Bone Marrow Transplantation (2004) 33, 165–169

www.nature.com/bmt

HLA typing

HistoCheck: rating of HLA class I and II mismatches by an internet-based software tool

H-A Elsner, D DeLuca, J Strub and R Blasczyk

Department of Transfusion Medicine, Hannover Medical School, Hannover, Germany

Summary:

HLA polymorphism is a major barrier for hematopoietic stem cell and solid organ transplantation. To estimate the allogeneic potential between HLA-mismatched stem cell donor/recipient pairs, we recently proposed a matching score (dissimilarity index) that is based on the structural data of HLA class I molecules, and on the functional similarity of amino acids (AA). This first approach revealed new features about presumptive subtype allogenicities within the HLA-A*23 and A*24 groups. We have now developed an internet-based software tool ('*HistoCheck*') that is capable to assess the allogenicity (matching score) between any pair of clinically relevant HLA class I, and also class II, alleles. Newly described HLA sequences will be regularly integrated into the database according to the nomenclature for factors of the HLA system updates. The software is intended to be a first step for estimating the allogenicity of HLA mismatches in peculiar clinical settings, as long as there are no reliable *in vitro* or clinical studies available. The algorithm can later be modified according to functional data, for example, peptide-binding specificities. With the extension of the sequence similarity concept to all clinically relevant HLA class I and II loci, *HistoCheck* may contribute to prevent HLA mismatching being a matter of chance.

Bone Marrow Transplantation (2004) **33,** 165–169. doi:10.1038/sj.bmt.1704301
Published online 1 December 2003
Keywords: *HistoCheck*; HLA; mismatch; rating; stem cell transplantation

The transplantation of hematopoietic stem cells is a potentially curative therapy for a variety of hematologic and nonhematologic diseases, whereby HLA matching of the donor and recipient is essential in order to reduce the risk of severe acute graft-versus-host disease (GVHD).[1–4]

Correspondence: Dr R Blasczyk, Department of Transfusion Medicine, Hannover Medical School, Carl-Neuberg-Str. 1, D-30625, Hannover, Germany; E-mail: blasczyk.rainer@mh-hannover.de
Received 23 March 2003; accepted 03 July 2003;
published online 1 December 2003

Despite a continuously growing number of potential stem cell donors and cord blood units, it may be impossible to find an HLA-compatible donor for carriers of rare antigens or unusual haplotypes. Furthermore, the advance of HLA class I typing at the allelic level will probably reveal far more subtype mismatches presently not detected. When there is no genotypically identical sibling and there are several alternative potential donors that all have a mismatch at an HLA class I or II locus, the allogenicity of mismatches may be estimated using the sequence-similarity matching (SSM) concept previously described by our working group.[5] In this concept, the AA differences between HLA alleles are evaluated and rated with regard to position within the molecule (peptide binding; contact with the T-cell receptor, TCR)[6,7] and with regard to the functional similarity of AA within proteins.[8] This procedure leads to an SSM score (allogenicity index), whereby high values represent high dissimilarity. When there are several mismatched donors, SSM scores may be calculated for any of them, and the donor with the least may be preferred. The SSM concept originally focused on the AA variations of the highly polymorphic antigen-presenting α1 and α2 domains in the HLA-A*23 and A*24 molecular groups,[5] and has now been extended to all HLA class I genes (A, B, C), and to the HLA class II genes DRB1, DQB1, and DPB1 within *HistoCheck*.

Material and methods

HLA structural data

Due to the strong structural homology between HLA class I molecules, the assignment of AA residues to functional regions, also of HLA-B and C, is based on X-ray crystallography data (PDB file for HLA-A*0201 as reported by Madden *et al*; MMDB: 1178, PDB: 1HHH; http://www.ncbi.nlm.nih.gov/entrez/query.fcgi?db=structure) of HLA-A2.[6,7] Moreover, the software also performs comparisons for the antigen-presenting β1 domain of the HLA class II molecules HLA-DR, DQ, and DP as encoded by the genes DRB1, DQB1, and DPB1 respectively, using the structural data (PDB file for DRB1*0101 as reported by Stern *et al*; MMDB: 764, PDB: 1DLH) of HLA-DR1.[9–11] Residues polymorphic between the donor and recipient are evaluated, and assigned either to regions of major (peptide-binding groove and/or region contacting the T-cell receptor) or minor

Table 1A HLA class I residues at polymorphic positions that contact the antigenic peptide and/or the TCR

Residue #	*Domain*	*Peptide binding*	*TCR contact*	*Residue #*	*Domain*	*Peptide binding*	*TCR contact*
7	α1	X	–	97	α2	X	–
9	α1	X	–	99	α2	X	–
24	α1	X	–	114	α2	X	–
45	α1	X	–	116	α2	X	–
58	α1	–	X	143	α2	X	–
59	α1	X	–	147	α2	X	–
63	α1	X	–	149	α2	–	X
65	α1	–	X	150	α2	X	X
66	α1	X	X	151	α2	–	X
67	α1	X	–	152	α2	X	–
69	α1	X	X	156	α2	X	–
70	α1	X	–	158	α2	–	X
72	α1	–	X	159	α2	X	X
73	α1	X	–	162	α2	–	–
74	α1	X	–	163	α2	X	X
76	α1	X	–	166	α2	–	X
77	α1	X	–	167	α2	X	X
80	α1	X	–	170	α2	–	X
81	α1	X	–	171	α2	X	–
84	α1	X	–				

allogeneic potential (remaining AA residues). The assignment of positions is given in Tables 1A and 1B (HLA class I) and in Table 2 (HLA class II).

Functional similarity of amino acids

The similarity between the single pairs of exchanged AA is measured by the distance matrix, as proposed by Risler *et al*. The basic idea of Risler's score was that two distinct AA are less dissimilar, the more often they substitute each other in functionally related proteins. Accordingly, the fewer a pair of AA substitutes each other, thus representing functional dissimilarity, the higher the score yielded, with the maximum value of 100.[8]

Algorithm of the SSM score

In order to obtain the SSM score, the scores representing the single AA exchanges are added up, as shown below:

$$\sum_{i}^{n1}\frac{Ri}{100}+\sum_{j}^{n2}\left[1+\frac{Rj}{100}\right]$$

where $i,\ldots,n1$ are mismatching AA without a critical funtion, with R_i being the corresponding Risler score. Similarly, $j,\ldots,n2$ are mismatching AA which are important for peptide binding and/or TCR contact, with R_j being the corresponding Risler's score. Residues that contribute to both peptide binding and contact with the TCR are rated as if they possess only one function. The described algorithm results in higher SSM scores even for low disparity at functionally relevant positions compared to any disparity at positions that are probably of functional minor importance.[5]

Results and discussion

In *HistoCheck*'s input menu (Figure 1), the HLA gene of interest can be chosen via a pull-down menu, and the alleles to be compared may be selected for the recipient, and for up to two donors. *HistoCheck*'s databases contain HLA sequences from the HLA Informatics Group (http://www.anthonynolan.org.uk/HIG) when doing compatibility checks. Newly described HLA sequences will be regularly integrated into the database according to the updates of the nomenclature for factors of the HLA system.

On the results page (Figure 2), *HistoCheck* indicates the positions of exchanges, and verifies whether these positions can be assigned to the peptide-binding site or to the region

Table 1B Polymorphic positions composing the HLA class I specificity pockets

Pocket A
7, 59, 63, 66, 70, 99, 159, 163, 167, 171
Pocket B
7, 9, 24, 45, 63, 66, 67, 70, 99
Pocket C
9, 70, 73, 74, 97
Pocket D
99, 114, 156, 159
Pocket E
97, 114, 147, 152, 156
Pocket F
77, 80, 81, 84, 116, 143, 147

Tables 1A and 1B show the assignment of residues polymorphic in HLA-A, B or C, to the α1 or α2 domains, to the peptide-binding site, the specificity pockets, and to the contact region for the TCR. In the column 'peptide binding' (Table 1A), those residues are indicated, which are solvent-accessible, and which thus potentially have contact with bound peptides. It corresponds to Table 14 of the study published by Saper *et al*.[6] In Table 1B, the assignment of residues to the respective pockets is given. Although that study contained no assignment for the polymorphic residues 69, 76 and 150, for *HistoCheck* these residues were nevertheless weighted as peptide-binding since they are potentially involved in peptide binding due to their solvent accessibility. The assignment to the contact region for the TCR is based on the study published by Garboczi *et al*.[7] In addition, it has to be noted that a residue may be assigned to a pocket, and at the same time also to the contact region for the TCR.

Table 2 HLA-DR1 residues of the β chain that contact the antigenic peptide and/or the TCR

Residue #	*Reference data*			*HistoCheck*	
	Peptide-binding site[a]	*Peptide-binding site*[b]	*TCR-binding site*[c]	*Peptide-binding site*	*TCR-binding site*
9	X	X	–	X	–
11	X	X	–	X	–
13	X	X	–	X	–
26	–	X	–	X	–
28	X	X	–	X	–
30	X	X	–	X	–
32	X	–	–	X	–
37	X	–	–	X	–
38	X	X	–	X	–
47	X	X	–	X	–
57	–	X	–	X	–
60	X	X	–	X	–
64	–	–	X	–	X
66	–	–	X	–	X
67	–	X	–	X	–
69	–	–	X	–	X
70	X	X	X	X	X
71	X	X	–	X	–
74	X	X	–	X	–
77	–	X	X	X	X
78	X	X	–	X	–
81	X	X	X	X	X
85	X	X	–	X	–
86	X	X	–	X	–
88	X	–	–	X	–
89	X	X	–	X	–
90	–	X	–	X	–

[a]Assignment according to the study by Brown *et al*.[9]
[b]Assignment according to the study by Stern *et al*.[10]
[c]Assignment according to the study by Hennecke *et al*.[11]
The assignment to the peptide or to the TCR-binding site, as given in the literature, is shown in the columns 'Reference data', whereby 'X' denotes assignment, and '–' denotes no assignment. When a residue was assigned to be part of the peptide-binding site in only one of the above studies,[9,10] for *HistoCheck* it was nevertheless assigned to this region (column '*HistoCheck*/peptide-binding site'). Since the assignment to the TCR-binding site was only based on a single study,[11] the study's and *HistoCheck*'s assignments are identical (columns 'Reference data/TCR-binding site' and '*HistoCheck*/TCR-binding site').

likely to have contact with the TCR. In the example given in Figure 2, for a recipient carrying the allele DRB1*110101, the donor with the allele DRB1*1103 may be preferred to a donor with DRB1*1102, since DRB1*1103 has a lower SSM score (2.60) than DRB1*1102 (3.24). In addition to the SSM calculations, *HistoCheck* visualizes the AA differences between the HLA molecules of the donor and recipient in a graph that is shown by the integrated RasTop (http://www.geneinfinity.org/rastop/) software, which is based on the computer program RasMol.[12] After the query, the RasMol (RSM) scripts containing the highlighted differences can be downloaded for further studies.

However, it is important to stress the present limitations of the software: The impact of single amino-acid mismatches may be underestimated by the present algorithm, since clinical data suggest that single mismatches, particularly at position 116, may have a disproportionately larger impact on the risk of the development of GVHD.[13] Moreover, since the assignment of positions is based on crystallographic data of HLA-A2 for class I and of HLA-DR1 for class II, in other HLA molecular variants, small pertubations in the main-chain dihedral angles can significantly alter molecular interactions at residues located several positions from the modified residue. Though not reflected in the algorithm, these alterations, in turn, can influence T-cell recognition, and thus have impact on the development of GVHD.

HistoCheck is intended to provide an easy-to-administer tool as a first step for comparing the allogenicity of different HLA class I and class II mismatches. Studies focusing on the systematic analysis of mismatches will be helpful in order to further improve the theoretical model proposed here. However, as long as such studies are not available, the present knowledge on the structure and function of MHC molecules, as in part reflected by the SSM score, may be utilized in order to bring HLA mismatching a step away from being a matter of chance. *HistoCheck* is available online under the address http://www.histocheck.de.

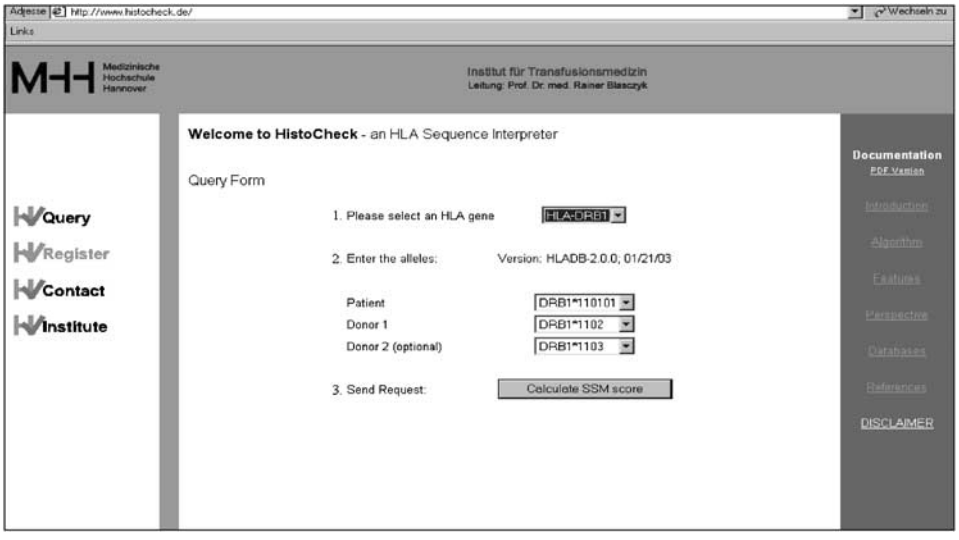

Figure 1 Screenshot of the input menu.

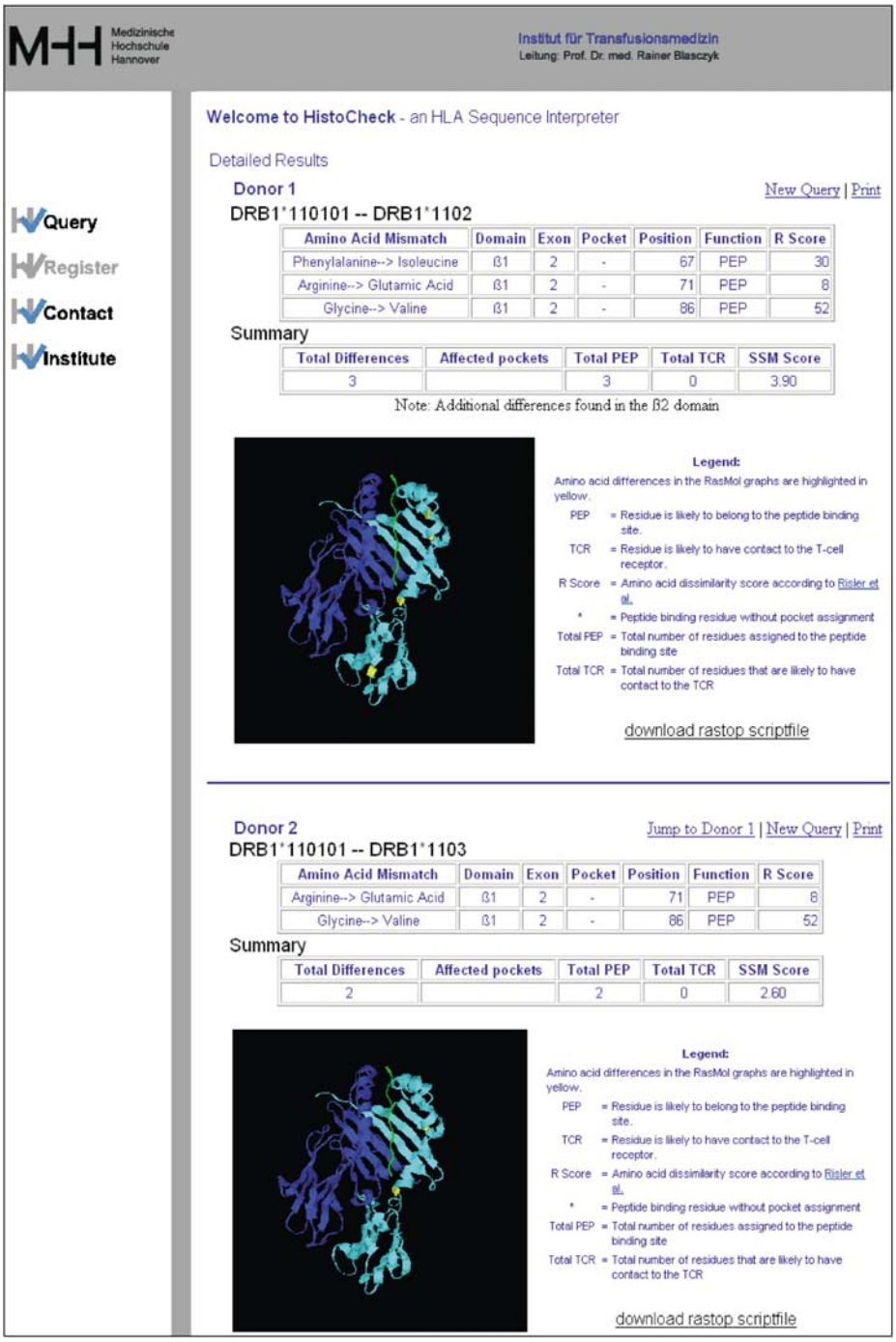

Figure 2 Screenshot of the results page. The figure has been constructed by superposition of three screenshots of the Results page.

References

1 Saba N, Flaig T. Bone marrow transplantation for nonmalignant diseases. *J Hematother Stem Cell Res* 2002; **11**: 377–387.
2 Sierra J, Storer B, Hansen JA *et al*. Unrelated donor marrow transplantation for acute myeloid leukemia: an update of the Seattle experience. *Bone Marrow Transplant* 2000; **26**: 397–404.
3 Petersdorf EW, Kollman C, Hurley CK *et al*. Effect of HLA class II gene disparity on clinical outcome in unrelated donor hematopoietic cell transplantation for chronic myeloid leukemia: the US National Marrow Donor Program Experience. *Blood* 2001; **98**: 2922–2929.
4 Hansen JA, Gooley TA, Martin PJ *et al*. Bone marrow transplants from unrelated donors for patients with chronic myeloid leukemia. *N Engl J Med* 1998; **338**: 962–968.
5 Elsner HA, Blasczyk R. Sequence similarity matching: proposal of a structure-based rating system for bone marrow transplantation. *Eur J Immunogenet* 2002; **29**: 229–236.

6 Saper MA, Bjorkman PJ, Wiley DC. Refined structure of the human histocompatibility antigen HLA-A2 at 2.6 A resolution. *J Mol Biol* 1991; **219**: 277–319.
7 Garboczi DN, Ghosh P, Utz U *et al*. Structure of the complex between human T-cell receptor, viral peptide and HLA-A2. *Nature* 1996; **384**: 134–141.
8 Risler JL, Delorme MO, Delacroix H, Henaut A. Amino acid substitutions in structurally related proteins. A pattern recognition approach. Determination of a new and efficient scoring matrix. *J Mol Biol* 1988; **204**: 1019–1029.
9 Brown JH, Jardetzky TS, Gorga JC *et al*. Three-dimensional structure of the human class II histocompatibility antigen HLA-DR1. *Nature* 1993; **364**: 33–39.
10 Stern LJ, Brown JH, Jardetzky TS *et al*. Crystal structure of the human class II MHC protein HLA-DR1 complexed with an influenza virus peptide. *Nature* 1994; **368**: 215–221.
11 Hennecke J, Carfi A, Wiley DC. Structure of a covalently stabilized complex of a human alphabeta T-cell receptor, influenza HA peptide and MHC class II molecule, HLA-DR1. *EMBO J* 2000; **19**: 5611–5624.
12 Sayle RA, Milner-White EJ. RASMOL: biomolecular graphics for all. *Trends Biochem Sci* 1995; **20**: 374.
13 Ferrara GB, Bacigalupo A, Lamparelli T *et al*. Bone marrow transplantation from unrelated donors: the impact of mismatches with substitutions at position 116 of the human leukocyte antigen class I heavy chain. *Blood* 2001; **98**: 3150–3155.