HLA-G (major histocompatibility complex, class I, G)

Shang-Mian Yie

The Second Medical College/Teaching Hospital, Chengdu University of Traditional Chinese Medicine, Chengdu, Sichuan, PR China (SMY)

Identity

Other names: MHC-G
HGNC (Hugo): HLA-G
Location: 6p22.1

DNA/RNA

Description

HLA-G is one of the non-classical class I (Ib) HLA molecules. The HLA-G gene is located at the short arm of chromosome 6 in the HLA region (6p21.2-21.3) between HLA-A and HLA-F genes (figure 1A). The gene structure of HLA-G is homologous to other HLA class I (Ia) genes consisting of 7 introns and 8 exons coding the heavy chain of the molecule. Exon 1 encodes the peptide signal, while exons 2, 3 and 4 encode the extracellular α1, α2 and α3 domains, respectively. Exons 5 and 6 encode the transmembrane and cytoplasmic domains of the heavy chain. Exon 7 is always absent from mature mRNA and due to the stop codon in exon 6; exon 8 is not translated (figure 1B).

Transcription

The functional mRNA level of a particular gene is regulated by the rate of synthesis, mainly driven by the promoter region (5'-UTR) of the given gene, as well as by the rate of degradation, stability, localization and translatability of the specific mRNA (Kuersten and Goodwin, 2003). The HLA-G gene promoter has a modified enhancer A (enha), S and X1 sequence, and a few alternative regulatory elements to regulate HLA-G gene transcription (figure 1C). So far, these regulatory elements include the locus control region (LCR) (Schmidt et al., 1993; Yelavarthi et al., 1993), which is located approximately 1.2 kb from the ATG initiation codon of the HLA-G gene. The CREB1 factor binds to this region (-1380/-1370), as well as to two additional cAMP response elements (CRE) dispersed through the promoter region at positions -934 and -770 from the ATG. An interferon-sensitive response element (ISRE) is located at position -744 from the ATG (Lefebvre et al., 2001). A heat shock element (HSE) is located within the HLA-G promoter at position -459/-454 that binds heat shock factor 1 (HSF-1) (Ibrahim et al., 2000). A progesterone receptor response element (PRE) is located at position -37 from the ATG (Yie et al., 2006), and three ras response elements (RRES) are situated along the HLA-G gene promoter (-1356, -142, -53) (Flajollet et al., 2009).

The 3'-UTR of the HLA-G gene also exhibits several regulatory elements including AU-rich motifs and a poly-A signal to influence mRNA stability, turnover, mobility and splicing pattern (Donadi et al., 2011). The primary transcript of HLA-G can be spliced into 7 alternative mRNAs (figure 1D) that encode membrane-bound (HLA-G1, -G2, -G3, -G4) and soluble (HLA-G5, -G6, -G7) protein isoforms (Carosella et al., 2003). HLA-G1 is the full-length HLA-G molecule, HLA-G2 lacks exon 3, HLA-G3 lacks exons 3 and 4, and HLA-G4 lacks exon 4. HLA-G1 to -G4 are membrane-bound molecules due to the presence of the transmembrane and cytoplasmic tail encoded by exons 5 and 6. HLA-G5 is similar to HLA-G1 but retains intron 4, HLA-G6 lacks exon 3 but retains intron 4, and HLA-G7 lacks exon 3 but retains intron 2. HLA-G5 and -G6 are soluble forms due to the presence of intron 4, which contains a premature stop codon to prevent the translation of the transmembrane and cytoplasmic tail. HLA-G7 is soluble due to the presence of intron 2, which contains a premature stop codon.
Figure 1. The human leukocyte antigen-G (HLA-G) gene and transcription. A. HLA-G gene location. B. HLA-G gene structure consisting of 7 introns (white color) and 8 exons (green color). C. HLA-G gene promoter exhibiting regulatory elements to regulate HLA-G gene transcription and 3'-UTR of the HLA-G gene exhibiting several regulatory elements including AU-rich motifs and a Poly-A signal to influence mRNA stability, turnover, mobility and splicing pattern. CRE stands for cAMP responsive element, RRES for ras response elements, ISRE for interferon-sensitive response element, HSE for heat shock response element, PRE for progesterone response element, and X1 for X1 box. D. HLA-G primary transcript can be spliced into 7 alternative mRNAs ranging from HLA-G1 to -G7.
Figure 2. HLA-G protein and its function. **A.** HLA-G protein exhibits a heterodimer consisting of globular domains (α1, α2 and α3 domains, transmembrane and cytoplasmic domains) and a light chain (β2-microglobulin). **B.** Alternative splicing of the primary transcript yields 7 protein isoforms: truncated isoforms are generated by excision of one or two exons encoding the globular domains, whereas translation of intron 4 or intron 2 yields soluble isoforms that lack the transmembrane domain. HLA-G molecules can form homomultimers through the generation of Cys42-Cys42 or Cys42-Cys147 disulphide bonds. **C.** Immunoregulatory activities mediated by HLA-G, where the involved target cells and receptors are indicated.

### Protein

#### Description

The HLA-G protein exhibits a heterodimer consisting of globular domains (α1, α2 and α3 domains, transmembrane and cytoplasmic domains) and a light chain (β2-microglobulin) called monomer (figure 2A). However, there may be 7 protein isoforms, generated by alternative splicing of the primary transcript: four of them being membrane-bound (HLA-G1, -G2, -G3 and -G4) and three soluble (HLA-G5, -G6 and -G7) (figure 2B). HLA-G1 is the complete isoform associated with β2-microglobulin. The HLA-G2 isoform has no α2 domain, while HLA-G3 has no α2 and α3 domains, and HLA-G4 has no α3 domain. The soluble HLA-G5 and -G6 isoforms contain the same extra globular domains as HLA-G1 and -G2, respectively, generated by transcripts conserving intron 4, which blocks the translation of the transmembrane domain. The 5’-region of intron 4 is translated until the generation of a stop codon, which gives the HLA-G5 and -G6 isoforms a tail of 21 amino acids responsible for their solubility.

The HLA-G7 isoform has only the α1 domain linked to two amino acids encoded by intron 2, which is retained in the corresponding transcript (Fujii et al., 1994; Ishitani and Geraghty, 1992; Paul et al., 2000).

In addition, HLA-G molecules can form dimers through the creation of disulphide bonds between two unique cysteine residues at positions 42 (Cys42-Cys42 bonds) and 147 (Cys42-Cys147 bonds) of the HLA-G heavy chain (figure 2B) (Boyson et al., 2002; Gonen-Gross et al., 2003). The dimerization has an oblique orientation that exposes the HLA-G receptor binding sites of the α3 domain upwards, making them more accessible to the receptors. Consequently, HLA-G dimers bind receptors with higher affinity and slower dissociation rates than monomers, and signal more efficiently than monomers as well (Shiroishi et al., 2006).

#### Expression

Classical class Ia antigens are ubiquitously expressed, whereas the expression of HLA-G is restrictive. HLA-G protein expression is only found in trophoblast cells in the placenta, certain immune cells (in most cases monocytes), thymus, cornea, proximal nail matrix,
erythroblasts and mesenchymal stem cells (Kovats et al., 1990; Crisa et al., 1997; Lila et al., 2001; Ishitani et al., 2003; Rebmann et al., 2003; Morandi et al., 2008). The reasons for HLA-G expression in some but not other tissues have not been fully elucidated. However, soluble HLA-G (sHLA-G) can be detected in the serum/plasma of both men and women. The main source of sHLA-G in the blood of men and non-pregnant women is most likely monocytes. Both CD4+ and CD8+ T cells and B cells also seem to be able to secrete HLA-G5 although to a lesser extent (Rebmann et al., 2003). The presence or level of sHLA-G in serum/plasma samples is associated with HLA-G polymorphism (Rebmann et al., 2001; Hvid et al., 2004a; Hvid et al., 2004b; Rizzo et al., 2005).

As mentioned earlier, HLA-G expression is mainly controlled at the transcriptional level by a unique gene promoter and at the post-transcriptional level by alternative splicing, mRNA stability, translation and protein transport to the cell surface. Many factors have been described that can potentially affect transcriptional and post-transcriptional mechanisms responsible for HLA-G regulation (Moreau et al., 2009).

**Function**

It has been demonstrated that HLA-G is an immune tolerogenic molecule, which plays an important role in the suppression of the immune responses (Carosella et al., 2008) (figure 2C). The immune-inhibitory function of HLA-G is realized by interacting with leukocyte receptors including leukocyte immunoglobulin-like receptor subfamily B member 1 (LILRB1) and member 2 (LILRB2), and killer cell immunoglobulin-like receptor 2DL4 (KIR2DL4) (Shiroishi et al., 2006; Gao et al., 2000).

LILRB1s are expressed on the surface of several leukocytes, such as NK and lymphomononuclear cells, while LILRB2s are primarily expressed on the surface of a restricted set of cells, including monocytes and dendritic cells (Brown et al., 2004). Both LILRB1 and LILRB2 have several inhibitory motifs (ITIM) receptors based on the immunoreceptor tyrosine in their cytoplasmic tail, which inhibits signaling events triggered by stimulatory receptors (Dietrich et al., 2001). LILRB1 and LILRB2 both interact with classical HLA class I molecules. However, their binding with HLA-G presents three- to four-fold higher affinity when compared to classical molecules (Shiroishi et al., 2003). LILRB1 and LILRB2 also bind to the α3 domain and β2-microglobulin of the HLA-G molecule, although LILRB2 binds with higher affinity than LILRB1. LILRB2 binds more to the α3 domain than to the β2-microglobulin domain. The binding sites of LILRB1 and LILRB2 are distinct. The Tyr36 and Arg38 residues of LILRB2 bind to the 195-197 loop of the α3 domain of HLA-G, whereas the Tyr38 and Tyr76 residues of LILRB1 bind to the Phe195 of HLA-G (Shiroishi et al., 2006).

KIRs are transmembrane glycoproteins expressed by natural killer cells and subsets of T cells, exhibiting two (KIR2D) or three (KIR3D) extracellular immunoglobulin-like domains, which also contain ITIMs. The KIR2DL4 binds to the α1 domain of the HLA-G molecule (Yan and Fan, 2005). However, the binding site of KIR2DL4 to HLA-G remains unknown (Donadi et al., 2011). Moreover, since an array of activator and inhibitor receptors is expressed on the surface of most NK cells and macrophages, and the final effector function is dependent on the balance between activator and inhibitor receptors (Hsu et al., 2002), the role of the interaction between KIR2DL4 and HLA-G in the modulation of the immune response has been a matter of much debate (LeMaoult et al., 2003; Apps et al., 2008).

HLA-G has an inhibitory effect on cytotoxic cells exhibiting CD8 on their surface through the interaction of the α3 domain of HLA-G with the CD8 a/α molecule. The α3 domain of HLA-G is the same site of interaction with CD8 a/α and LILIRBs. Even though the CD8 a/α and LILIRBs binding sites overlap, LILIRBs inhibit the binding of CD8 a/α to HLA molecules by displacing CD8 a/α and activating ITIMs (Shiroishi et al., 2003). Beside the extracellular domains of HLA-G, all segments of the molecule may contribute to its function. The short cytoplasmic tail retains HLA-G longer in the endoplasmic reticulum and prolongs the half-life of the molecule on the cell surface because of the lack of an endocytosis motif (Park et al., 2001; Park and Ahn, 2003). This permits multiple interactions with cells of the immune system.

Soluble HLA-G molecule can induce apoptosis in CD8+ activated T lymphocytes as well as in CD8+ NK cells (lacking the T cell receptor) at similar rates. The binding of soluble HLA molecules to CD8 leads to apoptosis upregulating Fas production and Fas/FasL interaction (Puppo et al., 2002; Contini et al., 2003). This mechanism represents an additional immunomodulatory effect of HLA-G.

**Mutations**

**Somatic**

44 HLA-G coding alleles have been defined based upon 72 single nucleotide polymorphisms (SNP) observed between exon 1 and intron 6. Nucleotide variability in the coding region of the HLA-G gene is evenly distributed throughout exons 2, 3 and 4, as well as in introns (Donadi et al., 2011). The heavy chain encoding region exhibits 33 SNPs but only 13 amino acid variations are observed, 4 of them in α1, 6 in α2 and 3 in the α3 domain (Donadi et al., 2011). The amino acid substitutions may account for the biological function of HLA-G, including peptide binding, isoform production, and ability to polymerize and modulate immune system cells.
The HLA-G promoter exhibits 29 SNPs to date. Since many of these polymorphisms either coincide with or are closed to the known regulatory elements, they may affect the binding of the corresponding regulatory factors (Tan et al., 2005; Hviid et al., 2006; Hviid et al., 2004a; Hviid et al., 2004b). Polymorphisms located at CpG sites may affect promoter methylation (Ober et al., 2006).

In some cases, polymorphism in the promoter region may be in linkage disequilibrium with 3'-UTR variants (Nicolae et al., 2005), and some of them could influence alternative splicing (Auboeuf et al., 2002) and mRNA stability (Rousseau et al., 2003).

In contrast to the coding region, the 3'-UTR of the HLA-G locus presents a high degree of variation. Since the 3'-UTR of HLA-G gene exhibits several regulatory elements including AU-rich motifs, a poly-A signal, as well as signals that regulate the spatial and temporal expression of an mRNA, the polymorphic sites may influence mRNA stability, turnover, mobility and splicing pattern. The polymorphic sites at the 3'-UTR seem to bear ranged in several haplotypes (Alvarez et al., 2009; Castelli et al., 2010; Donadi et al., 2011), their influence seems to occur simultaneously.

Figure 3. The role of HLA-G in tumor immunoediting. HLA-G is involved in every phase of tumor immunoediting to inhibit host immune response. During the tumor immunoediting process, HLA-G can be activated and up-regulated by many factors. Aberrant expression of HLA-G can: 1) disable effectors of innate and adaptive immunity in the elimination phase; 2) alter antigen presentation and contribute less immunogenic phenotype in the equilibrium phase; and 3) induce immunosuppressive cytokines and peripheral tolerance to the tumor (plus processes described in the two previous phases) in the escape phase.
Implicated in

Various cancers

Disease
Cancer is essentially considered a complex cell disease caused by abnormalities in the genetic material of transformed cells. However, cancer development is a complicated progressive process that involves a sequence of gene-environment interactions with dysfunctions in multiple systems, including immune functions. The immune system can specifically identify and eliminate tumor cells based on their expression of tumor-specific antigens or molecules induced during malignant cell transformation. This process is referred to as tumor immune surveillance (Swann and Smyth, 2007). Despite tumor immune surveillance, tumors can still develop in the presence of a functioning immune system. This occurs through tumor immunoediting, a process that comprises three major phases (Urosevic and Dummer, 2008): 1) the elimination phase in which most immuno-genic tumor cells are eliminated by cytotoxic T and NK cells; 2) the equilibrium phase in which tumor cells with reduced immunogenicity are selected; and 3) the escape phase in which variants that no longer respond to the host immune system are maintained (Urosevic and Dummer, 2008). HLA-G is involved in every phase of tumor immunoediting by decreasing the elimination of tumor cells, by inhibiting the cytotoxic function of T and NK cells, and by trogocytosis, (i.e. the intercell transference of viable HLA-G molecules), which renders competent cytotoxic cells unresponsive to tumor antigens (LeMaoult et al., 2007; Caumartin et al., 2007) (figure 3).

Prognosis
HLA-G is aberrantly expressed in many human solid malignant tumors in situ and malignant hematopoietic diseases including breast, ovarian, clear renal cell, colorectal, gastric, esophageal, lung, and hepatocellular cancers, as well as acute myeloid leukemia and chronic lymphocytic leukemia (B-CLL) (Carosella et al., 2008; Yie and Hu, 2011). The aberrant expression of HLA-G in malignant neoplasm is significantly correlated with poor clinical outcome of patients with colorectal cancer (CRC) (Ye et al., 2007), gastric cancer (GC) (Yie et al., 2007b), non-small cell lung cancer (NSCLC) (Yie et al., 2007c), esophageal squamous cell cancer (ESCC) (Yie et al., 2007a), breast cancer (He et al., 2010), hepatocellular cancers (Cai et al., 2009), and B-CLL (Nuckel et al., 2005). This is due to the fact that HLA-G expression favors tumor development and metastasis in every phase of cancer immunoediting by impairing anti-tumor immune responses (Urosevic and Dummer, 2008). A reverse correlation between HLA-G expression in tumors and the degree of tumor-infiltrating lymphocytes (TILs) has been demonstrated in a variety of cancer types (Yie and Hu, 2011). The presence of TILs indicates an anti-tumor cellular immune response (Yu and Fu, 2006). Also, it has been documented that HLA-G can increase regulatory T cells within TILs in breast and hepatocellular cancers (Chen et al., 2010; Cai et al., 2009). The regulatory T cells are a subset of immune T cells that inhibit the anti-tumor functions of tumor-specific T cells (Shevach, 2002). Therefore, current data suggest that the estimation of HLA-G expression is a novel prognostic marker useful in assessing host immune response since cancer can be explained, at least in part, as an abnormal immune system tolerance to uncontrolled cells (de la Cruz-Merino et al., 2008).

Furthermore, serum soluble HLA-G is increased in various types of cancer patients (Pistoia et al., 2007; Yie and Hu, 2011) including patients with melanoma (Ugurel et al., 2001), acute leukemia (Gros et al., 2006), multiple myeloma (Leleu et al., 2005), neuroblastoma (Morandi et al., 2007), lymphoproliferative disorders (Sebti et al., 2007), breast or ovarian cancer (Singer et al., 2003; Chen et al., 2010; He et al., 2010), non-small cell lung cancer (Cao et al., 2011), esophageal cancer (Cao et al., 2011), colorectal cancer (Zhu et al., 2011; Cao et al., 2011), gastric cancer (Cao et al., 2011) and hepatocellular carcinoma (Wang et al., 2011), when compared to normal healthy controls or benign disease cases. Although numerous and different cancer studies show preferential up-regulation of HLA-G in advanced diseases rather than in initial tumor lesions, currently available HLA-G data do support the notion that HLA-G can be used as a potential biomarker in the diagnosis of human carcinomas (Yie and Hu, 2011).

References


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


Gao GF, Wilcox BE, Wyer JR, Boulter JM, O'Callaghan CA, Maenaka K, Stuart DI, Jones EY, Van Der Merwe PA, Bell JI, Jakobsen BK. Classical and nonclassical class I major
HLA-G (major histocompatibility complex, class I, G)

Yie SM

...
HLA-G (major histocompatibility complex, class I, G)  
Yie SM


Hvid TV, Rizzo R, Melchiorri L, Stignani M, Baricordi OR. Polymorphism in the 5' upstream regulatory and 3' untranslated regions of the HLA-G gene in relation to soluble HLA-G and IL-10 expression. Hum Immunol. 2006 Jan-Feb;67(1-2):53-62


Flajollet S, Poras I, Carosella ED, Moreau P. RREB-1 is a transcriptional repressor of HLA-G. J Immunol. 2009 Dec 1;183(11):6948-59


This article should be referenced as such: