A Proposed Algorithm Predictive for Cytotoxic T Cell Alloreactivity

Monique M. Jöris, Jon J. van Rood, Dave L. Roelen, Machteld Oudshoorn and Frans H. J. Claas

*J Immunol* published online 18 January 2012
http://www.jimmunol.org/content/early/2012/01/18/jimmunol.1102086

**Subscription**
Information about subscribing to *The Journal of Immunology* is online at: http://jimmunol.org/subscription

**Permissions**
Submit copyright permission requests at: http://www.aai.org/About/Publications/JI/copyright.html

**Email Alerts**
Receive free email-alerts when new articles cite this article. Sign up at: http://jimmunol.org/alerts
A Proposed Algorithm Predictive for Cytotoxic T Cell Alloreactivity

Monique M. Jöris,*† Jon J. van Rood,*† Dave L. Roelen,† Machteld Oudshoorn,*† and Frans H. J. Claas†

Previously, we showed that with an increasing number of amino acid differences in single HLA class I-mismatched molecules, the probability of T cell alloreactivity decreases. It is unlikely that every amino acid difference will affect T cell alloreactivity in a similar way; we hypothesized that the effect of an amino acid difference may depend on its position and/or physicochemical properties. We selected 131 patient/donor pairs with either a single HLA-A or -C mismatch in the graft-versus-host direction and that were compatible for HLA-B, -DRB1, and -DQB1. The alloreactive CTL precursor (CTLp) frequency was determined and associated with the amino acid differences between the single HLA class I mismatches. In the β sheet, only amino acids that are noncompatible in their physicochemical properties affect T cell alloreactivity, whereas in the α helices, both compatible and noncompatible amino acids affect CTLp outcome. Positions 62, 63, 73, 80, 99, 116, 138, 144, 147, and 163 were bivariately associated with CTLp outcome, irrespective of the total number of amino acid differences. In multivariate analysis, positions 62, 73, 80, 116, 138, 144, and 163 were found to be most predictive for negative CTLp outcome. These results formed the basis for a weighted predictive mismatch score; pairs with the highest mismatch scores are estimated to be 13 times more likely to have a negative CTLp. This new algorithm may be a tool in donor selection for hematopoietic stem cell transplantation. The Journal of Immunology, 2012, 188: 000–000.

One of the major issues in hematopoietic stem cell transplantation (HSCT) is selection of a suitable donor when a fully HLA-matched donor is not available. A significant number of patients have to be transplanted with hematopoietic stem cells from a mismatched donor. The immunogenicity of MHC molecules may differ; therefore, the challenge is to select a donor for whom the HLA mismatch of the recipient is not very immunogenic. This is feasible, as confrontation with foreign MHC will not always lead to an alloreactive immune response (1). Previously, we have shown that a low CTL precursor (CTLp) frequency of the donor against the HLA mismatch of the recipient is associated with good clinical results of HLA class I-mismatched HSCT (2). Some MHC class I mismatches are indeed associated with low or undetectable CTLp frequencies, whereas high CTLp frequencies can be seen in other HLA class I-matched donor/patient pairs (1, 3, 4).

Previously, we investigated whether T cell alloreactivity could be predicted by the number of amino acid differences on the α helices and the β sheet of single MHC class I-mismatched molecules in 74 donor/patient pairs (5). Quantification of the alloimmune CTL response in vitro was obtained by the CTLp assay (6), because it has been proven to be clinically relevant in predicting alloimmune HSCT outcome (7–13). These studies focused on the amino acid differences on those parts of MHC class I molecules that are important for TCR contact and/or peptide binding. Motifs relevant for MHC–TCR interaction are predominantly located in the α helices, and peptide-binding residues are predominantly located in the β sheet (14–19). The preliminary results showed that T cell alloreactivity could be predicted to a certain extent. Single HLA-C mismatches with five or more amino acid differences on the α helices and five or more amino acid differences on the β sheet did not lead to T cell-mediated alloreactivity in vitro. So, with an increasing number of amino acid differences, the probability of a negative CTLp assay increased significantly. In that pilot study, the group of single HLA-A and -B mismatches was too small and included too many mismatches with few amino acid differences to enable generalization of these findings to all MHC class I molecules.

The humoral alloimmune response to HLA class I mismatches can be successfully predicted by the HLAMatchmaker algorithm (20, 21). As HLAMatchmaker takes only Ab-accessible sites of the HLA molecules into consideration, it is not a suitable tool to predict T cell alloreactivity (22). Previous attempts to associate predictive levels of cellular alloimmunity with HSCT outcome have been unsuccessful (23–25). So far, the actual outcome of the CTLp assay is used for donor selection in our center. The real challenge is to predict cellular alloimmune responses, especially because the CTLp assay is a complicated and time-consuming test. In the current study, we investigated a larger population of single HLA class I-mismatched pairs and attempt to extend our previous findings to other HLA class I molecules.

It is logical to assume that not every amino acid difference among MHC class I molecules may affect T cell alloreactivity in a similar way. Both position (specific positions, α helices, or β sheet) (26) and physicochemical properties (size, polarity, and charge) of the amino acids involved may play a role (27–29). We aim to develop an algorithm in which we incorporate all aspects of the amino acid differences to predict T cell alloreactivity against single MHC class I mismatches and to use this as a tool for single HLA class I-mismatched unrelated or related donor selection in HSCT.
Materials and Methods

**HLA genotyping and amino acid sequencing**

All donors and patients were typed at high resolution for HLA-A, -B, -C, -DRB1, and -DBQ1 as described previously (3). Briefly, PCR sequence-specific primer for high-resolution allele typing and sequence-based typing, for part of the HLA-C alleles, were used. Amino acid sequences were obtained by using the European Bioinformatics Institute Web site (30). MHC class I mismatches were examined for amino acid differences in the α1/2 domain, with specific interest in positions 50–85 and 138–179 in the α helices and positions 4–12, 21–28, 32–37, 94–102, 112–118, and 123–126 in the β sheet (5).

**Donor–patient pairs**

In total 164 donor/patient pairs registered by the Europondonor Foundation from 1990 to 2008, for whom successful CTLp assays were performed, were available for this study. The patients were treated at Dutch HSCC centers. The 150 unrelated donors originated from national or international donor registries, and 14 donors were related. All donor/patient pairs had a single HLA-A, -B, or -C mismatch in the graft-versus-host (GvH) direction and were compatible for HLA-DRB1 and -DQB1. Donor/patient pairs with single HLA class I mismatches without amino acid differences at the previously specified positions in the α helices or the β sheet (HLA-C*03:03–03:04 (n = 11), HLA-C*07:01–07:18 (n = 11), and HLA-B*35:02–35:04 (n = 1)) were excluded from analysis. The reason for this is that the amino acid differences in these combinations do not affect T cell recognition, as they are located on parts of the molecule that are not involved in the interaction with the TCR.

This study included only HLA-A (n = 55) and -C (n = 76) mismatched pairs, because single HLA-B (n = 20) mismatched pairs were rare and distinctive in the number and position of amino acid differences. Finally, 131 single MHC class I-mismatched donor/patient pairs were included in the analysis; the specific HLA mismatches are shown in Table I.

### CTLp assay

The CTLp assays were performed as described by Zhang et al. (6), with minor modifications as described by Oudshoorn et al. (3). CTLp assays were performed in the GvH direction. A negative CTLp outcome was defined as ≤1 recipient-specific CTLs/106 PBLs.

### Physicochemical properties of amino acids

We divided the amino acids into five groups according to their physicochemical properties (as listed in Table II). All amino acid differences of the single MHC class I-mismatched pairs were categorized into compatible (within-group) or noncompatible (between-group) amino acid differences. We hypothesized that compatible amino acid differences may have a different influence on T cell alloreactivity than noncompatible amino acid differences.

### Statistical analysis

We investigated the association of the number of all, noncompatible, and compatible amino acid differences on the α helices and β sheet (combined and separately) with the occurrence of a negative CTLp outcome. The

---

**Table I.** HLA class I mismatches (GvH direction) in the study population

<table>
<thead>
<tr>
<th>HLA-A (n = 55)</th>
<th>HLA-C (n = 76)</th>
</tr>
</thead>
<tbody>
<tr>
<td>01:01–03:01</td>
<td>02:01–03:01</td>
</tr>
<tr>
<td>01:01–24:02</td>
<td>02:01–68:01</td>
</tr>
<tr>
<td>01:01–68:01 (n = 2)</td>
<td>02:03–02:07</td>
</tr>
<tr>
<td>02:01–01:01</td>
<td>02:05–01:02 (n = 2)</td>
</tr>
<tr>
<td>02:01–02:05 (n = 3)</td>
<td>02:05–24:02</td>
</tr>
<tr>
<td>02:01–02:06 (n = 5)</td>
<td>03:01–11:01</td>
</tr>
<tr>
<td>02:01–02:11</td>
<td>11:01–03:01</td>
</tr>
<tr>
<td>02:01–03:01</td>
<td>11:01–26:01</td>
</tr>
<tr>
<td>02:01–11:01</td>
<td>11:01–68:01</td>
</tr>
<tr>
<td>02:01–23:01</td>
<td>11:01–03:01</td>
</tr>
<tr>
<td>02:01–26:01 (n = 2)</td>
<td>11:01–26:01</td>
</tr>
<tr>
<td>02:01–31:01</td>
<td>11:01–03:01</td>
</tr>
<tr>
<td>02:01–32:01 (n = 2)</td>
<td>24:02–23:01</td>
</tr>
<tr>
<td>02:01–33:01</td>
<td>24:02–24:01</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>HLA-C*03:03–03:04</th>
</tr>
</thead>
<tbody>
<tr>
<td>01:02–11:01</td>
</tr>
<tr>
<td>01:02–03:03 (n = 2)</td>
</tr>
<tr>
<td>01:02–03:04 (n = 2)</td>
</tr>
<tr>
<td>01:02–11:01</td>
</tr>
<tr>
<td>01:02–03:03 (n = 2)</td>
</tr>
<tr>
<td>01:02–03:04 (n = 2)</td>
</tr>
<tr>
<td>01:02–11:01</td>
</tr>
<tr>
<td>01:02–03:03 (n = 2)</td>
</tr>
<tr>
<td>01:02–03:04 (n = 2)</td>
</tr>
<tr>
<td>01:02–11:01</td>
</tr>
<tr>
<td>01:02–03:03 (n = 2)</td>
</tr>
</tbody>
</table>

---

**Table II.** Five groups of amino acids according to their physicochemical properties

<table>
<thead>
<tr>
<th>Group</th>
<th>Amino Acids</th>
</tr>
</thead>
<tbody>
<tr>
<td>Negatively charged</td>
<td>Asp (D), Glu (E)</td>
</tr>
<tr>
<td>Positively charged</td>
<td>Lys (K), His (H), Arg (R)</td>
</tr>
<tr>
<td>Polar/neutral</td>
<td>Gly (G), Thr (T), Ser (S), Gln (Q), Asn (N)</td>
</tr>
<tr>
<td>Apolar</td>
<td>Leu (L), Ile (I), Val (V), Met (M), Cys (C), Ala (A), Pro (P)</td>
</tr>
<tr>
<td>Large/apolar</td>
<td>Trp (W), Phe (F), Tyr (Y)</td>
</tr>
</tbody>
</table>

---

**FIGURE 1.** Statistical decision making.
amino acid difference variables were categorized into three groups (ter-
tiles) based on the distribution over the HLA mismatches. We chose to use
tertiles to ensure objectivity and even-sized categories. In all variables, the
group with the lowest number of amino acid differences was set as ref-
erence. Differences in discrete variables (all amino acid difference vari-
b[44x388][44x388]ables [in tertiles], HLA-DPB1 mismatching [yes/no], and killer cell Ig-like
receptor [KIR] ligand mismatching [yes/no]) between positive and nega-
tive CTLp outcome were examined with Pearson’s
x
x
2 test.
Odds ratios (OR), 95% confidence intervals (CI), and
p
p
values were
obtained using logistic regression analysis to quantify the association be-
tween the number of amino acid differences and negative CTLp outcome.
In addition, we analyzed each specific different amino acid position
(yes/no) separately to test for association with negative CTLp outcome. All
amino acid positions with a
p
p
value
#
0.200 were tested in bivariate
analysis correcting for the total number of amino acid differences. Those
that remained associated with negative CTLp outcome at a significant level
of
p
p
value
#
0.200 were selected for multivariate analysis.
Multivariate analysis using a backward stepwise approach (based on the
likelihood ratio test) was conducted to identify the most predictive model
for negative CTLp outcome. Under this approach, we start with fitting
a model with all of the amino acid positions of interest (identified in
univariate and bivariate analysis). Then, the least significant amino acid
position is dropped. We continue by successively refitting reduced models
and then reconsidering all dropped amino acid positions for reintroduction
into the model. This means that two separate significance levels must be
chosen for deletion from the model and for adding to the model (the entry
p
p
value is set at 0.050 and the removal
p
p
value at 0.100).
The
b
estimates (regression coefficients) of the amino acid positions
remaining in the final prediction model were used to define a weighted
predictive mismatch score. This mismatch score was also categorized into
three groups (tertiles) based on the distribution over the donor/patient pairs,
and the group with the lowest scores was set as reference. For both the
multivariate prediction and the final model, we included model evaluation
statistics and Somer’s D and C-statistics as a measure of association.
See Fig. 1 for statistical decision making. Two-sided
p
p
values
#
0.050 were considered statistically significant, and all analyses were performed
using SPSS 17.0 for Windows (SPSS Inc.).

Results
We tested whether T cell alloreactivity is dependent on position
and/or physicochemical properties of amino acid differences in
mismatched HLA class I molecules (Fig. 1). The study population
consisted of 55 donor/patient combinations mismatched for HLA-
A and 76 for HLA-C (Table I), of which 79 had a positive and 52
had a negative CTLp assay. The polymorphic amino acid positions are very different among these mismatched molecules. HLA-A molecules are more polymorphic in the α helices, and the HLA-C molecules are more polymorphic in the β sheet. Also, the total number of amino acid differences varies between the mismatches (median [range]: 11 [1–22]).

Of the 131 donor/patient pairs, 97 had either one (n = 66) or two (n = 31) mismatched HLA-DPB1 allele(s) in the GvH direction. For three donor/patient pairs, the HLA-DPB1 match grade was unclear due to missing typing of the donor. Of the 76 HLA-C mismatched donor/patient pairs, 19 had a KIR ligand mismatch in the GvH direction, and 57 were matched for their KIR ligands. Neither HLA-DPB1 nor KIR ligand mismatching was associated with CTLp outcome (data not shown).

With increasing numbers of amino acid differences in both the α helices and β sheet, the probability of a negative CTLp assay was statistically significant. In the α helices, all amino acid differences contribute to this effect (Fig. 2), whereas in the β sheet, only amino acid differences with noncompatible physicochemical properties do (Fig. 3). Therefore, only amino acid differences in the β sheet are categorized in compatible and noncompatible according to their physicochemical properties (Table II). Next, we combined the number of all amino acid differences in the α helices and only the number of noncompatible amino acid differences in the β sheet; the odds that the CTLp outcome is negative is almost five times larger for mismatches with ≥9 aa differences compared with mismatches with 0–5 aa differences (Fig. 4, Table III).

Twelve amino acid positions (62, 63, 73, 76, 77, 80, 99, 116, 138, 144, 147, and 163) in both α helices and β sheet were identified to be associated with negative CTLp outcome, even after correction for the total number of amino acid differences (data not shown). Eight amino acid positions (62, 63, 73, 80, 138, 144, 163, and 116) remained in the final model (after backward stepwise selection) and are therefore most predictive for negative CTLp outcome (Table IV). Their β estimates (regression coefficients) formed the basis of the weighted predictive mismatch score for negative CTLp outcome. Mismatched pairs with the highest mismatch scores are 13 times more likely to have a negative outcome of the CTLp assay (Fig. 5, Table V). Correction for the total number of amino acid differences by including this variable in the model did not alter the association.

### Discussion

Attempts to predict molecular interactions that lead to T cell alloreactivity have not yet resulted in clear-cut results. We retrospectively analyzed the association between T cell alloreactivity in vitro and single HLA-A and -C mismatches in 131 donor/patient pairs. Our results show that, compared with HLA alleles that are more similar to self-HLA alleles, divergent HLA alleles more often do not lead to T cell-mediated alloreactivity in vitro. This is in line with the previously proposed hypothesis that the possibility for TCR–MHC binding decreases if allogenic MHC differs too much from autologous MHC (5).

Next, we looked at amino acids that are (non)compatible in their physicochemical properties. In the β sheet, only amino acids that are distinctive in their physicochemical properties affect T cell alloreactivity negatively. This suggests that in contrast to compatible amino acids, noncompatible amino acid differences influence the peptide-binding repertoire in such a way that an MHC–peptide complex that is very distinct from self-MHC–peptide complexes prevents allorecognition. In the α helices, both compatible and noncompatible amino acids affect CTLp outcome, implying that in MHC–TCR interaction, all amino acid differences in the α helices have similar effects on allorecognition.

![Predictive mismatch score](http://www.jimmunol.org/DownloadedFrom/350x67-to-493x181.png)

**FIGURE 5.** Predicted association between the weighted predictive mismatch score and the probability of a negative CTLp outcome ($\chi^2 = 28.770; p = 0.000$).
Furthermore, we were able to identify a differential importance of amino acid differences at specific positions for the T cell-mediated alloresponse to HLA class I mismatches. Amino acids at position 62, 76, and 163 point up from the Ag binding site and are believed to be in direct contact with the TCR (14, 16). Amino acids at position 63, 73, 77, 80, 149, 99, and 116 point toward the Ag binding site and are believed to be involved in peptide binding (14). Position 63 and 99 form part of pocket A and B, 73 of pocket C, 147 of pocket E, and 77, 80, and 116 of pocket F (19, 31, 32), and residue 116 plays a key role in determining the specificity of the F pocket (32). Some of these positions and specific allele mismatches have been associated with transplant outcome, although not a single position or allele mismatch was consistently identified to be detrimental or favorable (23, 33–38). Differences in study design and/or population may be responsible for the inconsistent results.

The association between amino acid polymorphism and CTLp outcome that we found is not absolute and may be diluted because of the additional impact of noninherited maternal Ags (39), minor histocompatibility Ags, or indirect recognition of a peptide derived from the mismatched HLA molecule in one of the shared HLA molecules. Furthermore, the TCR repertoire of an individual is based on self-MHC and is therefore unique. Also, alloreactivity is often based on cross reactivity by virus-specific memory T cells (40). Together, this may explain the unpredictability of T cell responses.

The basis of T cell-mediated alloreactivity has been debated extensively and has generally resulted in two possible models, an MHC- and a peptide-driven mechanism. The first proposes that the alloreactive TCR directly recognizes polymorphisms in the allo-MHC molecules independently of the bound peptide and thereby adopting novel docking modes (41). The second suggests that responses to bound peptides that differ in sequence as well as peptides adopting different conformations when bound to the allo- or self-MHC molecules are most important in triggering alloreactivity (15, 42–44). If molecular mimicry is the basis of the alloimmune response, this would explain why increasingly divergent allo-MHC molecules are less likely to be targets for T cell-mediated cross-reactivity (45). Probably both mechanisms play a role (45–47), which would explain why the strongest effect is found when combining both the number of amino acid differences in α helices (TCR contact) and the β sheet (peptide-binding region).

In this study, we were able to reproduce previous findings (5) and extend them to single mismatched HLA-A molecules. Unfortunately, the number of single HLA-B mismatches was too low to show that a similar algorithm applies for HLA-B as well. Due to the relative low numbers, we were also not able to stratify for HLA-A and -C. Therefore, we could not test whether the mode of interaction with TCRs is similar for these two types of molecules. Although promising, these results need to be prospectively validated in a larger and extended population before the algorithm can be used for single HLA class I mismatched unrelated or related donor selection in HSCT. What these results do clearly show is that an HLA class I mismatch with few or small amino acid differences is not necessarily better than an HLA class I mismatch with numerous or large amino acid differences. The next step is to evaluate the clinical value of the weighted prediction mismatch score for transplant outcome.

**Acknowledgments**

We thank the employees of the Immunogenetics and Transplantation Immunology section of Leiden University Medical Center for HLA typing all donors and patients and performing the CTLp assays. We also thank Frits Koning and Jan Wouter Drijfhout for sharing expertise and advice regarding physicochemical properties of amino acids and Frits Koning for reading the manuscript and giving helpful comments.

**Disclosures**

The authors have no financial conflicts of interest.

**References**


frequency analysis in the selection of HLA matched unrelated donors for hematopoietic stem cell transplantation: the correlation of CTLp frequency with HLA class I genotyping and aGVHD development. Neoplasia 49: 26–32.


