Identifying MHC Class I Epitopes by Predicting the TAP Transport Efficiency of Epitope Precursors

Björn Peters,[1]* Sascha Bulik,* Robert Tampe,† Peter M. van Endert,‡ and Hermann-Georg Holzhütter*

We are able to make reliable predictions of the efficiency with which peptides of arbitrary lengths will be transported by TAP. The pressure exerted by TAP on Ag presentation thus can be assessed by checking to what extent MHC class I (MHC-I)-presented epitopes can be discriminated from random peptides on the basis of predicted TAP transport efficiencies alone. Best discriminations were obtained when N-terminally prolonged epitope precursor peptides were included and the contribution of the N-terminal residues to the score were down-weighted in comparison with the contribution of the C terminus. We provide evidence that two factors may account for this N-terminal down-weighting: 1) the uncertainty as to which precursors are used in vivo and 2) the coevolution in the C-terminal sequence specificities of TAP and other agents in the pathway, which may vary among the various MHC-I alleles. Combining predictions of MHC-I binding affinities with predictions of TAP transport efficiency led to an improved identification of epitopes, which was not the case when predictions of MHC-I binding affinities were combined with predictions of C-terminal cleavages made by the proteasome. ***The Journal of Immunology,*** **2003, 171: 1741–1749.**

Cytotoxic T cells of the vertebrate immune system are able to discriminate between normal and abnormal cells (possessing an altered gene expression or containing viral proteins) on the basis of the repertoire of peptides (T cell epitopes) presented on the cell surface by MHC class I (MHC-I).[2] Before being presented as a T cell epitope, a peptide has to pass a number of consecutive filtering processes involved in the MHC-I presentation pathway (reviewed in Ref. 1). Most of the MHC-I-presented peptides, or longer peptide precursors of them, are cut out from intracellular proteins by proteasomes. Peptides generated by the proteasome have to escape from proteolytic attack of various cytosolic proteases before they are transported into the endoplasmic reticulum (ER) by TAP. Within the ER, peptides may undergo N-terminal trimming, whereas their C terminus is kept intact (2, 3). The ER aminopeptidase associated with Ag processing responsible for this trimming was recently identified (4–6). Peptides with correct size and proper amino acid sequence motifs bind to MHC-I, and the receptor-peptide complex is transferred to the cell surface. Peptides in the ER with less efficient MHC-I binding are either degraded there or exported for rapid degradation in the cytosol.

It is a challenging goal to provide computational methods that allow predicting T cell epitopes from the large pool of peptides with 8–11 residues that in principle can be derived from intracellular proteins. Such prediction tools would be useful for several immunological applications, including the intelligent design of peptide vaccines. The first prediction tools developed for individual steps of the MHC-I presentation pathway were established for MHC-I binding (7), which appears to be the most selective step. Moreover, several algorithms to predict proteasomal cleavages exist (8–10).

The experimental basis for previous attempts to predict TAP transports were the in vitro affinity assays (11, 12), in which the affinity has been shown to correspond closely with the transport rate of TAP (12). Thus, in this study we took the correspondence of TAP transport and affinity to be exact, allowing us to equate predictions of TAP affinity with predictions of TAP transport. Predictions of TAP affinity (13) were limited to 9-mers and exhibited large allele-specific differences in their ability to discriminate presented epitopes from random 9-mers. This has led to the conclusion that either MHC-I alleles are loaded by different degrees of TAP-independent transport (14) or varying amounts of epitopes are transported as N-terminal prolonged precursors (15). The second reasoning has been shown to be true for several epitopes (15–17) and receives further support by the identification of the aminopeptidase responsible for the N-terminal trimming of precursors in the ER.

To extend the TAP affinity predictions to longer peptides, we introduce a new scoring method using only the C terminus and the tree N-terminal residues of a peptide. The scores at these positions are taken from scoring matrices measured for 9-mer peptides. The validity of this procedure is demonstrated for a set of longer peptides. We then examine the possible in vivo selective effect of TAP by comparing the predicted TAP transport efficiency of known human MHC-I epitopes with that of random 9-mers encoded by the same set of host proteins. Because the actual precursors generated in vivo are only known in a few cases, the effective TAP transport rate of a peptide is estimated by taking the average score of all of its potential N-terminal precursors up to a maximal length L. This provides a significantly better discrimination between epitopes and random sequences than does looking at the transport of the 9-mer alone. Surprisingly, the discrimination further improves when the TAP score for the N terminus is down-weighted by a factor of 0.2 in comparison with the score of the C terminus. This finding was confirmed by an independent dataset consisting

*Institut für Biochemie, Charite, Humboldt Universität, Berlin, Germany; †Institut für Biochemie, Biozentrum, Goethe-Universität Frankfurt, Frankfurt, Germany; and ‡Institut National de la Santé et de la Recherche Médicale, Institut Necker, Paris, France

Received for publication January 28, 2003. Accepted for publication May 19, 2003.

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] Address correspondence and reprint requests to Dr. Björn Peters, Institut für Biochemie, Charite, Humboldt Universität, Monbijoustrasse 2, 10117 Berlin, Germany. E-mail address: bjoern_peters@gmx.net

[2] Abbreviations used in this paper: MHC-I, MHC class I; ER, endoplasmic reticulum; SMM, stabilized matrix method; AUC, area under curve; ROC, receiver operating characteristic.

Copyright © 2003 by The American Association of Immunologists, Inc. 0022-1767/03/$02.00

of mouse epitopes. We propose two main explanations for the necessary down-weighting of the N terminus in TAP predictions: 1) the uncertainty about which epitope precursors are actually present in vivo and 2) the possible coevolution of TAP with other agents in the presentation pathway. Simple computer simulations of the pathway suggest that the first hypothesis can only partially account for the down-weighting of the N terminus.

Finally, we tested the identification of epitopes for a single MHC-I allele (HLA-A0201) with a combined prediction of TAP- and MHC-I-binding affinities. These combined TAP and MHC-I predictions give better results than the already very accurate MHC-I binding predictions alone. In comparison, combining proteasomal cleavage predictions with MHC-I binding leads to worse results than does relying on MHC-I binding predictions alone.

Materials and Methods

Stabilized matrix method (SMM)

The matrix element $s_{i,a}$ denotes the score of amino acid a at sequence position i of the peptide. The total score S_k for a given peptide k with the amino acids $a_k(i)$ at position i is given by the following summation:

$$S_k = s_0 + \sum_i s_{i,a_k(i)}$$

where s_0 is a constant offset. The values for $s_{i,a}$ and s_0 are determined by minimizing the distance between predicted scores S_k and measured values for a set of training peptides:

$$\Phi(\{s_{i,a}\}) = \sum_k (S_k - \text{measured}_k)^2$$

To avoid overfitting, a second term is added to the minimization function:

$$\Psi(\{s_{i,a}\}, \lambda) = \Phi(\{s_{i,a}\}) + \lambda \sum_{i,a} s_{i,a}^2$$

By minimizing this objective function with a nonzero λ value, a tradeoff is introduced between optimally reproducing the experimental values (including their inevitable experimental error) and minimizing the parameters $s_{i,a}$. This forces all parameters that do not significantly lower the distance ϕ toward small values. The optimal value λ_{opt} for λ is determined by cross-validation on the training set.

This mathematical concept is quite commonly used when solving inverse problems, where λ is called the regularization parameter (a short introduction to inverse problems is given in Ref. 18, chapter 18).

Datasets

The epitopes were extracted from the SYFPEITHI database (19). All human and mouse 9-mer MHC-I epitopes for which a source sequence was available and for which the epitope is found exactly once in that sequence were included in the database. We did not include MHC-I ligands, which are known to bind but which are not presented naturally, or epitopes derived from signal sequences. The database was subdivided into three separate datasets. The HLA-X dataset consists of 203 human 9-mer epitopes from 87 sequences that are not presented by HLA-A0201. The H2-X dataset consists of 67 9-mer mouse epitope from 56 sequences, and the HLA-A0201 dataset consists of 87 human epitopes from 51 sequences.

The TAP affinities of peptides were measured as described in Ref. 13. We removed all 9-mer peptides that were present in any of the epitope datasets described above, with 430 9-mer peptides and 64 longer peptides remaining.

Statistical significance for differences in area under curve (AUC)

To assess whether one prediction is significantly better than another, we resampled the set of peptides for which predictions were made. Using bootstrapping with replacement, 50 new datasets were generated, with a constant ratio of epitopes to random 9-mers. We then calculated the difference in AUC for the two predictions on each new dataset. One prediction is significantly better than another if the distribution of the differences in AUC values is significantly above 0, which we measure using a standard t test with a p value of 0.001.

MHC-I pathway simulations

Using the protein sequences from which the epitopes of the HLA-X dataset originate, we randomly generated a set of m fragments per sequence obeying a log-normal length distribution as observed for the cleavage products of the proteasome (20). These m fragments per sequence are considered to be the pool of peptides generated by the proteasome that contain a C-terminal 9-mer, which can bind to an MHC-I molecule. Which of these fragments becomes an epitope is decided by their affinity to TAP, which we calculate using Equation 1. The fragment with the highest affinity per sequence is considered to be an epitope precursor, defining an epitope with its last nine down-stream residues. The other $m - 1$ fragments are discarded. We then tried to identify these artificially generated epitopes among all other 9-mers contained in the protein sequences by applying the TAP transport score (Equation 3) at varying values of α. In all simulations, the best prediction of epitopes was achieved with α values smaller than one.

There are three free parameters in the simulation: the number m of different fragments used to define a single epitope and the mean and SD of the log-normal length distribution of peptides generated. The larger the value of m, the higher the selective power that TAP has in the pathway in comparison with the proteasome and the MHC-I molecules. By systematically increasing the value of m, we found that with $m = 10$ the AUC value for the discrimination between epitopes and nonepitopes on the basis of the TAP score for the C terminus alone was close to those AUC values in Figs. 4 and 5 observed with real experimental data. The length dependence of the AUC values was in good concordance with that shown in Figs. 4 and 5 when choosing the mean of the log-normal length distribution in the range 9–11.

Proteasomal cleavage prediction

The NetChop algorithm is implemented online at www.cbs.dtu.dk/services/NetChop. There are different NetChop prediction methods available, trained on different datasets. We used NetChop 20S, which is trained on in vitro protein digests. PaProc is also available online at www.paproc.de. We used the wild type III method, which was trained on the largest dataset. The FragPredict method currently is not available online, but it is as a computer program distributed on request by H.-G. Holzhütter. We only used its cleavage site prediction algorithm and neglected its capability to predict fragment formation to make it comparable with the other two methods.

Results

Prediction of TAP affinities for 9-mers

We have used four scoring matrices like the one shown in Table I, which have been designed to predict TAP affinities of 9-mer peptides (expressed in terms of log(IC_{50}) values). Two of the scoring matrices were taken from the literature and have been derived directly from experiments. The Ala-matrix was constructed by using the peptide AAASAAAAY as a reference and measuring IC_{50} values for the peptides possessing an exchanged amino acid at one of the nine sequence positions (12, 21). The mix-matrix was generated using libraries of 9-meric peptides $X_1X_2 \ldots Y \ldots X_9$, where X_i stands for a mixture of different amino acids and Y is a specific amino acid occupying a fixed sequence position (11). These libraries compete in binding with the totally randomized peptide library $X_1X_2 \ldots\ldots X_9$. The third scoring matrix (SMM matrix) was established by means of the recently developed SMM (32). The matrix was determined using a set of peptides by minimizing the distance between their measured IC_{50} values and those predicted by the scoring matrix. In contrast with previous approaches, this method takes into account the inevitable noise contained in the experimental data by preferring a "smoother" matrix to a matrix perfectly (over)fitting the training data. A description of SMM is given in *Materials and Methods*. By averaging over the three scoring matrices, we generated the consensus matrix (Table I), which is expected to give better predictions of TAP IC_{50} values than the individual matrices because their prediction errors can partially compensate for each other.

To evaluate the quality of these four scoring matrices, we compared their predictions for a set of 430 peptides with experimentally known IC_{50} values. The resulting scatter plots are shown in

Table I. *Consensus scoring matrix*[a]

Amino Acid	(N1) Position 1	(N2) Position 2	(N3) Position 3	Position 4	Position 5	Position 6	Position 7	Position 8	(C) Position 9	Amino Acid
A	−1.56	−0.25	−0.10	0.24	−0.10	0.17	0.27	0.00	0.55	A
C	0.05	−0.01	−0.02	0.11	0.09	0.05	0.00	−0.13	0.00	C
D	1.37	1.42	1.83	−0.23	0.33	0.32	1.07	0.32	1.83	D
E	1.65	0.02	1.51	0.08	0.54	−0.13	0.64	0.44	1.58	E
F	1.03	−0.45	−1.05	−0.50	−0.26	0.08	−0.50	0.17	−2.52	F
G	0.28	1.14	1.70	0.45	0.66	0.12	1.41	−0.38	1.41	G
H	0.21	0.33	−0.23	−0.21	−0.11	−0.06	−0.19	0.39	0.55	H
I	−0.11	−0.49	−0.62	−0.09	−0.42	−0.75	−0.94	0.45	−0.52	I
K	−1.03	−0.41	0.09	−0.23	−0.08	−0.26	0.44	0.12	−0.45	K
L	−0.50	0.09	−0.11	0.11	−0.34	0.02	−0.73	0.01	−0.94	L
M	−0.38	−0.46	−0.58	−0.35	−0.26	0.30	−0.64	−0.11	−0.29	M
N	−1.43	0.69	1.01	0.38	0.49	−0.27	0.16	0.33	1.33	N
P	1.43	3.00	0.22	−0.04	−0.72	−0.13	−0.84	0.03	−0.09	P
Q	0.47	−0.97	0.39	0.15	0.15	−0.07	0.34	0.26	0.12	Q
R	−1.34	−1.47	−0.42	−0.27	−0.32	−0.75	−0.09	−0.42	−1.47	R
S	−0.56	−0.34	0.11	0.27	0.45	0.31	0.87	−0.51	2.26	S
T	−0.12	−0.04	0.43	0.23	0.43	0.49	0.39	−0.46	0.72	T
V	−0.49	−0.50	−0.71	0.27	0.37	−0.02	−0.29	0.10	−0.30	V
W	0.54	−0.64	−1.65	−0.18	−0.78	0.31	−0.50	−0.63	−0.87	W
Y	0.50	−0.67	−1.80	−0.18	−0.13	0.28	−0.87	0.02	−2.91	Y

[a] The matrix was generated by averaging over the Ala-, Mix-, and SMM-matrix described in the text and normalized by setting the mean of each column to zero. The matrix elements represent $\log(IC_{50})$ values for TAP binding that add up to the $\log(IC_{50})$ value of the peptide. A low matrix entry at a given position corresponds to an amino acid well suited for TAP binding.

Fig. 1. With all matrices, a significant correlation between predicted and measured IC_{50} values was obtained. As expected, the consensus matrix gives the best results, although the SMM-matrix is only marginally worse.

Prediction of TAP affinities for longer peptides

The above scoring matrices have been derived on data for 9-mers. To predict IC_{50} values for peptides with more than nine residues, we took advantage of the fact that binding of peptides to TAP is mainly determined by the C-terminal and the three N-terminal residues (11, 13, 22, 23). Hence, we neglected the influence of "inner" residues and calculated TAP affinities of peptides with arbitrary length by scoring only the C terminus and the three N-terminal residues using the four corresponding columns of the 9-mer matrices. Thus, for a peptide with the amino acid sequence N1, N2, N3, N4, . . . , C, the TAP score t is given by

$$t = mat_{1,N1} + mat_{2,N2} + mat_{3,N3} + mat_{9,C} \quad (1)$$

where $mat_{i,X}$ denotes the score of residue X at sequence position i. To test how well Equation 1 predicts TAP affinities of peptides with more than nine residues, we applied it to 64 peptides with lengths between 10 and 18 aa. As shown in Fig. 2, the correlation between predicted and measured affinity values is lower than for the 9-mers, but it is still significant. The consensus matrix again provided higher correlation than all other matrices, so that it was used in all further applications.

Using TAP affinity predictions for the identification of epitopes

To assess the selective role of TAP within the MHC-I presentation pathway, we tested how well the TAP affinity scores can separate known 9-meric MHC-I epitopes from random 9-mers. We included all 9-meric epitopes contained in the SYFPEITHI database (19) that are presented naturally by any human MHC-I allele, except those presented by HLA-A0201 (which are used later on), and for which the sequence of the source protein is available. We did not include MHC-I ligands (which are known to bind but which are not presented naturally) or epitopes derived from signal sequences. All other 9-mers contained in the protein sequences from which the epitopes originated were taken as random control peptides (nonepitopes). We will refer to this set of epitopes and random 9-mers as the HLA-X dataset.

This definition of epitopes and nonepitopes makes the SYFPEITHI database the gold standard in this work. One has to be aware that several 9-mers will be falsely classified as nonepitopes because the SYFPEITHI database is not complete. This fact could influence the results of our analyses only if the rate of unidentified epitopes (which is expected to be low) becomes comparable with the misclassification rates of the prediction methods used. In the case that future experimental work unexpectedly reveals this premise to be true, a critical revision of our results will be required.

To measure prediction quality, we used receiver operating characteristic (ROC) curves (24). For a given cutoff value that separates peptides by their predicted TAP affinity into potential epitopes and nonepitopes, the two variables *sensitivity* (true positives/total positives) and *1-specificity* (false positives/total negatives = false alarm rate) are calculated. By systematically varying the cutoff value from the lowest to the highest predicted score, a ROC curve like the one shown in Fig. 3 is plotted. Prediction quality is measured by the AUC, which is 0.5 for random predictions and 1.0 for perfect predictions. The AUC is equivalent to the probability that the score of a randomly chosen epitope is lower (better) than that of a randomly chosen nonepitope. This measure has the advantage of not relying on a single arbitrarily chosen cutoff value for the prediction score. Using the complete 9-mer consensus matrix, the resulting AUC value for the HLA-X dataset is 0.702 (Fig. 3, *curve a*), indicating a relevant but not very good prediction.

We repeated this analysis by including potential epitope precursors carrying N-terminal extensions. TAP affinities for N-terminal precursors of length 9, 10, . . . , L were calculated for all epitopes and nonepitopes by means of Equation 1. The TAP transport score $\bar{t}$ of a potential 9-mer epitope is obtained by averaging

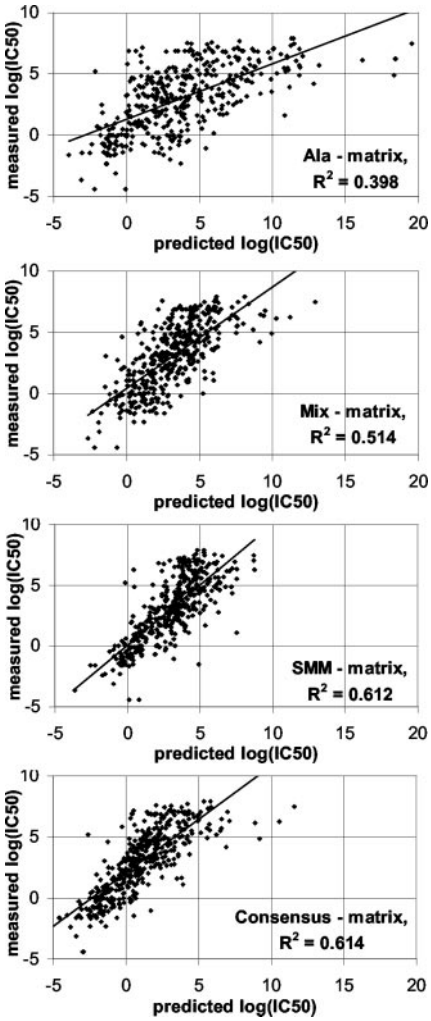

FIGURE 1. Comparison of predicted and measured in vitro TAP affinity values of 9-mer peptides. The scatterplots depict the observed log(IC_{50}) values of 430 9-meric peptides vs predicted log(IC_{50}) values using the scoring matrix indicated at the *bottom right* of each panel. The solid curves represent linear regression lines. Because the SMM matrix is calculated from the IC_{50} values of the same 430 peptides, its scores were determined using fivefold cross-validation, i.e., splitting the total set of peptides into five subsets, establishing the scoring matrix on only four of these subsets, and then using the matrix to make predictions for the subset that was left out.

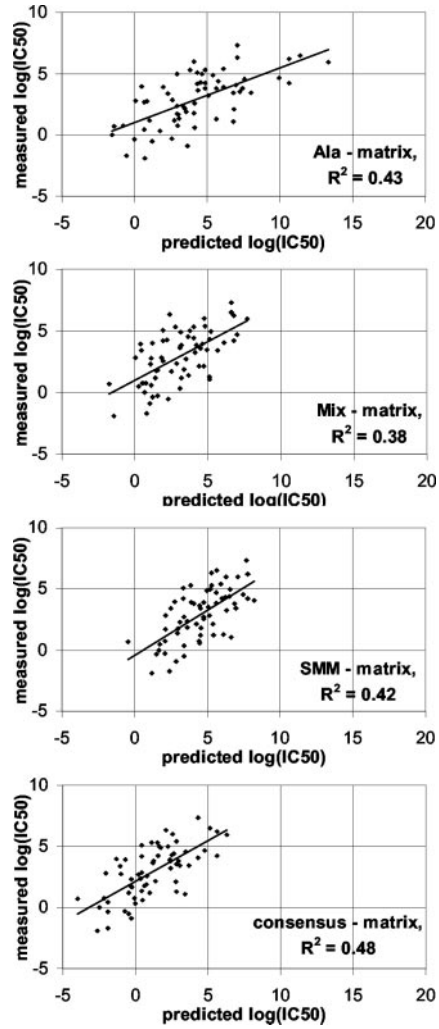

FIGURE 2. Comparison of predicted and measured in vitro TAP affinity values for peptides longer than 9 aa. The scatterplots depict the observed log(IC_{50}) values of 64 peptides vs theoretical log(IC_{50}) values predicted by using the scoring matrix indicated at the *bottom right* of each panel. The length distribution of peptides was as follows: 36 10-mers, 18 11-mers, six 12-mers, and one each of 13-, 15-, 16-, and 18-mers. The solid curves represent linear regression lines. The predicted log(IC_{50}) values were calculated by applying the scoring matrices derived for the 9-meric peptides (cf Fig. 1) to the three N-terminal residues and the C-terminal residue of each peptide as described by Equation 1.

over the calculated affinities of all these precursors up to a maximal length L:

$$\bar{t}_L = \frac{1}{L-8}\sum_{l=9}^{L} t(precursor_l)$$

$$= mat_{9,C} + \frac{1}{L-8}\sum_{l=9}^{L} mat_{1,N1} + mat_{2,N2} + mat_{3,N3} \qquad (2)$$

Note that all precursors contribute to the transport score with identical C termini, whereas the N-terminal contributions vary. Successively increasing the maximal number L of allowed N-terminal extensions and using the corresponding TAP transport scores to discriminate between epitopes and nonepitopes, we obtained the AUC values depicted in Fig. 4 (*curve a*). For $L = 9$ (no N-terminal extension), Equation 2 is equivalent with Equation 1 and the AUC value amounts to 0.700, which is marginally lower than the value 0.702 obtained when using the complete scoring matrix. This finding further justifies the usage of Equation 1. As shown in Fig. 4, the AUC values increase significantly for values of $L > 9$, demonstrating that inclusion of precursor peptides into TAP scoring improves the prediction of epitopes. (In *Materials and Methods*, we explain how we evaluate the significance of differences in AUC values.)

The increase of AUC values for $L > 18$ was not expected, because the TAP transport efficiency for peptides exceeding this length has been shown to drop off significantly (25). Evidently, increasing step by step the possible length L of epitope precursors, the statistical average across their N-terminal scores will converge against a stable limit value, thus rendering N-terminal scoring more and more meaningless for the prediction of TAP affinities. Hence, in the limit $L \to \infty$, only the C terminus will account for differences in the TAP scores of different 9-mers. To see how close we are to this limit, we calculated the AUC using the C terminus

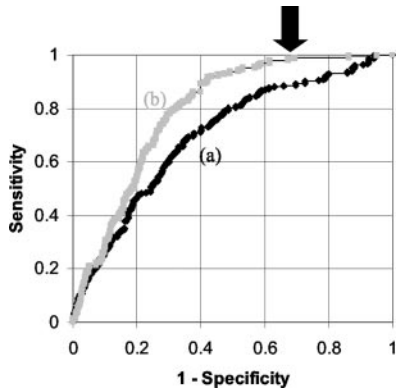

FIGURE 3. ROC curves for the HLA-X dataset. Each point of the curve belongs to a cutoff value for the TAP transport score, separating the peptides of the HLA-X dataset into good TAP substrates and bad TAP substrates. Equating good TAP substrates with epitopes and bad TAP substrates with nonepitopes, one may set up a two-by-two contingency table. Sensitivity is defined as the fraction of epitopes predicted to be good TAP substrates. Specificity is defined as the fraction of nonepitopes (random peptides in our studies) predicted to be bad TAP substrates. *Curve a* uses the TAP affinities predicted by the consensus matrix for the 9-mers themselves as the transport score, resulting in an AUC value of 0.702. *Curve b* uses Equation 3 as the transport score with $\alpha = 0.2$ and $L = 10$, i.e., allowing for precursors up to length 10 and down-weighting the contribution of the N-terminal residues (AUC = 0.791). The improvement is almost completely in the high sensitivity region. The arrow indicates the point in *curve b* that corresponds to the sensitivity and specificity reached when choosing the cutoff value = 1, which is used later in the combined TAP and MHC-I predictions.

for scoring only (Fig. 4, *curve b*). To our surprise, the AUC value of 0.782 is higher than the AUC values obtained before. This finding raises the question of whether the rise in AUC values seen with increasing length L of precursors does really reflect the usage of longer precursors in Ag production or whether the N-terminal scores are just adding noise to the prediction, which is smoothed out with increasing L. To check this, we weighted TAP transport scores of the N-terminal residues by a factor α:

$$\bar{t}_{L,\alpha} = mat_{9,C} + \frac{\alpha}{L-8}\sum_{l=9}^{L} mat_{1,N1} + mat_{2,N2} + mat_{3,N3} \qquad (3)$$

In Fig. 4, *curve a* corresponds with $\alpha = 1$ and *curve b* corresponds with $\alpha = 0$. If the increase in AUC values obtained with precursors $L > 9$ is only an artifact, one would expect the AUC for all values of L to grow monotonously when decreasing α from 1 to 0. If not, one would expect to find the optimal value of α somewhere between 1 and 0. The latter case is true: a maximum value of AUC was obtained for $\alpha = 0.2$ (Fig. 4, *curve c*), which was significantly above the AUC value obtained when only scoring the C terminus. *Curve b* in Fig. 3 depicts the ROC obtained when choosing $L = 10$ (i.e., one N-terminal extension) and $\alpha = 0.2$. Hence, predicting TAP affinities of N-terminally extended epitope precursors by down-weighting their N-terminal scores in comparison with their C-terminal scores significantly improves the discrimination between epitopes and nonepitopes. This indicates that such precursors indeed play an important role in Ag presentation. We will analyze possible explanations for the down-weighting of the N terminus below.

To exclude that the improved scoring obtained when choosing $\alpha < 1$ is a specific property of the HLA-X dataset, we applied the same scoring procedure to a completely independent set of mouse

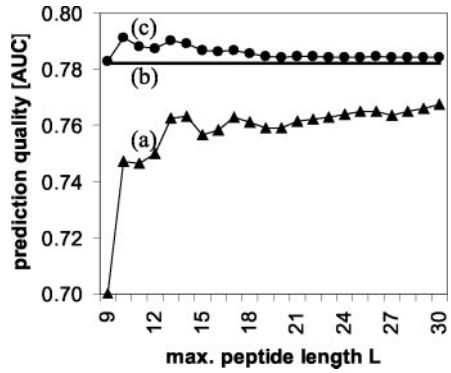

FIGURE 4. Prediction quality for the HLA-X dataset as a function of the maximal precursor lengths. Plotted is the prediction quality measured by the AUC of the TAP transport score given in Equation 3 for different predictions. *a*, Equal weight for N and C terminus ($\alpha = 1$); *b*, C terminus score only ($\alpha = 0$); *c*, optimal prediction with down-weighted N terminus ($\alpha = 0.2$).

epitopes (H2-X dataset). We used the human TAP matrices to separate epitopes from random 9-mers (cf Fig. 5). It has been shown that there are significant differences between the murine and human TAP specificities (26), because human TAP translocates peptides with hydrophobic and basic C termini, whereas mouse TAP prefers only peptides with hydrophobic C termini. As expected, this results in generally lower AUC values than those for the HLA-X dataset. Nevertheless, qualitatively the three curves *a–c* are related to each other in exactly the same way as those shown in Fig. 4 for the HLA-X dataset. Using the scores for the N and C termini with equal weights ($\alpha = 1$) for the prediction of TAP affinities results in a worse discrimination between epitopes and nonepitopes than that from neglecting the N terminus completely ($\alpha = 0$). Again, a better prediction of epitopes on the basis of their TAP affinities is achieved when the scores for the N terminus are down-weighted with $\alpha = 0.2$.

TAP affinity predictions for individual MHC-I alleles

We repeated the calculations from the previous section for the individual MHC-I alleles that make up the HLA-X dataset to see how the results vary. We restricted this analysis to those alleles for which at least 10 epitopes are present in the HLA-X dataset. First, we studied how well the epitopes of each individual allele can be identified by TAP affinity scores computed without inclusion of

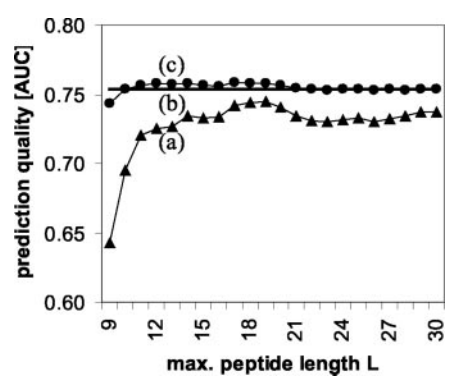

FIGURE 5. Prediction quality for the H2-X dataset as a function of the maximal precursor lengths. Plotted is the prediction quality measured by the AUC of the TAP transport score given in Equation 3 for different predictions. *a*, Equal weight for N and C terminus ($\alpha = 1$); *b*, C terminus score only ($\alpha = 0$); *c*, better prediction with down-weighted N terminus ($\alpha = 0.2$)

possible precursors or down-weighting of the N-terminal residues (i.e., putting $L = 9$ and $\alpha = 1$ in Equation 3). The resulting AUC values (Table II) show huge variations from 0.39 to 0.89. The differences in prediction quality for the individual alleles correspond very well with those in Refs. 13 and 14, where the alleles HLA-B27, -A3, and -A24 were classified as efficient for TAP loading and the alleles HLA-B07, B08, and A0201 were classified as inefficient for TAP loading.

Repeating the AUC calculations with the optimal parameters $L = 10$ and $\alpha = 0.2$ obtained for the entire HLA-X dataset, the AUC values fell in a much narrower range between 0.71 and 0.88, i.e., a subdivision into TAP-efficient and TAP-inefficient alleles is no longer preserved. These results provide evidence that TAP plays an equally important role for peptide loading of all alleles considered. Intriguingly, some alleles such as HLA-B27 or HLA-A3 seem to be preferentially loaded with peptides directly imported from the cytosol, whereas other alleles such as HLA-0201 are preferentially loaded with peptides entering the ER as N-terminally extended precursors where they are cut to final size.

Finally, we calculated the optimal value of α for each individual allele when setting $L = 10$. The resulting values varied between 0 and 4, showing that the optimal value of α is extremely allele specific: the better the C-terminal residues required for effective TAP transport agree with those C-terminal residues enabling effective MHC-I binding, the lower the weight that has to be put on the N-terminal residues. The optimal value of $\alpha = 0.2$ for the whole HLA-X dataset shows that, on the average, C-terminal amino acid motives required for effective TAP transport and MHC-I binding overlap more strongly than the corresponding N-terminal motives do. This is probably due to the stronger force for coevolution on the C terminus, which undergoes no change from TAP transport to MHC-I binding, whereas the N terminus can be trimmed.

Consequences of the uncertainty about which N-terminally extended precursors are generated in vivo

Another explanation of why better epitope predictions were achieved with $\alpha < 1$ is the uncertainty about which epitope precursors are actually transported in vivo to liberate the definitive epitope in the ER by N-terminal trimming. Equation 3 is based on the unrealistic assumption that up to a critical length L, all N-terminally prolonged precursors of an epitope are present in comparable abundance. Given that several precursor are not generated in vivo, their score for the N terminus will dilute that of the existent precursors. From the statistical point of view, this would favor putting a higher weight on the score of the C terminus or, equivalently, down-weighting scores of the N-terminal residues.

To estimate the implications of precursor uncertainty for the choice of α, we have performed simplified simulations of the MHC-I pathway (a detailed description of these simulations is given in *Materials and Methods*). We randomly generated an artificial set of epitopes possessing N-terminally prolonged precursors with good TAP affinities and obeying a log-normal length distribution as reported for fragments generated by the 20S and 26S proteasomes (20). We then tried to reidentify these epitopes among a set of random peptides using the TAP score described above. The highest AUC values were indeed obtained when choosing $\alpha < 1$. Fig. 6 shows the AUC values for such a simulated dataset. In this case, the highest AUC value was obtained for $L = 11$ and $\alpha = 0.6$. Varying the width of the hypothetical length distribution in these simulations, the optimal α values were always between 0.6 and 0.9, i.e., larger than the value $\alpha \approx 0.2$ yielding the best prediction of epitopes on real experimental datasets but always <1.

Table II. *Individual alleles*

Allele	No. of Epitopes	AUC ($L = 9$, $\alpha = 1$)	AUC ($L = 10$, $\alpha = 0.2$)	Optimal α for $L = 10$
HLA-B35	10	0.39	0.80	0.0
HLA-B07	11	0.43	0.71	0.0
HLA-B08	10	0.69	0.80	0.0
HLA-B44	11	0.78	0.88	0.0
HLA-A24	37	0.81	0.87	1.0
HLA-A3	11	0.82	0.75	1.2
HLA-B27	20	0.89	0.77	4.0
HLA-A0201	87	0.65	0.70	0.4

Identification of epitopes by combining predictions of TAP transport efficiency with predictions of MHC-I affinity

We tested whether the combination of predictions for two main steps of the presentation pathway, TAP transport and MHC-I binding, may improve the identification of epitopes. These calculations were performed on a set of 87 HLA-A0201-presented epitopes that had been omitted from the HLA-X dataset. For the prediction of peptide binding to HLA-A0201, an SMM-type scoring matrix was recently established on a dataset comprising the IC_{50} values of 533 nonepitope peptides.[3] This scoring matrix possesses a high capacity to identify epitopes (AUC = 0.919; cf Fig. 7, *curve a*). To combine predictions of MHC-I binding scores with predictions of TAP transport, we first calculated TAP binding scores for all 9-mers contained in the source sequences of the HLA-A0201 epitopes using Equation 3 with the parameters $L = 10$ and $\alpha = 0.2$, which have given the best results for the HLA-X dataset. We then classified all 9-mers with TAP scores above the threshold value 1 as not transportable and excluded them from the set of epitope candidates. This cutoff value was chosen by examining the ROC curve for the HLA-X dataset: only 1.5% of epitopes, but 32% of random 9-mers, have a higher (worse) TAP score (Fig. 3, arrow).

In the second step, the predicted MHC-I binding scores of the remaining peptides (having TAP scores <1) were used to discriminate between epitopes and nonepitopes. Based on this two-step prediction protocol, the AUC value increases significantly to 0.932 (Fig. 7, *curve b*). As elsewhere in the manuscript, the significance of the difference between the AUC values was assessed using a t test with $p = 0.001$, as described in *Materials and Methods*. The improvement is largest in the high sensitivity region. Demanding

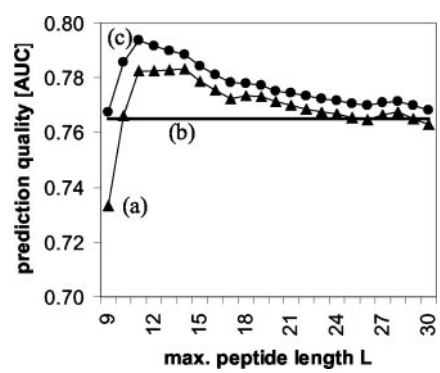

FIGURE 6. Prediction quality on a simulated dataset. The simulated dataset was generated as described in *Materials and Methods* with $m = 10$ and a log normal fragment length distribution centered around 11. Plotted is the prediction quality measured by the AUC of the TAP transport score given in Equation 3 for different predictions. *a*, Equal weight for N and C terminus ($\alpha = 1$); *b*, C terminus score only ($\alpha = 0$); *c*, optimal prediction with down-weighted N terminus ($\alpha = 0.6$).

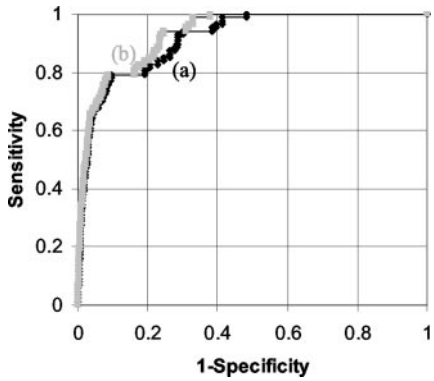

FIGURE 7. ROC curves for the combined TAP and MHC-I prediction on the HLA-A0201 dataset. The curve with black diamonds (*a*; AUC = 0.919) shows the (very high) level of the MHC-I prediction alone. The consistently better curve with gray squares (*b*; AUC = 0.932) is made by classifying all 9-mers with a TAP transport score worse than 1 as not transported ($\alpha = 0.2$, $L = 10$) and limiting the MHC-I prediction to the transported peptides.

100% sensitivity, specificity is increased from 52 to 62% when using the combined prediction instead of MHC-I affinity prediction alone, which is equivalent to a 20% drop in the number of false positives. This drop in the rate of false positives increases from 5 to 25% with rising sensitivity over the range of the ROC curve. This systematic increase can be explained by the fact that, due to similar peptide preferences of TAP and MHC, the likelihood for a peptide with high affinity to HLA-A0201 to possess at least a decent TAP transport rate is higher than for a peptide with low affinity to HLA-A0201.

We repeated the same two-step prediction for several mouse MHC-I alleles using scoring matrices for the MHC-I affinity prediction that were measured by Udaka et al. (27). Unfortunately, the number of epitopes available per allele is small (ranging from 9 to 21). For three of the alleles, the combined predictions gave better AUC values than did MHC-I affinity predictions alone (data not shown); for one allele the combined prediction was worse. This shows that the combined MHC-I + TAP prediction using a human TAP matrix works for mouse epitopes, even though there are significant differences between the murine and human TAP specificities. This should improve further when using a scoring matrix based on experimental data for murine TAP.

Identification of epitopes by combining predictions of C-terminal proteasomal cleavages with predictions of MHC-I affinity

As demonstrated above, predictions of epitopes buried within a set of protein sequences can be improved by using calculated TAP transport efficiencies as a filter that rules out poorly transported peptides without notably reducing the number of true epitopes in the remaining dataset. Hence, it was obvious to check whether the same strategy can be used to filter out those epitope candidates that are unlikely to be generated by the proteasome. Hitherto, there are no indications that the C terminus of proteasomal fragments undergoes further trimming along the MHC-I presentation pathway (28). Therefore, selecting potential epitopes and their N-terminally prolonged precursors by the probability that their C terminus is generated by the proteasome should single out false epitope candidates without losing true epitope candidates. To this end, we have used the three publicly available methods, NetChop (10), PaProc (29), and FragPredict (8) to predict C-terminal proteasomal cleavages. The PaProc method gives four different discrete scores (−, +, ++, and +++), whereas the other methods have a continuous output. First, we tested whether predictions of C-terminal cleavages alone may provide a reasonable discrimination between epitopes and nonepitopes. Fig. 8*a* shows the ROC curves for the three methods when applied to the HLA-X dataset. According to the AUC values, the best discriminations were achieved with NetChop, closely followed by FragPredict, and PaProc was significantly inferior to the other two prediction methods. Comparing the ROC curves of Fig. 8*a* with those of Fig. 3, it can be inferred that the discriminating power of existing prediction methods for proteasomal cleavage sites is far below that of TAP transport scoring developed in this paper.

We also tested combined predictions of C-terminal proteasomal cleavage and MHC-I binding, using the same two-step prediction protocol as described for TAP. For each of the three prediction methods of proteasomal cleavages, the cutoff value was chosen such that 30% of peptides on the HLA-X dataset had lower cleavage scores for their C-terminal residue and thus were singled out as "not generated." The same fraction of peptides is classified as nontransported by TAP. For PaProc, the fraction of omitted peptides was larger, in that this method predicted ~60% of peptide bonds to have the lowest score (−, not cleaved). The ROC curves for the combined predictions are shown in Fig. 8*b*, all of them indicating that the combined predictions are significantly worse than those based on predictions of MHC-I binding affinities alone.

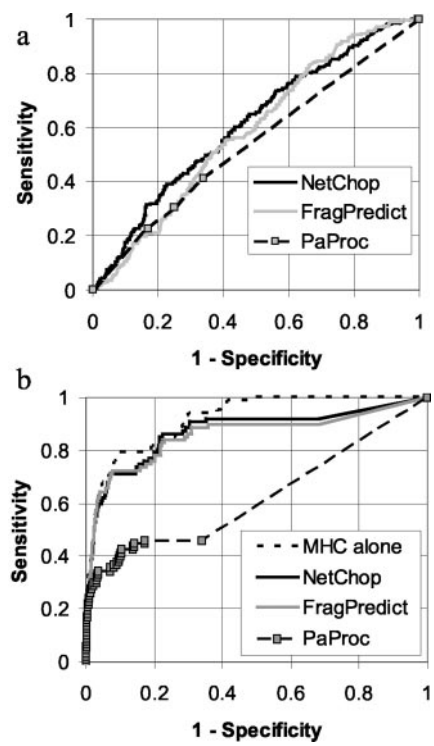

FIGURE 8. ROC curves for proteasomal cleavage prediction. *a*, For the three proteasomal cleavage prediction methods NetChop (AUC = 0.61), FragPredict (AUC = 0.59), and PaProc (AUC = 0.54), the score for C-terminal cleavage is used to predict epitopes from the HLA-X dataset. *b*, Combined proteasomal cleavage and MHC-I binding prediction. Potential epitopes with a score for proteasomal cleavage of their C terminus worse than a fixed cutoff are considered not cleavable; the others are scored according to their predicted MHC-I binding affinity. The best results are again obtained with NetChop (cutoff = 0.1, AUC = 0.872), followed by FragPredict (cutoff = 0.5, AUC = 0.858) and PaProc (cutoff = +, AUC = 0.623), all of which are worse than using MHC-I prediction alone (AUC = 0.919).

Discussion

In this paper, we have developed a method for predicting the TAP transport efficiency for peptides of arbitrary length. To this end, we exploited the fact that the terminal residues of a peptide determine most of its binding ability and accordingly used the scoring matrices derived from 9-mer peptide TAP affinity assays to score only the C terminus and the three outmost N-terminal residues. As demonstrated in Fig. 2, this procedure gives reasonably good predictions of TAP affinities for peptides of size 10–18.

Our interest in TAP affinities of peptides longer than 9 aa was raised because it recently became clear that several MHC-I epitopes are generated by N-terminal trimming of precursor peptides that are likely to be transported into the ER by TAP. The significant improvement of epitope predictions we found when considering transport of precursors strongly suggests that a large portion of epitopes is indeed processed this way. This underlines the pressure exerted by TAP on the selection of MHC-I epitopes.

Because the true in vivo precursors of an epitope are not generally known, we have used an effective TAP score by averaging across the scores of all precursors up to a certain length. Interestingly, the highest prediction quality was achieved when the scores of the N-terminal residues were down-weighted. We reasoned on the basis of simulations and results from scoring for individual MHC-I alleles that this down-weighting partially reflects coevolution of TAP and the average MHC-I allele as to the preference for certain C-terminal residues which epitope precursors are used in vivo.

Using predicted TAP transport efficiencies as a filter before prediction of MHC-I binding affinities, it was possible to further improve the already very accurate classification quality achieved using MHC-I affinity predictions alone. For example, a researcher interested in a list including all possible HLA-A0201 epitopes stemming from a given protein (100% sensitivity required) can expect a 20% decrease in the number of false positive suggestions when using the combined TAP + MHC predictions rather than the MHC predictions alone. Such a two-step prediction protocol, however, failed when predictions of C-terminal proteasomal cleavages were used as filter; i.e., relying on MHC-I affinity predictions alone gave better results than did combining them with proteasomal cleavage predictions. This disappointing result has three possible causes. One is that the selective power of the proteasome is weak, in that it generates nearly every possible peptide. Second, there might be other proteases serving as suppliers of antigenic peptides besides the proteasome. Finally, existing prediction algorithms of proteasomal cleavage sites might not be accurate enough. We favor the last explanation because in vitro digests of epitope-containing model substrates by the proteasome provide with very few exceptions the epitope or one N-terminally prolonged precursor (30). The poor quality of prediction algorithms for proteasomal cleavage sites is also evidenced by contradictory results obtained when applying them to the same set of test protein sequences. We think that the poor prediction quality of proteasomal cleavages is mainly caused by the lack of a sufficiently large set of quantitative and consistent experimental data on cleavage rates, which are more difficult to measure and interpret (31) than the affinity assays used to characterize peptide binding to TAP and MHC-I.

Summarizing our attempts to improve epitope predictions by a combination of different prediction tools, we have to conclude that currently the only reliable strategy for reducing a given set of epitope candidates without the risk of losing true epitopes is to filter out those peptides exhibiting poor TAP transport scores.

Another approach for predicting potential MHC-I epitopes by their amino acid sequence is to identify sequence motifs common to all epitopes presented by a specific MHC-I allele, as realized in the SYFPEITHI database (19). This approach does not differentiate between the influences of the proteasome, TAP or MHC-I on epitope selection, but has been shown to work well in practice. However, it has a principal drawback, because epitope sequences do not contain the full information used in the presentation pathway; the epitope may originate from a group of N-terminal prolonged precursors, generated by the proteasome, partially trimmed by cytosolic peptidases, transported by TAP into the ER, and then cut to final size. These steps preceding binding to the MHC-I receptor may depend on sequence motifs in the flanking regions up- and downstream of the generated peptide. Hence, we believe our approach to developing prediction algorithms for each individual step of the MHC-I presentation pathway and combining them to be superior. However, high quality experimental data for each step and advanced prediction techniques are needed to rival the prediction quality currently achieved by SYFPEITHI. Unfortunately, we cannot compare the predictive quality of the two approaches, because there is no independent blind set available. SYFPEITHI is trained on the data we use as test sets. For a neutral comparison, a significantly large set of newly identified naturally presented epitopes would be needed, or an older version of the SYFPEITHI prediction algorithm would have to be used and tested on more recently included epitopes. As a consequence, no conclusions about which method is currently better at identifying epitopes can be drawn here.

The improvements achieved when including TAP transport of precursors into epitope predictions are in the high sensitivity regime of the ROC curve (cf Fig. 7). It is often argued that high sensitivity of epitope predictions is of less practical relevance than having a high specificity, i.e., to end up with a short list of high probability epitope candidates for a given protein sequence. This view is wrong for two reasons. First, from the medical point of view, it can be equally interesting to identify all possible epitopes within a given protein sequence, requiring a high sensitivity of predictions. Second, when combining predictions for several steps of the MHC-I pathway whereby predictions of one step are used as a filter for the input to the next, it is very important to throw out as few true epitopes in each step as possible. Such a multistep prediction protocol automatically increases specificity from one step to the next.

There are two free parameters in our prediction of TAP transport scores: α and L. Throughout this paper, we used the optimal values $\alpha = 0.2$ and $L = 10$ determined for the HLA-X dataset containing epitopes from all human MHC-I alleles except HLA-A0201. These parameters show large variations when calculated for the individual alleles that make up the HLA-X dataset (Table II), because they are heavily influenced by each individual allele's binding preference. The parameter values for the entire HLA-X dataset, which averages out the individual allele's binding preferences, should reflect the true effect of TAP more accurately. The optimal parameter values calculated for the H2-X dataset ($\alpha_{opt} = 0.02$ and $L_{opt} = 11$) or the combined MHC-I and TAP predictions for HLA-A0201 ($\alpha_{opt} = 0.6$ and $L_{opt} = 18$), which should also reflect the true effect of TAP, are considerably different from those for the HLA-X set. However, the decrease in prediction quality when using $\alpha = 0.2$ and $L = 10$ instead of the optimal parameters for the individual datasets is quite small ($\Delta AUC < 0.006$) compared with the loss in prediction quality of $\Delta AUC \approx 0.100$ when making predictions without down-weighting and neglecting precursors (i.e., $\alpha = 1$, $L = 9$). Apparently, whereas down-weighting of the N terminus ($\alpha < 1$) and inclusion of precursors longer than the presented 9-mers ($L > 9$) improve prediction quality, the optimal values for α and L are not natural constants. Because the optimal

parameters $\alpha = 0.2$ and $L = 10$ determined on the largest dataset (HLA-X) lead to predictions close to the optimum on all datasets, we recommend their usage in all cases, although from a biological perspective $L = 10$ seems to be too small because longer precursors are known to be used in vivo.

When applying an epitope prediction protocol that is based on algorithms for several individual steps of the MHC-I presentation pathway, it is of utmost importance that each prediction algorithm is trained on a database containing only information on that specific step. For example, prediction methods that are supposed to predict MHC-I binding, but have been trained on data including epitope presentation implicitly take into account side effects produced by TAP and the proteasome. A combination of such an "impure" MHC-I binding prediction with a prediction of TAP transport or proteasomal cleavage thus bears the risk of overestimating the role of TAP or the proteasome for the presentation pathway. For that reason, we consider scoring matrices derived from randomized peptide libraries to be the best estimation of MHC-I binding, unless a large amount of individual peptide binding data is available, as is the case for the HLA-A0201 allele. For the same reason, examining the cleavage determining amino acid motifs of the proteasomal cleavages by analyzing average epitope sequences and their flanking regions (NetChop 2.0) can be problematic. If a strong preference for a favorable residue at the C terminus of an epitope is detected, this could be the result of proteasomal cleavages, TAP preference, or being an average trait of MHC-I molecules.

Outlook

Our work shows that including N-terminally prolonged precursors of an epitope in TAP transport may greatly improve epitope predictions. This also makes clear that epitope prediction tools using only reported epitope sequences as an input, in the long run, will be inferior to prediction tools based on prediction algorithms for all steps of the entire presentation pathway that take into consideration the generation and successive procession of epitope precursors. The next step along this line is obviously to include the proteasome, for which no prediction algorithms with sufficiently high reliability are currently available. Accurate prediction of proteasomal fragments would lead to a further improvement of TAP transport predictions which then, instead of considering all precursors up to length L as equally probable, can be restricted to those precursors actually generated. Eventually, this should also make the parameter α obsolete, because there would be no uncertainty as to which precursors are generated, and coevolution between peptide specificities of the proteasome, TAP, and MHC-I would be included in the model.

Acknowledgments

We thank K. Udaka for supplying the affinity matrices for the mouse MHC alleles.

References

1. Shastri, N., S. Schwab, and T. Serwold. 2002. Producing nature's gene-chips: the generation of peptides for display by MHC class I molecules. *Annu. Rev. Immunol. 20:463.*
2. Craiu, A., T. Akopian, A. Goldberg, and K. L. Rock. 1997. Two distinct proteolytic processes in the generation of a major histocompatibility complex class I-presented peptide. *Proc. Natl. Acad. Sci. USA 94:10850.*
3. Serwold, T., S. Gaw, and N. Shastri. 2001. ER aminopeptidases generate a unique pool of peptides for MHC class I molecules. *Nat. Immunol. 2:644.*
4. Serwold, T., F. Gonzalez, J. Kim, R. Jacob, and N. Shastri. 2002. ERAAP customizes peptides for MHC class I molecules in the endoplasmic reticulum. *Nature 419:480.*
5. York, I. A., S. C. Chang, T. Saric, J. A. Keys, J. M. Favreau, A. L. Goldberg, and K. L. Rock. 2002. The ER aminopeptidase ERAP1 enhances or limits antigen presentation by trimming epitopes to 8–9 residues. *Nat. Immunol. 3:1177.*
6. Saric, T., S. C. Chang, A. Hattori, I. A. York, S. Markant, K. L. Rock, M. Tsujimoto, and A. L. Goldberg. 2002. An IFN-γ-induced aminopeptidase in the ER, ERAP1, trims precursors to MHC class I-presented peptides. *Nat. Immunol. 3:1169.*
7. Parker, K. C., M. A. 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.*
8. Holzhutter, H. G., C. Frommel, and P. M. Kloetzel. 1999. A theoretical approach towards the identification of cleavage-determining amino acid motifs of the 20 S proteasome. *J. Mol. Biol. 286:1251.*
9. Kuttler, C., A. K. Nussbaum, T. P. Dick, H. G. Rammensee, H. Schild, and K. P. Hadeler. 2000. An algorithm for the prediction of proteasomal cleavages. *J. Mol. Biol. 298:417.*
10. Kesmir, C., A. K. Nussbaum, H. Schild, V. Detours, and S. Brunak. 2002. Prediction of proteasome cleavage motifs by neural networks. *Protein Eng. 15:287.*
11. Uebel, S., W. Kraas, S. Kienle, K. H. Wiesmuller, G. Jung, and R. Tampe. 1997. Recognition principle of the TAP transporter disclosed by combinatorial peptide libraries. *Proc. Natl. Acad. Sci. USA 94:8976.*
12. Gubler, B., S. Daniel, E. A. Armandola, J. Hammer, S. Caillat-Zucman, and P. M. van Endert. 1998. Substrate selection by transporters associated with antigen processing occurs during peptide binding to TAP. *Mol. Immunol. 35:427.*
13. Daniel, S., V. Brusic, S. Caillat-Zucman, N. Petrovsky, L. Harrison, D. Riganelli, F. Sinigaglia, F. Gallazzi, J. Hammer, and P. M. van Endert. 1998. Relationship between peptide selectivities of human transporters associated with antigen processing and HLA class I molecules. *J. Immunol. 161:617.*
14. Brusic, V., P. van Endert, J. Zeleznikow, S. Daniel, J. Hammer, and N. Petrovsky. 1999. A neural network model approach to the study of human TAP transporter. *In Silico Biol. 1:109.*
15. Lauvau, G., K. Kakimi, G. Niedermann, M. Ostankovitch, P. Yotnda, H. Firat, F. V. Chisari, and P. M. van Endert. 1999. Human transporters associated with antigen processing (TAPs) select epitope precursor peptides for processing in the endoplasmic reticulum and presentation to T cells. *J. Exp. Med. 190:1227.*
16. Paz, P., N. Brouwenstijn, R. Perry, and N. Shastri. 1999. Discrete proteolytic intermediates in the MHC class I antigen processing pathway and MHC I-dependent peptide trimming in the ER. *Immunity 11:241.*
17. Goldberg, A. L., P. Cascio, T. Saric, and K. L. Rock. 2002. The importance of the proteasome and subsequent proteolytic steps in the generation of antigenic peptides. *Mol. Immunol. 39:147.*
18. Press, W. H., S. A. Teukolsky, W. T. Vetterling, and B. P. Flannery. 1992. *Numerical Recipes in C.* Cambridge University Press, Cambridge, U.K.
19. Rammensee, H., J. Bachmann, N. P. Emmerich, O. A. Bachor, and S. Stevanovic. 1999. SYFPEITHI: database for MHC ligands and peptide motifs. *Immunogenetics 50:213.*
20. Kisselev, A. F., T. N. Akopian, K. M. Woo, and A. L. Goldberg. 1999. The sizes of peptides generated from protein by mammalian 26 and 20 S proteasomes: implications for understanding the degradative mechanism and antigen presentation. *J. Biol. Chem. 274:3363.*
21. Daniel, S., S. Caillat-Zucman, J. Hammer, J. F. Bach, and P. M. van Endert. 1997. Absence of functional relevance of human transporter associated with antigen processing polymorphism for peptide selection. *J. Immunol. 159:2350.*
22. Uebel, S., and R. Tampe. 1999. Specificity of the proteasome and the TAP transporter. *Curr. Opin. Immunol. 11:203.*
23. van Endert, P. M., D. Riganelli, G. Greco, K. Fleischhauer, J. Sidney, A. Sette, and J. F. Bach. 1995. The peptide-binding motif for the human transporter associated with antigen processing. *J. Exp. Med. 182:1883.*
24. Bradley, A. P. 1997. The use of the area under the ROC curve in the evaluation of machine learning algorithms. *Pattern Recognition 30:1145.*
25. van Endert, P. M., R. Tampe, T. H. Meyer, R. Tisch, J. F. Bach, and H. O. McDevitt. 1994. A sequential model for peptide binding and transport by the transporters associated with antigen processing. *Immunity 1:491.*
26. Momburg, F., J. Roelse, J. C. Howard, G. W. Butcher, G. J. Hammerling, and J. J. Neefjes. 1994. Selectivity of MHC-encoded peptide transporters from human, mouse and rat. *Nature 367:648.*
27. 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.*
28. Rock, K. L., and A. L. Goldberg. 1999. Degradation of cell proteins and the generation of MHC class I-presented peptides. *Annu. Rev. Immunol. 17:739.*
29. Nussbaum, A. K., C. Kuttler, K. P. Hadeler, H. G. Rammensee, and H. Schild. 2001. PAProC: a prediction algorithm for proteasomal cleavages available on the WWW. *Immunogenetics 53:87.*
30. Kessler, J. H., N. J. Beekman, S. A. Bres-Vloemans, P. Verdijk, P. A. van Veelen, A. M. Kloosterman-Joosten, D. C. Vissers, G. J. ten Bosch, M. G. Kester, A. Sijts, et al. 2001. Efficient identification of novel HLA-A(*)0201-presented cytotoxic T lymphocyte epitopes in the widely expressed tumor antigen PRAME by proteasome-mediated digestion analysis. *J. Exp. Med. 193:73.*
31. Peters, B., K. Janek, U. Kuckelkorn, and H. G. Holzhutter. 2002. Assessment of proteasomal cleavage probabilities from kinetic analysis of time-dependent product formation. *J. Mol. Biol. 318:847.*
32. Peters, B., W. Tong, J. Sidney, A. Sette, and Z. Weng. Examining the independent binding assumption for binding of peptide epitopes to MHC-I molecules. *Bioinformatics. In press.*