, sampling was biased and not equal for the two methods. Therefore, rather than conducting a comparative analysis of the type performed for the random peptides, we examined the variance of predictions by the respective methods and whether the two methods were complementary.

We first examined how prediction relates to actual binding. As seen in Fig. 6, *b* and *c*, standard error along the *y*-axis (SE*y* 0.75) is smaller for Qbag than for Lib (SE*y* 0.93). There are two outliers in Fig. 6*b*, but SE*y* is still greater without them. Sampling of the Que^{high} peptides was more biased than that of the Lib^{high} peptides. Que^{high} peptides with intermediate Lib scores were underrepresented (Fig. 6*a*). If these points were included, however, they would most likely settle in the middle of Fig. 6*c* and would not greatly affect the above interpretation.

We next examined how the prediction could be refined by combining the two methods. The Lib scores and the Que scores are plotted for all the peptides shown in Fig. 6, *b* and *c* (Fig. 6*d*). Closed circles represent peptides for which predictions by the Lib and Qbag methods were consistent. Although peptides that scored high by only one method (○, △) tend to bind more weakly than peptides represented by the closed circles, they are still better binders among peptides that are scored in the same range by the alternative method. Thus, the open symbols (△ in *e,* ○ in *f*) tend to settle below the line of correlation (i.e., better than expected binders). This suggests that the two methods can be used complementarily, especially when the predictions by the two methods differ considerably.

Discussion

Dynamic experiment design by Qbag has been shown to be quite effective in sequence analysis of MHC-binding peptides. As we have shown, Qbag exhibited a superior predictive performance as compared with existing methods. In addition, we emphasize that Qbag enables capturing of sequence patterns, which is difficult to do by existing positional scanning type of analysis. Qbag is also only the second after Lib (14) that can predict the binding capacities relative to the real number K_d values. For most investigators looking for MHC-binding peptides, it is not easy to measure the peptide binding. However, with this level of precision attained by our methodology, direct screening of T cell responses may become feasible. In principle, it is possible to apply this method to MHC class II-binding peptides also. Once the HMMs are trained, prediction can be automated for high throughput analyses. Recently, DNA/protein microchip technology has been developing rapidly, and genome-wide sequencing efforts are in progress. Consequently, the number of target proteins for immunotherapy against tumors, infectious agents, and autoimmune disorders will increase considerably. Thus, automatic methods for evaluating them will be in high demand, and a more accurate, faster prediction method will be desirable.

Prediction of peptide binding in real number K_d values has a critical importance in strategic planning for antitumor immunotherapy (28, 29) as well as in understanding the pathogenesis of cryptic epitopes in autoimmune disorders (30–32). In the cases of tumor-specific Ags, it often happens that peptides that have marginal affinities to the host's MHC molecules serve as tumor Ags, especially when the Ags are also expressed in the normal tissues (33, 34). Focusing the peptide search in the marginal affinity range is easily attainable with real number prediction.

The present study is the first attempt of its kind, and there is ample room for improvement regarding the number of training peptides and the choice and specific configurations of the component algorithm. Although training of the algorithm still requires costly materials and labor-intensive experiments, once it is done for an MHC molecule, the information obtained will be a shared resource that can be widely exploited forever. The large number of MHC alleles is another hurdle in compiling the comprehensive information that would cover most people. To achieve this goal, organizing a global collaboration would be necessary to exploit precious peptide resources and standardize the measurements.

Thus far, a major obstacle to computational analysis has been the insufficient number of binding data. Reaching the level of performance achieved here with active learning of 181 peptides (in addition to 186 initial peptides) by Qbag seems promising. One might argue that if these informative 181 peptide binding data were available, other algorithms might perform as well. This may be so, but the essence of this study lies not in the performance of HMM per se but in the fact that Qbag has enabled the selection of a set of most informative peptides for empirical measurements. We happened to use as many as 186 peptides for initial training in this study, but now that the proposed approach has been proved effective, in subsequent analyses it may be possible to start from a smaller

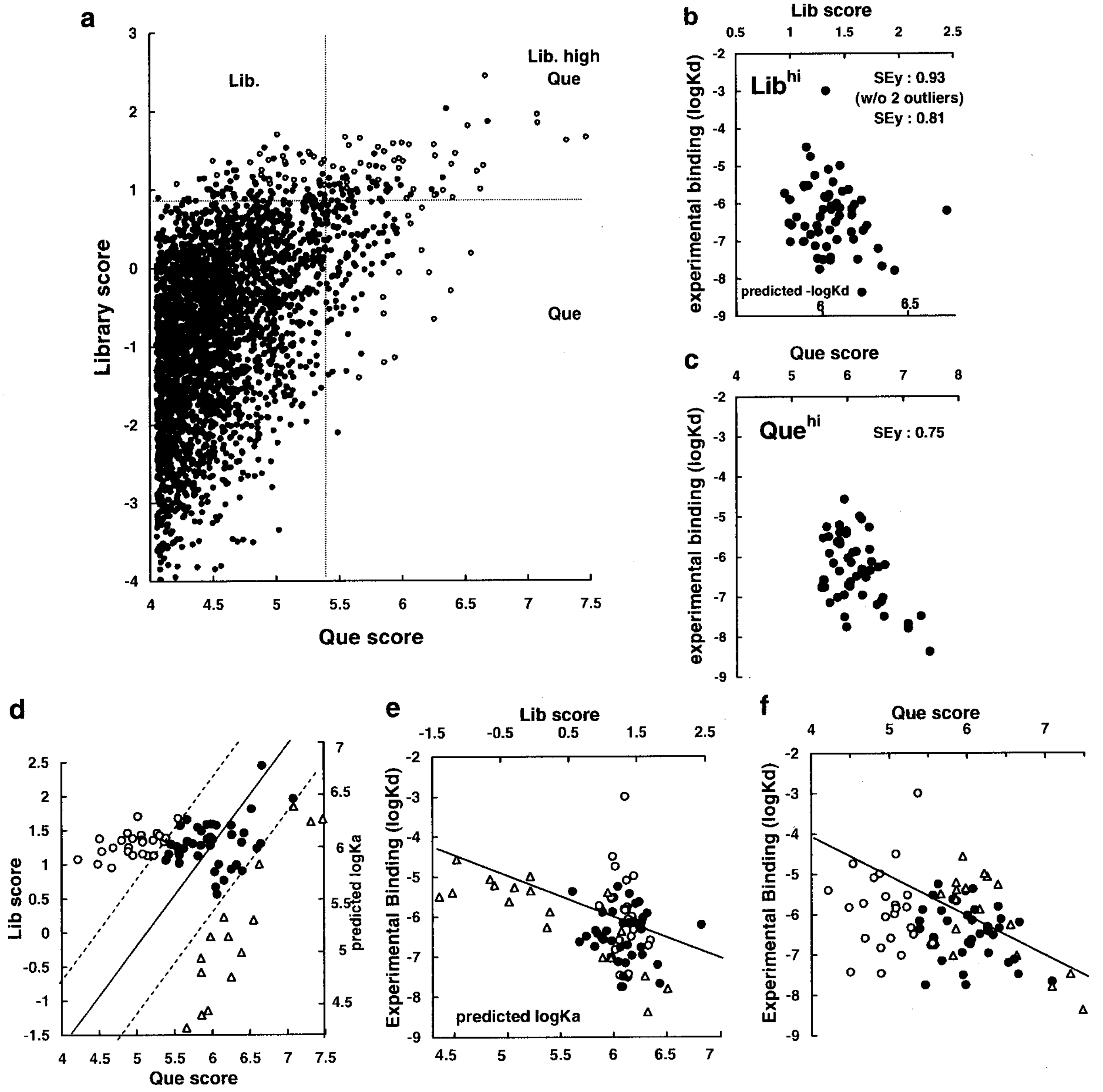


FIGURE 6. Prediction of D^b binding peptides in existing protein sequences. *a,* All 9-mer peptides existing in 21 proteins were scored by the Qbag and Lib methods. The relationship of 2 scores is shown for the top 3000 peptides by the Que score. Eighty peptides that were top ranked by the Lib or Qbag method were arbitrarily chosen (*a,* ○) and subjected to binding measurement (*b* and *c*). A Lib score at 2 SD (0.95) from the mean score corresponds roughly to the affinity threshold of the T cell epitope peptides (14) (closely dotted line in *a*). A Que score of 5.4 (closely dotted line in *a*) or above covers a comparable number of the top ranked peptides. The peptides shown in *b* and *c* were combined, and the relationship between their Que and Lib scores was examined in *d*. Compared with the peptides with predictions that are consistent by the two methods (●, arbitrary thresholds are shown by broken lines), peptides with binding that was estimated to be strong by the Lib score but not as strong by the Que score (○) had a tendency to bind less (*e*). Likewise, peptides with high Que but low Lib scores (△) bound less (*f*). However, these peptides (○, △) bound better than what was expected by their Lib (for △ in *e*) and Que (for ○ in *f*) scores alone. The second axes indicating the scale of predicted log K_d values by the Lib score are inserted

number of peptides. In practice, the number of peptides for initial training is not necessarily a serious problem because peptides used for analysis of other MHC molecules can also serve the purpose.

We used HMM because of the advantages mentioned above, but other algorithms can well replace HMM. Qbag is a general method that can take any algorithm as its component. Which algorithm would serve best as the component is an open question not specifically addressed in the present study. We deposited all the peptide binding data as supplemental material. They should offer an excellent test bed for researchers to test their arsenal algorithms. Comparing different approaches for computational prediction published to date is not easy at present. The previously developed methods mostly use binary output. Although there are a few algorithms dealing with real number predictions, in their evaluation the data pool was split; part of it was used for training and the rest for assessment (9, 10). Because the peptides used for assessment were mostly a biased population of peptides intended for the authors' own aims, the training peptides and those used for performance evaluation consisted more or less of a similar type. Comparing the performance on a common, unbiased population of peptides would be necessary for fair judgment on competing prediction methods.

This time, we did not test peptides in the low affinity range. Instead, we used the P-C curve and the measure of linear correlation between the predicted and actual binding. The former indicated a better ranking ability of Qbag, compared with Lib. Smaller variance of prediction values with respect to the actual binding was also confirmed for Qbag by linear regression analysis in Fig. 6, *b* and *c*.

We found that the Qbag and Lib methods are complementary to each other and that combining them enhances the predictive performance of either method in isolation. It appears that with a limited number of binding data, although HMMs capture some sequence patterns of binding peptides, they do not capture as much position-specific information as Lib. This seems to indicate that the technique of randomizing the amino acid positions except for the position of interest, used in Lib, is in fact quite effective for extracting position-specific information (6, 7). Library scanning also requires a set of 172 sublibraries (19 × 9 + 1, for a random reference library). However, the advantage of Lib lies in the fact that the same set can be used for any 9-mer binding MHC class I molecule. Therefore, for analysis of a new MHC molecule, combining the two methods and using a couple of hundred new peptides for the Qbag learning would be a good option.

Acknowledgments

We thank R. Yamamoto for synthesis of some of the peptides. We also thank Drs. Klaus Kärre and Hans-Georg Rammensee for RMAS cells.

References

1. Falk, K., O. Roetzschke, S. Stevanovic, G. Jung, and H. Rammensee. 1991. Allele-specific motifs revealed by sequencing of self-peptides eluted from MHC molecules. *Nature 351:290.*
2. Rammensee, H., T. Friede, and S. Stevanovic. 1995. MHC ligands and peptide motifs: first listing. *Immunogenetics 41:178.*
3. Ruppert, J., J. Sidney, E. Celis, R. Kubo, H. Grey, and A. Sette. 1993. Prominent role of secondary anchor residues in peptide binding to HLA-A2.1 molecules. *Cell 74:929.*
4. Parker, K., M. Bednarek, and J. E. Coligan. 1994. Scheme for ranking potential HLA-A2 binding peptides based on independent binding of individual peptide side-chains. *J. Immunol. 152:163.*
5. Hammer, J., E. Bono, F. Gallazzi, C. Belunis, Z. Nazy, and F. Sinigaglia. 1994. Precise prediction of MHC class II-peptide interaction based on peptide side chain scanning. *J. Exp. Med. 180:2353.*
6. Stryhn, A., L. Pedersen, T. Romme, C. Holm, A. Holm, and S. Buus. 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.*
7. Udaka, K., K.-H. Wiesmuller, S. Kienle, G. Jung, and P. Walden. 1995. Tolerance to amino acid variations in peptides binding to the MHC class I proetin H-2Kb. *J. Biol. Chem. 270:24130.*
8. Udaka, K., K.-H. Wiesmueller, S. Kienle, G. Jung, and P. Walden. 1995. Decrypting the structure of MHC-I restricted CTL epitopes with complex peptide libraries. *J. Exp. Med. 181:2097.*
9. Brusic, V., C. Schoenbach, M. Takiguchi, V. Ciesielski, and L. Harrison. 1997. Application of genetic search in derivation of matrix models of peptide binding to MHC molecules. In *International Conference on Intelligent Systems for Molecular Biology*, p. 75.
10. Gulukota, K., J. Sidney, A. Sette, and C. DeLisi. 1997. Two complementary methods for predicting peptides binding major histocompatibility complex molecules. *J. Mol. Biol. 267:1258.*
11. Honeyman, M., V. Brusic, N. Stone, and L. Harrison. 1998. Neural network-based prediction of candidate T-cell epitopes. *Nat. Biotechnol. 16:966.*
12. Rammensee, H., J. Bachmann, N. Emmerich, O. Bachor, and S. Stevanovic. 1999. SYFPEITHI: database for MHC ligands and peptide motifs. *Immunogenetics 50:213.*
13. Andersen, M., L. Tan, I. Sondergaard, J. Zeuthen, T. Elliott, and J. Haurum. 2000. Poor correspondence between predicted and experimental binding of peptides to class I MHC molecules. *Tissue Antigens 55:519.*
14. Udaka, K., K.-H. Wiesmuller, S. Kienle, G. Jung, H. Tamamura, H. Yamagishi, K. Okumura, P. Walden, T. Suto, and T. Kawasaki. 2000. An automated prediction of MHC class I-binding peptides based on positional scanning with peptide libraries. *Immunogenetics 51:816.*
15. Madden, D., D. Garboczi, and D. Wiley. 1993. The antigenic identity of peptide-MHC complexes: a comparison of the conformations of five viral peptides by HLA-A2. *Cell 75:693.*
16. Mamitsuka, H. 1998. Predicting peptides that bind to MHC molecules using supervised learning of hidden Markov models. *Proteins 33:460.*
17. Brusic, V., G. Rudy, M. Honeyman, J. Hammer, and L. Harrison. 1998. Prediction of MHC class II-binding peptides using an evolutionary algorithm and artificial neural network. *Bioinformatics 14:121.*
18. Brusic, V., G. Rudy, A. Kyne, and L. Harrison. 1996. MHCPEP: a database of MHC-binding peptides: update 1995. *Nucleic Acids Res. 24:242.*
19. Mamitsuka, H. 1996. A learning method of hidden Markov models for sequence discrimination. *J. Comput. Biol. 3:361.*
20. Abe, N., and H. Mamitsuka. 1998. Query learning strategies using boosting and bagging. In *Fifteenth International Conference on Machine Learning*. Morgan Kaufmann, San Francisco, p. 1.
21. Mamitsuka, H., and N. Abe. 2000. Efficient mining from large databases by Query learning. In *Seventeenth International Conference on Machine Learning*. Morgan Kaufmann, San Francisco.
22. Baldi, P., Y. Chauvin, T. Hunkapiler, and M. A. McClure. 1994. Hidden Markov models of biological primary sequence information. *Proc. Natl. Acad. Sci. USA 91:1059.*
23. Krogh, A., M. Brown, I. S. Mian, K. Sjolander, and D. Haussler. 1994. Hidden Markov models in computational biology: applications to protein modeling. *J. Mol. Biol. 235:1501.*
24. Park, J., K. Karplus, C. Barrett, R. Hughey, D. Haussler, T. Hubbard, and C. Chothia. 1998. Sequence comparisons using multiple sequences detect twice as many remote homologues as pairwise methods. *J. Mol. Biol. 284:1201.*
25. Krogh, A., I. S. Mian, and D. Haussler. 1994. A hidden Markov model that finds genes in *E. coli* DNA. *Nucleic Acids Res. 22:4768.*
26. Breiman, L. 1996. Bagging predictors. *Machine Learning 24:123.*
27. Kearns, M., and U. Vazirani. 1994. *An Introduction to Computational Learning Theory.* MIT Press, Cambridge.
28. Apostolopoulos, V., M. Yu, A. Corper, L. Teyton, G. Peietersz, I. McKenzie, and I. Wilson. 2002. Crystal structure of a non-canonical low-affinity peptide complexed with MHC class I: a new approach for vaccine design. *J. Mol. Biol. 318:1293.*
29. Dyall, R., W. Bowne, L. Weber, J. LeMaoult, P. Szabo, Y. Moroi, G. Piskun, J. Lewis, A. Houghton, and J. Nikolic-Zugic. 1998. Heteroclitic immunization induces tumor immunity. *J. Exp. Med. 188:1553.*
30. Fairchild, P., R. Wildgoose, E. Atherton, S. Webb, and D. Wraith. 1993. An autoantigenic T cell epitope forms unstable complexes with class II MHC: a novel route for escape from tolerance induction. *Int. Immunol. 5:1151.*
31. Fairchild, P. 1999. Reversal of immunodominance among autoantigenic T-cell epitopes. *Autoimmunity 30:209.*
32. Anderton, S., N. Viner, P. Matharu, P. Lowrey, and D. Wraith. 2002. Influence of a dominant cryptic epitope on autoimmune T cell tolerance. *Nat. Immunol. 3:175.*
33. Tourdot, S., A. Scardino, E. Saloustrou, D. Gross, S. Pascolo, P. Cprdopatis, F. Lemonnier, and K. Kosmatopoulos. 2000. A general strategy to enhance immunogenicity of low-affinity HLA-A2.1-associated peptides: implication in the identification of cryptic tumor epitopes. *Eur. J. Immunol. 30:3411.*
34. Foss, F. 2002. Immunologic mechanisms of antitumor activity. *Semin. Oncol. 29:5.*