Human Immunology xxx (2023) xxx–xxx

Contents lists available at ScienceDirect

journal homepage: www.elsevier.com/locate/humimm

Review

Computational and atomistic studies applied to the understanding of the structural and behavioral features of the immune checkpoint HLA-G molecule and gene

Cinthia C. Alves [a], Thaís Arns [b], Maria L. Oliveira [c], Philippe Moreau [d,e], Dinler A. Antunes [f], Erick C. Castelli [g], Celso T. Mendes-Junior [h], Silvana Giuliatti [c], Eduardo A. Donadi [a,*]

[a] Department of Medicine, Division of Clinical Immunology, Ribeirão Preto Medical School, University of São Paulo, SP, Brazil
[b] Luxembourg Centre for Systems Biomedicine, Luxembourg
[c] Department of Genetics, Ribeirão Preto Medical School, University of São Paulo, SP, Brazil
[d] CEA, DRF-Institut François Jacob, Service de Recherches en Hémato-Immunologie, Hôpital Saint-Louis, Paris, France
[e] U976 HIPI Unit, IRSL, Université Paris-Cité, Paris, France
[f] Department of Biology and Biochemistry, University of Houston, Houston, USA
[g] Department of Pathology, School of Medicine, São Paulo State University (UNESP), Botucatu, SP, Brazil
[h] Departamento de Química, Faculdade de Filosofia, Ciências e Letras de Ribeirão Preto, Universidade de São Paulo, Ribeirão Preto, SP, Brazil

ARTICLE INFO

Keywords:
HLA-G
Isoforms
LILRB1
LILRB2
Hypoxia-inducible factor 1

ABSTRACT

We took advantage of the increasingly evolving approaches for *in silico* studies concerning protein structures, protein molecular dynamics (MD), protein-protein and protein-DNA docking to evaluate: (i) the structure and MD characteristics of the HLA-G well-recognized isoforms, (ii) the impact of missense mutations at HLA-G receptor genes (*LILRB1/2*), and (iii) the differential binding of the hypoxia-inducible factor 1 (HIF1) to hypoxia-responsive elements (HRE) at the *HLA-G* gene. Besides reviewing these topics, they were revisited including the following novel results: (i) the HLA-G6 isoforms were unstable docked or not with β_2-microglobulin or peptide, (ii) missense mutations at *LILRB1/2* genes, exchanging amino acids at the intracellular domain, particularly those located within and around the ITIM motifs, may impact the HLA-G binding strength, and (iii) HREs motifs at the *HLA-G* promoter or exon 2 regions exhibiting a guanine at their third position present a higher affinity for HIF1 when compared to an adenine at the same position. These data shed some light into the functional aspects of HLA-G, particularly how polymorphisms may influence the role of the molecule. Computational and atomistic studies have provided alternative tools for experimental physical methodologies, which are time-consuming, expensive, demanding large quantities of purified proteins, and exhibit low output.

Contents

Abbreviations: PDB, Protein data bank; PSP, protein structure prediction; MD, molecular dynamic simulation; HIF1, hypoxia-inducible factor 1; HLA-G, human leukocyte antigen-G; LILRB1 and LILRB2, leukocyte Immunoglobulin-like receptor 1 and 2; ITIM, immunoreceptor tyrosine-based inhibitory motif; SNVs, single nucleotide variants; LD, linkage disequilibrium; HRE, hypoxia regulatory elements; SNP, single nucleotide polymorphism; RMSD, Root Mean Square Deviation; RMSF, Root Mean Square Fluctuations; FUBAR, Fast-Unconstrained Bayesian Approximation; DBD, DNA binding domain; HS, HADDOCK score; Hbond, Hydrogen bond; C-TAD, C-terminus transcriptional activation domain.

* Corresponding author at: Department of Medicine, Division of Clinical Immunology, Ribeirão Preto Medical School, University of São Paulo, Avenida Bandeirantes 3900, Monte Alegre, 14049-900 Ribeirão Preto, SP, Brazil.
E-mail address: eadonadi@fmrp.usp.br (E.A. Donadi).

https://doi.org/10.1016/j.humimm.2023.01.004
Received 11 October 2022; Revised 12 January 2023; Accepted 16 January 2023
Available online xxxx

1. Introduction

Massive parallel sequencing has provided data regarding the complete structure of many genes, supplying information to unravel the amino acid sequence of proteins and the regulatory nucleotide sequence to understand the putative translational and posttranslational gene control. Because protein function is related to its three-dimensional (3D) structure, the relationship between amino acid sequence and protein structure may facilitate the understanding of the biological function of proteins [1]. Seminal procedures to identify 3D protein structures are primarily focused on experimental methods, such as X-ray, nuclear magnetic resonance, spectroscopy and cryo-electron microscopy, which are time-consuming, expensive, demand relatively large quantities of purified proteins, and exhibit low-output [2]. An alternative to experimental procedures is the use of computational (*in silico*) approaches using public datasets, such as the protein data bank (PDB, https://www.rcsb.org/) [3] PDB is a comprehensive archive of information regarding the 3D shapes of proteins, nucleic acids, and complex assemblies solved by experimental methods, containing more than 194,000 entries [3]. In addition to static structures, Molecular Dynamic (MD) simulations may provide the behavior of the modeled structures.

Overall, tertiary protein structure prediction (PSP) algorithms are classified into two main groups: (i) template-based modeling (homology and threading methods) and (ii) template-free modeling (a*b initio*). The first one uses solved PDB protein structures as a template to model the target protein according to the amino acid sequence or folding similarity. When a target-template similarity is <20 %, the second modeling method is applied, which considers the physical–chemical properties of each amino acid in the query polypeptide chain to predict tertiary structures [4]. Besides PSP methods, molecular docking algorithms have been used to predict the "fit" between three-dimensional structures; i. e., how the structure of one molecule binds to the structure of another molecule to form a molecular complex, including proteins, RNAs, DNAs, peptides, or other small molecules [5]. The result of PSPs is an arrangement of protein atoms in coordinates of a Cartesian plane [3].

Proteins cannot be exclusively considered as a set of amino acids delimited in a certain spatial orientation; on the contrary, proteins can adopt different conformations according to the environment in which they play their functions or interact with other molecules. In these cases, it is necessary to carry out MD simulation studies to understand the behavior of the molecules [6]. The system to be simulated consists of a collection of interacting particles represented as atoms, encompassing the biological molecule inside a water box in conditions of controlled volume, temperature, pH, pressure, and ion content, over a specific time. This strategy has permitted to simulate a biological system in a physiological environment with the atoms in motion to verify whether the 3D structure remains stable during its movement along a determined time [7].

In this review, we revisited our previous computational findings regarding: (i) the structural aspects of the human leukocyte antigen-G (HLA-G) isoforms (complete membrane-bound and some isoforms) [8], (ii) the consequences on protein stability of gene missense mutations at the major HLA-G receptors (LILRB1 and LILRB2) [9], and (iii) the atomistic mechanisms that may explain the differential action of the hypoxia-inducible factor 1 (HIF1) on the *HLA-G* gene [10] and complemented these approaches adding novel unpublished data.

1.1. Structural aspects of the HLA-G isoforms

The first crystallographic structure of the HLA-G molecule coupled with β_2-microglobulin and peptide [11] exhibited some gaps, leaving aside important residues sequences associated with the interaction of HLA-G with its receptors, the final portion of the α3 domain, and the transmembrane and intracytoplasmic domains. To fill these gaps, we evaluated the structure and MD of the membrane-bound HLA-G1 molecule, the soluble HLA-G1 dimer, and the soluble HLA-G5 isoform. The membrane-bound HLA-G1 and the soluble HLA-G1 dimer coupled with β_2-microglobulin and the histone H2A-derived RIIPRHLQL nonapeptide [12] were stable. Similarly, the HLA-G5 isoform (α1-3 domains and 21 residues encoded by intron 4) coupled with the β_2-microglobulin and with the RIIPRHLQL nonapeptide was stable, whereas HLA-G5 structures uncoupled with the β_2-microglobulin or with the RIIPRHLQL nonapeptide were unstable because of the random movement produced by the 21-residue tail retained from intron 4, which opened the cleft permitting the escape of the peptide [8].

In Supplementary results and discussion section, we complemented the modeling of HLA-G isoforms, adding the structural analysis and MD of the HLA-G6 monomer, which is encoded by exon 2 (α1 domain) and exon 4 (α3 domain) of the *HLA-G* gene, eliminating exon 2, but conserving intron 4, retaining the same 21-residue tail also present in the HLA-G5 isoform.

1.2. Missense mutations at the LILRB1/LILRB2 genes

LILRB1 (LIR-1/CD85j/ILT2) and LILRB2 (LIR-2/CD85d/ILT4) belong to the leukocyte receptor complex (LRC) located at the long arm of chromosome 19. Among the functional members of the leukocyte Ig-like receptor (LILR) family, five are inhibitory (LILRB1-5), five activators (LILRA1,2,4–6), and one is a soluble (LILRA3) receptor [13]. LILRB1/2 contain four extracellular Ig-like domains (D1-D4), a transmembrane domain, and 3–4 immunoreceptor tyrosine-based inhibitory motifs (ITIMs) in the cytoplasmic tails, which recruit tyrosine phosphatases, inhibiting activating signals [14,15]. Ligand binding to receptor yields the phosphorylation of the ITIM tyrosine residue by Src tyrosine kinases that recruit two cytoplasmic tyrosine phosphatases, referred to as SHP-1 and SHP-2, which dephosphorylate signaling molecules [16]. These receptors interact with HLA class I classical and non-classical molecules, as well as with proteins from viruses and parasites [17–19] transducing negative signals through their ITIMs, inhibiting the function of natural killer, T/B-cells, and myelomonocytic cells [20,21]. Overall, the interactions with HLA class I occur through the binding of the two LILRB1/2N-terminal extracellular domains (D1/D2) with the α3 and $\beta_2$microglobulin domains of the HLA molecules.

LILRB1 interacts with HLA-A2 molecules through: (i) six D1 (Q18, K42, W67, E68, G97, A98) and seven D2 (Y99, I100, Q125, V126, A127, D184, L187) residues with fourteen $\beta_2$microglobulin residues (I1, Q2, *R*3, T4, V85, T86, L87, S88, Q89, K91, I92, V93, K94, D96), and (ii) six D1 residues (Y36, Y38, R39, K41, T43, Y76) with six α3 domain residues (A193, V194, S195, D196, E198, V248) [22]. In con-

trast, LILRB2 employs a lower number of residues for the interaction with HLAs, which results in a predominant recognition of α3 and almost independence of $\beta_2$microglobulin. LILRB2 interacts through (i) four D2 residues (W67, D177, N179, V183) with two residues from $\beta_2$microglobulin (K6, K91) and (ii) five D1 residues (R36, Y38, K42, I47, T48) that bind to three HLA-G α3 residues (F195, Y197, E229) [23]. Compared to other LRCs, *LILRs* show a fair degree of polymorphism [13]; however, the *LILRB1/2* family has an accelerated evolution, with increased interspecies differences compared to the genome average [24,25].

Little attention has been devoted to the *LILRB1/2* diversity [26–29]. Despite the high nucleotide conservation, especially at the *LILRB1/2* coding regions, a recent study conducted by our group highlighted the presence of several polymorphisms affecting the protein stability and, consequently, the ligand recognition and function of these receptors [9]. In total, 58 single nucleotide variants (SNVs), arranged in 13 *LILRB1* haplotypes, were identified. The majority (68.6 %) of these SNVs is located at the D1-D4 domains and may affect the binding to HLA class I molecules. Additionally, we and other authors identified variants that may affect the LILRB1 stability, such as P68L (rs1061679), T93A (rs12460501), L114R (rs778057754), T142I (rs1061680), I155S (rs1061681) at the D1/D2 domains, and A309V (rs61737895), G350R (rs61739173), Q401L rs61739175), at the D3/D4 domains [26,29,30]. Nevertheless, only P68L (D1 domain) contacts the HLA class I molecules, disturbing protein folding and reducing the binding with HLA-G [26,31]. Additionally, P68L is in complete linkage disequilibrium (LD) with ten other variants in *LILRB1* gene [9]. The rs1061679 SNV shows a differential frequency in worldwide populations, where the rs1061679*T allele exhibits a frequency of 0.56 in Brazilians, 0.37 in Africans, 0.40 in East Asians, 0.49 in admixed Americans, 0.30 in European, and 0.29 in South Asians, according to the 1000 Genome Project Consortium [32].

For the *LILRB2* gene, we identified 41 SNVs, arranged in 11 haplotypes. Compared to the *LILRB1* gene, a lower number of variations at Ig-like domains (51.2 %) was detected, with only 5 SNVs (12.19 %) at D1/D2 domains. Natural selection signatures at residue level revealed positive selection signals only at the D3/D4 domains; however, structural analysis identified a unique residue at D1/D2 (M235V-rs386056) domains exhibiting a protein destabilizing effect. Noteworthy, missense mutations at the D3/D4 domains may alter the interaction with HLA molecules because these domains serve as a scaffold for the D1/D2 domains [33]. Tajima's D Test results support these findings, showing a signature of purifying selection at the complete *LILRB2* gene regions [9].

In addition to the myriad of destabilizing missense mutations observed at the *LILRB1/2* regions that encode the D1/D2 domains, we also observed missense mutations at the *LILRB1/2* genes that encode the intracellular domain and specifically at the ITIM domains, which are presented in the Results/Discussion section.

1.3. HLA-G mutations that may influence the binding of the hypoxia-inducible factor 1 (HIF1) to hypoxia-responsive elements (HRE)

HLA-G transcription can be induced by the binding of HIF1 with HREs under hypoxic conditions in physiological and non-physiological situations. For instance, hypoxia is an important physiological microenvironment for placental development and formation of the maternal-fetal interface [34,35], increasing HLA-G expression in fetal endothelial cells, cytotrophoblasts and amniotic fluid, contributing to the immune tolerance of the fetus by the mother. In addition, HIF1-HRE binding targeting *HLA-G* may have a role in individual adaptation mechanisms to high altitudes, since increased soluble *HLA-G* was observed in mountain climbers [36] and *HLA-G* genetic variability may impact high altitude adaptation [37]. In non-physiological situations, the low level of cellular and tissue oxygen has been reported to increase transcription and production of HLA-G in several HLA-G-negative tumor lineages, such as melanoma and glioma [38–40], in which tumor cells use HLA-G immunomodulatory mechanisms to immune system escape.

HIF1 is a key hypoxia transcription factor regulator that belongs to the Helix-Loop-Helix-PER-ARNT-SIM protein family. HIF1 functions as a heterodimer, composed of an oxygen-regulated α subunit (HIF1a) and a stable β subunit (HIF1b) [41]. Under hypoxic conditions, HIF1a dimerizes with the HIF1b subunit and binds to hypoxia regulatory elements (HRE:5′RCGTG'3) in the target genes, stimulating gene transcription in association with coactivator protein complexes, like CBP/p300 [42].

DNA binding residues at the HIF1 α and β subunits may bind to several potential HREs motifs identified at the *HLA-G* promoter and coding region. At the promoter region, two HREs located 1.4 kb region upstream the translation start codon (ATG) have been described. The first is at positions −242 to −238 base pairs (bp) upstream the first translated ATG, at the genomic region chr6:29827603–29827607 (hg38), with the sequence 5′GCGTG'3. The second is at positions −966 to −962 bp, genomic region chr6:29826879–29826883, also with sequence 5′GCGTG'3. The first site is non-functional [40] and no known polymorphism was observed, while the second present one frequent variant, at position −964 G > A (rs1632947) [38]. Two additional HREs have been described at positions + 281 to + 285 bp (5′ACGTG'3, chr6:29828125–29828129) and + 291 to + 295 bp (downstream the 5′CACGC3′ sequence, chr6:29828135–29828139) at exon 2 [35]. Despite no polymorphism has been reported at the + 281 HRE, two variants were identified in the second HRE (sequence + 291 to + 295): +292 A > T (rs41551813) and + 293C > T (rs72558173). The first SNV causes a non-synonymous amino acid change from threonine to serine (T31S) and the second from threonine to methionine (T31M) in the α1 domain of the protein [39,40] (Fig. 1).

The −964G > A SNP at the promoter region is the most frequent in worldwide populations, exhibiting an allele frequency of 0.5369 for the −964A allele, while the exon 2 SNPs present a frequency of 0.0609 for the + 292 T allele and of 0.0009 for the + 293 T allele [43,44]. The −964A variant is observed in the *HLA-G**01:04, *G**01:01:02, *G**01:05 N, and *G**01:06 alleles, while other alleles exhibit the −964G variant. The exon 2 + 292 T variant is present in the *HLA-G**01:03 allele and the + 293 T variant in the *HLA-G**01:10 and *G**01:11 alleles (Fig. 1). Among the exon 2 HRE haplotypes, the *G**01:03, *G**01:10 and *G**01:11 alleles have a world population frequency lower than 1 % [43,44].

Yaghi et al [40] first described the impact of *HLA-G* gene variability on HIF1-HRE interaction and *HLA-G* expression under hypoxic conditions, using the hypoxia-mimicking agent deferoxamine. The expression occurred due a synergic action of the HIF1 binding to the promoter −966 HRE together with the exon 2 HREs (HREs + 281 and + 291). The −964G allele exhibited increased *HLA-G* expression in comparison to −964A allele, revealing the impact of the −964G > A SNP on gene regulation by HIF1 [40].

To understand the atomistic mechanisms associated with the differential action of HIF1 on the HRE promoter region, we constructed protein-DNA (HIF1-HRE) complexes, containing the −964G and −964A alleles and compared the physical–chemical properties and the complex stability, using chemistry computational approaches and MD, respectively. The −964A allele exhibited a decreased surface interaction and binding affinity to the HIF1/HRE complex and presented lower stability when compared to the complex containing the G allele, under hypoxic conditions [40]. In addition, the −964A allele induced a deformation in the double helix structure that led to a stacked base-pair opening in the DNA strand, negatively influencing the HIF1/HRE interaction and complex stability [10]. To understand the role of the HIF1/HLA-G interaction at the exon 2 HRE, in this review we further studied the atomistic features of this DNA-protein interaction.

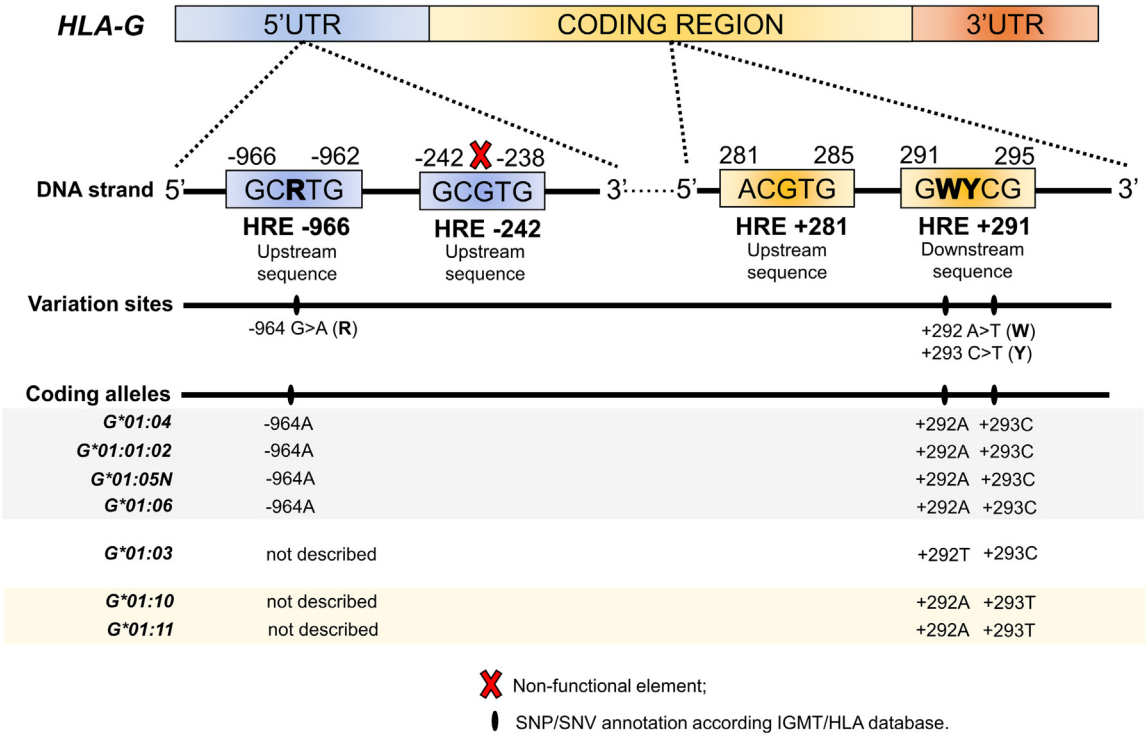

Fig. 1. Hypoxia regulatory elements (HRE) at the promoter and coding region (exon 2) of the *HLA-G* gene. A summary of the haplotypes that present at least one *HLA-G* allele located into HREs at the promoter and coding region is shown. 5′UTR: 5′untranslated region; 3′UTR: 3′untranslated region; SNP: Single Nucleotide Polymorphism: population frequency ≥ 1 %; SNV: Single Nucleotide Variation; population frequency < 1 %. The **R/W/Y** at the DNA strand represent the IUPAC nucleotide code, where R refers to G > A; W refers to A > T; and Y refers to C > T.

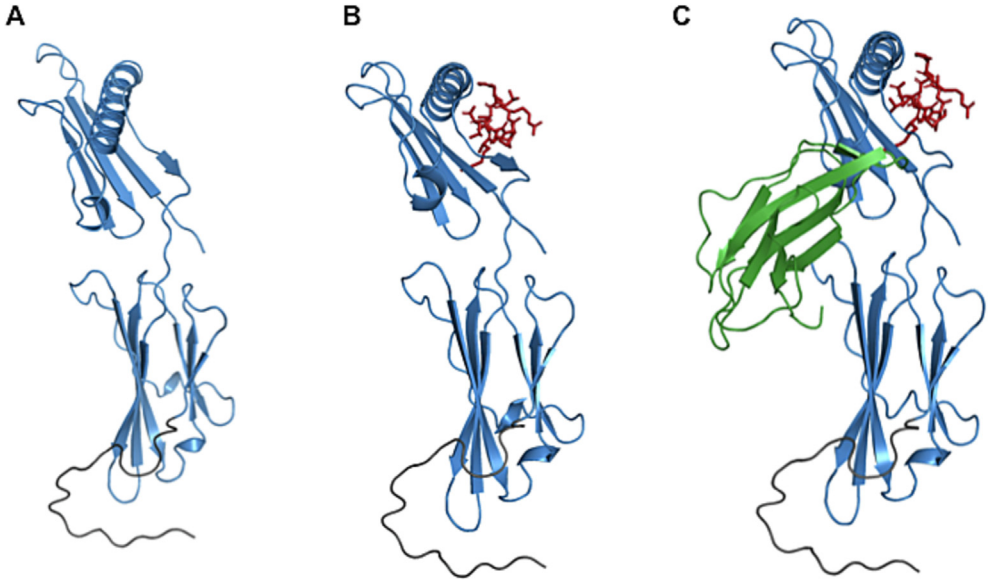

Fig. 2. Three-dimensional representation of the HLA-G6 isoform. (A) Free HLA-G6 monomer (blue); (B) HLA-G6 isoform coupled with the histone H2A RIIPRHLQL nonapeptide (red); (C) HLA-G6 isoform coupled with the β_2-microglobulin (green) and the RIIPRHLQL nonapeptide (red). In black is represented the 21 residues retained from intron 4. (For interpretation of the references to colour in this figure legend, the reader is referred to the web version of this article.)

2. Supplementary results and discussion

2.1. Structural aspects of the HLA-G6 isoform

The fact that HLA-G6 acts through LILRB2 is of particular interest because this receptor is expressed only by monocytes, dendritic cells, and macrophages [45]. The structural aspects of the HLA-G6 monomer are described here for the first time. Considering that HLA-G6 shares the α1/α3 domains with the complete HLA-G protein, a homology modeling strategy was used for this portion, while the residues retained from intron 4 were modeled by an *ab initio* methodology, as described for the HLA-G5 isoform [8]. The complete membrane-bound HLA-G1 model (based on the HLA-G*01:01 molecule that is the most common worldwide) was then applied as template for three

possible HLA-G6 isoform structures: monomer, monomer containing the nonapeptide in the cleft, and monomer containing the nonapeptide in the cleft and coupled with β_2-microglobulin. Validation software, image and structure visualization, residue interaction, and the Cα Root Mean Square Deviation (RMSD) and Root Mean Square Fluctuations (RMSF) values, calculated using the initial structures as reference, were performed as previously described [8] Three MD of 300 ns were performed, using GROMACS v5.1.4 [46] and CHARMM36m force field [47]. MD simulations were also performed in triplicate using GROMACS v4.6.5 package and the G54a7 force field, for a total of 2.2 μs, following closely similar approaches as described for the HLA-G5 model [8].

The best models obtained for the HLA-G6 isoform are shown in Fig. 2, illustrating the HLA-G6 alone (Fig. 2A), coupled with the RIIPRHLQL nonapeptide (Fig. 2B) or with peptide and β_2-microglobulin (Fig. 2C). The 21-residue tail retained from intron 4 assumes a loop conformation (Fig. 2 A/B/C).

The molecular dynamics of the HLA-G6 isoforms showed that two replicates (r1 and r2) assumed a relatively stable behavior when compared to replicate r3, which after 50 ns of simulation reached very high levels of RMSD (7–17 Å). The loop from the intron translation is also largely responsible for destabilizing the model, interacting with the α1 and α3 domains, destabilizing its whole structure. The RMSF showed that the structural fluctuation was more concentrated in the α1 domain and loop from the intron. The initial and final moments of the simulation demonstrated that the interaction of the loop with the α1 domain was able to lead to the partial unfolding of the α-helix found in this region. The RMSD value calculated between the structures removed from the trajectory at the initial and final moments is like the values obtained previously for unstable structures.

Soluble HLA-G6 isoform containing the peptide coupled to the α1 domain: the same structure of the monomer of the HLA-G6 isoform was used for the docking of the RIIPRHLQL nonapeptide to the α1 domain of the monomer to assess whether the presence of the peptide in the groove would be able to provide greater stability to the structure. The structure selected from the docking demonstrated strong binding energy of the RIIPRHLQL nonapeptide to the α1 domain (Fig. 2B). This structure was subjected to MD to examine the stability of the model. Even during the most stable simulations (r1 and r2), the RMSD values were high (10–27.7 Å), which correlates with the instability of the structure. The structure referring to the r3 replicate loses the necessary contacts with the peptide at 200 ns of the simulation when the peptide escapes from the α1 domain and leads to a sudden increase in the RMSD values. In addition, the RMSF showed very large structural variation throughout the HLA-G6 isoform and the peptide dynamics, contrary to what was observed for the HLA-G6 monomer, in which the fluctuation of residues was more concentrated in the residues that make up the α1 domain. The comparison of the initial and final structures of the trajectory demonstrated extensive deformations in the entire HLA-G6 isoform, such as an inversion of the α-helix and β-sheets that make up the α1 domain.

Soluble HLA-G6 isoform containing the peptide coupled to the α1 domain and to β_2-microgobulin: literature reports indicate the lack of the α2 domain as a reason for the impossibility of the presence of β_2-microglobulin in isoforms composed of α1-α3 domains, such as HLA-G6 [48]. Based on the structure of the HLA-G5 [8] isoform coupled to β_2-microglobulin and the interaction positions found in these structures, it was possible to obtain the docking of the HLA-G6 isoform coupled with peptide and β_2-microglobulin (Fig. 2C). The comparison of residues that are part of the HLA-G6 and β_2-microglobulin interaction demonstrated a smaller number of residues than those found in the HLA-G5 and β_2-microglobulin interaction interface [8], even considering the absence of the α2 domain in the HLA-G6 isoform.

MD for the system containing HLA-G6 with the nonapeptide RIIPRHLQL and coupled to the β2-microglobulin protein reached a total time of approximately 60 ns. Upon reaching this point of the trajectory, the software used for the dynamics simulations indicated the great instability of the system and the impossibility of continuing the test. All replicates of the HLA-G6 simulations reached this point of instability and ended abruptly. The RMSD values for replicates r1 and r2 were relatively stable compared to previous simulations for the other structural possibilities of HLA-G6. The initial and final moments of the trajectory demonstrated torsion of the α-helix that make up the α1 domain, but the other structural alterations were much less aggressive than the alterations seen in the previous simulations for the other structural possibilities of HLA- G6.

We believe that the impossibility of proceeding with the MD simulations is due to the positioning of the β_2-microglobulin, its distance from the α3 domain and, consequently, the lack of essential interactions between the α1-α3 and β_2-microglobulin domains. The RMSD values (Fig. 3) achieved during the HLA-G6 trajectories permitted to observe that the HLA-G6 monomer was the most stable construction. Even the HLA-G6 isoform containing peptide and β_2-microglobulin reached values close to RMSD, and its molecular dynamics system proved to be so unstable that it was not possible to continue the simulations.

As none of our presented structural possibilities were stable when subjected to MD, a possible stable structural conformation for these soluble HLA-G6 structures that have the α1 and α3 domains could be dimers composed of α1–α3: α1–α3, as in the work published by Kuroki et al. [49], where HLA-G2 (α1–α3 membrane) naturally forms a free β_2-microglobulin homodimer and does not have disulfide bonds joining the dimer monomers, resembling the arrangement of class II HLA heterodimers (α1-α2:β1-β2), an idea proposed by Ishitani and colleagues in 1992 [50]. One might think that perhaps the HLA-G6 isoform cannot exist as a soluble monomer and needs to be in a dimeric conformation to be stable.

The HLA-G6 and all studied isoform models/structures are available through GitHub (https://github.com/thaisarns/HLA-G6), permitting the use of these models as templates for modeling other alleles and isoforms.

2.2. Missense mutations at the major HLA-G receptors (LILRB1/LILRB2)

ITIM polymorphisms may impact the LILRB1/2 inhibitory signaling, interfering with the immune response regulation [51]. We reported two previously unpublished missense variants (rs41308744 and rs202204734) at the ITIM motif 2 of *LILRB1* gene [9]. These variants are in complete LD with each other but have no disequilibrium with any other variant evaluated (Fig. 4). The alternative C allele shows a frequency distribution in 1000 Genomes Project populations varying from 0.01 in East/South Asians to 0.03 in Africans, while the variant is monomorphic in American and European populations. In Brazilians, its frequency is like that found for Africans included in the 1000 Genomes Project. Despite the non-availability of ITIM structural analysis, codon-based signatures of selection analysis using the Fast-Unconstrained Bayesian Approximation (FUBAR) method [24] identified a positive selection signature under the residue E564K/Q or E564A. Other two missense variations showing positive selection signals (rs28409473-G583E and rs16985478-E625K) were indeed identified in the cytoplasmatic residues, around the ITIM regions. The rs28409473 is in complete LD with three other D1 polymorphisms (Fig. 4), including the T93A (rs12460501), identified by FoldX analysis [52,53] as the second more destabilizing variation. Noteworthy, T93A (rs12460501) and P68L (rs1061679) are in complete LD with each other and when evaluating only missense polymorphisms from cytoplasmic and D1/D2 domains, rs1061679 shows complete LD with four variants located at the D1 domain (Fig. 4). About 24 % of the LILRB1SNVs observed in the study encoded amino acids in the cytoplasmatic region. The Tajima's D Test results [54] agree with these results. Interestingly, a significant signal of balancing selection, or heterozygous advantage, was found for the final portion of the gene.

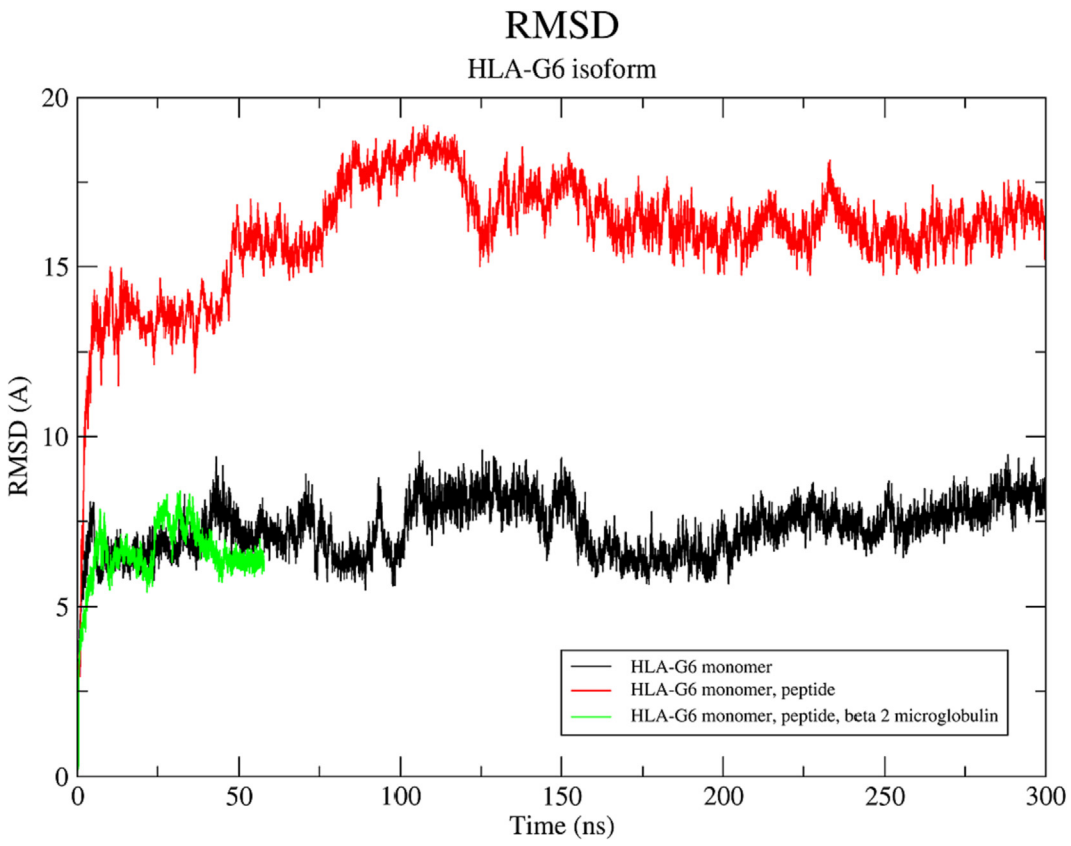

Fig. 3. RMSD (Root Mean Square Deviation) values obtained for the HLA-G6 isoforms. The molecular dynamics of the HLA-G monomer coupled with the β2-microglobulin and the histone H2A RIIPRHLQL nonapeptide shows the instability of the complex at 60 s.

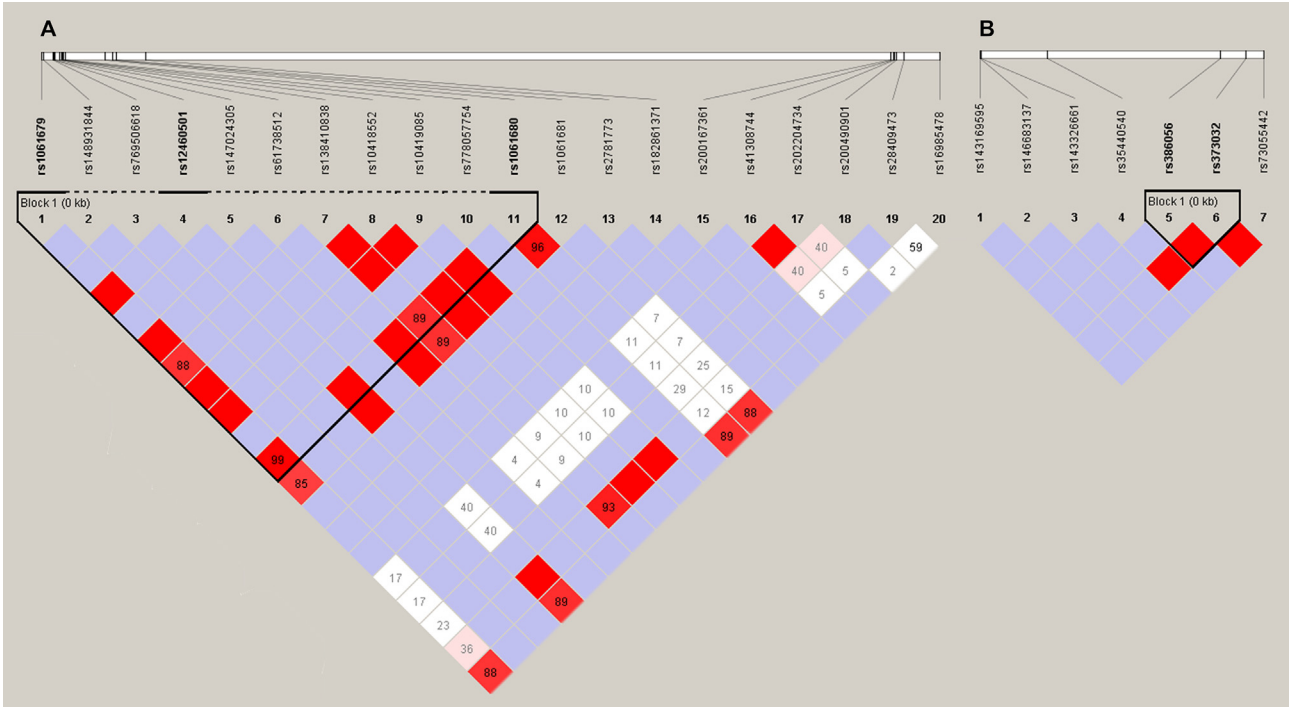

Fig. 4. Linkage disequilibrium (LD) among the missense single nucleotide variations observed at the nucleotide sequences that encode the D1/D2, ITIM-2, and other cytoplasmic regions of the *LILRB1* (A) and *LILRB2* (B) genes. Tagger SNPs are highlighted in bold. Areas in red indicate strong LD (LOD ≥ 2, D' = 1); areas in light red indicate moderate LD (LOD ≥ 2, D' < 1); areas in blue indicate LD with a lack of statistical evidence (LOD ≤ 2, D' = 1); while areas in white indicate no LD (LOD ≤ 2, D' < 1). D' values different from 1.00 are represented inside the squares as percentages. LOD, log of the odds; D', pairwise correlation between single-nucleotide polymorphisms. (For interpretation of the references to colour in this figure legend, the reader is referred to the web version of this article.)

Immunity-related genes correspond to one of the main classes enriched with either positive or balancing selection signatures [55]. Balancing selection is also an explanation for the high genetic diversity found in the HLA complex and is associated with the maintenance of genetic variability in the population [56]. In this case, the observed selection signatures highlight the importance of the immune signaling plasticity, i.e., the intensity of inhibitory signals transduced by the ITIM-2 motif may be alternated according to the polymorphisms observed, impacting on the LILRB1-mediated regulation of immune response. Therefore, although LILRB1 binds to conserved domains in HLA class I molecules, distinct LILRB1 proteins may have differential functional properties.

A more restricted genetic diversity is observed for *LILRB2* coding residue variation at the intra-cytoplasmatic region. Only a synonymous variation T593T (rs144888744) was observed at the ITIM-3 motif, showing no functional significance. Additionally, the alternative frequency is low in almost all populations from the 1000 Genomes Project, varying from 0.001 in Africans to 0.093 in South Asians, and in the Brazilian population, in which the frequency of rs144888744*G is 0.001. When LD is evaluated, the first observation concerns to the high level of SNVs with so low frequencies of the alternative allele that do not even are included in the LD plot. Among the variants that give rise to LD blocks, rs386056 and rs373032 are in complete LD with each other (Fig. 4). Both are at the D2 domain, with rs386056 differentiating haplotypes from group C, while rs373032 helps to identify haplotypes from group B [9]. The SNV rs386056 is also in complete LD with rs35440540, located in the cytoplasmic region, while rs373032 is in complete LD with rs73055442, situated at the D1 domain (Fig. 4). Moreover, both polymorphisms show expressive frequency variability. The rs386056*T frequency varies from 0.11 to 0.23 in Africans, Brazilians, Europeans, and Americans to 0.45 in the East Asians from the 1000 Genomes Project, while the rs373032*T frequency varies from 0.03 to 0.08 in Africans and East Asians to 0.14–0.31 in American, Brazilians, Europeans, and South Asians [9,31].

Purifying selection removes deleterious variations in a process known as background selection that is essential for preserving biological function [57].Therefore, given the important role of LILRB2 in immune regulation of lymphocytic cells, it would be expected a selective regimen supporting the maintenance of a proper molecule functioning. In conclusion, these results demonstrate the understanding of the impact of polymorphic variations and natural selection forces in *LILRB1/2* regions, given that changes in the molecules encoded by different alleles might disturb these protein–protein interactions, which have a central role in immunological response regulation.

2.3. The differential action of the HIF1 on the HLA-G gene

To further explore the role of HIF1 protein on the *HLA-G* HRE variability, we complemented our previous *in silico* analysis that evaluated the impact of the promoter region −964G > A HRE [10] now studying the polymorphic site at exon 2 (+292A > T and + 293C > T) within HRE. These studies were performed using the same DNA/protein docking strategy, as previously described [10]. DNA double helix 3D structures were constructed spanning from the + 274 to + 302 nucleotide [58], encompassing the + 292A > T (Figs. 1 and 5). The + 291 HRE polymorphic downstream sequence was used as active residues to drive the docking [59,60] of the free DNA molecules with HIF1 3D structure [10]. Three protein-DNA complexes encompassing the + 291 HRE were constructed and named as follows: (i) HIF1-HRE + 291AC (+292A; +293C), (ii) HIF1-HRE + 291TC (+292 T; +293C), and (iii) HIF1-HRE + 291AT (+292A; +293 T), as shown in Figs. 1 and 5.

Docked complexes presented similar 3D structures (data not shown), in which the HIF1 was attached to HRE in accordance with experimental/computational protein/DNA complexes [10,61–63]. The physical–chemical properties of the complex were used to calculate the HADDOCK score (HS) that estimates the binding affinity between the two docked molecules. Negative HS values mean that molecule partners interact easily and exhibit great binding affinity [5,60]. The HIF1-HRE + 291AC complex exhibited higher binding affinity (HS = -167.26) than the other two complexes, i.e., HIF-HRE + 291TC (HS = -161.15), and HIF1-HRE + 291AT (HS = -158.19). These values corroborate those observed for the docking of HIF1 to the −964G > A SNP, in which the HIF1-HRE-964G (-966 upstream HRE sequence: 5′GC**G**TG’3) and HIF1-HRE-964A (-966 upstream HRE sequence: 5′GC**A**TG’3) exhibited HS values of −163.53 and −147.90, respectively [10]. Similarly, the HIF1-HRE + 291AC (+291 downstream HRE sequence 3′GT**G**CG’5) shares the same HRE sequence of the HIF1-HRE-964G; and the HIF1-HRE + 291AT (+291 downstream HRE sequence: 3′GT**A**CG’5) has the same regulatory sequence of the HIF1-HRE-964A. This homology occurs because the promoter region and the exon 2 + 291 HRE shares the same third nucleotide (guanine/adenine) at the HIF1 HRE sequences of the studied complexes. In conclusion, we can infer that a guanine at the third residue position of the HRE has higher binding affinity for HIF1 than an adenine at the same position, irrespective of the location of the HRE in the *HLA-G* gene.

Besides evaluating the influence of HRE variability on the HIF1 binding, we studied the atomic interactions between amino acid side chains and nucleotides in the major and minor grooves of the DNA

Fig. 5. DNA double helix three-dimensional structure exhibiting the exon 2 HREs of the *HLA-G* gene. Sequence differences observed at the + 291 HRE polymorphisms are highlighted in the gray box, which were used to model the DNA molecules. DNA double helixes exhibited a length of 35 base pairs.

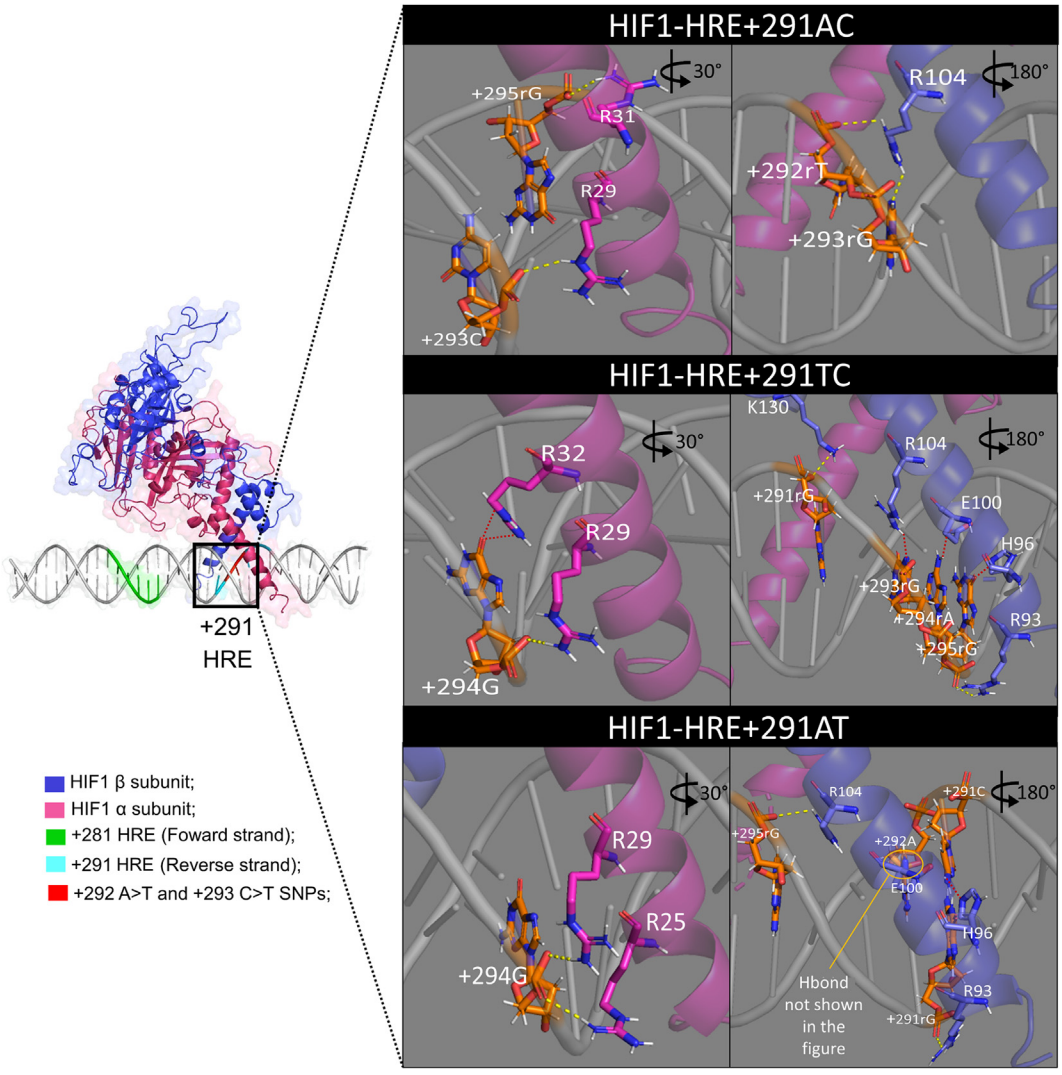

Fig. 6. Intermolecular interactions between + 291 hypoxia-responsive element (HRE) and HIF1 atom-pair, emphasizing the + 292 A > T and + 293C > T polymorphism at the *HLA-G* gene. On the left is shown a three-dimensional representation of the protein-DNA structure. The central box represents the + 291 HRE region. On the right are shown details of the hydrogen bonds (distance 2.5–3.2 Å) formed between residue atoms of α (pink) and β (blue) subunits with the atoms of the nucleotides. Hydrogen bond indicated by yellow dotted lines refers to amino acid atoms that interact with the sugar-phosphate backbone DNA atoms. Dotted dark red lines refer to amino acid atoms interacting with nitrogen base atoms. Nucleotides numbered as + 291G, +292A/T, +293C/T, +294C, +295G are upstream sequence and nucleotides numbered as + 291rC, +292rT/A, +293rG/A, +294rG, +295rC are downstream sequence of the DNA strand. (For interpretation of the references to colour in this figure legend, the reader is referred to the web version of this article.)

double helix, emphasizing the role of hydrogen bonds (Hbonds) [64]. Considering the interaction between HIF1 amino acids with the HRE + 291 nucleotides, we observed that Hbond formation was observed between amino acid atoms with: i) sugar-phosphate backbone DNA atoms (4 Hbonds for the HIF1-HRE + 291AC; 3 Hbonds for the HIF1-HRE + 291TC; and 5 Hbonds to the HIF1-HRE + 291AT, and ii) nitrogen base DNA atoms (none for the HIF1-HRE + 291AC; 5 Hbonds to the HIF1-HRE + 291TC; and 1 Hbond to the HIF1-HRE + 291AT) (Fig. 6).

Hydrogen bonds formed between atoms pairs of the amino acid side chains with the sugar-phosphate backbone are important interactions to stabilize the molecular complex, while interactions between atoms pairs of protein residues with DNA nitrogenous bases are responsible for the specificity of the binding of the transcription factor to the target regulatory element [65]. In this context, HIF1 binds at the + 291HRE containing + 292 T and + 293C allele combination in a more specific manner than the other complexes, because the HIF1-HRE + 291TC complex has more amino acids interacting with nitrogen base atoms (Fig. 6).

Although the distance from the HLA-G −966 to the + 281 HREs covers more than thousand nucleotides, the synergic interaction with HIF1 may be possible through CBP/p300 mediator proteins (co-factors), which may induce the formation of a DNA loop that allows the proximity of the transcription factor with more than one HRE, potentiating the transcriptional gene activity [40,66]. The HIF1 gene transcription regulation is dependent on the interaction between CBP/p300 with the HIF1a C-terminus transcriptional activation domain (C-TAD) [67]. Despite CBP/p300 does not bind directly to DNA, it may act as a bridge to connect one or more transcription factors, such as HIF1 with their respective regulatory elements, interacting with components of the basal transcription complex [68]. Therefore, *in silico* methods bring important insights regarding *HLA-G* genomic variability and the interaction between HIF1 with its regulatory element. Unfortunately, due to the absence of important components in these theoretical models, such as the C-TAD of the HIF1 α subunit and of the CBP/p300 cofactor, which are important to gene transactivation, we are unable to fully understand of HIF1 role in the transcriptional regulation of the *HLA-G* under hypoxic

conditions at an atomistic level. Further studies are needed to clarify *HLA-G* regulation by HIF1.

3. Conclusion

Although protein 3D structures as well as their partner molecules may provide information regarding functional activity, they also present limitations. For instance, very large protein systems require a powerful computer to calculate atoms dynamic behavior, requiring the use of supercomputers [69]. Furthermore, molecular docking predicts the molecular orientation of the docked complex using a scoring-function to estimate molecule affinity, based on physical–chemical features of solved molecule datasets [70,71], which may vary according to the underlying situation in physiological or non-physiological conditions.

Concluding, the atomistic theoretical models and MD studies may: (i) improve the understanding of the static HLA-G crystallographic structures; (ii) define the stability of the complete HLA-G molecule and of its isoforms; (iii) provide the basis for the modeling of other HLA-G molecules in addition to the HLA-G*01:01; (iv) help to explain experimental findings regarding the HLA-G binding stability to the major LILRB1/2 receptors; (v) recognize the putative differential immunosuppressive role of LILRB1/2 receptors caused by intracellular residue variations, particularly at the ITIM motifs; and (vi) understand the atomistic features of functional studies regarding the differential role of transcription factors on the *HLA-G* gene.

Declaration of Competing Interest

The authors declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper.

Acknowledgements

The authors acknowledge the support from the Brazilian Research Support Foundations CNPq (projects #302060/2019-7 for EAD and #151177/2022-8 for CCA), CAPES (Finance code 001), and FAPESP (projects INCTC #465539/2014-9 and # 2017/10780-4 for EAD).

References

[1] B. Kuhlman, P. Bradley, Advances in protein structure prediction and design, Nat. Rev. Mol. Cell Biol. 20 (2019) 681–697, https://doi.org/10.1038/s41580-019-0163-x.

[2] L. Slabinski, L. Jaroszewski, A.P.C. Rodrigues, L. Rychlewski, I.A. Wilson, S.A. Lesley, A. Godzik, The challenge of protein structure determination-lessons from structural genomics, Prot. Sci. 16 (2007) 2472–2482, https://doi.org/10.1110/ps.073037907.

[3] H.M. Berman, J. Westbrook, Z. Feng, G. Gilliland, T.N. Bhat, H. Weissig, I.N. Shindyalov, P.E. Bourne, The Protein Data Bank, Nucl. Acids Res. 28 (2000), https://doi.org/10.1093/nar/28.1.235.

[4] D. Cozzetto, A. Tramontano, Advances and pitfalls in protein structure prediction, Curr. Prot. Pept. Sci. 9 (2008), https://doi.org/10.2174/138920308786733958.

[5] X.-Y. Meng, H.-X. Zhang, M. Mezei, M. Cui, Molecular docking: a powerful approach for structure-based drug discovery, Curr. Comput. Aided-Drug Des. 7 (2011) 146–157, https://doi.org/10.2174/157340911795677602.

[6] A. Hospital, J.R. Goñi, M. Orozco, J.L. Gelpí, Molecular dynamics simulations: advances and applications, Adv. Appl. Bioinforma. Chem. 8 (2015), https://doi.org/10.2147/AABC.S70333.

[7] S.A. Hollingsworth, R.O. Dror, Molecular dynamics simulation for all, Neuron 99 (2018) 1129–1143, https://doi.org/10.1016/j.neuron.2018.08.011.

[8] T. Arns, D.A. Antunes, J.R. Abella, M.M. Rigo, L.E. Kavraki, S. Giuliatti, E.A. Donadi, Structural modeling and molecular dynamics of the immune checkpoint molecule HLA-G, Front. Immunol. 11 (2020), https://doi.org/10.3389/fimmu.2020.575076.

[9] M.L.G. Oliveira, E.C. Castelli, L.C. Veiga-Castelli, A.L.E. Pereira, L. Marcorin, T.M. T. Carratto, A.S. Souza, H.S. Andrade, A.L. Simões, E.A. Donadi, D. Courtin, A. Sabbagh, S. Giuliatti, C.T. Mendes-Junior, Genetic diversity of the LILRB1 and LILRB2 coding regions in an admixed Brazilian population sample, HLA. 100 (2022) 325–348, https://doi.org/10.1111/tan.14725.

[10] C.C. Alves, E.A. Donadi, S. Giuliatti, Structural characterization of the interaction of hypoxia inducible factor-1 with its hypoxia responsive element at the −964g > a variation site of the hla-g promoter region, Int. J. Mol. Sci. 22 (2021), https://doi.org/10.3390/ijms222313046.

[11] C.S. Clements, L. Kjer-Nielsen, L. Kostenko, H.L. Hoare, M.A. Dunstone, E. Moses, K. Freed, A.G. Brooks, J. Rossjohn, J. McCluskey, Crystal structure of HLA-G: A nonclassical MHC class I molecule expressed at the fetal–maternal interface, Proc. Natl. Acad. Sci. 102 (2005) 3360–3365. https://doi.org/10.1073/pnas.0409676102.

[12] M. Diehl, C. Münz, W. Keilholz, S. Stevanović, N. Holmes, Y.W. Loke, H.-G. Rammensee, Nonclassical HLA-G molecules are classical peptide presenters, Curr. Biol. 6 (1996) 305–314, https://doi.org/10.1016/S0960-9822(02)00481-5.

[13] D. Brown, J. Trowsdale, R. Allen, The LILR family: modulators of innate and adaptive immune pathways in health and disease, Tissue Antigens 64 (2004) 215–225, https://doi.org/10.1111/j.0001-2815.2004.00290.x.

[14] A.M. Martin, J.K. Kulski, C. Witt, P. Pontarotti, F.T. Christiansen, Leukocyte Ig-like receptor complex (LRC) in mice and men, Trends Immunol. 23 (2002) 81–88, https://doi.org/10.1016/S1471-4906(01)02155-X.

[15] X. Kang, J. Kim, M. Deng, S. John, H. Chen, G. Wu, H. Phan, C.C. Zhang, Inhibitory leukocyte immunoglobulin-like receptors: Immune checkpoint proteins and tumor sustaining factors, Cell Cycle 15 (2016) 25–40, https://doi.org/10.1080/15384101.2015.1121324.

[16] B.A. Kane, K.J. Bryant, H.P. McNeil, N.T. Tedla, Termination of immune activation: an essential component of healthy host immune responses, J. Innate Immun. 6 (2014) 727–738, https://doi.org/10.1159/000363449.

[17] M. Colonna, J. Samaridis, M. Cella, L. Angman, R.L. Allen, C.A. O'Callaghan, R. Dunbar, G.S. Ogg, V. Cerundolo, A. Rolink, Human myelomonocytic cells express an inhibitory receptor for classical and nonclassical MHC class I molecules, J. Immunol. 160 (1998) 3096–3100.

[18] D. Cosman, N. Fanger, L. Borges, M. Kubin, W. Chin, L. Peterson, M.-L. Hsu, A novel immunoglobulin superfamily receptor for cellular and viral MHC Class I Molecules, Immunity 7 (1997) 273–282, https://doi.org/10.1016/S1074-7613(00)80529-4.

[19] K.R. Chan, E.Z. Ong, H.C. Tan, S.L.-X. Zhang, Q. Zhang, K.F. Tang, N. Kaliaperumal, A.P.C. Lim, M.L. Hibberd, S.H. Chan, J.E. Connolly, M.N. Krishnan, S.M. Lok, B.J. Hanson, C.-N. Lin, E.E. Ooi, Leukocyte immunoglobulin-like receptor B1 is critical for antibody-dependent dengue, Proc. Natl. Acad. Sci. 111 (2014) 2722–2727. https://doi.org/10.1073/pnas.1317454111.

[20] K.J. Anderson, R.L. Allen, Regulation of T-cell immunity by leucocyte immunoglobulin-like receptors: innate immune receptors for self on antigen-presenting cells, Immunology 127 (2009) 8–17, https://doi.org/10.1111/j.1365-2567.2009.03097.x.

[21] W. van der Touw, H.-M. Chen, P.-Y. Pan, S.-H. Chen, LILRB receptor-mediated regulation of myeloid cell maturation and function, Cancer Immunol. Immunother. 66 (2017) 1079–1087, https://doi.org/10.1007/s00262-017-2023-x.

[22] B.E. Willcox, L.M. Thomas, P.J. Bjorkman, Crystal structure of HLA-A2 bound to LIR-1, a host and viral major histocompatibility complex receptor, Nat. Immunol. 4 (2003) 913–919, https://doi.org/10.1038/ni961.

[23] M. Shiroishi, K. Kuroki, L. Rasubala, K. Tsumoto, I. Kumagai, E. Kurimoto, K. Kato, D. Kohda, K. Maenaka, Structural basis for recognition of the nonclassical MHC molecule HLA-G by the leukocyte Ig-like receptor B2 (LILRB2/LIR2/ILT4/CD85d), Proc. Natl. Acad. Sci. 103 (2006) 16412–16417. https://doi.org/10.1073/pnas.0605228103.

[24] F. Canavez, N.T. Young, L.A. Guethlein, R. Rajalingam, S.I. Khakoo, B.P. Shum, P. Parham, Comparison of Chimpanzee and human Leukocyte Ig-like receptor genes reveals framework and rapidly evolving genes, J. Immunol. 167 (2001) 5786–5794, https://doi.org/10.4049/jimmunol.167.10.5786.

[25] N.T. Young, F. Canavez, M. Uhrberg, B.P. Shum, P. Parham, Conserved organization of the ILT / LIR gene family within the polymorphic human leukocyte receptor complex, Immunogenetics 53 (2001) 270–278, https://doi.org/10.1007/s002510100332.

[26] C.L. Davidson, N.L. Li, D.N. Burshtyn, LILRB1 polymorphism and surface phenotypes of natural killer cells, Hum. Immunol. 71 (2010) 942–949, https://doi.org/10.1016/j.humimm.2010.06.015.

[27] K. Kuroki, N. Tsuchiya, M. Shiroishi, L. Rasubala, Y. Yamashita, K. Matsuta, T. Fukazawa, M. Kusaoi, Y. Murakami, M. Takiguchi, T. Juji, H. Hashimoto, D. Kohda, K. Maenaka, K. Tokunaga, Extensive polymorphisms of LILRB1 (ILT2, LIR1) and their association with HLA-DRB1 shared epitope negative rheumatoid arthritis, Hum. Mol. Genet. 14 (2005) 2469–2480, https://doi.org/10.1093/hmg/ddi247.

[28] K. Hirayasu, J. Ohashi, H. Tanaka, K. Kashiwase, A. Ogawa, M. Takanashi, M. Satake, G.J. Jia, N.-O. Chimge, E.W. Sideltseva, K. Tokunaga, T. Yabe, Evidence for natural selection on leukocyte immunoglobulin-like receptors for HLA Class I in Northeast Asians, Am. J. Hum. Genet. 82 (2008) 1075–1083, https://doi.org/10.1016/j.ajhg.2008.03.012.

[29] N.A. Papanikolaou, E.R. Vasilescu, N. Suciu-Foca, Novel single nucleotide polymorphisms in the human immune inhibitory immunoglobulin-like T cell receptor type 4, Hum. Immunol. 65 (2004) 700–705, https://doi.org/10.1016/j.humimm.2004.04.003.

[30] J.S. Affandi, Z.K.A. Aghafar, B. Rodriguez, M.M. Lederman, S. Burrows, D. Senitzer, P. Price, Can immune-related genotypes illuminate the immunopathogenesis of cytomegalovirus disease in human immunodeficiency virus–infected patients?, Hum Immunol. 73 (2012) 168–174, https://doi.org/10.1016/j.humimm.2011.11.005.

[31] K. Yu, C.L. Davidson, A. Wójtowicz, L. Lisboa, T. Wang, A.M. Airo, J. Villard, J. Buratto, T. Sandalova, A. Achour, A. Humar, K. Boggian, A. Cusini, C. van Delden, A. Egli, O. Manuel, N. Mueller, P.-Y. Bochud, D.N. Burshtyn, LILRB1 polymorphisms influence posttransplant HCMV susceptibility and ligand

interactions, J. Clin. Investig. 128 (2018) 1523–1537, https://doi.org/10.1172/JCI96174.
[32] A. Auton et al, A global reference for human genetic variation, Nature 526 (2015) 68–74, https://doi.org/10.1038/nature15393.
[33] Q. Wang, H. Song, H. Cheng, J. Qi, G. Nam, S. Tan, J. Wang, M. Fang, Y. Shi, Z. Tian, X. Cao, Z. An, J. Yan, G.F. Gao, Structures of the four Ig-like domain LILRB2 and the four-domain LILRB1 and HLA-G1 complex, Cell Mol. Immunol. 17 (2020) 966–975, https://doi.org/10.1038/s41423-019-0258-5.
[34] B.A. Kilburn, J. Wang, Z.M. Duniec-Dmuchkowski, R.E. Leach, R. Romero, D.R. Armant, Extracellular matrix composition and hypoxia regulate the expression of HLA-G and integrins in a human trophoblast cell line1, Biol. Reprod. 62 (2000) 739–747, https://doi.org/10.1095/biolreprod62.3.739.
[35] C. Gu, S. Park, J. Seok, H.Y. Jang, Y.J. Bang, G.I.J. Kim, Altered expression of ADM and ADM2 by hypoxia regulates migration of trophoblast and HLA-G expression†, Biol. Reprod. 104 (2021) 159–169, https://doi.org/10.1093/biolre/ioaa178.
[36] M. Bourguignon, L. Yaghi, S. Flajollet, I. Radanne-Krawice, N. Rouas-Freiss, D. Lugrin, J.-P. Richalet, E.D. Carosella, P. Moreau, Increased soluble human leukocyte antigen–G levels in peripheral blood from climbers on Mount Everest, Hum. Immunol. 71 (2010) 1105–1108, https://doi.org/10.1016/j.humimm.2010.08.011.
[37] V.C. Jacovas, R.T. Michita, R. Bisso-Machado, G. Reales, E.M. Tarazona-Santos, J. R. Sandoval, A. Salazar-Granara, J.A.B. Chies, M.C. Bortolini, HLA-G 3'UTR haplotype frequencies in highland and lowland South Native American populations, Hum. Immunol. 83 (2022) 27–38, https://doi.org/10.1016/j.humimm.2021.09.002.
[38] C.-C. Chang, S. Ferrone, HLA-G in melanoma: can the current controversies be solved?, Semin Cancer Biol. 13 (2003) 361–369, https://doi.org/10.1016/S1044-579X(03)00027-0.
[39] G. Mouillot, C. Marcou, I. Zidi, C. Guillard, D. Sangrouber, E.D. Carosella, P. Moreau, Hypoxia Modulates HLA-G Gene Expression in Tumor Cells, Hum. Immunol. 68 (2007) 277–285, https://doi.org/10.1016/j.humimm.2006.10.016.
[40] L. Yaghi, I. Poras, R.T. Simoes, E.A. Donadi, J. Tost, A. Daunay, B.S. de Almeida, E. D. Carosella, P. Moreau, Hypoxia inducible factor-1 mediates the expression of the immune checkpoint HLA-G in glioma cells through hypoxia response element located in exon 2, Oncotarget 7 (2016) 63690–63707, https://doi.org/10.18632/oncotarget.11628.
[41] V.L. Dengler, M.D. Galbraith, J.M. Espinosa, Transcriptional regulation by hypoxia inducible factors, Crit. Rev. Biochem. Mol. Biol. 49 (2014) 1–15, https://doi.org/10.3109/10409238.2013.838205.
[42] G.L. Semenza, Hypoxia-inducible factors in physiology and medicine, Cell 148 (2012), https://doi.org/10.1016/j.cell.2012.01.021.
[43] E.C. Castelli, J. Ramalho, I.O.P. Porto, T.H.A. Lima, L.P. FelÃcio, A. Sabbagh, E.A. Donadi, C.T. Mendes-Junior, Insights into HLA-G genetics provided by worldwide haplotype diversity, Front. Immunol. 5 (2014), https://doi.org/10.3389/fimmu.2014.00476.
[44] E.C. Castelli, B.S. de Almeida, Y.C.N. Muniz, N.S.B. Silva, M.R.S. Passos, A.S. Souza, A.E. Page, M. Dyble, D. Smith, G. Aguileta, J. Bertranpetit, A.B. Migliano, Y. A.O. Duarte, M.O. Scliar, J. Wang, M.R. Passos-Bueno, M.S. Naslavsky, M. Zatz, C. T. Mendes-Junior, E.A. Donadi, HLA-G genetic diversity and evolutive aspects in worldwide populations, Sci. Rep. 11 (2021) 23070, https://doi.org/10.1038/s41598-021-02106-4.
[45] K.-Y. HoWangYin, M. Loustau, J. Wu, E. Alegre, M. Daouya, J. Caumartin, S. Sousa, A. Horuzsko, E.D. Carosella, J. LeMaoult, Multimeric structures of HLA-G isoforms function through differential binding to LILRB receptors, Cell. Mol. Life Sci. 69 (2012) 4041–4049, https://doi.org/10.1007/s00018-012-1069-3.
[46] M.J. Abraham, T. Murtola, R. Schulz, S. Páll, J.C. Smith, B. Hess, E. Lindahl, GROMACS: High performance molecular simulations through multi-level parallelism from laptops to supercomputers, SoftwareX. 1–2 (2015) 19–25, https://doi.org/10.1016/j.softx.2015.06.001.
[47] J. Huang, S. Rauscher, G. Nawrocki, T. Ran, M. Feig, B.L. de Groot, H. Grubmüller, A.D. MacKerell, CHARMM36m: an improved force field for folded and intrinsically disordered proteins, Nat. Methods. 14 (2017) 71–73, https://doi.org/10.1038/nmeth.4067.
[48] V.A. Tysoe-Calnon, J.E. Grundy, S.J. Perkins, Molecular comparisons of the β2-microglobulin-binding site in class I major-histocompatibility-complex α-chains and proteins of related sequences, Biochem. J 277 (1991) 359–369, https://doi.org/10.1042/bj2770359.
[49] K. Kuroki, K. Mio, A. Takahashi, H. Matsubara, Y. Kasai, S. Manaka, M. Kikkawa, D. Hamada, C. Sato, K. Maenaka, Cutting Edge: Class II–like Structural Features and Strong Receptor Binding of the Nonclassical HLA-G2 Isoform Homodimer, J. Immunol. 198 (2017) 3399–3403, https://doi.org/10.4049/jimmunol.1601296.
[50] A. Ishitani, D.E. Geraghty, Alternative splicing of HLA-G transcripts yields proteins with primary structures resembling both class I and class II antigens, Proc. Natl. Acad. Sci. 89 (1992) 3947–3951, https://doi.org/10.1073/pnas.89.9.3947.
[51] A. Wiśniewski, A. Kowal, E. Wyrodek, I. Nowak, E. Majorczyk, M. Wagner, E. Pawlak-Adamska, R. Jankowska, B. Ślesak, I. Frydecka, P. Kuśnierczyk, Genetic polymorphisms and expression of HLA-G and its receptors, KIR2DL4 and LILRB1, in non-small cell lung cancer, Tissue Antigens 85 (2015) 466–475, https://doi.org/10.1111/tan.12561.
[52] J. Schymkowitz, J. Borg, F. Stricher, R. Nys, F. Rousseau, L. Serrano, The FoldX web server: an online force field, Nucleic Acids Res. 33 (2005) W382–W388. https://doi.org/10.1093/nar/gki387.
[53] R. Guerois, J.E. Nielsen, L. Serrano, Predicting changes in the stability of proteins and protein complexes: a study of more than 1000 mutations, J. Mol. Biol. 320 (2002) 369–387, https://doi.org/10.1016/S0022-2836(02)00442-4.
[54] F. Tajima, Statistical method for testing the neutral mutation hypothesis by DNA polymorphism, Genetics 123 (1989) 585–595, https://doi.org/10.1093/genetics/123.3.585.
[55] P.C. Sabeti, S.F. Schaffner, B. Fry, J. Lohmueller, P. Varilly, O. Shamovsky, A. Palma, T.S. Mikkelsen, D. Altshuler, E.S. Lander, Positive natural selection in the human lineage, Science 312 (2006) (1979) 1614–1620, https://doi.org/10.1126/science.1124309.
[56] D. Garrigan, P.W. Hedrick, Perspective: Detecting adaptive molecular polymorphism: Lessons from the MHC, Evolution (N Y). 57 (2003) 1707–1722, https://doi.org/10.1111/j.0014-3820.2003.tb00580.x.
[57] I. Cvijović, B.H. Good, M.M. Desai, The Effect of Strong Purifying Selection on Genetic Diversity, Genetics 209 (2018) 1235–1278, https://doi.org/10.1534/genetics.118.301058.
[58] S. Li, W.K. Olson, X.-J. Lu, Web 3DNA 2.0 for the analysis, visualization, and modeling of 3D nucleic acid structures, Nucl. Acids Res. 47 (2019) W26–W34, https://doi.org/10.1093/nar/gkz394.
[59] G.C.P. van Zundert, J.P.G.L.M. Rodrigues, M. Trellet, C. Schmitz, P.L. Kastritis, E. Karaca, A.S.J. Melquiond, M. van Dijk, S.J. de Vries, A.M.J.J. Bonvin, The HADDOCK2.2 Web Server: User-Friendly Integrative Modeling of Biomolecular Complexes, J Mol Biol. 428 (2016) 720–725, https://doi.org/10.1016/j.jmb.2015.09.014.
[60] S.J. de Vries, M. van Dijk, A.M.J.J. Bonvin, The HADDOCK web server for data-driven biomolecular docking, Nat Protoc. 5 (2010) 883–897, https://doi.org/10.1038/nprot.2010.32.
[61] K.W. Schulte, E. Green, A. Wilz, M. Platten, O. Daumke, Structural basis for aryl hydrocarbon receptor-mediated gene activation, Structure 25 (2017) 1025–1033. e3, https://doi.org/10.1016/j.str.2017.05.008.
[62] D. Wu, N. Potluri, J. Lu, Y. Kim, F. Rastinejad, Structural integration in hypoxia-inducible factors, Nature 524 (2015) 303–308, https://doi.org/10.1038/nature14883.
[63] D. Wu, X. Su, N. Potluri, Y. Kim, F. Rastinejad, NPAS1-ARNT and NPAS3-ARNT crystal structures implicate the bHLH-PAS family as multi-ligand binding transcription factors, Elife 5 (2016), https://doi.org/10.7554/eLife.18790.
[64] M. Michael Gromiha, J.G. Siebers, S. Selvaraj, H. Kono, A. Sarai, Intermolecular and Intramolecular Readout Mechanisms in Protein–DNA Recognition, J Mol Biol. 337 (2004) 285–294, https://doi.org/10.1016/j.jmb.2004.01.033.
[65] N.M. Luscombe, J.M. Thornton, Protein–DNA Interactions: Amino Acid Conservation and the Effects of Mutations on Binding Specificity, J Mol Biol. 320 (2002) 991–1009, https://doi.org/10.1016/S0022-2836(02)00571-5.
[66] M. Garziera, L. Scarabel, G. Toffoli, Hypoxic Modulation of HLA-G Expression through the Metabolic Sensor HIF-1 in Human Cancer Cells, J Immunol Res. 2017 (2017) 1–13, https://doi.org/10.1155/2017/4587520.
[67] S.J. Freedman, Z.-Y.-J. Sun, F. Poy, A.L. Kung, D.M. Livingston, G. Wagner, M.J. Eck, Structural basis for recruitment of CBP/p300 by hypoxia-inducible factor-1α, Proc. Natl. Acad. Sci. 99 (2002) 5367–5372, https://doi.org/10.1073/pnas.082117899.
[68] F. Wang, C.B. Marshall, M. Ikura, Transcriptional/epigenetic regulator CBP/p300 in tumorigenesis: structural and functional versatility in target recognition, Cellular and Molecular, Life Sci. 70 (2013) 3989–4008, https://doi.org/10.1007/s00018-012-1254-4.
[69] A. Heinecke, W. Eckhardt, M. Horsch, H.-J. Bungartz, Supercomputing for Molecular Dynamics Simulations (2015), https://doi.org/10.1007/978-3-319-17148-7.
[70] L. Pinzi, G. Rastelli, Molecular Docking: Shifting Paradigms in Drug Discovery, Int J Mol Sci. 20 (2019) 4331, https://doi.org/10.3390/ijms20184331.
[71] F.D. Prieto-Martínez, M. Arciniega, J.L. Medina-Franco, Molecular docking: current advances and challenges, TIP Revista Especializada En Ciencias Químico-Biológicas. 21 (2018), https://doi.org/10.22201/fesz.23958723e.2018.0.143.