The basis of HLA-G mediated dysregulation of immune effector cells
Angenommen durch den Senat: 10.08.2018

Präsident: Prof. Dr. med. Christopher Baum

Wissenschaftliche Betreuung: Prof. Dr. med. Rainer Blasczyk

Wissenschaftliche Zweitbetreuung: Prof. Dr. rer. nat. Reinhard Schwinzer

1. Referent: Prof. Dr. med. Rainer Blasczyk
2. Referent: Prof. Dr. rer. nat. Reinhard Schwinzer
3. Referent: Prof. Dr. rer. biol. hum. Roland Jacobs

Tag der mündlichen Prüfung: 10.08.2018

Prüfungsausschuss

Vorsitz: Prof. Dr. rer. biol. hum. Roland Jacobs

1. Prüfer: Prof. Dr. med. Rainer Blasczyk
2. Prüfer: Prof. Dr. rer. nat. Reinhard Schwinzer
3. Prüfer: Prof. Dr. rer. biol. hum. Roland Jacobs
meiner Familie
Abstract

Alexander Celik: “The basis of HLA-G mediated dysregulation of immune effector cells”

The non-classical HLA class Ib molecule HLA-G adopts a specialized role in the immune system, especially during pregnancy where HLA-G expression prevents rejection of the embryo. Because the embryo constitutes a semi-allograft, localized suppression of the immune system is needed to allow for embryonic development. This is facilitated in part through the down regulation of classical HLA class Ia molecules and the simultaneous upregulation of the non-classical molecules HLA-E and predominantly HLA-G on extravillous trophoblasts in the placenta. HLA-G acts as a potent immune suppressor by inhibiting NK cell mediated lysis triggered by the lack of HLA class I surface expression, inducing T regulatory cells and apoptosis in CD8+ T cells as well as inhibiting the function of B and dendritic cells. Potentially due to these broad immune inhibitory functions, ectopic HLA-G expression was observed in many different tumor entities; particularly in Hodgkin’s lymphoma HLA-G expression is discussed as a potent immune evasion mechanism. HLA-G is the most comprehensive characterized HLA class Ib molecule, however the importance of allelic subtypes has not been systematically analyzed, yet. HLA-G*01:01, G*01:03 and G*01:04 are the most common alleles in Europe and these variants are differentiated from one another by single nucleotide polymorphisms that lead to the exchange of amino acids. The aim of this study was to characterize the determinants of HLA-G mediated dysregulation of immune effector cells. To understand the functional differences, cytotoxicity assays were performed that revealed an increased protective potential for HLA-G*01:04 when compared to the other two alleles. Sequence analysis of the presented peptides of all three variants demonstrated unique repertoires with marginal overlap between alleles. The observed functional differences can be clearly attributed to the alterations in peptide anchors that appear similar for HLA-G*01:01 and G*01:03, yet exhibit the absence of the auxiliary anchor Proline at peptide position p3 for G*01:04-restricted peptides. Modelling of identified peptides with and without Proline at p3 highlighted the constraints of HLA-G*01:01 induced peptides in comparison to the less confined conformation that G*01:04-restricted peptides can achieve. Because of its physiologically restricted tissue distribution, we additionally addressed the question whether the HLA-G*01:01 repertoire is affected by ectopic expression in a Hodgkin’s lymphoma cell line that exhibits similar features to the Hodgkin-Reed-Sternberg phenotype of classic Hodgkin’s lymphoma. Mass-spectrometric peptide analysis revealed a
bias within the range of possible amino acids at the p1 anchoring position depending on the parent cell line while the C-terminal anchor remained unaffected. This bias, however, could not be attributed to limitations of the proteomic content; the proteome of the cell lines was comparable suggesting that the recruitment of the respective HLA-G ligandome indicates a tissue specific capacity. To ensure that peptide loading was not impaired in these transformed cells, immunoprecipitation experiments were carried out that revealed comparable association with components of the peptide loading complex, however, association with TAP was absent. Transduction of a TAP-deficient cell line showed the expression of stable HLA-G molecules on the cell surface further reinforcing the capability of HLA-G to select peptides independently of TAP. The ability of HLA-G to recruit and present peptides in the absence of such integral parts of the peptide loading complex makes HLA-G an exquisite modulator during tumor immune evasion episodes. Combined with its differential inhibition against NK cells depending on the specific HLA-G allele, future typing of HLA-G might improve the outcome for HLA-G positive lymphoma patients and will guide towards individualized autologous treatment strategies.
Alexander Celik: “The basis of HLA-G mediated dysregulation of immune effector cells”

Das nicht-klassische HLA Klasse Ib Molekül HLA-G übernimmt vor allem während der Schwangerschaft eine besondere Rolle innerhalb des Immunsystems, da seine Expression die Abstoßung des Embryos verhindert. Der Embryo stellt für das Immunsystem einen Semi-Allograft dar, weshalb eine lokalisierte Immunsuppression notwendig ist, um die Entwicklung des Embryos zu gewährleisten. Dies geschieht zum Teil durch die Herabregulation von HLA Klasse Ia Molekülen zusammen mit der gleichzeitigen Hochregulation von nicht-klassischen HLA Klasse Ib Molekülen wie HLA-E und vor allem auch HLA-G auf extravillären Trophoblasten der Plazenta. HLA-G fungiert hier als potenter Immunsuppressor durch die Inhibition der NK-Zell-vermittelte Lyse, ausgelöst durch die fehlende Oberflächenexpression von HLA Klasse I Molekülen, aber auch durch die Induktion T regulatorischer Zellen und der Apoptoseinduktion in CD8⁺ Zellen. Darüber hinaus inhibiert HLA-G die Funktion von B- und dendritischen Zellen. Vermutlich auch aufgrund dieser breiten immun-inhibitorischen Wirkungen wurde eine ektopische HLA-G Expression in unterschiedlichen Tumorentitäten beobachtet und insbesondere im Hodgkin Lymphom wird die Expression von HLA-G als wichtiger Immunevasionsmechanismus diskutiert. HLA-G ist das am umfangreichsten charakterisierte HLA Klasse Ib Molekül, jedoch wurde der Einfluss allelischer Subtypen bisher nicht systematisch untersucht. HLA-G*01:01, G*01:03 und G*01:04 sind die häufigsten Allele in Europa, wobei sich die verschiedenen Varianten durch Einzelnukleotid-Polymorphismen unterscheiden die zum Austausch von Aminosäuren führen. Das Ziel dieser Arbeit war die Charakterisierung der Faktoren der HLA-G vermittelten Deregulation von Immunoeffektorzellen. Um die funktionellen Unterschiede zu verstehen wurden Zytotoxizitätsassays durchgeführt, die ein erhöhtes protektives Potential von HLA-G*01:04 im Vergleich zu den anderen beiden Alllen zeigten. Die beobachteten, funktionellen Unterschiede können eindeutig Änderungen in den Peptidankern zugeschrieben werden, welche für HLA-G*01:01 und G*01:03 Ähnlichkeiten aufweisen, in G*01:04-restringierten Peptiden jedoch durch die Abwesenheit des Hilfsankers Prolin an Peptidposition p3 gekennzeichnet sind. Die Modellierung der identifizierten Peptide, mit und ohne Prolin an p3, heben die konformellen Einschränkungen von HLA-G*01:01-restringierten Peptiden hervor im Vergleich zu der offeneren Konformation, die G*01:04-restringierte Peptide einnehmen können. Aufgrund der physiologisch eingeschränkten Gewebeverteilung wurde zusätzlich
untersucht in wie weit das Peptidrepertoire von HLA-G*01:01 bei ektopischer Expression in einer Hodgkin Lymphom-Zelllinie, welche ähnliche Eigenschaften zu einem Hodgkin-Reed-Sternberg-Phänotyp in klassischem Hodgkin Lymphom aufzeigt, beeinflusst wird. Durch massenspektrometrische Peptidanalyse wurde gezeigt, dass es zu einer präferenziellen Selektion von Ankeraminosäuren an p1 in Abhängigkeit der parentalen Zelllinie kommt, wohingegen der C-terminale Anker unbeeinflusst blieb. Da das Proteom beider Zelllinien vergleichbar ist, kann eine Präferenz der Ankerposition nicht durch Limitierungen im verfügbaren Proteom erklärt werden, was den Schluss nahelegt, dass die Rekrutierung der jeweiligen HLA-G-Ligandome gewebespezifisch erfolgt. Um weiter sicher zu stellen, dass die Ankerpräferenz nicht durch eine Beeinträchtigung bei der Peptidbeladung innerhalb dieser transformierten Zellen zustande kommt, wurden Immunpräzipitationsexperimente durchgeführt. Diese Experimente zeigten eine vergleichbare Assoziation mit den einzelnen Komponenten des Peptidladungskomplexes in beiden Zelllinien. Interessanterweise konnte in beiden Zelllinien keine Assoziation mit TAP nachgewiesen werden. Durch Transduktion einer TAP-defizienten Zelllinie konnte die Expression stabiler HLA-G Moleküle auf der Zelloberfläche gezeigt werden, was die Annahme verstärkt, dass HLA-G das Potential besitzt Peptide unabhängig von TAP zu selektionieren. Die Fähigkeit Peptide in der Abwesenheit eines solch wichtigen Teils des Peptidladungskomplexes zu rekrutieren und zu präsentieren, hebt die Bedeutsamkeit von HLA-G als wichtigen Modulator während Tumor-Immunevasionsepisoden hervor. In Kombination mit der vom spezifischen HLA-G Allel abhängigen, unterschiedlich starken Inhibition von NK-Zellen, könnte eine Typisierung von HLA-G das Behandlungsergebnis HLA-G-positiver Lymphompatienten verbessern und sollte auch im Hinblick auf individualisierte autologe Behandlungsstrategien als wichtiger Faktor berücksichtigt werden.
# TABLE OF CONTENTS

List of figures........................................................................................................................................ix

Abbreviations ..........................................................................................................................................x

1 Introduction .........................................................................................................................................1

1.1 Immune surveillance is enabled by the HLA system.................................................................1

1.1.1 Genomic organization of the HLA gene cluster.................................................................1

1.1.2 CD8\(^+\) T cells and NK cells monitor HLA class I molecules ................................3

1.1.3 The structure of HLA class I molecules determines antigen presentation.................4

1.1.4 HLA class I molecules present peptides of cellular origin ...........................................4

1.1.5 Peptide binding motifs are specific for HLA class I alleles ...........................................6

1.2 Non-classical HLA class Ib molecules .....................................................................................7

1.2.1 HLA class I histocompatibility antigen, alpha chain G .................................................7

1.2.2 HLA-G as a potent inhibitor of immune cell functions during pregnancy ..................9

1.2.3 Peptide presentation is restricted in HLA-G ..................................................................10

1.3 The role of HLA-G in malignancies .........................................................................................11

1.3.1 HLA-G expression is elevated in different tumor entities ..............................................11

1.3.2 Hodgkin’s lymphoma as a model of HLA-G mediated immune evasion ......................11

2 Project Aims .....................................................................................................................................13

3 Publications .......................................................................................................................................14

3.1 Research Paper 1 .......................................................................................................................14

3.2 Research Paper 2 .......................................................................................................................24

4 Results and Discussion .....................................................................................................................34

5 References .......................................................................................................................................38

6 Danksagung .....................................................................................................................................50

7 Curriculum Vitae .............................................................................................................................51
LIST OF FIGURES

Figure 1: Genomic organization of the HLA gene cluster. ...................................................... 2
Figure 2: HLA class I peptide loading is facilitated by the PLC. ............................................. 5
Figure 3: Structural features of HLA-G. ................................................................................. 8
# Abbreviations

<table>
<thead>
<tr>
<th>Abbreviation</th>
<th>Description</th>
</tr>
</thead>
<tbody>
<tr>
<td>AA</td>
<td>amino acid</td>
</tr>
<tr>
<td>APC</td>
<td>antigen presenting cell</td>
</tr>
<tr>
<td>β2m</td>
<td>beta-2-microglobulin</td>
</tr>
<tr>
<td>CD160</td>
<td>cluster of differentiation 160</td>
</tr>
<tr>
<td>CD4</td>
<td>cluster of differentiation 4</td>
</tr>
<tr>
<td>CD8</td>
<td>cluster of differentiation 8</td>
</tr>
<tr>
<td>CD94</td>
<td>cluster of differentiation 94</td>
</tr>
<tr>
<td>CDR</td>
<td>complementary determining region</td>
</tr>
<tr>
<td>CLL</td>
<td>chronic lymphatic leukemia</td>
</tr>
<tr>
<td>CNX</td>
<td>calnexin</td>
</tr>
<tr>
<td>CRT</td>
<td>calreticulin</td>
</tr>
<tr>
<td>CTL</td>
<td>cytotoxic T lymphocyte</td>
</tr>
<tr>
<td>DC</td>
<td>dendritic cell</td>
</tr>
<tr>
<td>DRiP</td>
<td>defective ribosomal protein</td>
</tr>
<tr>
<td>e.g.</td>
<td>exempli gratia</td>
</tr>
<tr>
<td>et al.</td>
<td>et alii</td>
</tr>
<tr>
<td>ER</td>
<td>endoplasmic reticulum</td>
</tr>
<tr>
<td>ERp57</td>
<td>endoplasmic reticulum resident protein 57</td>
</tr>
<tr>
<td>HL</td>
<td>Hodgkin’s lymphoma</td>
</tr>
<tr>
<td>HLA</td>
<td>human leukocyte antigen</td>
</tr>
<tr>
<td>sHLA</td>
<td>soluble HLA</td>
</tr>
<tr>
<td>HRS</td>
<td>Hodgkin-Reed-Sternberg</td>
</tr>
<tr>
<td>HSP</td>
<td>heat shock protein</td>
</tr>
<tr>
<td>i.e.</td>
<td>id est</td>
</tr>
<tr>
<td>IL</td>
<td>interleukin</td>
</tr>
<tr>
<td>ILT</td>
<td>Ig-like transcript</td>
</tr>
<tr>
<td>KIR</td>
<td>killer cell immunoglobulin-like</td>
</tr>
<tr>
<td>MHC</td>
<td>major histocompatibility complex</td>
</tr>
<tr>
<td>MM</td>
<td>multiple myeloma</td>
</tr>
<tr>
<td>NHL</td>
<td>non-Hodgkin lymphoma</td>
</tr>
<tr>
<td>NK</td>
<td>natural killer</td>
</tr>
<tr>
<td>NKC</td>
<td>natural killer complex</td>
</tr>
<tr>
<td>PBR</td>
<td>peptide binding region</td>
</tr>
<tr>
<td>PLC</td>
<td>peptide loading complex</td>
</tr>
</tbody>
</table>
ABBREVIATIONS

TAP  transporter associated with antigen processing
TCR  T cell receptor
TPN  tapasin
UGT1 UDP-glucuronosyltransferase 1
UTR  untranslated region

List of amino acids
A  Alanine
C  Cysteine
D  Aspartic acid
E  Glutamic acid
F  Phenylalanine
G  Glycine
H  Histidine
I  Isoleucine
K  Lysine
L  Leucine
M  Methionine
N  Asparagine
P  Proline
Q  Glutamate
R  Arginine
S  Sulfur
T  Threonine
V  Valine
W  Tryptophan
Y  Tyrosine
1 INTRODUCTION

1.1 Immune surveillance is enabled by the HLA system

The major histocompatibility complex (MHC) was first described as the genomic region that is responsible for the rejection of transplanted tissue grafts between different strains of mice. Further investigations revealed that the MHC encodes surface glycoproteins that present antigens in the form of small peptides. These peptides can be of self- or non-self-origin and are recognized by different subsets of T cells that are part of the adaptive immune system. The antigen recognition by T cells, however, is further restricted to a distinct set of MHC molecules, an observation that could eventually explain the graft rejection that is observed after transplantation between incompatible donors-recipient pairs (1, 2). In humans, this system was later named the human leukocyte antigen (HLA) system (3). It became apparent that the same system is pivotal in aiding the immune system with the recognition and subsequent clearing of aberrantly behaving cells, for instance, during pathogenic episodes when foreign pathogens invade the cell or in case of cellular mutations that lead to anomalous cell proliferation. The immune system differentiates between two classes of cell surface glycoproteins: HLA class I and HLA class II. HLA class I molecules are expressed on nearly all nucleated cells and platelets and are recognized by natural killer (NK) cells and cytotoxic T lymphocytes (CTL). CTLs are T cells expressing a T cell receptor (TCR) and the CD8 co-receptor that recognizes the α3 domain of HLA class I molecules. In accordance to their function, HLA class I molecules are continuously scanned by CD8+ T cells thus enabling constant monitoring of any change occurring inside the cells. HLA class II expression is limited to antigen presenting cells (APCs) such as dendritic cells (DCs), macrophages or B cells that process and present pathogenic antigens to CD4+ T cells. The β-subunit of HLA class II molecules is recognized by the CD4 co-receptor of T helper cells.

1.1.1 Genomic organization of the HLA gene cluster

The HLA gene cluster is located on the short arm of chromosome 6 (6p21.1- 6p21.3) and contains over 220 genes (4, 5), many of which encode genes related to the immune system. Based on their main effector types, the HLA gene cluster is divided into three classes (Figure 1). About 20 genes are encoded in the HLA class I region, encompassing the alpha chain of the classical HLA class Ia molecules HLA-A, HLA-B and HLA-C. Notably, the HLA class I light chain β₂microglobulin (β₂m) is encoded on chromosome 15 (15q21.1). Additionally, within the HLA class I locus, the non-classical HLA class Ib proteins HLA-E, HLA-F and HLA-
G are encoded alongside genes for the HLA class I-like genes MICA and MICB. The HLA class II region contains the α and β chains of the class II proteins HLA-DQ, HLA-DM, HLA-DR, HLA-DP as well as further proteins associated with the HLA system such as components of the peptide loading complex (PLC) (TPN, TAPBP, TAP1 or TAP2, LMP2, LMP7). Finally, at the HLA class III region, proteins of the complement system (C2, C4A/B, and CFB), heat shock proteins (HSP), immune receptors (G6b, G6d, G6f) and TNF genes (TNFA, LTA/B) as well as genes related to inflammation are encoded (5, 6). The HLA genes are among the most polymorphic in the human genome, with HLA class I molecules being especially polymorphic counting currently 12,893 different HLA class I alleles (5). Among these genes, greatest variability is observed within the HLA class Ia loci whereas HLA class Ib genes show very few polymorphisms.

Figure 1: Genomic organization of the HLA gene cluster. The HLA gene cluster on chromosome 6 is divided into three regions: class I, II and III. Reproduced with permission from (Klein and Sato, 2000), Copyright Massachusetts Medical Society (4).
1.1.2 CD8$^+$ T cells and NK cells monitor HLA class I molecules

To achieve effective immune surveillance HLA class I molecules can cover different immunological functions by interacting with CD8$^+$ T cells of the adaptive immune system and through interactions with NK cells that are part of the innate immune system. The presentation of antigenic peptides is recognized by both NK cells (7, 8) and CD8$^+$ T cells (9, 10). On NK cells, two major gene clusters are distinguished: the natural killer complex (NKC) and the leukocyte receptor complex (LRC). Both contain receptors that are known to interact primarily with HLA class I molecules (6). In the LRC region, two important classes of receptors are the killer-cell immunoglobulin (Ig)-like receptors (KIR) and the leukocyte Ig-like receptors (LILR, also known as Ig-like transcripts ILT). Although many KIRs have an inhibiting function, activating KIRs are concurrently expressed and in combination with other, different kinds of NK cell receptors (e.g. gene products of the NKC) a fine balance between activating and inhibiting signals will determine if the NK cells’ inherent cytotoxic activity is suppressed (11, 12). KIRs are able to distinguish HLA class I molecules complexed to self-peptides from HLA class I molecules complexed to pathogenic peptides. More importantly, KIRs become activated in case the inhibiting signal is lacking due to the missing self-signal provided by the peptide-HLA class I complex (pHLA) (13, 14), e.g. when HLA class I expression on the cell surface is down-regulated during pathogenic events.

In contrast to NK cell recognition, CD8$^+$ T cells become activated by specific recognition of presented non-self-peptides (10). To achieve that, CD8$^+$ T cells have to be able to differentiate many different possible antigens. Recognition is facilitated through the T cell receptor (TCR) that is comprised of two subunits. Most T cells express the $\alpha\beta$ TCR and about 2 - 15% of T cells express the $\gamma\delta$ TCR (15-17). Each receptor consists of a constant and a variable part that is important for antigen recognition. The variable part is encoded in the V, D and J gene segments. To build a receptor with a high degree of variability in the antigen recognition site, the specificity is modulated through somatic recombination of the V, D and J gene segments during T cell maturation in the thymus (18). The rearrangement of the VDJ results in 3 complementary determining regions (CDR) that interact with the pHLA complex (19, 20). The CDR1 of the $\alpha$-chain detects the N-terminus of the antigen and the CDR1 of the $\beta$-chain detects the C-terminus of the antigen. The CDR2 recognizes the $\alpha$-helices of the HLA class I molecule and the CDR3 region detects mostly the presented antigen (19, 21, 22). T cells with a functional and stable TCR are CD4$^+$CD8$^+$ double positive and are then positively selected for either HLA class I or class II recognition. This will decide their co-receptor (CD8, CD4) and additionally, in case the TCR binds too strongly to an HLA that presents a self-peptide it will undergo apoptosis because it is self-reactive (negative selection) (23, 24). This ultimately leads to T cells exhibiting a great variety of very specific
TCRs capable of recognizing distinct pHLA complexes that allow for the discrimination of self-peptides and peptides of foreign or unusual origin.

1.1.3 **The structure of HLA class I molecules determines antigen presentation**

The crystal structure of the HLA class I molecule HLA-A2, first solved in 1987 by Bjorkman et al. (10), greatly expanded our understanding of HLA antigen presentation. Mature HLA class I molecules are trimeric complexes composed of the HLA heavy chain non-covalently associated with the \( \beta_2m \) light chain and a peptide bound by the extracellular part of the HLA molecule. The peptides presented on HLA class I molecules are usually 8 – 10 amino acids (AAs) in length (25, 26). The HLA heavy chain is made up of 3 extracellular \( \alpha \)-domains, one transmembrane domain and a cytoplasmic tail. The \( \alpha_1 \)- and \( \alpha_2 \)-domains are encoded by exon 2 and 3, respectively, and correspond to the part forming the peptide binding region (PBR). This region is enclosed by two \( \alpha \)-helices that lie on top of an anti-parallel \( \beta \)-sheet building a cleft. The biochemical properties of this cleft allow for the binding of short peptides in a way that makes them in part accessible from the top (10, 27). Due to the nature of the binding, the peptide is anchored at the N-terminus and the C-terminus, whereas the middle part has the potential to bulge out of the PBR. This exposes the peptides primarily to the CDR3 region of the TCR (21, 28). The Ig-like \( \alpha_3 \)-domain is relatively conserved between different HLA alleles and is recognized by the CD8 co-receptor of CD8\(^+\) T cells (29). To facilitate the presentation of a wide range of peptides from different sources, and therefore the potential to present pathogenic antigens, the greatest variability between different HLA class I alleles is found in exon 2 and 3. The abundance of polymorphic HLA class Ia variants provides the feasibility to detect many different pathogens.

1.1.4 **HLA class I molecules present peptides of cellular origin**

The peptides presented on HLA class I molecules constitute an important differentiator for the immune system to separate healthy cells, presenting self-peptides, from infected cells that are invaded by pathogens. To sample a broad scope of the intracellular proteome, peptides for HLA class I presentation are sourced from many different proteins (30). In the cytosol, protein degradation continuously takes place by the proteasome that generates peptides of 3 – 22 AAs in length (31). Proteins are labelled for degradation via ubiquitination and comprise misfolded or aged proteins as well as defective ribosomal products (DRiPs) (30, 32, 33). The variety of targets for peptide presentation helps to minimize potential restrictions for accessible substrates and also allows for detection of any pathogenic proteins or degradation products that may be present in the cytosol. Of these, many are not able to
cross the membrane barrier of the endoplasmic reticulum (ER) by themselves and need to be actively transported into the ER lumen where peptide loading takes place. Most of the peptides are not suitable for HLA class I presentation, hence those are destroyed before translocation (34) or are either too small or too large to be effectively transported into the ER (35). Only about 15% of the peptides show the appropriate length for presentation on HLA class I molecules (36, 37). Peptides significantly longer than 12 AAs are still translocated to the ER by the transporter associated with antigen processing (TAP). However, further trimming is needed once inside the ER lumen because longer peptides greatly reduce the stability of the pHLA complexes. Inside the ER, this is mainly accomplished by N-terminal trimming through ERAP1 and ERAP 2 (38, 39). As long as sufficient peptide antigens are provided, peptide loading and formation of the pHLA-β2m complex can be facilitated.

Figure 2: HLA class I peptide loading is facilitated by the PLC. Nascent HLA class I heavy chain is translocated into the ER where association with β2m and subsequent recruitment into the PLC is realized through chaperons CNX and CRT. Through coordinated interaction of CRT and ERp57 and direct interaction with TPN, empty β2m/HLA class I complexes are localized to TAP and by further trimming through ERAP peptides suitable for peptide loading are provided. Stable pHLA complexes are released from the PLC and reach the cell surface through the Golgi route. Adapted from “HLA Class I Polymorphism and Tapasin Dependency” by Badrinath et al., 2014, “HLA and Associated Important Diseases”, p. 44. CC BY 3.0 (40).
The HLA heavy chain is first translated at the ER and subsequent assembly of the pHLA complex is facilitated using components of the ER quality control that are available in the ER (40, 41). Upon translation, the nascent heavy chain is immediately trimmed by glycosidase I and II leaving a single terminal glucose residue. This allows for interaction with the chaperon calnexin (CNX) (42) and the subsequent association of β2m (43). Monoglycosylated, peptide-free HLA class I-β2m dimers are recruited into the peptide loading complex (PLC) by calreticulin (CRT) where direct interaction with tapasin (TPN) is established (44, 45). There are two TPN molecules one associated with each TAP1 and TAP2 in the PLC (46) that localize empty HLA heavy chain-β2m dimers to the source of peptides (Figure 2). Additionally, TPN is linked to ERP57 by a disulfide bond between Cysteine 95 of the TPN and Cysteine 57 of the ERP57 (47). ERP57 then supports the folding by disulfide bond isomerization. Following successful loading of the HLA-β2m complex, it finally becomes the trimeric pHLA-β2m complex. The PLC components are sensitive to the HLA class I molecule holding a peptide: if a peptide of high affinity is bound, affinity with TPN is reduced releasing the trimeric HLA complex from the PLC. Additionally, the terminal glucose is cleaved off after dissociation from CRT. The mature pHLA then migrate through the Golgi to the plasma membrane. Low affinity peptides dissociate from the binding groove after release from the PLC. In that case, the HLA molecule can become a substrate for UDP-glucuronosyltransferase 1 (UGT1), and re-glycosylated HLA can enter the PLC again by association through CRT (41, 48). This will ultimately lead to binding of high affinity peptides and thus the presentation of stable pHLA complexes on the cell surface.

1.1.5 Peptide binding motifs are specific for HLA class I alleles

While the PLC ensures the loading of optimized peptides, the nature of the peptide binding groove naturally determines the sequence of the bound peptide. These peptides are bound by non-covalent interactions such as van der Waals forces, ionic interactions or hydrogen bonds within the binding groove. High resolution crystallographic studies identified six specificity pockets (A – F) alongside the binding cleft that can host distinct AA side chains (27, 49, 50). The properties of those pockets greatly depend on the composition of proximate AAs. The region corresponding to the peptide binding groove is encoded by exon 2 and 3, which are the most polymorphic regions within a specific HLA gene (5) and determine immunogenic variability. For most HLA class I alleles, peptides of canonical length, i.e. 9 AAs, are anchored at fixed positions within the peptide sequence. Preferred anchoring positions for most alleles include the peptides p2 and C-terminal pΩ position where the AA side chains extend into the B and F pocket, respectively (25, 27, 50, 51). While most HLA class I molecules exhibit a strong preference for stringent peptide anchors, distinct alleles
exhibit secondary anchoring positions that enable certain alleles to present a broader scope of peptides. The nature of the bound peptides is further dictated by the available proteome and although the protein sources can shift corresponding to the tissue where the HLA molecule is expressed, the pattern of the peptide binding motif remains unaltered.

1.2 Non-classical HLA class Ib molecules

The non-classical HLA class Ib molecules HLA-E, HLA-F and HLA-G differ from their classical HLA class Ia counterparts in many key aspects. First, the highly polymorphic nature of class Ia molecules is greatly reduced in HLA class Ib molecules. To date, 27, 30 and 58 different alleles are known for HLA-E, HLA-F and HLA-G, respectively (5). Because many of these polymorphisms contain synonymous substitutions or single nucleotide polymorphisms (SNP) located in introns, the amount of different, functional proteins is even more decreased. The low number of variance appears to correlate with the observation that these non-classical HLA molecules adopt very specialized immune functions and usually interact with conserved receptors on specialized immune cells. HLA-E, for instance, is known to present a very narrow set of peptides originated from the signal sequence of other HLA class I molecules (52). Additionally, HLA-E appears to be the main interaction partner for the NKG2A/CD94 heterodimer expressed on NK cells. Thus, HLA-E expression allows for indirect surveillance of the cellular HLA expression through NK cells (53), potentially as a second line of defense against pathogens that try to evade the immune system by interfering with HLA class I expression. Nevertheless, in recent years it became apparent that HLA-E is able to select and present non-canonical peptides of diverse origin depending on the cellular condition and that the immune system is equally capable to recognize this shift in the presented peptide repertoire (54-57). HLA-F is the most enigmatic of the non-classical HLA molecules, however, it was implicated that HLA-F is protective during pregnancy (58), as well as during the antiviral response to HIV-I (59). However, until recently the biochemical framework and mode of peptide presentation were unclear (60). Latest structural evidence points to different modes of operation for peptide bound HLA-F and HLA-F in an open conformation that could potentially support the presentation of unusual long peptides (60).

1.2.1 HLA class I histocompatibility antigen, alpha chain G

HLA-G differs in many regards not only from other HLA class Ia molecules but also from HLA-E and HLA-F. With 58 different alleles encoding for 17 distinct proteins, HLA-G is the most polymorphic of the non-classical HLA molecules. Allele distribution in various
populations worldwide favors HLA-G*01:01 as the most prevalent allele. Additionally, G*01:04 and G*01:03 are the second and third most common alleles across different European populations (61). In contrast to HLA-E and other HLA class Ia molecules, HLA-G expression appears localized at immune privileged sites and its expression is temporarily elevated during pregnancy. The HLA-G peptide repertoire appears quite narrow, potentially due to its stringent tissue specific expression. In comparison to the very specific subset of peptides presented by HLA-E, however, the peptides presented by HLA-G are sampled from a broader range of proteins.

Figure 3: Structural features of HLA-G. A) The crystal structure of HLA-G*01:01 shows similarity to other HLA class I molecules. Domains α1 and α2 form a peptide binding cleft capable of presenting nonameric peptides. β2m is non-covalently associated with the Ig-like α3-domain. B) Top view of HLA-G shows the binding of the nonameric peptide RIIPRHLQL by the α1- and α2-domain. The graphical illustration is based on the available crystallographic data PDB: 1YDP. C) Splicing of HLA-G mRNA leads to seven different isoforms. Membrane-bound isoforms contain HLA-G1 that constitutes the full length membrane bound HLA glycoprotein. Isoforms HLA-G2, -G3 and -G4 are missing one or more α-domain. By retaining a stop codon after exon 4 or exon 2, the soluble forms HLA-G5, -G6 and -G7 are produced. Adapted from “HLA-E, HLA-F and HLA-G — The Non-Classical Side of the MHC Cluster” by Foroni et al., 2014, "HLA and Associated Important Diseases", p. 69. CC BY 3.0 (62).
HLA-G interacts with different types of receptors and thus with different types of immune effector cells (63-65). HLA-G was proposed to be a classical peptide presenter (66), featuring nonameric peptides bound in the peptide binding groove of the full length, membrane-bound HLA-G molecule (Figure 3). However, several alternative splice forms of the HLA-G mRNA were discovered (67, 68). These variants lead to four membrane-bound isoforms termed HLA-G1, that constitutes a full-length trimeric HLA molecule; HLA-G2, -G3 and -G4 are characterized by the loss of one or more α-domain. Corresponding soluble forms are generated through the retention of a stop codon after exon 4 or exon 2 that leads to the corresponding HLA-G5, -G6 and -G7 isoforms (Figure 3). Additionally, through a highly specific cleavage site present in the α3-domain, HLA-G1 is also a target for the matrix metalloproteinase MMP-2 that can cleave off HLA-G1 from the cell surface (69). While the occurrence of different HLA-G splice variants and their potential specifications is intriguing, very little data exists on the trimeric protein forms of these variants. Based on mRNA expression data and immunohistochemical stainings the HLA-G1, cleaved soluble HLA-G1 and HLA-G5 isoforms appear to be the most prevalent forms in human tissue (70).

1.2.2 HLA-G as a potent inhibitor of immune cell functions during pregnancy

Pregnancy poses the immunological dilemma that the embryo has to be tolerated by the maternal immune system even though it basically constitutes a semi-allograft. To avoid rejection, a T helper 2 (Th2) response characterized by secretion of IL-4, IL-5 and IL-10 aides local immune suppression at the maternal-fetal interface in order to manage the plethora of immune cells such as CD56⁺CD16⁻ NK cells, macrophages, CD4⁺ and CD8⁺ T cells that are present in the placenta (71-73). In addition, to induce a state of perceived self, fetal tissue down regulates HLA-A and HLA-B expression and villous trophoblasts start to express membrane-bound as well as soluble HLA-G molecules (74, 75). Additionally, an upregulation of HLA-E expression and low amounts of HLA-C is observed (76). In the placenta, HLA-G is thought to be a key player for the induction of a localized immunosuppressive environment because of its ability to induce immune tolerance and modulate the local environment through induction of cytokine secretion (75, 77, 78). To achieve this, HLA-G has to interact with many different types of immune effector cells that are present at the fetal-maternal interface. Through receptor binding, HLA-G can directly inhibit NK and T cells. It could be observed early on that HLA-G alone is sufficient to inhibit NK cell mediated lysis (79). To facilitate NK inhibition, HLA-G interacts with the KIR2DL4 receptor (63). Similar to the exclusive HLA-E-NKG2A/CD94 interaction, HLA-G is the only known ligand for this NK cell receptor. Additionally, HLA-G was also shown to bind ILT-2 and ILT-4. ILT-2 is expressed on most immune cells such as NK, T, B and also dendritic cells
(80-83). ILT-4 is expressed mainly on APCs of the myelomonocytic lineage (83-85). Both receptors are able to bind other HLA class I molecules, however, affinity of both ILT-2 and ILT-4 appears highest when ligated to HLA-G (65). It could be demonstrated that HLA-G itself induces expression of these receptors on effector cells (86). HLA-G also triggers long lasting effects through immune regulation by inducing CD4⁺CD25⁺FOXP3⁺ T regulatory cells (87) as well as type 1 T regulatory cells (86) and HLA-G is able to arrest maturation of ILT-4 positive dendritic cells (88). HLA-G is further capable to induce apoptosis of CD8⁺ T cells via CD8 and of endothelial cells via interaction with CD160 (65). Most interactions with the ILT receptors are thought to be facilitated by the α3-domain, however, in congruence to the binding of KIRs to HLA class Ia molecules, an interaction of KIR2DL4 with the α1 domain is assumed (89, 90).

1.2.3 Peptide presentation is restricted in HLA-G

First observations about peptide presentation in non-classical HLA molecules were made in the mid-1990s. Lee et al. (91) and Diehl et al. (66) demonstrated through peptide elution studies from transfected LCL721.221 cells that HLA-G*01:01 presents a restricted peptide repertoire. Identified peptides were mainly anchored by Leucine at pΩ and Isoleucine or Leucine at p2. Proline was identified as an auxiliary anchor at p3 as well as a preference for hydrophobic side chains at p7. Ishitani et al. (92) expanded the knowledge of the peptide repertoire by showing that in term placenta, peptides presented on HLA-G*01:01 appear to be even more restricted being derived practically only from a cytokine receptor-like molecule. However, due to technical limitations, peptide sequencing was focused on the identification of nonameric peptide ligands. The crystal structure of HLA-G complexed with the peptide RIIPRHLQL was solved in 2005 (90) and showed a peptide binding groove with very similar properties to that of HLA-E. Similar to other HLA class I molecules, AA side chains of peptide anchors p2 and pΩ are nested inside pockets B and F, respectively. However, the rest of the peptide binding groove appears more stringent resulting in an additional interaction of p2 with pocket D as well as interactions between p6 and p7 with pocket D and E, respectively. Additional structures using different peptides highlighted a greater conformational flexibility of the peptide binding region that can even lead to a widening of the cleft depending on the bound antigen. Additionally, the bound peptide directly influences the stability of the pHLA complex, meaning that the bound peptide not only directly affects recognition but also impacts the half-life of the molecule on the cell surface (93).
1.3 The role of HLA-G in malignancies

Due to its unique immune modulatory features, HLA-G was implicated in a wide range of different malignancies. HLA-G polymorphisms were shown to be associated with pregnancy complications (94-96), implicated in rheumatoid arthritis (97, 98) and multiple sclerosis (99, 100), and more recently it was observed that in end-stage renal disease higher levels of HLA-G5 and -G6 are present, although lower levels of -G5 are found in acute rejection (101). After lung transplantation higher levels of HLA-G expression were found in patient plasma, potentially mediating graft acceptance (102). Since HLA-G molecules are readily detectable in human plasma and serum, many attempts were made to utilize HLA-G as a biomarker; however, due to the complex nature of HLA-G biology, HLA-G levels are not yet used comprehensively.

1.3.1 HLA-G expression is elevated in different tumor entities

Ectopic expression of HLA-G was found in many different tumor entities, ranging from hematological to solid organ malignancies. Elevated levels of soluble HLA-G associated with extracellular vesicles were found in the plasma of breast cancer patients and were related to disease progression; however, free soluble HLA-G appears to be correlated with better outcome after neoadjuvant chemotherapy (103). Additionally, it was observed that HLA-G is inducible by IL-10 (94) and IL-10 is also elevated in non-Hodgkin Lymphoma (NHL) (104, 105) and lung cancer (106). In esophageal squamous cell carcinoma elevated HLA-G and IL-10 levels were observed in the plasma, and here, HLA-G expression was associated with poor prognosis (107). In colorectal cancer, Guo et al. observed elevated HLA-G expression in 70.6% of the analyzed tumor samples and concluded that HLA-G serves as an independent factor for overall survival (108). Soluble HLA-G levels were increased in plasma of patients with different hematological malignancies such as chronic lymphatic leukemia (CLL), T-NHL, B-NHL (104) and in multiple myeloma (109). It was further shown that in CLL a 14bp polymorphism correlates with increased HLA-G expression and that a higher percentage of T regulatory cells is present in 14bp del/del patients (110).

1.3.2 Hodgkin’s lymphoma as a model of HLA-G mediated immune evasion

Hodgkin’s lymphoma (HL) is a cancer that results from the malignant transformation of B cells. Often, these transformed cells are detected first in lymph nodes from where they spread throughout the body. Most cases are categorized as classic HL, in which
multinucleated cells, the so called Hodgkin-Reed-Sternberg (HRS) cells, are observed in the lymph node. Notably about this cancer is that inside the reactive infiltrate within the lymph nodes there are many different immune effector cells surrounding a comparatively small amount of HRS cancer cells (111, 112). Down regulation of HLA class I expression is an important step in tumor immune evasion, however, at the same time NK cells should recognize the loss of HLA expression and become activated. HRS cells lack HLA class Ia surface expression but at the same time Diepstra et al. (113) found HLA-G expression in 54% of HRS cells in lymph nodes of HL patients. These findings were further expanded by Caocci et al. (114) who could show similar expression patterns in HRS cells from cHL patients, where HLA-G expression also correlated with disease progression. Little is known about the genotype of the expressed HLA-G variants, however, De Re et al. could recently link a specific HLA-G genotype in the 3'UTR region (+3027-C/A) to reduced event-free survival in comparison to a C/C genotype, even though actual HLA-G expression in the HL cells was only found in about 20% of C/C genotyped patients (115), suggesting that a strong linkage disequilibrium with other critical genes that impact prognosis could exist (115, 116).
2 PROJECT AIMS

Over the last years, HLA-G came into the focus of tumor biology due to its broad immune regulatory features and although HLA-G is the best characterized of the non-classical HLA class I molecules, many of these features and their mode of action remain elusive. HLA-G splicing results in different isoforms, even soluble HLA-G molecules. Since the HLA-G heavy chain appears in large parts invariant due to the low degree of polymorphisms and, in addition, certain features of the HLA-G genomic sequence located at the 5’ and 3’ UTRs were shown to influence expression and mRNA stability, the importance of individual HLA-G alleles have not been comprehensively investigated, yet. Cellular viability is governed by dynamics between immune effector cell receptors and their cognate ligands; therefore the presentation of peptides complexed with a certain HLA molecule determines the individual fate of a cell. Understanding the mechanism of how such HLA-mediated recognition events translate to cellular immunity will guide towards safe and personalized cellular therapies.

To understand how HLA-G interacts with a diverse set of conserved receptors the first aim was to understand the properties of the antigens presented by HLA-G. Therefore, the focus was put on the three most common alleles in Europe (HLA-G*01:01, G*01:03, G*01:04) and a model system that allows for the analysis of a single HLA class I allele on NK cell recognition was used. Additionally, the presented peptide repertoire of each allele was analyzed and exemplary peptides were modelled by using available crystallographic structures to elucidate the constrained peptide features (research paper 1).

Following the discovery of an unexpected peptide diversity bound to the mostly invariant HLA-G molecule, the aim was to further investigate whether the peptide repertoire would also be subjected to alterations when expressed in a cancerous malignancy. HLA-G is characterized as an immune evasion mechanism in classic Hodgkin’s lymphoma, hence the HL cell line HDLM-2 was used for peptide fishing and compared to the peptide repertoire of HLA-G*01:01 derived from the erythroleukemic cell line K562. The results exhibit a clear shift in the preferred anchor residue at p1 for HDLM-2 derived HLA-G-restricted peptides. To exclude whether this confinement would be imposed by the cellular proteome, the proteomic data from both cell lines were compared to the identified peptide sources and additionally, to exclude impairment of PLC components in transformed cells, association of the HLA-G heavy chain with components of the PLC were investigated (research paper 2).
HLA-G mediated immune regulation is impaired by a single amino acid exchange in the alpha 2 domain

Alexander A. Celik, Gwendolin S. Simper, Trevor Huyton, Rainer Blasczyk, Christina Bade-Dödinger

*Institute for Transfusion Medicine, Hannover Medical School, Carl-Nielsen-Str. 1, 30625 Hannover, Germany
**Department of Cellular Logistics, Max Planck Institute for Biophysical Chemistry, Am Fassberg 11, 37077 Göttingen, Germany

** Abstract **

The trade-off from HLA class I expression to HLA-G expression support the immune evasion of malignant cells. The essential role of the virtually invariant HLA-G in immune tolerance, tumor immunology and its expression frequency in immune privileged tissues is known; however the specific importance of allelic subtypes in immune responses is still not well understood. HLA-G*01:01, *01:03 and *01:04 are the most prevalent allelic variants differing at residues 31 and 110, respectively. In cytotoxicity assays applying K562 cells transduced with the HLA-G variants as targets and NK cells as effectors the differential protective potential of HLA-G variants was analyzed. Their peptide profiles were determined utilizing soluble HLA technology. An increased protective potential of HLA-G*01:04 could be observed. All variants exhibit a unique peptide repertoire with marginal overlap, while G*01:04 differs in its peptide anchor profile substantially. The functional differences between HLA-G subtypes could be explained by the constraint of the bound peptides, modifying the pHLA-G accessible surface. For the first time a contribution of amino acid alterations within the HLA-G heavy chain for peptide selection and NK cell recognition could be observed. These results will be a step towards understanding immune tolerance and will guide towards personalized immune therapeutic strategies.

1. Introduction

HLA class I molecules (HLA-I) are expressed on almost all cell types and occupy a crucial role in regards to the human immune system because HLA-I molecules provide a way for immune cells to surveille the health status of every individual cell. This screening is performed by CD8 T cells of the adaptive immune system that are able to distinguish not only different HLA alleles but also their presented peptides [1,2]. A presented peptide mediates the specificity to the TCR, however, classical HLA-I molecules HLA-A/B/C are also capable of binding NK cells thus providing a self-signal to inhibit NK cell mediated cytoxicity. Of the non-classical HLA-I, HLA-E specifically presents peptides derived from the leader sequence of HLA-I molecules to provide a specific signal to the NK cell receptor complex NK2G2x/CD94 allowing for rudimentary distinction between self and non-self by NK cells [3]. In the case of HLA-E, it was shown that the substitution of one AA is sufficient to change the presented self-peptide [4,5] and thus potentially influences the interaction with the NK cell receptor subunits NKG2A and NKG2C. HLA-G is also known to interact with NK cells; however, in contrast to HLA-E it was also shown that HLA-G interacts with most cells of the adaptive immune system as well. Compared to HLA-I, expression of HLA-G is restricted to certain immune privileged tissues or induced during specific time frames primarily during pregnancy in the placenta. Here, HLA-G confers immune tolerance against foreign antigens expressed on fetal tissue [6]. The induction of a perceived self is facilitated by specialized cells present in the placenta via expression of HLA-G and subsequent interference with the local immune system. Common to hematological malignancies is the ability to elude tumor immune surveillance, however, the underlying disease mechanisms are quite diverse. This lead to a plethora of treatment regiments tailored to the specific diseases as well as new therapy options that are currently explored (e.g. treatment using CAR T cells [7] or allogeneic NK cells [8]). Common to most tumors is the downregulation of classical HLA class I molecules (HLA-Ia) in order to escape immune surveillance. However, the loss of HLA-Ia can trigger NK cell activation [9,10] and subsequent elimination of aberrant cells. To evade NK recognition, providing a self-inhibitory signal by expression of non-classical HLA class I (HLA-b) molecules HLA-E and HLA-G may pose an evasion strategy not unlike viral immune evasion [11]. This comprises a simultaneous down-regulation of HLA-I, hence recognition through the
TCR is impaired whereas HLA-E remain expressed in virtually all cells in order to present peptides derived from the signal sequence of HLA-I molecules. It was shown that HLA-E can present peptides independently of tapasin [4], a crucial part of the peptide loading complex responsible for successful loading of high affinity peptides. Tapasin, as well as other proteins of the peptide loading complex are often targets for viral interference. In the case of HCMV it was shown that a leader sequence was provided by gpUL40 [12] stabilizes HLA-E surface expression and provides an inhibitory signal to NK cells. In contrast, it appears that even though HLA-G is capable of tapasin-independent [13] peptide loading, HLA-G carrying peptides of low affinity are cycled between CD168 and the ER [13,14] for additional processing. Upregulation of HLA-Ib molecules may also be facilitated by various types of cancer in order to gain immune tolerance. Due to its primary function to interfere in regulatory immune processes [15], it is hardly surprising that, among other factors, upregulation of HLA-G was found in diverse myeloproiferative malignancies such as multiple myeloma [16], chronic lymphocytic leukemia [17], Hodgkin lymphoma [18,19] and non-Hodgkin lymphoma [17] as well as in breast cancer [20] and glioblastoma [21]. HLA-G is the most polymorphic of the HLA-Ib molecules and in contrast to HLA-E, where two alleles are balanced in worldwide populations [22], several HLA-G alleles are maintained [23]. However, sequence data is limited because the HLA-G locus is not routinely analyzed in tissue typing laboratories. The HLA-Ia locus belongs to the most polymorphic region in the human genome [24]. The increase in diversity impacts the variety of antigenic peptides that can be presented by different populations. Due to much lower polymorphisms, peptide repertoire appears much more restricted in HLA-Ib. However, it was shown that non-canonical peptides can be presented by HLA-E and that immunological reactions can be mediated by such ligands due to the peptide mediated nature of the TCR-HLA interaction. The preservation of different HLA-G alleles in different populations suggests that the amino acid (AA) substitutions could exert different biological functions. Expression of HLA-Ia and its effects on peptide and T cell engagement are well characterized, however, the underlying mechanisms of tolerance induced by HLA-G are not well understood, especially in the context of cancers such as breast cancer [20]. Similar to HLA-Ia molecules, HLA-G presents a conserved peptide repertoire of noncanonical peptides [25-27].

A unique feature of HLA-G is the occurrence of splice variants leading to soluble forms as well as truncated molecules [28,29] that further expand the variability of this conserved molecule. To date, three main receptors are identified to react with HLA-G: KIR2DL4, ILT-2 and ILT-4 and expression of these receptors can also be induced by HLA-G [30]. Known interactions include inhibitory effects on NK cell mediated cytolysis via KIR2DL4 [31], induction of Tregs [32] as well as induction of Tc1 cells [33] and maturation arrest of ILT-4 positive DCs [34]. Additionally, interaction of soluble HLA-G was shown to induce apoptosis of T cells via CD95 and epithelial cells by interaction with CD69.

While these interactions are facilitated mostly through the conserved αααα domain [35], the potential influence of the bound peptide remains elusive, however, consistent with the binding of other KIRs to HLA-I, an interaction of the αααα domain with the KIR2DL4 receptor was proposed [36,37]. Even though TCR engagement does not appear likely, MHC loci homologues to human HLA-G in New and Old World monkeys appear to exhibit functions similar to HLA-Ia [38].

In this work we aimed to systematically analyze the three most prevalent HLA-G variants in Europe (HLA-G*01:01, G*01:03, G*01:04) that are differentiated from each other by the exchange of single AAs and elucidate the impact of these polymorphisms on NK cell immune recognition.

2. Material and methods

2.1. Cell lines & isolation of NK cells

HLA class I negative K562 cell lines transduced with HLA-G*01:0x (exon 1-6) or sHLA-G*01:0x (soluble, exon 1-4) were maintained in RPMI1640 (Lonza, Basel, Switzerland) supplemented with 10% heat inactivated FCS (Lonza, Basel, Switzerland), 2 mM L-glutamine (c. pro, Oberndorla, Germany), 100 U/ml penicillin and 100 μg/ml streptomycin (c. pro, Oberndorla, Germany). The NKL cell line was maintained in RPMI1640, 2 mM L-glutamine, 100 U/ml penicillin and 100 μg/ml streptomycin with the addition of 15% heat inactivated FCS, 1 mM sodium pyruvate (c. pro, Oberndorla, Germany) and 200 U/ml IL-2 (PeproTech, Rocky Hill, NJ, USA). HEK293T cells, used for production of lentiviral particles, were cultured in DMEM (Lonza, Basel, Switzerland) supplemented with 10% heat inactivated FCS, 2 mM L-glutamine, 100 U/ml penicillin, 100 μg/ml streptomycin and 1 mg/ml Geneticin® (Life Technologies, Carlsbad, USA). All cell lines were maintained at 37 °C and 5% CO2. NK cells were isolated from PBMCs from healthy blood donors using the Human NK Enrichment kit (STEMCELL Technologies, Vancouver, Canada) according to the manufacturer's instructions.

2.2. Cloning of HLA-G constructs

 Constructs encoding for soluble HLA-G (shHLA-G*01:0x, exon 1-4) were generated from JEG-3 cDNA, subcloned into the lentiviral vector pBRUPTST.mcs pre, as previously described [29]. Site-directed mutagenesis, as described previously [4], was used to generate shHLA-G*01:03 and shHLA-G*01:04 encoding constructs by introducing a single point mutation at position c.162A > T for HLA*01:03 or c.400C > A for HLA-G*01:04. The respective inserts were verified through sequencing.

2.3. Stable transduction of K562 cells with HLA-G*01:0x

According to the method described by Bade-Dooding et al. [39], K562 cells were stably transduced with different HLA-G variants. The expression HLA-G or shHLA-G was confirmed by flow cytometry, ELISA and Western blot. Secretion of soluble HLA-G molecules in the supernatant was verified via Western blot of the supernatant of K562/ G*01:0x cells using a mAb recognizing the V5-tag (Life Technologies, Darmstadt, Germany).

2.4. Large-scale production of shHLA-G and mass spectrometric analysis of shHLA-G derived peptides

Large scale production of recombinant shHLA-G*01:0x was performed according to soluble HLA technology [40] using bioreactors (Integra Biosciences, Bietigheim, Germany). The shHLA-G molecules were purified on NHS-activated HitTrap columns (GE Healthcare, Freiburg, Germany) coupled to the mAb W6/32. Quality and quantity of purified protein was verified by ELISA and western blot. 3 mg of each purified shHLA-G*01:0x sample was treated with trichloroacetic acid (TFA) (J.T. Baker, Phillipsburg, USA) to a final concentration of 1% to separate the peptides from heavy and light chain. Following size exclusion chromatography using Amicon Ultra filter units (Merck Millipore, Billerica, USA) with a 10kD cutoff, the peptide solution was vacuum dried to a volume of 1 ml and treated with ammonium bicarbonate (NH4HCO3), dichloroethane (DTE), iodoacetamide (JAA) and trichloroacetic acid (TCA) (Merck, Darmstadt, Germany). The scissile sample was purified on a C18 Tip (Thermo Fisher Scientific, Walhalm, USA) with 0.1% TFA and eluted with 60% acetonitrile/0.1% TFA. The eluate was vacuum dried and solubilized in 30 μl 2% acetonitrile/0.1% TFA. The LC/MS analysis was performed with a Dionex UltiMate 3000 high-performance LC system and a LTQ Orbitrap Velos Hybrid IT Mass Spectrometer. Data were analyzed using the protein alignment tool BLAST and the Uniprot database. Of the identified 9mer peptides, preferred AA occurrence at each position have been analyzed using Weblogos [41]. Peptide preparation and analysis has been performed twice, first and second datasets corresponded.
2.5. Cytotoxicity assay

A flow cytometry based cytotoxicity assay was used to assess protective activity of K562-HLA-G 01:0x, similar to the method described by Kremer et al. [5]. Each experiment was performed using two technical replicates. NK cells used in the assay were freshly isolated from PBMCs of HLA-G typed donors. Briefly, K562 cells were lentivirally transduced with vectors encoding for HLA-G 01:0x variants and subsequently sorted for equal HLA-G expression. Following CFSE-labeling, target cells were co-incubated for 4 hr with isolated NK cells from three different donors as well as with NKL cells at a ratio of 1:10 (target:effector). Finally, cell death was evaluated using 7-AAD staining. CFSE positive cells were gated and 10,000 events were recorded. Target cells stained with CFSE but not co-incubated with NK cells were used to measure the amount of spontaneous dead target cells. Cells without staining for 7-AAD were used as additional gating controls. Subsequently, specific lysis was calculated according to the following formula:

\[
\text{Specific lysis} = \left(\frac{\% \text{ dead target cells} - \% \text{ spontaneous dead target cells}}{100 - \% \text{ spontaneous dead target cells}}\right) \times 100
\]

2.6. Modelling

The HLA-G 01:01 structures (2001, 2012) and 01:04 structures (1YDP, 3KYN, 3KYO) were taken for structure analysis and additional non-crystallographic symmetry molecules or receptors were removed. The structures were then superposed on each other in Coot [42], all figures were made with CCP4MG [43].

3. Results

3.1. Features of HLA-G restricted peptides

The HLA class I negative K562 cell line was lentivirally transduced with lentiviral vectors encoding for soluble HLA-G variants (HLA-G 01:0x Exon 1 through 4) and a C-terminal V5-tag. Successful recombinant shHLA-G 01:0x-V5 expression was verified by Western blot using an anti-V5 mAb (Fig. 1A). shHLA-G containing supernatant was harvested from bioreactors, affinity purified and HLA-G bound peptides were eluted and analyzed utilizing mass spectrometry. A total of 91, 54 or 37 peptides restricted to HLA-G 01:01, G 01:03 or G 01:04, respectively, were sequenced. The respective peptide sequences are given in Table 1. Within these peptide pools (Fig. 1B), the majority of HLA-G restricted peptides exhibited canonical length for all allotypes. Among these were 25.3%, 14.8 or 27.0% 8mer, 45.1%, 53.7% or 48.7% 9mer and 14.3%, 14.9% or 13.5% 10mer peptides. Longer peptides (> 10 AA) were observed in more than 15.4%, 16.7% or 10.8% for HLA-G 01:01, G 01:03 or G 01:04, respectively.

3.2. The invariant peptide repertoire of HLA-G 01:04 is not shared with other allotypes

Peptide overlap between the analyzed peptide repertoires (Fig. 1C) showed a total of 5 peptides shared between HLA-G 01:01, G 01:03 and G 01:04, conforming to a length of 6-10 AA. Peptide sequences are given in Table 1. Between G 01:01 and G 01:04, 2 overlapping peptides were found, whereas between G 01:03 and G 01:04, a total of 3 identical peptides were observed. Sequences of overlapping peptides are given in Table 4. Shared peptides were found to be predominantly 8-9 AA in length.

Analysis of the anchoring positions of the 9mer peptides shows that all allotypes exhibit Leucine as the dominant C-terminal anchor (Fig. 2). For p2, no anchor could be determined, however, auxiliary anchors at positions p3 and p4 are most likely Lysine and Proline, respectively, for HLA-G 01:01 and G 01:03.

3.3. Protection from NK cell mediated killing is conferred by HLA-G expression

HLA-G confers protection against NK cell mediated killing when expressed on the cell surface. To assess the different alleles, the HLA class I negative K562 cell line was lentivirally transduced with vectors encoding for membrane-bound HLA-G 01:01, G 01:03 or G 01:04. HLA-G expressing K562 cells were sorted using the anti-HLA-G mAb MEM-G/9 (Suppl. Fig. 1B) and used for cytotoxicity assays with human NK cells isolated from peripheral blood. The specific lysis was calculated for each recombinant HLA-G 01:0x expressing K562 target cell line as well as for untransduced K562 cells using NK cells from three different donors as well as the NK cell line NKL. As depicted in Fig. 3, the specific lysis of the HLA-G 01:04 cells when using NK cells derived from donor 1 (G 01:01/G 01:04) is 41.1%, whereas expression of different HLA-G alleles confers different protection levels against NK cell mediated killing. For HLA-G 01:01- or G 01:03-expressing K562 cells, killing is reduced to 27.1% or 29.4%, however for G 01:04-expressing
### Table 1

<table>
<thead>
<tr>
<th>Sequence</th>
<th>Length</th>
<th>Source</th>
</tr>
</thead>
<tbody>
<tr>
<td>继续在下一页</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>
K562 cells specific lysis through NK cells is decreased to 8.1%. For NK cell derived from Donor 2 (G’01:01/G’01:04), 36.2% of K562 cell lysis could be observed, whereas HLA-G 01:01 - G’01:05 or G’01:04 expression reduced the specific lysis to 30.4%, 30.2% or 9.4%. Similar results could be observed when using NK cells from Donor 3 (G’01:01/G’01:01) as effectors, as lysis of 32.0% with K562 cells as targets was reduced to 26.0%, 15.6% or 11.0% when G’01:01, G’01:03 or G’01:04 are expressed. As proof of principle, we used the NK cell line NR2 (G’01:01/G’01:01) that did not undergo any steps of purification and has not been exposed to any pathogens assuming reproducibility of the results on HLA-G soluble receptor. Similarly, the specific lysis was reduced from 41.6% with K562 cells as targets to 19.7% for K562/G’01:01 cells, 20.0% or 17.5% for K562/G’01:03 or K562/G’01:04 cells.

3.4. Pro α at p3 facilitates terminal limitations on the peptide backbone

To understand the implications of the absent Pro at p3 in the peptide binding motif of HLA-G 01:04, selected peptides were modeled into the binding groove of available HLA-G’01:01 and HLA-G’01:04 structures (2DYP, G’01:01; 3X4N, G’01:04), accordingly the conformation of an identified G’01:04-specific peptide (FLMGQLNGLG) was compared to two Pro-containing peptides eluted from G’01:01 (RQPYAVSELS, RSPGPGNGLG). Depicted in Fig. 4 are the modeled peptides and their most likely conformation. Peptide constraints are observed creating a “kinked” peptide around positions 3–6 for peptides with a P at p3.

4. Discussion

In this study we investigated the endogenous peptide repertoire of HLA-G’01:01, G’01:03 and G’01:04 that are the most common HLA-G alleles in the European population [23,44]. These alleles are separated by a conservative AA substitution exchanging either small neutral residues (p.31T > S, G’01:02) or hydrophobic residues (p.110L > I, G’01:04) at positions that do not directly interact with peptide binding, peptide selection or binding of known immune receptors. Even though HLA-G is the most polymorphic molecule of the non-classical HLA-i [24], HLA-G’01:03 is by far the most common allele found in different human populations [23]. Similar to classical HLA class 1 molecules, HLA-G presents peptides of 9 AA in length [23,27] featuring defined peptide anchors at position p2 (Ile, Leu) and p2 (Leu) as well as an auxiliary anchor at p3 (Pro). Polymorphisms in HLA-G’01:03 and G’01:04 are located outside the peptide binding groove either at the beginning of the β-sheet floor of the groove or in a loop region outside the α2-helix. Analyzing the presented peptide repertoire revealed a dominant peptide length of 9 AA for all alleles that is consistent with the previously published peptides, however, comparison of the different peptides showed very little overlap between the alleles and especially between G’01:04 and G’01:02 or G’01:03. Furthermore, G’01:04 exhibits an altered peptide anchor motif compared to the other two alleles and to the HLA-G peptide binding motif previously published. These findings are reminiscent of HLA-E, were an AA exchange in the outer loop region of HLA-E’01:03 was shown to impact peptide presentation and thus potentially influencing the interaction with the NKGD2/CD94 heterodimeric receptor on NK cells [5].

To analyze the immunological impact of these polymorphisms, we investigated the potential to inhibit NK cell mediated cytotoxicity. Stable expression of HLA-G’01:04 on K562 cells conferred protection against NK cells, however, in case of G’01:04 an increase in protection was observed. To explain this observation, the structures available in the PDB database were analyzed. A number of structures exist of HLA-G’01:04 (3XKN, 3XKN, 3Y0P) with HLA-G’01:04 and in complex with its inhibitory receptor ILT4 (2DYP). Superposition of the HLA-G’01:01 and HLA-G’01:04 structures shows no obvious differences that might suggest a reason for the auxiliary anchor at p3 (Pro) for HLA-G’01:01 and HLA-G’01:03 but not for HLA-G’01:04. The regions around the polymorphic residues (p.31T > S, G’01:03) and (p.110L > I, G’01:04) are outside the peptide binding region (PBR) and have seemingly no observable influence on the structure. These regions are also distal from the ILT4 receptor binding site (2DYP) and it is inconceivable that they could affect the function of this receptor. The region around residue 110 appears highly mobile in all the structures and this is reflected in increased B-factors for atoms in this region.

The mutational analysis of the HLA-G’01:01 structures 110Leu > Ile and increased HLA-G’01:04 110Leu > Ile structures using FOLDX [45] illustrated no overall trend in the differences for the dG of unfolding (data not shown) but highlights that there are no obvious effects on the overall stability of HLA-G molecules irrespective of the polymorphism. This however might not be the case when we consider each HLA-G molecule in complex with the peptide loading complex. The different alleles clearly present a varied subset of peptides and it is likely that this polymorphism may impact the dynamics of HLA-G allele interaction with the PLa. Interestingly tapasin is predicted to bind a loop comprising residues 128–136 below the 2–1 helix of the MHC class I heterodimer [46,47], that is structurally nearby.

Data mining of the extensive number of PMHC-I structures in the protein data bank has demonstrated that many peptides show
interresidue interactions that enforce specific peptide conformations. Three subsets of peptide conformations (types I, II, and III) were described and >50% have been estimated to exhibit internal constraints that shape the selection of presented peptides [48]. Critically the presence of proline residues increases the propensity of such constrained peptides due to its imposition of significant torsional limitations on the peptide backbone. These constraints can also be suggested for peptides with an auxiliary anchor at p5 (Pro) as observed in the HLA-G 0.01 and HLA-G 0.03 peptides. KIR2DL4 was proposed to be the primary receptor for HLA-G, and crystallization studies on KIR2DL4 with HLA-Cw3 could previously show that the receptor stands in direct contact with the α1- and α2-helix as well as with positions p7 and p8 of the presented peptide [49]. Nevertheless, using surface plasmon resonance studies, an interaction between KIR2DL4 and HLA-G (refolded with either previously described peptides KIR3DL1 [25, 27] or KIR2DL1 [27]) could not be observed [50]. However, evidence suggests that binding of soluble HLA-G leads to internalization of KIR2DL4 and subsequent induction of a proinflammatory response through the NK cell [51], opening up the possibility that soluble HLA-G is the main interaction partner. Also, influence of the different splice variants on KIR2DL4 interaction is not known. The other two known receptors for HLA-G, ILT-2 and ILT-4, generally bind HLA-A at the α3-domain and β2m with their D3D2 domains [52]. Nevertheless, both ILT-2 and ILT-4 have a higher affinity to HLA-G than to other HLA-A molecules [53]. Studies on the structure of the D1D2-region of ILT-4 in complex with HLA-G revealed that the receptor interacts primarily with the α3-domain utilizing the N-terminal Ig-like domain with additional contact points between β2m and the interdomain region. Nevertheless, the model suggests that there is very little interaction with the α1- and α2-domains of the HLA molecule, suggesting a minor role for the presented peptide for the recognition [36,53] similar to the interaction of ILT-2 with HLA-A2 [54]. After solving the structure of the D3D4 domain, Nam et al. [55] proposed a model whereby the D3D4 domain is able to angle towards the α1-α2-helices where a peptide exchange could potentially impact the

<table>
<thead>
<tr>
<th>Table 2</th>
<th>HLA-G 0.03restricted peptides.1</th>
<th>Length</th>
<th>Source</th>
</tr>
</thead>
<tbody>
<tr>
<td>KRPFGAQKL</td>
<td>8</td>
<td>PDZ domain-containing protein 11</td>
<td></td>
</tr>
<tr>
<td>KRPFGML</td>
<td>8</td>
<td>PDZ domain-containing protein 11</td>
<td></td>
</tr>
<tr>
<td>KRPFGML</td>
<td>8</td>
<td>Tyrosine-RNA ligase, mitochondrial</td>
<td></td>
</tr>
<tr>
<td>SPQKSTKL</td>
<td>8</td>
<td>Myosin-specific enhancer factor 3C</td>
<td></td>
</tr>
<tr>
<td>SINLSEL</td>
<td>8</td>
<td>Hbk asparagine polypeptide</td>
<td></td>
</tr>
<tr>
<td>TPSNLKL</td>
<td>8</td>
<td>Unknown</td>
<td></td>
</tr>
<tr>
<td>TPQPRQKK</td>
<td>8</td>
<td>High affinity immunoglobulin alpha and immunoglobulin mu Fc receptor</td>
<td></td>
</tr>
<tr>
<td>TPVSNQKL</td>
<td>8</td>
<td>DNA repair protein HUS1</td>
<td></td>
</tr>
<tr>
<td>ALQQLNLQ</td>
<td>9</td>
<td>Retinoic-acid-like protein 1</td>
<td></td>
</tr>
<tr>
<td>HPPSYEKL</td>
<td>9</td>
<td>PDZ domain-containing protein 11</td>
<td></td>
</tr>
<tr>
<td>KAPGSLNL</td>
<td>9</td>
<td>G2 and 5 phase-expressed protein 1</td>
<td></td>
</tr>
<tr>
<td>KAPGSLNL</td>
<td>9</td>
<td>260 protein regulatory subunit 4</td>
<td></td>
</tr>
<tr>
<td>KAPPYKL</td>
<td>9</td>
<td>Nodal modulator 1</td>
<td></td>
</tr>
<tr>
<td>KRRSPSRL</td>
<td>9</td>
<td>Pyrroline-5-carboxylate reductase</td>
<td></td>
</tr>
<tr>
<td>KRRSPSRL</td>
<td>9</td>
<td>Unknown</td>
<td></td>
</tr>
<tr>
<td>KRRSPSRL</td>
<td>9</td>
<td>Zink finger CCHL domain-containing protein</td>
<td></td>
</tr>
<tr>
<td>KRRSPSRL</td>
<td>9</td>
<td>Protein PRX1C</td>
<td></td>
</tr>
<tr>
<td>KRRSPSRL</td>
<td>9</td>
<td>Isocitrate 3 of Oxidized-binding protein-related protein 9</td>
<td></td>
</tr>
<tr>
<td>KRRSPSRL</td>
<td>9</td>
<td>Eukaryotic peptide chain release factor subunit 1</td>
<td></td>
</tr>
<tr>
<td>KRRSPSRL</td>
<td>9</td>
<td>Mitochondrial inner membrane protein</td>
<td></td>
</tr>
<tr>
<td>KRRSPSRL</td>
<td>9</td>
<td>Mitochondrial inner membrane protein</td>
<td></td>
</tr>
<tr>
<td>KRRSPSRL</td>
<td>9</td>
<td>7 complex protein 1 subunit delta</td>
<td></td>
</tr>
<tr>
<td>KRRSPSRL</td>
<td>9</td>
<td>Nuclear protein 56</td>
<td></td>
</tr>
<tr>
<td>KRRSPSRL</td>
<td>9</td>
<td>Soluble center family 10 (Thiamine transporter)</td>
<td></td>
</tr>
<tr>
<td>KRRSPSRL</td>
<td>9</td>
<td>7EMT-like protein</td>
<td></td>
</tr>
<tr>
<td>KRRSPSRL</td>
<td>9</td>
<td>Unknown</td>
<td></td>
</tr>
<tr>
<td>KRRSPSRL</td>
<td>9</td>
<td>Unknown</td>
<td></td>
</tr>
<tr>
<td>KRRSPSRL</td>
<td>9</td>
<td>Peroxisome proliferator-activated receptor gamma coactivator-related protein 1</td>
<td></td>
</tr>
<tr>
<td>KRRSPSRL</td>
<td>9</td>
<td>Vesnatin</td>
<td></td>
</tr>
<tr>
<td>KRRSPSRL</td>
<td>9</td>
<td>Coiled-coil domain-containing proteins 186 (HAXXH20 protein)</td>
<td></td>
</tr>
<tr>
<td>QLSVQPLNL</td>
<td>9</td>
<td>Chp/p300 interacting transcriptional regulator 2</td>
<td></td>
</tr>
<tr>
<td>RPEPTGLRL</td>
<td>9</td>
<td>Peroxomembrane protein PXE16</td>
<td></td>
</tr>
<tr>
<td>RPEPTGLRL</td>
<td>9</td>
<td>Peroxomembrane protein PXE16</td>
<td></td>
</tr>
<tr>
<td>RPEPTGLRL</td>
<td>9</td>
<td>Pre-mRNA-processing factor 19</td>
<td></td>
</tr>
<tr>
<td>RQVAYGVLQ</td>
<td>9</td>
<td>Pre-mRNA-processing factor 19</td>
<td></td>
</tr>
<tr>
<td>SRIQGELK</td>
<td>9</td>
<td>Cytoplasmic domain 1 intermediate chain 2</td>
<td></td>
</tr>
<tr>
<td>STPSTNL</td>
<td>9</td>
<td>Vascular endothelial growth factor receptor 1</td>
<td></td>
</tr>
<tr>
<td>STPSTNL</td>
<td>9</td>
<td>Exobrinoreceptor microdomain-associated protein 8</td>
<td></td>
</tr>
<tr>
<td>STPSTNL</td>
<td>9</td>
<td>NOD-like receptor NOD1 and coiled-body phosphopeptidase 1</td>
<td></td>
</tr>
<tr>
<td>STPSTNL</td>
<td>9</td>
<td>N-acetyltransferase 10</td>
<td></td>
</tr>
<tr>
<td>STPSTNL</td>
<td>9</td>
<td>Nectinophil</td>
<td></td>
</tr>
<tr>
<td>STPSTNL</td>
<td>9</td>
<td>GCN5-related NAD-dependent histone deacylase</td>
<td></td>
</tr>
<tr>
<td>STPSTNL</td>
<td>9</td>
<td>Ras-related protein 1</td>
<td></td>
</tr>
<tr>
<td>STPSTNL</td>
<td>9</td>
<td>Lamine-rich repeat-containing protein 16A</td>
<td></td>
</tr>
<tr>
<td>STPSTNL</td>
<td>9</td>
<td>Laminin receptor 2</td>
<td></td>
</tr>
<tr>
<td>STPSTNL</td>
<td>9</td>
<td>Laminin chain 1 domain</td>
<td></td>
</tr>
<tr>
<td>STPSTNL</td>
<td>9</td>
<td>Transmembrane protein</td>
<td></td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Table 3</th>
<th>HLA-G 0.04restricted peptides.1</th>
<th>Length</th>
<th>Source</th>
</tr>
</thead>
<tbody>
<tr>
<td>KLLVQGEK</td>
<td>8</td>
<td>Protein FAM18B</td>
<td></td>
</tr>
<tr>
<td>KLLVQGEK</td>
<td>8</td>
<td>Myosin-specific enhancer factor 3C</td>
<td></td>
</tr>
<tr>
<td>KLLVQGEK</td>
<td>8</td>
<td>Unknown</td>
<td></td>
</tr>
<tr>
<td>KLLVQGEK</td>
<td>8</td>
<td>Zink finger protein 469</td>
<td></td>
</tr>
<tr>
<td>KLLVQGEK</td>
<td>8</td>
<td>Serine/threonine-protein kinase WNK2</td>
<td></td>
</tr>
<tr>
<td>KLLVQGEK</td>
<td>8</td>
<td>Zink finger protein 536</td>
<td></td>
</tr>
<tr>
<td>KLLVQGEK</td>
<td>8</td>
<td>Unknown</td>
<td></td>
</tr>
<tr>
<td>KLLVQGEK</td>
<td>8</td>
<td>Nodal modulator 1</td>
<td></td>
</tr>
<tr>
<td>KLLVQGEK</td>
<td>8</td>
<td>Polypeptide-1-c2-microglobin</td>
<td></td>
</tr>
<tr>
<td>KLLVQGEK</td>
<td>8</td>
<td>Unknown</td>
<td></td>
</tr>
<tr>
<td>KLLVQGEK</td>
<td>8</td>
<td>Glycine-RNA ligase</td>
<td></td>
</tr>
<tr>
<td>KLLVQGEK</td>
<td>8</td>
<td>DNA replication licensing factor MO33</td>
<td></td>
</tr>
<tr>
<td>KLLVQGEK</td>
<td>8</td>
<td>Transcription factor AP-4 (Activating enhancer binding protein 4)</td>
<td></td>
</tr>
<tr>
<td>KLLVQGEK</td>
<td>8</td>
<td>TATA-box-binding protein PTF00</td>
<td></td>
</tr>
<tr>
<td>KLLVQGEK</td>
<td>8</td>
<td>Thrombospondin-4</td>
<td></td>
</tr>
<tr>
<td>KLLVQGEK</td>
<td>8</td>
<td>Unknown</td>
<td></td>
</tr>
<tr>
<td>KLLVQGEK</td>
<td>8</td>
<td>Claquiquin 2</td>
<td></td>
</tr>
<tr>
<td>KLLVQGEK</td>
<td>8</td>
<td>NACHT, LIIR and PYD domains-containing protein 3</td>
<td></td>
</tr>
<tr>
<td>KLLVQGEK</td>
<td>8</td>
<td>Alpha-actinin-2</td>
<td></td>
</tr>
<tr>
<td>KLLVQGEK</td>
<td>8</td>
<td>Proteins FAM20B</td>
<td></td>
</tr>
<tr>
<td>KLLVQGEK</td>
<td>8</td>
<td>Follistatin homology domain-containing protein H</td>
<td></td>
</tr>
<tr>
<td>KLLVQGEK</td>
<td>8</td>
<td>Tenascin C protein 195 (Tenascin-C)</td>
<td></td>
</tr>
<tr>
<td>KLLVQGEK</td>
<td>8</td>
<td>Neurin-2</td>
<td></td>
</tr>
<tr>
<td>KLLVQGEK</td>
<td>8</td>
<td>Zink finger protein 669</td>
<td></td>
</tr>
<tr>
<td>KLLVQGEK</td>
<td>8</td>
<td>Unknown</td>
<td></td>
</tr>
<tr>
<td>KLLVQGEK</td>
<td>8</td>
<td>Leucine-rich repeat-containing protein 16A</td>
<td></td>
</tr>
<tr>
<td>KLLVQGEK</td>
<td>8</td>
<td>Kinesin-like protein 907</td>
<td></td>
</tr>
<tr>
<td>KLLVQGEK</td>
<td>8</td>
<td>Protein phosphatase 1 regulatory subunit 128</td>
<td></td>
</tr>
<tr>
<td>KLLVQGEK</td>
<td>8</td>
<td>Ig lambda chain V1 region NEWM</td>
<td></td>
</tr>
</tbody>
</table>

1 Peptides are presented in N-terminal to C-terminal orientation.

2 Peptides are presented in N-terminal to C-terminal orientation.
Table 4
Shared peptides between HLA-G*01:01, G*01:03 and G*01:04.a

<table>
<thead>
<tr>
<th>Sequence</th>
<th>#AA</th>
<th>G*01:01</th>
<th>G*01:03</th>
<th>G*01:04</th>
<th>Source</th>
</tr>
</thead>
<tbody>
<tr>
<td>KRRPGQQL</td>
<td>8</td>
<td>x</td>
<td>x</td>
<td>x</td>
<td>PDI domain-containing protein 11</td>
</tr>
<tr>
<td>KRRPGQNL</td>
<td>8</td>
<td>x</td>
<td>x</td>
<td>x</td>
<td>Myocyte-specific enhancer factor 2C</td>
</tr>
<tr>
<td>KPFFASM</td>
<td>8</td>
<td>x</td>
<td>x</td>
<td>x</td>
<td>Tumor-suppressor</td>
</tr>
<tr>
<td>TPYVEES</td>
<td>8</td>
<td>x</td>
<td>x</td>
<td>x</td>
<td>L-2B29-restricted peptide</td>
</tr>
<tr>
<td>TPYVEQSR</td>
<td>8</td>
<td>x</td>
<td>x</td>
<td>x</td>
<td>DNA repair protein REV1</td>
</tr>
<tr>
<td>KAAAPYQQL</td>
<td>9</td>
<td>x</td>
<td>x</td>
<td>x</td>
<td>Neoglucosidase</td>
</tr>
<tr>
<td>RPPQEGING</td>
<td>9</td>
<td>x</td>
<td>x</td>
<td>x</td>
<td>Chop-p300-interacting transactivator</td>
</tr>
<tr>
<td>AEQUGDVE</td>
<td>9</td>
<td>x</td>
<td>x</td>
<td>x</td>
<td>Zinc finger protein 469</td>
</tr>
<tr>
<td>ALOQNELD</td>
<td>9</td>
<td>x</td>
<td>x</td>
<td>x</td>
<td>Retinoic-acid-like protein 1</td>
</tr>
<tr>
<td>KSFPAASEL</td>
<td>9</td>
<td>x</td>
<td>x</td>
<td>x</td>
<td>Proteasome 20S ATPase subunit 1 variant</td>
</tr>
<tr>
<td>KSBASFM</td>
<td>9</td>
<td>x</td>
<td>x</td>
<td>x</td>
<td>Pyrimidine-5-carboxylate reductase 2</td>
</tr>
<tr>
<td>KNTYEREF</td>
<td>9</td>
<td>x</td>
<td>x</td>
<td>x</td>
<td>Zinc finger CCH domain-containing protein</td>
</tr>
<tr>
<td>KPYTTEESL</td>
<td>9</td>
<td>x</td>
<td>x</td>
<td>x</td>
<td>Ocytocin-binding protein-related protein 9</td>
</tr>
<tr>
<td>KVQAVVQL</td>
<td>9</td>
<td>x</td>
<td>x</td>
<td>x</td>
<td>Eukaryotic peptide chain release factor subunit 1</td>
</tr>
<tr>
<td>KVQYQYSL</td>
<td>9</td>
<td>x</td>
<td>x</td>
<td>x</td>
<td>T-complex protein 1 subunit delta 1</td>
</tr>
<tr>
<td>KPYTVYQI</td>
<td>9</td>
<td>x</td>
<td>x</td>
<td>x</td>
<td>Nuclear protein 50</td>
</tr>
<tr>
<td>MGVRDAMF</td>
<td>9</td>
<td>x</td>
<td>x</td>
<td>x</td>
<td>Unknown</td>
</tr>
<tr>
<td>PKLELASE</td>
<td>9</td>
<td>x</td>
<td>x</td>
<td>x</td>
<td>Peroxisome proliferation-activated receptor gamma coactivator-related protein 1</td>
</tr>
<tr>
<td>QLSRTPNL</td>
<td>9</td>
<td>x</td>
<td>x</td>
<td>x</td>
<td>Coiled-coil domain-containing protein 180</td>
</tr>
<tr>
<td>SBBQSHSM</td>
<td>9</td>
<td>x</td>
<td>x</td>
<td>x</td>
<td>Cytoplasmic dynein 1 intermediate chain 2</td>
</tr>
<tr>
<td>KEEQLYL</td>
<td>9</td>
<td>x</td>
<td>x</td>
<td>x</td>
<td>Peptidyl-prolyl cis-trans isomerase</td>
</tr>
<tr>
<td>KSPYTLNL</td>
<td>9</td>
<td>x</td>
<td>x</td>
<td>x</td>
<td>Glycine-rich RNA binding</td>
</tr>
<tr>
<td>KTFPQENGL</td>
<td>9</td>
<td>x</td>
<td>x</td>
<td>x</td>
<td>Eukaryotic peptide chain release factor subunit 1</td>
</tr>
<tr>
<td>YQNYRPOQDD</td>
<td>10</td>
<td>x</td>
<td>x</td>
<td>x</td>
<td>Ig kappa chain V-III region</td>
</tr>
<tr>
<td>KRPWQGQFE</td>
<td>10</td>
<td>x</td>
<td>x</td>
<td>x</td>
<td>NOLC1 protein (CDE) Nucleolar and coiled-body phosphoprotein</td>
</tr>
<tr>
<td>PGEVSDKD</td>
<td>10</td>
<td>x</td>
<td>x</td>
<td>x</td>
<td>Nucleolar</td>
</tr>
<tr>
<td>CPREDRAAED</td>
<td>12</td>
<td>x</td>
<td>x</td>
<td>x</td>
<td>Larval-rich repeat-containing protein 16A</td>
</tr>
<tr>
<td>TQRRSAGAQIQ</td>
<td>12</td>
<td>x</td>
<td>x</td>
<td>x</td>
<td>Ig lambda chain V-1 region NEMH</td>
</tr>
<tr>
<td>SYTAYTVFPR</td>
<td>13</td>
<td>x</td>
<td>x</td>
<td>x</td>
<td>Transhydrolase</td>
</tr>
</tbody>
</table>

a Presentation by allotype is marked with "x". Peptides are presented in N-terminal to C-terminal orientation and sorted by length (#AA).

Fig. 2. Peptide anchor positions. Anchoring positions of 9mer peptides depicted as sequence logos. Peptides presented by all alleles were preferentially anchored by Leucine at the C-terminus. (A and B) Peptides derived from HLA-G*01:01 or G*01:03 were additionally anchored by Proline at p3, respectively. (C) Depicts the sequence logo for HLA-G*01:04. (D) Model of HLA-G illustrating the AA exchange of G*01:03 (red) and G*01:04 (blue) respective to G*01:01. (PDB: 1YO9). (For interpretation of the references to colour in this figure legend, the reader is referred to the web version of this article.)

interaction with the receptor.
Polymorphism in immunologically engaged molecules offer a broad range of functions as indicating pathogenic invasions or the existence of foreign/non-self antigens. Classical HLA molecules represent the most polymorphic system with specialized population specific functions, while the non-classical HLA molecules usually operate as protective backups for the immune system. HLA-G is a classic example by offering protection for the fetus that constitutes an allograft for the maternal body by deceiving autologous consistency. Since the polymorphisms of non-classical HLA molecules are marginal it becomes obvious that these molecules are ligands for the innate immune response. However, studies revealed that peptide presentation of HLA-E is divers and pH2A-E molecules compete with certain immune receptors from the innate system. This means in turn that the peptide selection and presentation of HLA non-classical molecules are the key to immunology and tolerance. The unexpected diversity of HLA-G allele-restricted peptide
Fig. 3. Cytotoxicity assay. HLA-G01:0x expression on K562 cells confers protection against NK cell mediated killing. (A) CFSE labelled target cells were analyzed for 7-AAD staining and the percentage of specific lysis was calculated. (B) Specific lysis of K562 and K562/HLA-G’01:0x cells through isolated human NK cells from three different donors as well as the NK cell line NKL. Depicted are representative FACS plots and the calculated specific lysis is given in %.

selection and HLA-G allelic variability in immunogenic tolerance seems to be caused by structural alterations of the accessible pHLA-G surface for the immune receptors. This compelling evidence for HLA-G involvement in peptide selection suggest the need for HLA-G typing in patient management strategies for NK cell based anti-tumor treatment. HLA-G has an essential role in immune tolerance; the impact of polymorphic variants on immune function has been systematically analyzed. Allelic HLA-G subtypes differ in peptide binding motifs,
Funding

This work was supported by the Deutsche José Carreras Leukämie-Stiftung and the Deutsche Gesellschaft für Hämatologie und medizinische Onkologie [FS16/05].

Appendix A. Supplementary data

Supplementary data associated with this article can be found, in the online version, at https://doi.org/10.1016/j.jhimimm.2018.03.010.

References


3.2 Research Paper 2

**HLA-G peptide preferences change in transformed cells: impact on the binding motif**

Alexander A. Celik¹ · Gwendolin S. Simper¹ · Wiebke Hiemisch¹ · Rainer Blasczyk¹ · Christina Bade-Döding¹

Received: 13 February 2018 / Accepted: 23 March 2018 © The Author(s) 2018

**Abstract**

HLA-G is known for its strictly restricted tissue distribution. HLA-G expression could be detected in immune privileged organs and many tumor entities such as leukemia, multiple myeloma, and non-Hodgkin and Hodgkin’s lymphoma. This functional variability from mediation of immune tolerance to facilitation of tumor immune evasion strategies might translate to a differential NK cell inhibition between immune-privileged organs and tumor cells. The biophysical invariance of the HLA-G heavy chain and its contrary diversity in immunity implicates a strong influence of the bound peptides on the pHLA-G structure. The aim was to determine if HLA-G displays a tissue-specific peptide repertoire. Therefore, using soluble sHLA-G technology, we analyzed the K562 and HDLM-2 peptide repertoires. Although both cell lines possess a comparable proteome and recruit HLA-G-restricted peptides through the same peptide-loading pathway, the peptide features appear to be cell specific. HDLM-2 derived HLA-G peptides are anchored by an Arg at p1 and K562-derived peptides are anchored by a Lys. At p2, no anchor motif could be determined while peptides were anchored at p2 with a Leu and showed an auxiliary anchor motif Pro at p3. To appreciate if the peptide anchoring alterations are due to a cell-specific differential peptidome, we performed analysis of peptide availability within the different cell types. Yet, the comparison of the cell-specific proteome and HLA-G-restricted ligandome clearly demonstrates a tissue-specific peptide selection by HLA-G molecules. This exclusive and unexpected observation suggests an exquisite immune function of HLA-G.

**Keywords** HLA-G · Hodgkin’s lymphoma · Peptide presentation · Tumor immune escape

**Introduction**

Human leukocyte antigen (HLA)-G, a non-classical HLA class Ib molecule, plays an important role in immune protection and is in contrast to HLA class Ia molecules exclusively expressed at immune privileged sites (Kovats et al. 1990). Unlike for classical HLA class Ia alleles, only 58 alleles of HLA-G are described (Robinson et al. 2015), which result in even less protein isoforms. Among these, the HLA-G*/A01:01 allele is the most prevalent in different European populations (Castelli et al. 2014; Matte et al. 2000). Additionally, various splice variants of HLA-G can be processed (Hvid et al. 1998; Ishitani and Geraghty 1992) resulting in four membrane-bound forms (HLA-G1–G4) and three soluble forms (HLA-G5–G7). HLA-G1 constitutes the full-length HLA-G molecule that closely resembles other HLA class I molecules; however, alternative splicing leads to the exclusion of one or more α-domain in the HLA-G2, G3, or G4 isoforms. In addition to these, soluble isoforms corresponding to the membranous counterparts (e.g., HLA-G5 constitutes the soluble form of HLA-G1) are achieved by the retention of a stop codon after exon 4 (Dong et al. 2003; Gonen-Gross et al. 2005) or by cleavage of membranous HLA-G1 from the cell surface (soluble HLA-G1). However, while it is possible to detect these mRNA variants, the protein expression and biological function of these truncated forms remain elusive, while the HLA-G1, soluble HLA-G1, and HLA-G5 isoforms appear to be the most expressed forms in healthy tissue (Dahl et al. 2014; Paul et al. 2000). In the placenta (Kovats et al. 1990), immune tolerance is conferred by the expression of HLA-G on extravillous trophoblasts (Apps et al. 2009) and through secretion of soluble HLA-G isoforms (Fournel et al. 2000). HLA-G interacts with different subsets of immune effector cells (NK, T, B, macrophages), usually resulting in inhibition of
these cells (Bainbridge et al. 2000; Li et al. 2009; Naji et al.
2014; Rouas-Freiss et al. 1997). Inhibition of NK cell mediated
cytolysis is facilitated through KIR2DL4 activation
(Rajagopalan and Long 1999). T cell activity is inhibited by
interaction of HLA-G with ILT-2 and ILT-4 (Riteau et al.
1999) and additionally leads to unresponsive and suppressive
CD4+ T cell phenotypes (LeMoult et al. 2004). Purified sol-
uble HLA-G was shown to lead to apoptosis in activated
CD8+ T cells (Fournel et al. 2000) and suppression of prolif-
eration in CD4+ T cells (Bainbridge et al. 2000). Soluble
HLA-G also interferes with immune regulation by induction of
CD4+CD25+Foxp3+ Tregs (Selmani et al. 2008) and
Tr1 cells; induction by HLA-G+ IL-10 producing DCs was
shown in vitro (Gregori et al. 2010). In addition, it was shown
that interaction of ILT-4 can also occur with β2-m free forms
of HLA-G (Gomis-Gross et al. 2005).

Structurally, HLA-G appears similar to other HLA class I
molecules. The α1 and α2 domain form the peptide-binding
cleft that presents self-peptides of preferably 9 amino acid
(AA) residues in length (Diehl et al. 1996; Lee et al. 1995)
to the immune system. HLA-G-restricted peptides feature an-
chor motifs at peptide position p2 in the form of Isoleucine (I)
and Leucine (L), as well as Leucine (L) at p2L with a strong
Proline (P) auxiliary anchor at p3. X-ray analyses of peptide/
HLA-Gβ2m complexes showed that peptide presentation is
facilitated analogous to other HLA class I proteins, although
the peptide is positioned deeper in the cleft similar to the
peptide presentation on HLA-E (Clements et al. 2005;
Clements et al. 2007). Structural insight into KIR2DL4 com-
plexed to HLA-G illustrated that Methionine76 (M76) and
Glutamine79 (Q79) in the α1 domain are critical for
KIR2DL4 recognition, although specific binding sites of the
receptor remain unknown (Yan and Fan 2005) and any poten-
tial peptide interactions remain intangible. In contrast, similar
to the interaction of other KIRs with MHC-I, interaction
between HLA-G and the ILT-2 or ILT-4 receptor is thought to be
facilitated mainly through the α3 domain.

Hodgkin's lymphoma (HL) is a malignant lymphatic disor-
der characterized by an abnormal type of mature B cells called
Hodgkin-Reed-Sternberg (HRS) cells (Swerdlow et al. 2017).
Classic HL (cHL) is designated by small amounts of HRS cells
that are surrounded by normal immune cells. Similar to other B
cell lymphomas, HRS cells tend to loose many of the typical B
cell lineage factors and the expression of the B cell receptor
(Kanzer et al. 1996; Schwering et al. 2003). Concurrently,
expression of immune checkpoint proteins such as PD-L1
(Green et al. 2010; Roenner et al. 2016) and CD80
(Kosmaczewski et al. 2002; Murray et al. 1995) helps to evade
immune recognition. At the same time, downregulation of
HLA class I expression is a common feature especially in
EBV-negative HL (Diepstra et al. 2008), thus preventing the
presentation of neo-antigens. Studies on HRS cell lines per-
formed by Liu et al. (2017) showed that each cell line affected
overall HLA expression by mutations in either genes encoding
for β2m, HLA- A, or the class II major histocompatibility com-
plex transactivator (CIITA). Notably, lack of HLA surface ex-
pression should induce an NK cell response of the HLA class
I HRS cells; however, the reactive infiltrate surrounding the
cancer cells does not show increased presence of NK cells and
absence of NK cell-mediated cytotoxicity. Potent inhibitors
of NK cell activation include HLA class Ib molecules HLA-E and
HLA-G. Indeed, the expression of HLA-G was suggested to be
a potential immune evasion mechanism in HL. Diepstra et al.
(2008) found HLA-G expression in 54% of HRS cells of
lymphoma patients, and notably, HLA-G expression was asso-
ciated with a lack of HLA class Ia surface expression. Further
studies by Caocci et al. (2016) demonstrated that 55% of
HRS cells from lymph nodes of cHL patients were positive
for HLA-G expression. Interestingly, the surrounding lympho-
cytes and histocytes were HLA-G positive and correlation
with disease progression suggests an association of these
immunoreactive patterns of HRS cells and the tumor microenvi-
nronment on disease outcome. Abrerrant HLA-G expression
was also described in other hematological malignancies such as
non-Hodgkin lymphoma (Sebti et al. 2007), chronic lymphatic
leukemia (Sebti et al. 2007), and multiple myeloma (Leleu
et al. 2005) as well as in solid tumors, such as breast cancer
(Konig et al. 2016; Rebmam et al. 2003), and non-small cell
lung cancer, particularly in advanced disease stages (Ben
Amor et al. 2016; Yan et al. 2015).

For HLA class Ia proteins, the variability of the heavy chain
dictates the pathway of peptide recruitment and restricts the
amount of presentable peptides from the available peptidome.
The HLA class Ia molecule and/or the bound peptide together
dictate the fate of a presenting cell. The nature of HLA-bound
peptides influences the overall stability of a peptide-HLA com-
plex (pHLA). Due to increased stability, prolonged surface
expression of pHLA complexes implements enhanced T cell
immunogenicity (Trujillo et al. 2014; Yewdell and Bennik
1999). The present peptide repertoire is dependent on the
specific HLA allele and the available peptidome, e.g., in
HLA-B27 alleles; it was suggested that depending on the
subtype, differential interaction with tapasin (TPN) or the
transporter associated with antigen processing (TAP) influ-
ences the presented peptide (Lopez de Castro et al. 2004),
and in the case of HLA-B*44:28, it was shown that position
156 influences HLA/TPN association (Badrinarth et al. 2012).
However, of the potential tens of thousands of different pHLA
combinations available in the cell, the presentation of certain
self-peptides is favored (Hunt et al. 1992) and of these, only
few are immunogenic (Assarsson et al. 2007; Hamdahl et al.
2012). Wang et al. (1997) further showed that alloergic re-
ponses against low-abundance epitopes are not limited by
lower affinity of T cells for non-self MHC molecules, yet, the
exchange of a single αα can dramatically alter the affinity of
specific T cell receptors (Uchtenhagen et al. 2013).
Immunogenetics

The described differential immune functions of HLA-G make a biological sense of the HLA-G invariability doubtful and a variability and involvement of bound peptides obvious. HLA-G is non-polymorphic, however, able to differentiate in certain protein products dependent on the tissue where it is expressed. In this study, we focused on HLA-G*01:01 as the most prevalent allele in different populations and compared the peptide features after expression in two different cell types. We utilized the erythroblastic HLA-neg. cell line K562, a standard cell line for HLA allele-specific peptide determination and the cell line HDLM-2 that expresses an HRS phenotype and is frequently used as an HLA model cell line (Berglund et al. 2003) for the comparison of HLA-G tissue-specific peptide selection. The tissue specificity of peptides and the involvement of peptide sequences to the overall available surface for immune receptors seem distinct and a comprehensive analysis of HLA-G-restricted peptides indispensable. A deep insight into the biochemistry of pHLA-G complexes will guide towards understanding immune escape mechanism in tumors.

Material and methods

Antibodies

Anti-β2m (polyclonal, #A0072, DukO), anti-CRT (polyclonal #ABR-01176, Dianova), anti-ERP57 (polyclonal #ADI-SPA-585, Enzo Life Sciences), anti-HLA-A/B/C (W6/32, AbD Serotec®), anti-HLA-A/B/C-PE (W6/32, eBioscience), anti-HLA-G (MEM-G/9, Thermo Fisher Scientific), anti-TAP1 (polyclonal #ADI-CSA-620, Enzo Life Sciences), anti-TAP2 (polyclonal #ADI-CSA-625 J, Enzo Life Sciences), anti-V5 (MCA1360, AbD Serotec), rabbit anti-mouse IgG-HRP (polyclonal, #P0161, DukO), goat anti-rabbit IgG-HRP (polyclonal, #P0448, DukO), and rat anti-mouse IgG-PE (RMG1-1, Biolegend).

Maintenance of cell lines

The cell lines HDLM-2, K562, and T2 were maintained in RPMI1640 (Lonza) supplemented with 10% heat inactivated FCS (Lonza), 2 mM L-glutamine (c. pro.), 100 U/ml penicillin, and 10 µg/ml streptomycin (c. pro.). For transduction, lentiviral particles were produced in HEK293T cells that were cultured in DMEM (Lonza) supplemented with 10% heat inactivated FCS, 2 mM L-glutamine, 100 U/ml penicillin, 100 µg/ml streptomycin, and 1 mg/ml Geneticin® (Life Technologies). Cell lines were maintained at 37 °C and 5% CO2.

Cloning of HLA-G variants and generation of HLA-G expressing cell lines

Constructs encoding for full-length HLA-G*01:01 (exon 1–6) were generated from JEG-3 cDNA and subcloned into the lentiviral vector pRRL.PPT.SFFV.mcs.pre as previously described (Bade-Doeding et al. 2011). Site-directed mutagenesis was used to generate soluble sHLA-G*01:01 (exon 1–4) vectors as described previously (Celik et al. 2016). The respective inserts were verified through sequencing. Using the methods established by Bade-Doeding et al. (2011), K562 and T2 cells were stably transduced with full-length HLA-G*01:01, K562 and HDLM-2 cells were stably transduced with sHLA-G*01:01. Expression and efficiency of recombinant protein expression were confirmed by FACS, ELISA, or Western blot.

Large-scale production of soluble HLA-G molecules and mass spectrometric analysis

Using soluble HLA technology (Kunze-Schumaecher et al. 2014), sHLA-G*01:01 molecules were produced in bioreactors (Integra Biosciences), supernatant was harvested weekly, pooled, and filtered using a 0.45 µm membrane (Merck Millipore). After purification at pH 8.0 using NHS-activated HiTrap columns (Life Technologies) coupled to the HLA-G-specific monoclonal antibody MEM-G/9 or HLA class I specific monoclonal antibody W6/32, sHLA-G molecules were eluted (100 mM glycine/HCl buffer pH 2.7) from the column and the functionality of trimeric complexes verified using an W6/32-based ELISA. For separation of peptides from the complexes, 3 mg of each purified sHLA-G sample was treated with trifluoroic acid (J. T. Baker) at a final concentration of 0.1% and separated from heavy chain and β2m using a 10 kDa cutoff membrane (Merck Millipore). The peptide solution was vacuum concentrated and treated with NH4HCO3, DTT, iodoacetamide, and trichloroacetic acid. After purification with 0.1% TFA on Pierce™ C18 Tips (Thermo Fisher Scientific), sample was eluted with 60% acetonitrile/0.1% TFA, vacuum-dried and solubilized in 30 µl 2% acetonitrile/0.1% TFA. Utilizing LC/MS (Dionex UltiMate 3000 high-performance LC system coupled to LTQ Orbitrap Velos Hybrid FT Mass Spectrometer), peptides were analyzed.

Enzyme-linked immunosorbent assay

Ninety-six-well plates (NUNC Maxisorp) were coated with 500 µl/well W6/32 antibody in PBS overnight at 4 °C. After washing the plate 2x with 250 µl PBS + 0.05% Tween/well, the plate was blocked for 1 h with 250 µl/well with blocking solution (PBS containing 2% BSA w/v). The plate was loaded with protein standard and sample diluted in blocking solution and incubated for 2 h at RT. After washing the plate 4x with 250 µl PBS/T, 100 µl/well anti-β2m antibody was added.
After 1 h incubation at RT, the plate was then washed 3×. One hundred microliter per well anti-rabbit HRP-conjugated antibody was added and incubated for 1 h at RT. The plate was washed 6× with PBS/T and 100 μl TMB One™ substrate (KEM-EN-TEC Diagnostics) was added. The reaction was stopped using 100 μl/well acidic stop solution (3 M sulfuric acid, 1 M HCl in ddH2O) and the plate analyzed using a Synergy 2 ELISA microplate reader.

**Western blotting**

Proteins were boiled at 95 °C for 15 min in SDS sample buffer containing reducing agent (Invitrogen by Life Technologies™) and separated on a 4–12% Bis-Tris gel (Invitrogen by Life Technologies™). After transfer to a PVDF membrane (Invitrogen by Life Technologies™), the membrane was blocked in 5% milk powder (Roth) in PBS (w/v). Incubation of primary antibodies was performed overnight at 4 °C. The next day, the membrane was washed 3× with PBS/T, incubated with secondary HRP-conjugated antibody for 1 h at RT, and finally washed again 3× with PBS/T. Clarity™ ECL substrate (BioRad) was added and the chemiluminescent signal detected using a FluorChem® FC2 imaging system.

**Immunoprecipitation of components of the peptide loading complex (PLC)**

Immunoprecipitation experiments were performed essentially as described by Badrinath et al. (2012). Briefly, 1 × 10⁷ cells were lysed for 1 h in 400 μl lysis buffer, TBS containing 5 mg/ml digitonin (Sigma Aldrich) and protease inhibitors (Roche Diagnostics), on ice. The lysate was centrifuged at 16000×g for 15 min at 4 °C. The supernatant was precleared for 1 h at 4 °C with protein A sepharose (GE Healthcare) and end-to-end mixing and centrifuged for 1 min at 4000×g at 4 °C; a lysis control was set aside. This was followed by immunoprecipitation of proteins for 1 h at 4 °C with protein A sepharose coupled to 3 μg anti-HLA-G antibody. Subsequently, the samples were centrifuged for 1 min at 4000×g and 4 °C, the beads were washed 2× with 10 mM TrisHCL pH 7.4 containing 1 mg/ml digitonin and 450 mM NaCl followed by one washing step with 10 mM TrisHCL pH 7.4 containing 450 mM NaCl and one with 10 mM TrisHCL pH 7.4. Precipitated protein was eluted by boiling the samples in SDS buffer and used for final analysis.

**Analysis of HLA-G expression on transduced cells**

5 × 10⁵ cells were washed 2× with 2 ml FACS buffer (PBS containing 2% FCS and 2 mM EDTA) at 500×g for 5 min at 4 °C. Samples were then incubated with 100 μl Fe-block (10% AB-serum in PBS) for 20 min at 4 °C followed by incubation with an anti-HLA-G antibody (MEM-G/9) directly in the Fe-block for 20 min at 4 °C. After washing 2× with 2 ml FACS buffer, the cells were incubated with fluorochrome-labeled secondary antibody for 20 min at 4 °C and washed 2× with 2 ml FACS buffer. Finally, the cell pellet was resuspended in 250 μl buffer and analyzed using a BD FACSCanto™ II.

**Results**

**HLA-G*01:01 restricted peptide features**

To determine the HLA-G*01:01 restricted peptide repertoire, HDLM-2 or K562 cells were stably transduced with sHLA-G*01:01 and, after cultivation in bioreactors, soluble HLA molecules were affinity purified. After isolation of HLA-G restricted peptides, peptide sequences were analyzed by mass spectrometry. A total of 85 or 93 individual peptides were found in HDLM-2 or K562 cells (Fig. 1a), respectively. Only two shared peptides from K562 and HDLM-2 cells, respectively, could be detected, PNLTHASL and RHPQPGAVEL (Fig. 1a). Peptides of 8 AA in length were observed in 17.0% (HDLM-2) or 23.2% (K562) of the cases; however, most of these peptides (Fig. 1b), 60.0% (HDLM-2) or 41.1% (K562), featured a length of 9 AA, while peptides of 10 AA in length could be virtually only detected in K562 cells (16.1%), constituting for only 2.4% in HDLM cells. Similarity, peptides longer than 10 AA were found in 17.1% (K562) or 1.3% (HDLM-2) of the cases. The protein source of most peptides from both cell lines derived from either nucleic proteins (24.1% HDLM-2, 30.8% K562), the cytosol (19.3% HDLM-2, 17.6% K562) or from proteins shuttling between the nucleus, and the cytosol or specialized compartments (45.8% HDLM-2, 41.8% K562, Fig. 1c). Compartment-specific peptides (ER, Golgi, membranes) were found in 6.0, 1.2, and 3.6% (HDLM-2) or 5.5, 1.1, and 3.3% (K562), respectively.

**The peptide anchor at p1 is tissue specific**

Comparing the frequencies of individual AAs at every position in the sequences of identified 9-meric peptides reveals unexpected differences. An overview of AA frequencies at every peptide position (Fig. 2a) reveals a pattern that appears different at position p1 and p2, whereas at position p3 and p4, similarities are observed. All peptides are anchored at position p1, p3, and p0 (Fig. 2b-e). A striking occurrence of Lysine (L) at p0 could be observed with 78.3% for HDLM-2-derived peptides and 60.2% for K562-derived peptides. An anchor motif at p2 could not be defined (Fig. 2c). Noticeable peptides are anchored at p1 were furthermore a tissue-specific anchor motif could be observed; HDLM-2-derived peptides feature an Arginine (R) at p1 (69.9%) and K562-derived peptides a Lysine (K) (36.1%). A strong anchor motif at p3 Proline (P)
for all peptides (39.8% HDLM-2, 29.0% K562) could be detected as previously specified (Diehl et al. 1996).

Fig. 1 HLA-G*01:01-restricted peptide features. Peptides presented on HLA-G*01:01 were eluted and sequenced after expression in K562 or HDLM-2 cells. a Depicted is the total of individual, single peptide sequences from both cell lines. Between individual peptide sequences, overlap was observed in only two instances (PNLTHLASL, RHPQPGAVEL). b Depicted are the frequencies of the peptide length (number of AAs) that were observed. Most peptides showed canonical length of 9 AA, although longer peptides were also observed in fewer instances. Peptide frequencies are given on the y-axis; length of AAs is given on the x-axis. c Depiction of the localization of the protein source of the sequenced peptides. In both cell lines, peptides are recruited mostly from proteins shuttling between compartments, the nucleus and the cytosol and loading pathway. Therefore, the association of HLA-G with major proteins of the PLC was analyzed by immune-precipitation experiments using an HLA-G-specific antibody. Immuno-precipitation of shHLA-G*01:01 in K562 and HDLM-2 transfectants showed association with (CRT), ERp57, and TPN; however, no interaction with TAP could be detected (Fig. 3a). To assess whether HLA-G molecules are still present on the surface in the case of TAP deficiency, recombinant expression of HLA-G on the surface of TAP-deficient T2 cells was analyzed (Fig. 3b). The expression of HLA-G on the surface of these cells supported the observation that peptide loading of HLA-G occurs in the absence of TAP. The unconventional peptide loading pathway of HLA-G in both cell lines leads to the conclusion that the differential peptide repertoire might not be due to differential pathways.

**Discussion**

HLA-G is a potent immune inhibitor and well known to induce tolerance during pregnancy (Hunt et al. 2006; Kovats.
A) HDLM-2 + K562 restricted peptide features

![Graph showing HDLM-2 restricted peptides and K562 restricted peptides]

B) AA residues at p1

![Graph showing frequency of AA residues at p1]

C) AA residues at p2

![Graph showing frequency of AA residues at p2]

D) AA residue at p3

![Graph showing frequency of AA residue at p3]

E) AA residue at pΩ

![Graph showing frequency of AA residue at pΩ]

Fig. 2 Difference in peptide features based on tissue derivation. a Frequencies of specific AAs at position 1 to 9 of identified 9-mers. AAs are given in one-letter code, higher frequency occurrences are depicted in more saturated color. b, c, d, e Depiction of the frequencies of AA residues at anchoring positions p1, pΩ, p3, and pΩ. AAs are given in one-letter code on the x-axis; frequencies are depicted on the y-axis. A 15% threshold is represented by a dashed line.

et al. 1990). In contrast to HLA class Ia, expression of HLA-G is usually confined to the placenta and certain immune-privileged tissues; however, ectopic expression was found in many different tumor entities (Ben Amor et al. 2016; Leclu et al. 2005; Rehmann et al. 2003; Sebti et al. 2007) where it is thought to feature a tolerogenic effect on different subsets of immune cells including CD56⁺ NK cells, CD4⁺CD8⁺ T cells, and B cells. Such non-reactive immune effectors are also present inside the reactive infiltrate surrounding HRS cells in cHL (Liu et al. 2014). HLA-G is discussed as an immune evasion mechanism in cHL (Caocci et al. 2016; Diepstra et al. 2008), since typically more than half of the HRS cells are positive for HLA-G expression, while HLA class Ia expression is downregulated at the same time appearing feasible that HLA-G constitutes an immune evasion mechanism in cHL. The downregulation of HLA class Ia molecules is a meaningful immune escape strategy since the polymorphic nature of HLA class Ia molecules enables them to present a wide variety of peptides covering plenty of a pathogenic peptidome (Wang et al. 2017). Additionally, through the limitation of HLA class Iib molecules to present a narrow variety of
Table 1 HDL-M-2-specific HLA-G-bound peptides from proteins with increased expression in K562 cells and K562-specific HLA-G-bound peptides from proteins with increased expression in HDL-M-2 cells

<table>
<thead>
<tr>
<th>Sequence</th>
<th>#AA</th>
<th>Origin</th>
<th>Source</th>
<th>TPM† in K562</th>
</tr>
</thead>
<tbody>
<tr>
<td>RGPGPRKPL</td>
<td>9</td>
<td>ELF4b</td>
<td>HDL-M-2</td>
<td>2.1×</td>
</tr>
<tr>
<td>RVDKAVAL</td>
<td>9</td>
<td>PDE4DIP</td>
<td>HDL-M-2</td>
<td>16.7×</td>
</tr>
<tr>
<td>RARQKIDKDL</td>
<td>9</td>
<td>ERGIC3</td>
<td>HDL-M-2</td>
<td>2.9×</td>
</tr>
<tr>
<td>RKKSADTL</td>
<td>9</td>
<td>LDHA</td>
<td>HDL-M-2</td>
<td>1.4×</td>
</tr>
</tbody>
</table>

Sequence: #AA Origin Source TPM† in HDL-M-2
| 8 | ME2C | K562 | 2.7× |
| 9 | MAPK3 | K562 | 4.6× |
| 9 | FAM91AI | K562 | 3.5× |
| 10 | SYNE2 | K562 | 24.0× |

mRNA expression data from the human protein atlas (https://www.proteinatlas.org) (Thul et al. 2017). TPM transcripts per million, † increase content available inside the cell and therefore in turn predetermine the available ligandome (de Verteuil et al. 2010; Fortier et al. 2008). Early peptide elution studies reported that HLA-G presents only a restricted set of peptides (Diehl et al. 1996); in term placenta, these are practically only derived from a cytokine receptor-like molecule (Ishihama et al. 2003). Apart from peptide invariance, HLA-G constitutes a classic peptide presenter presenting 9-mers as shown by pool sequencing studies of Diehl et al. (1996) and Lee et al. (1995). Most of the peptides we observed also displayed canonical length and were primarily derived from proteins available in the cytosol and the nucleas as well as proteins shuttling between these two compartments.

Considering the previously depicted conserved nature of HLA-G ligands, the variety of peptides and difference in the presented peptide repertoire from the two different sources, the erythroleukemic cell line K562 and the Hodgkin's lymphoma cell line HDL-M-2, was remarkable. The p3 anchor Leu as well as Pro as an auxiliary anchor at p3 (Diehl et al. 1996) was described previously, although in both cases, p1 was divided between Arg and Lys (Lee et al. 1995) when expressed in LCL721.221 cells. We could identify both motifs; however, in the present study, the peptide anchor at p1 could be stringently attributed to the cell type the HLA-G molecules were expressed in. It is known for other class I alleles that peptide presentation is tissue specific (Fortier et al. 2008); however, anchoring positions are usually allele specific (Madden 1995; Yewdell and Bennink 1999). It is well known and for many HLA class Ia alleles established that allelic features dictate the peptide-binding motif. Therefore, a mismatch can be weighted by magnitude on the peptide motif. In the background of these HLA G acknowledgments, it was unforeseeable that HLA-G features an alteration of the peptide motif depending on the tissue. The obvious question occurs whether the differential peptide motif arises by virtue of peptide availability in the different cell types. Hence, we analyzed the online available expression levels of protein sources from the human protein atlas (Thul et al. 2017) to define a tissue-specific ligandome. Most interestingly, we found that peptide selection and presentation by HLA-G*01:01 is not a matter of source availability, since K562 cells have the same or higher

**A) Immuno-precipitation**

**B) Recombinant HLA-G*01:01 expression**

**Fig. 3** Expression of recombinant HLA-G molecules in HDL-M-2 and K562 cells. a Immuno-precipitation of sHLA-G*01:01 was performed using an anti-HLA-G antibody (MEM-G/9) in HDL-M-2 as well as K562 cells and Western blot against proteins of the PLC was performed. Notably, no association with TAP was observed. LC lysate control, IP immune-precipitation. b Using an anti-HLA-G-specific antibody (MEM-G/9), recombinant expression of HLA-G*01:01 on the surface of TAP-negative T2 could be detected.
protein source availability for peptide presentation as HDM-L2 cells and vice versa. This strengthens the assumption that HLA-G selects and presents peptides of different anchor motifs tissue specifically. An explanation for the unexpected and unusual diverging peptide anchors for the same allele with virtually the same proteomic content might be the presence of different peptide selection and loading pathways. Therefore, we aimed to analyze the association of HLA-G*01:01 with major proteins from the PLC. Loading of optimized peptides is usually facilitated by the PLC. Here, TAP is important for peptide translocation and through association with TPN. TAP is allocated to the immediate vicinity of the PLC and thus the HLA molecule (Cresswell et al. 1999). This facilitates efficient loading of peptides that are 8–12 AA (Androlewicz and Cresswell 1996) in length and aides peptide optimization (Blum et al. 2013; Williams et al. 2002) by TPN. Nevertheless, it has been reported before that certain ER-derived antigens can be presented in the case of TAP impairment in mice and that these can present recognizable CTL epitopes (Durgeau et al. 2011). Additionally, it has been demonstrated that DCs are able to present antigens on MHC class Ia molecules TAP independently through an endolysosomal vesicular pathway, especially when stimulated by an TLR9 activator (Chen and Jondal 2009). In the present study, we could demonstrate that HLA-G*01:01 does not associate with TAP and additinally presents pHLA-G complexes on the surface of TAP-negative cells. A TAP-independent peptide loading might indicate a tissue-specific peptide selection of HLA-G, probably controlled through the presence of certain cytokines. Post-HLA class Ia downregulation and HLA class Ib upregulation, an increase in IL-10 production, could be traced in lung cancer (Urosevic et al. 2001). This is in line with findings that IL-10 expression is elevated in non-Hodgkin lymphoma (el-Far et al. 2004; Sebit et al. 2007) and findings that show that HLA-G expression is inducible by IL-10 (Moreau et al. 1999) in monocytes and human trophoblasts, highlighting the capabilities of prolonged HLA-G expression in immunosuppressive environments.

One of the limitations of our study is that it focused on the HLA-G*01:01 peptide presentation using transduced cell lines that does not take different HLA-G-specific regulatory factors into account. HLA-G transcription and also translation are known to be impacted by intron variations and variations in the UTR. Recently, Misra et al. (2014) observed a protective effect in end-stage renal disease (ESRD) and acute allograft rejection for HLA-G*01:01:01 and G*01:04:01 haplotypes whereas G*01:01:01:03, G*01:01:02, G*01:06, and G*01:05 N were risk associated. Additionally, higher levels of soluble HLA-G isoforms G5 and G6 were also present in ESRD cases suggesting that differences in the expression profile may modulate risk for ESRD and acute allograft rejection. Furthermore, findings by Agrawal et al. (2015) elucidated that certain HLA-G 5’ UTR SNPs increase the risk for idiopathic recurrent spontaneous abortion. Potentially, mutations that impact HLA-G transcript stability may also impact the generation of stable pHLA-G complexes. Additional studies, e.g., crystallographic analyses are further needed to aid understanding of the scope of such differential peptide presentations.

However, appreciating the results of peptide anchor alteration and tissue-specific peptide selection facilitates the understanding of the exquisite immune function of HLA-G and its amazing flexibility in the mediation of tolerance.

**Funding Information** The authors would like to thank the Deutsche José Carreras Leukämie-Stiftung and the Deutsche Gesellschaft für Hämatologie und medizinische Onkologie for funding.

**Open Access** This article is distributed under the terms of the Creative Commons Attribution 4.0 International License (http://creativecommons.org/licenses/by/4.0/), which permits unrestricted use, distribution, and reproduction in any medium, provided you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons license, and indicate if changes were made.

**References**


Braud V, Jones EY, McMichael A (1997) The human major histocompatibility complex class I molecule HLA-E binds signal sequence-


4 RESULTS AND DISCUSSION

HLA class I molecules are highly polymorphic and pivotal for the immune system to surveille the health status of every single cell in the body. T cell receptors scan the intracellular proteome through the presentation of antigens on the surface of HLA class Ia molecules (HLA-Ia). The simultaneous focus of the receptors interactions with host MHC class Ia and fragments of foreign antigens is unique in receptor-ligand interactions. The inability of cells during foreign invasion episodes to further present HLA-Ia/peptide complexes (117) support the expression of HLA-Ib/peptide complexes that depict ligands for the innate immune receptors (118). The requirement of a deep understanding of the biological and biophysical function of HLA-Ib molecules becomes clear. The lack of diversity in the AA composition of the heavy chain underpins the assumption that the tissue restriction of HLA-Ib molecules might be compensated through peptide diversity. We could previously demonstrate that the interaction between NK receptors and HLA-E molecules is based on the same peptide-mediated biological system as TCR/HLA-Ia interactions (57).

In comparison to HLA-E and HLA-F, HLA-G is the most polymorphic of the non-classical HLA molecules, while it simultaneously exhibits the most restricted tissue distribution (5, 77). Of the 58 alleles, HLA-G*01:01 is by far the most prevalent allele in any population worldwide, followed by HLA-G*01:03 and G*01:04 as most common alleles in European populations (61). Compared to HLA-E, where a very specialized role by presenting a constrained set of peptides derived from the signal sequence of other HLA molecules (62) could be assumed, HLA-G features a greater variability in its repertoire allowing the interaction with a diverse set of immune effector cells. However, almost any interaction leads to an inhibition of the immune system (77-79). It was proposed early on that HLA-G*01:01 constitutes a classical peptide presenter whose potential for antigen presentation is constrained through its restricted tissue distribution (66). Because of the low degree of polymorphisms between HLA-G*01:01, G*01:03 and G*01:04, each distinguished by a single AA difference located outside the peptide binding region (PBR), it was assumed that a functional difference between these variants is unlikely. Nevertheless, we could previously demonstrate that a single AA exchange in the loop region of the α2-domain of HLA-E*01:03 impacts peptide presentation (56, 57). HLA-Ib molecules are part of a conserved system including mostly invariant receptors; hence it becomes obvious that any structural peptide-mediated alteration has the potential to greatly impact the dynamics of interaction with their cognate receptors. The first part of this project was focused on the molecular basis of immune recognition mediated by HLA-G variants and the impact of single AA exchanges on NK cell recognition. Since HLA-G expression was proposed to be an important immune evasion mechanism in
Hodgkin’s lymphoma (113, 114), the second part of this study was focused on the expression of HLA-G in an Hodgkin’s lymphoma (HL) model cell line that exhibits the Hodgkin-Reed-Sternberg (HRS) phenotype.

To analyze HLA-G*01:01, G*01:03 and G*01:04, the HLA negative cell line K562 was stably transduced with lentiviral constructs encoding for either HLA-G variant. K562 cells represent the gold standard for the analysis of NK cell based cellular assays and are therefore suitable to analyze the protective impact of HLA-G allelic variants. K562 cells producing soluble HLA-G (sHLA-G) molecules were cultivated in bioreactors for large scale protein production. The gene construct for these sHLA-G molecules is missing the transmembrane domain and thus corresponds to the HLA-G1 isoform that is expressed on the cell surface. To analyze the presented peptide repertoire, sHLA-G molecules were harvested and trimeric complexes subsequently purified using affinity purification. After elution, the bound peptides were sequenced utilizing mass spectrometry. Peptide sequencing revealed that the majority of the peptides are 9 AAs in length, while the presentation of longer peptides could also be detected. The peptide binding motif of HLA-G*01:01 derived from K562 cells is comparable to the motif described by Diehl et al. (66), however strikingly, HLA-G*01:04 differed in its peptide properties from the other allelic variants, whereas G*01:01 and G*01:03 featured several similarities. All variants exhibit Leucine at pΩ and Lysine at p1. HLA-G exhibits a restricted repertoire, however, very little overlap is observed between HLA-G*01:04 and G*01:01 or G*01:03, even though the proteomic content available is the same for all variants. A potential reason for that is found in the dissimilarities of the anchor motif, where in HLA-G*01:04 derived peptides the Proline auxiliary anchor at p3 was absent. Modelling of the HLA-G*01:01 derived peptides RQPYAVSEL and RSPPPGMGL that contain Proline at p3 and the HLA-G*01:04 derived peptide FLNGQNLGI into the available crystal structures of HLA-G*01:01 (PDB: 2DYP) and G*01:04 (PDB: 3KYO) showed that Proline at p3 introduces a constraint to the peptide alignment from p3 to p6 that appears less stringent in the G*01:04 derived peptide conformations. Such conformational differences may also explain the observation that HLA-G*01:04 proved to be more protective against NK cell mediated lysis than G*01:03 or G*01:01, when using isolated NK cells from HLA-G typed donors (HLA-G*01:01/G*01:01, G*01:01/G*01:04) or a NK cell line (NKL, HLA-G*01:01/G*01:01). The biological function of HLA heavy chain variability is to extend the pool of available peptide antigens and thus the variability of accessible peptide/HLA surfaces for immune receptors (119). Therefore, DNA typing is compulsory for donor-recipient matching. In the field of cellular therapeutics, NK cell-based therapies for the treatment of certain leukemia (120, 121) are on the rise. The HLA-G allele-specific peptide binding motifs suggest that patient typing for HLA-G pre-treatment might be a way to improve NK cell-based immunotherapeutic strategies.
The marginal polymorphic nature of HLA-G in comparison to HLA-Ia molecules and the differentiated NK cell receptor capture raised the question if HLA-G would select and present peptides in a tissue specific manner. To address this question, the HLA-G*01:01-restricted peptide repertoire derived from K562 cells was compared to those of an HL model cell line that exhibits the typical HRS phenotype of classic HL. HDLM-2 cells were stably transduced with lentiviral constructs encoding for sHLA-G*01:01; HDLM-2- and K562-derived peptide data were compared with available peptide data deposited in the SYFPEITHI (122) database. The data demonstrated a bias for tissue specific peptide anchors. At pH, peptides from both cell lines were anchored by Leucine, however, at p1 a clear bias for either Arginine or Lysine depending on the parent cell line could be detected. The spectrum of presented peptides depends on the available proteome. To exclude the possibility that differences in the anchor motif are due to limited proteomic content, we correlated source protein availability to expression data available from the human protein atlas (123). Strikingly, peptides restricted to one cell line were often derived from proteins that show equal or even higher expression levels in the other cell line and vice versa. Yet, presentation inside the cell appeared to be restricted to a specific anchor at p1. This lead to the question whether association with the peptide loading complex (PLC) may be impacted in HDLM-2 cells and thus influences the selection of peptides, however, after performing immunoprecipitation experiments we found comparable associations with PLC components in both cell lines, although, in both cases TAP appears unused for peptide loading, suggesting that HLA-G has the capability to facilitate TAP independent peptide selection in these transformed cells.

The major findings from this study extend our understanding of how HLA-G impacts NK cell recognition, potentially by the presentation of a broader spectrum of peptides than previously thought, even though, the restricted tissue distribution would suggest a highly conserved peptide binding motif. This further results in functional differences, with an increased protective potential of HLA-G*01:04 in comparison to G*01:01 or G*01:03. Although HLA typing is routinely performed, typing of non-classical HLA molecules is not widely adopted, even though newer studies highlight the rising clinical importance of non-classical HLA molecules. Studies on HLA-E, for instance, strongly suggest that prospective typing in unrelated hematopoietic stem cell transplantation could improve post-transplant prognosis in acute leukemia patients (124, 125). HLA-G as well as soluble HLA-G variants are readily detected in many different tumor entities and combined with our findings that HLA-G*01:04 exerts a greater protective function against NK cell mediated lysis, HLA-G typing might provide an intelligent step forward to optimize treatment for lymphoma patients where NK cell based anti-tumor strategies are considered.
Furthermore, the finding that HLA-G selects different peptide pools in a tissue-specific manner may hint to the possibility that in transformed cells the selection of unconventional peptides is favored, increasing the chance that tumor specific antigen presentation occurs. Many attempts are made to predict and refine tumor neo-antigens using *in silico* prediction tools, however, these are usually based on the assumption that the biochemical properties of the PBR only allow for a rigid peptide binding motif (126). Newer studies combining experimental data from the sequenced individual RNA mutanome with peptide prediction were able to identify neo-epitope-specific T cells (127). Additionally, in recent years evidence also emerged that HLA class I peptide presentation is much less stringent than previously assumed (128). HLA-E, for instance, is capable to elicit CD8⁺ T cell responses by presenting peptides of non-canonical length derived from *Mycobacterium tuberculosis*, even though these constitute only a very small percentage of the peptide pool and do not adhere to the proposed binding motif established from leader peptides (129). Therefore, appreciation of unconventional peptide selection and presentation of not only non-classical but also classical HLA class I molecules may guide further efforts concerning the establishment of HLA tissue specific ligandomes. The present work allows for the consideration of new strategies for anti-tumor therapies particularly with regard to HLA-G interference and the necessity for HLA-G genomic typing.
REFERENCES


68. Ishitani A, Geraghty DE. Alternative splicing of HLA-G transcripts yields proteins with primary structures resembling both class I and class II antigens. Proc Natl Acad Sci U S A. 1992 May 1;89(9):3947-51.


6 DANKSAGUNG

Zu allererst möchte ich mich bei meinem Erstgutachter Herrn Professor Rainer Blasczyk bedanken, für die Möglichkeit meine Doktorarbeit am Institut für Transfusionsmedizin anzufertigen. Besonders bedanken möchte ich mich auch für seine Unterstützung, die es mir ermöglicht hat, mein Wissen auf der EFI Summer School zu vertiefen und meine Arbeit auf diversen Konferenzen einem Fachpublikum vorzustellen, was zu neuen Impulsen und auch Motivation für meine weitere wissenschaftliche Arbeit führte.

Von ganzem Herzen möchte ich mich auch bei meiner Arbeitsgruppenleiterin, Frau Dr. Christina Bade-Döding bedanken, die mich schon früh sehr stark gefordert und gefördert hat. Sie hat mir dabei viel Vertrauen entgegengebracht und Freiheiten eingeräumt, um eigenständig wissenschaftlichen Fragestellungen nachzugehen und stand mir immer mit Rat und Tat zur Seite.

Mein Dank gilt natürlich auch Herrn Professor Reinhard Schwinzer für die wissenschaftliche Zweitbetreuung und die freundliche Unterstützung und die Anregungen während der Supervisor Meetings.

Bedanken möchte ich mich ebenfalls bei Herrn Dr. Trevor Huyton, für seine weitreichende Expertise bei den Strukturanalysen sowie für die vielen hilfreichen Anregungen.

Mein Dank gilt auch der deutschen José Carreras Leukämiestiftung und der deutschen Gesellschaft für Hämatologie und Medizinische Onkologie, deren finanzielle Unterstützung des Projekts eine große Hilfe war.


Zu guter Letzt möchte ich mich noch bei meinen Liebsten bedanken:

Melina, vielen Dank für deine Geduld, deine Unterstützung und deinen grenzenlosen Rückhalt! Du hast mich mehr als nur einmal wieder aufgebaut.

Ich bin außerdem zutiefst dankbar für die Unterstützung meiner Familie, ohne die ich es nicht bis hierher geschafft hätte. Besonders auch meiner Schwester Alice, die schon während meiner Studienzeit immer ein offenes Ohr für Probleme jeglicher Art hatte.
Erklärung:

Hiermit erkläre ich, dass ich die Dissertation („The basis of HLA-G mediated dysregulation of immune effector cells“) selbstständig verfasst habe.


Die Dissertation wurde bisher nicht für eine Prüfung oder Promotion oder für einen ähnlichen Zweck zur Beurteilung eingereicht. Ich versichere, dass ich die vorstehenden Angaben nach bestem Wissen und der Wahrheit entsprechend gemacht habe.

Ort, Datum: Hannover 10.10.18 Unterschrift: A. Celik