UNIVERSITÀ DEGLI STUDI DI TRIESTE
e
UNIVERSITÀ CA' FOSCARI DI VENEZIA

XXXI CICLO DEL DOTTORATO DI RICERCA IN CHIMICA

Predictive and prognostic value of the HLA-G immunocheckpoint in metastatic colorectal cancer treated with FOLFIRI regimen

Settore scientifico-disciplinare: **BIO/12**

Ph.D. STUDENT
LUCIA SCARABEL

Ph.D. PROGRAM COORDINATOR
PROF. BARBARA MILANI

THESIS SUPERVISOR
PROF. GABRIELE GRASSI

CO-SUPERVISORS
DR. MARICA GARZIERA
PROF. SILVANO GEREMIA
PROF. MARIO GRASSI
DR. GIUSEPPE TOFFOLI

ACADEMIC YEAR 2017/2018

This thesis project was carried out at the
Joint Doctoral Program in Chemistry between
University of Trieste and Ca' Foscari University in Venice,
Ph.D. Coordinator Prof. Barbara Milani,
in collaboration with the Experimental and Clinical Pharmacology Unit of
the Centro di Riferimento Oncologico (CRO) IRCCS in Aviano,
Director Dott. Giuseppe Toffoli.

Questo lavoro di tesi è stato svolto presso
la Scuola di Dottorato in Chimica in convenzione fra
l'Università degli Studi di Trieste e l'Università Ca' Foscari di Venezia,
Coordinatrice Prof.ssa Barbara Milani,
in collaborazione con la Struttura Operativa Complessa di Farmacologia Sperimentale e Clinica
del Centro di Riferimento Oncologico (CRO) IRCCS *di Aviano,*
Direttore Dott. Giuseppe Toffoli.

Contents

ABSTRACT

In recent years, new and innovative immunotherapy drugs entered the market revolutionizing the clinical management of patients suffering of cancer diseases. These strategies aim to reinforce the immune system through the interaction with immune checkpoint inhibitors. In this contest, immunogenetics could give insights in the governance of genetics on the phenomenon of immune tolerance to tumor antigens, leading to a better understanding of the cancer natural history and to the identification of biomarkers of treatment response, toxicity and prognosis. In light of this, the analysis of new immune factors as human leukocyte antigen-g (HLA-G) could potentially offer new tools to be translated in the clinical practice.

In this thesis, the role of HLA-G in the metastatic colorectal cancer (mCRC) was here investigated in a retrospective cohort of patients. Despite the improvement of various strategies such as monoclonal antibodies combined with chemotherapy to treat these patients, the 5-year survival rate for mCRC is only slightly over 12% [1]. Moreover, nearly half of mCRC patients are resistant to 5-fluorouracil-based chemotherapies and unresponsive to immunotherapy. Then, novel predictive and prognostic immunobiomarkers are needed to better manage the patients with mCRC.

This thesis was divided in three sessions: PART 1, 2 and 3.

In PART 1, the predictive and prognostic effect of functional germline polymorphisms in the *HLA-G* 3'UTR was evaluated in 248 patients with mCRC homogeneously treated with first-line FOLFIRI (irinotecan (CPT-11) plus 5-fluorouracil (5-FU) and leucovorin) regimen. The primary endpoints were overall survival (OS) and time to progression (TTP). Secondary endpoints were tumor response and develop of severe toxicity. The study endpoints were also evaluated considering different HLA-G expression level classes, determined on the base of the *HLA-G* haplotypes. Two "HLA-G secretor models" were designed, on the base of the reported associations of *HLA-G* 3'UTR haplotypes with different expression levels of the HLA-G protein: the HLA-G 2 levels (M (Medium/High) level=UTR-1, UTR-3, UTR-4, UTR-6; L (Low) level=UTR-2, UTR-5, UTR-7); and 3 levels (H (High) level=UTR-1; I (Intermediate) level=UTR-2, UTR-3, UTR-4, UTR-6; S (Low) level=UTR-5, UTR-7) secretor models. To exclude a potential confounder factor in the results, we also considered the *UGT1A1*28* (rs8175347) polymorphism due to its known impact on irinotecan toxicity (decreased glucoronidation of the 7-ethyl-10-hydroxycamptothecin (SN-38), the active metabolite of the CPT-11).

In PART 2, the soluble HLA-G (sHLA-G) protein was measured in a subgroup of 40 patients with mCRC and studied for the associations with clinical characteristics of the patients and CPT-11 pharmacokinetic parameters. Furthermore, the possible association between sHLA-G expression and the non-coding regulatory (3'UTR) and coding *HLA-G* polymorphisms was analyzed to determine potential genotypic/phenotypic relationships. To our knowledge this is the first study to report the genotyping of coding exons for HLA-G mature protein in patients with mCRC. Finally, in PART 3, the intermolecular interactions between HLA-G and CPT-11 were explored. At the best of our knowledge this matter has never been analyzed before, even if it could explain interindividual variations in the clinical outcome.

In PART 1, the characterization of *HLA-G* 3'UTR polymorphisms was performed using the standard validated Sanger method. Specifically, the genetic analysis was realized to identify novel candidate prognostic and predictive *HLA-G*-related biomarkers that can be potentially translated in the personalization of mCRC treatment. We investigated the influence of *HLA-G* 3'UTR polymorphisms and related haplotypes in both OS and TTP, but we did not find any significant association. However, the evaluation of HLA-G secretor models in survival highlighted an association with a "high" secretor HLA-G with a favorable prognosis, and "low" secretor HLA-G, with a poor and unfavorable prognosis. Specifically, we found an association between the presence of the low-level S/S class and a significant reduced OS (HR=2.91, 95%CI: 1.03-8.18;

p=0.0432) compared to the intermediate I/I class, according to the HLA-G 3 levels secretor model, in multivariate analysis. Moreover, the negative impact on OS defined by the homozygous UTR-5/UTR-7 (S/S) low levels class in this model, seems to be independent from the effect exerted by *UGT1A1**28 polymorphism, as confirmed in the "strata" analyses. In TTP, candidate markers emerged only considering the contribution of both *HLA-G* 3'UTR and *UGT1A1**28 polymorphisms, with a similar trend of "high" secretor HLA-G with a favorable outcome. In particular, according to the HLA-G 2 levels secretor model, we found a significant survival advantage in TTP, in carriers heterozygous for the *UGT1A1**28 allele and the M "high" level HLA-G class (HR=0.71, 95%CI: 0.53-0.94; *p*=0.0231), in multivariate analysis.

The objective tumor response in patients with mCRC after first-line FOLFIRI treatment was evaluated according to: the response rate (RR) (stable disease (SD) + progression disease (PD) *vs* complete response (CR) + partial response (PR)); the clinical benefit (CB) (PD *vs* SD+ PR+CR), and the complete response (CR *vs* SD + PD + PR). A predictive value of *HLA-G* 3'UTR polymorphisms in tumor response was highlighted only for the CR and CB parameters. The main result was found in CR: patients homozygous for the *HLA-G* UTR-1 haplotype, related to high level of HLA-G expression, were associated with complete regression of malignant lesions (OR=0.1, 95%CI: 0.03-0.41; *p*=0.0013), in multivariate analysis. This data was confirmed when the single *HLA-G* 3'UTR alleles (rs371194629-*Del*, rs1710-*G*, rs1063320-*C*, rs9380142-*G*), included in UTR-1 haplotype, were analyzed (additive genetic model). Regarding the CB, one copy of *HLA-G* UTR-3 haplotype was associated in multivariate analysis with a reduced risk of CB (OR=0.47, 95%CI: 0.25-0.91; *p*=0.0253), and then was associated with PD. The association of UTR-3 haplotype in patients that not received a benefit from FOLFIRI treatment was concordant for the different allelic representation in some *HLA-G* 3'UTR alleles (rs1710-*C*, rs1063320-*G*, rs9380142-*A*) with respect to the UTR-1 haplotype. Nevertheless, the effect of UTR-1 in homozygous patients compared to UTR-3 in heterozygous patients was due to the "strength" of the two copies of *Del* allele, as confirmed in the high-level HLA-G classes, in both secretor models. Besides that, evaluating the possible influence of the *UGT1A1*28* polymorphism, previously reported to be associated in the same study cohort with a worse tumor response in carriers of the wild type **1* allele, we found that the protective effect of *Del* allele in CR, was confirmed only in patients with *UGT1A1*1/*1* genotype, in both high level HLA-G classes of secretor models, in particular in the H/H carriers (HR=0.09, 95%CI: 0.02-0.46; *p*=0.0037), in multivariate analysis.

Finally, we studied the risk of develop the main severe G3-4 toxicities of irinotecan-containing FOLFIRI regimen, considering first cycle (first cycle toxicities) and total amount of planned FOLFIRI chemotherapy cycles (cumulative toxicities). The main and severe toxicities are neutropenia, diarrhea, and other non-hematological toxicities of the GI tract that limit the clinical use of the treatment (nausea, vomiting), and are related to irinotecan. Hematological toxicity, non-hematological toxicity, total toxicity derived from the sum of the hematological and non-hematological events, and neutropenia, were evaluated in both first-cycle and cumulative option. The *UGT1A1*28/*28* genotype had showed an association with the G3-4 hematological and total toxicity only after the first cycle. The main significant associations in first cycle toxicities were identified in total toxicity for the functional *HLA-G* rs9380142-*(A/G+G/G)* genotype, according to the dominant model, that was associated with a reduced risk of G3-4 total toxicity (OR=0.38, 95%CI: 0.14-1.02; *p*=0.0447); this result was confirmed in heterozygous carriers for UTR-1 (OR=0.22, 95%CI: 0.06-0.78; *p*=0.0192), which contain the mutated rs9380142-*G* allele. The protective role of rs9380142-*(A/G+G/G)* genotype (OR=0.5, 95%CI: 0.27-0.89; *p*=0.0181) and heterozygous UTR-1 haplotype (OR=0.44, 95%CI: 0.23-0.82; *p*=0.0102) was also replicated in the cumulative total toxicity. On the contrary, the *HLA-G* heterozygous UTR-6 haplotype, characterized by the presence of rs9380142-*A* allele, was associated with an increased risk of G3-4 total cumulative toxicity (OR=3.34, 95%CI: 1.24-9.02; *p*=0.0173). The rs1707-*C* allele, according to the log-additive model, and the rs1710-*(C/C+C/G)* and rs1063320-*(G/G+G/C)* genotypes, both according to the recessive model, were found to be predictive of G3-4 cumulative neutropenia.

Patients homozygous for *HLA-G* UTR-4 haplotype, that includes the rs1707-*C*, rs1710-*G* and rs1063320-*C* mutated alleles, were associated with increased risk of G3-4 neutropenia, confirming the value of single SNPs, but with a weak significance. Similarly, these alleles and the UTR-4 haplotype were identified as predictors of G3-4 hematological toxicity, in the cumulative toxicity analysis. These findings are explained by the main contribution of neutropenia among the hematological toxicities.

Of note, in first cycle toxicity, a trend towards an association with develop of G3-4 toxicities was observed when low level HLA-G classes were coupled with the **28* allele, suggesting the influence of the *UGT1A1*28* allele. Moreover, the diminished risk of develop G3-4 total toxicity observed in presence of the *Del* allele in first cycle toxicity analysis, was confirmed (OR=0.26, 95%CI: 0.07-0.9; *p*=0.0343) in the heterozygous carriers of the high level H class and **1* wild type allele.

In cumulative toxicity analysis, heterozygous patients for S class were associated with an increased risk of G3-4 non-hematological toxicity (OR=2.52, 95%CI: 1.16-5.48; *p*=0.0196). We observed that low level HLA-G classes had an increased risk of G3-4 non-hematological toxicity, in particular in presence of the **1* allele, as evidenced in *UGT1A1*1/*1* stratum for L class (OR=2.09, 95%CI: 1.04-4.19; *p*=0.0333) according to the additive model. Finally, the different expression levels classes of HLA-G seemed to slightly modify the predictive toxicity profiles of patients with mCRC treated with FOLFIRI regimen, independently from the *UGT1A1*28* polymorphism, almost in the cumulative toxicities.

In PART 2, the sHLA-G protein (HLA-G1 and HLA-G5 isoforms) was measured in plasma samples from 40 patients with mCRC using a commercial ELISA assay. Plasma samples were retrospectively collected before the start of chemotherapy treatment with first-line FOLFIRI regimen. The median sHLA-G value was 116.4 U/ml, showing higher concentration values with respect to those previously reported in healthy controls and other CRC subpopulations studied. The correlation between sHLA-G and clinical-pathological characteristics, showed a significant association with the number of metastatic sites, comparing patients with 1 metastatic site and patients with more than 1 extra-regional lesion (*p*=0.0214). This result was also confirmed stratifying the patients for 1 metastatic site, 2 sites and >2 extra-regional lesions (*p*=0.0486). The study of genotypic-phenotypic correlation between non-coding/coding *HLA-G* polymorphisms and sHLA-G levels, showed no significant associations, probably due to the small sample size. However, the highest sHLA-G value (1552.7 U/ml) was identified in a patient carrier of the rs371194629-*Del/Del* genotype (and consequently UTR-1/UTR-1 haplotype), related to high HLA-G secretor, as we observed a trend of diminished sHLA-G in carriers for *Ins/Ins* genotype (91.5 U/ml) compared to the *Del* carriers (120.4 U/ml) in the recessive model (*Ins/Ins vs Del/Del*+*Ins/Del*). Intriguingly, for the synonymous *HLA-G* coding SNP rs1130355 (p.Pro57Pro) we observed a trend for sHLA-G median value in wild type *A/A* (135.4 U/ml), heterozygous *A/G* (120.4 U/ml) and mutated *G/G* (90.6 U/ml) genotypes, suggesting a possible mechanism influencing *HLA-G* transcription efficiency. Furthermore, the unexpected prevalence of the rs1130355-*A* allele, reported as the wild type allele in EUR subpopulation, may suggest an association with the CRC disease that should be explored and confirmed in a larger cohort of patients.

Pharmacokinetic parameters (CPT-11 AUC, SN38 AUC, SN38G, biliary Index, glucuronidation ratio) were available for 22 of 40 patients with mCRC, here considered for evaluation of circulating sHLA-G. In this subset of patients, median sHLA-G value was 86.4 U/ml: patients were stratified in sHLA-G High (≥86.4 U/ml) and sHLA-G Low (<86.4 U/ml) secretors. A significant (*p*=0.0216) distribution of CPT-11 concentration was found considering the two groups: patients with sHLA-G High were associated with lower levels of plasmatic CPT-11 AUC (median 16.3 µmol*h), and patients with sHLA-G Low with higher CPT-11 AUC (median 25.5 µmol*h). Similarly, a significant inverse correlation was found between sHLA-G and Biliary Index (*p*=0.0181): patients with sHLA-G High were associated with lower levels of BI (median value 3.6 U/ml) and patients with sHLA-G Low with increased levels of BI (median value 7.6 U/ml). Further, we observed that *Del* allele was prevalent in patients in the sHLA-G High subgroup (*Del*=69%, *Ins*=31%), while

Ins allele was the most abundant in patients in the sHLA-G Low (*Del*=46%, *Ins*=54%); the difference in the distribution of *Del* and *Ins* alleles and related genotypes comparing the subgroups was not statistically significant.

In PART 3, *in silico* interaction between the HLA-G protein and the CPT-11 drug was examined. We observed that the hydrophobic pocket of the HLA-G protein, constituted by the α1 and α2 domains, perfectly hosts the CPT-11 molecule with good complementary of the electrostatic surfaces. Moreover, the possible formation of the HLA-G/CPT-11 complex was investigated by spectrophotometric analysis using the fluorescence resonance energy transfer (FRET) signal between HLA-G and CPT-11 in a titration assay. Preliminary evidences indicate a strong protein-drug interaction with a dissociation constant of the HLA-G/CPT-11 complex in the micromolar range.

Lastly, our study suggests a very attractive prognostic and predictive role for the HLA-G molecule in mCRC and the evaluation of its levels could improve the management of patients treated with first-line FOLFIRI regimen. The unexplored role of irinotecan as immunotherapeutic drug that potentially inhibits HLA-G molecule was for the first time investigated and needs further studies. In conclusion, these findings show novel applications for HLA-G molecule as predictive and prognostic immunobiomarker that could have implications also from a therapeutic point of view, suggesting the development of a biological drug based on HLA-G structure, that could open new scenarios for the future of the immunotherapy in the fight against cancer.

1. RATIONALE

Immunotherapy was identified as Clinical Cancer Advances of the 2016 in the American Society of Clinical Oncology (ASCO) Annual Meeting [2] and still represent a breakthrough in the international debate focused on fight cancer [3]. Its relevance it was demonstrated on patients with several type of cancer through the clinical effects of the novel immunotherapy drugs commercialized in these recent years. The based-on strategy is to reinforce the immune system targeting immune checkpoint inhibitors such as cytotoxic T-lymphocyte antigen 4 (CTLA-4), programmed death ligand 1 (PD-L1) and programmed death ligand 1 (PD-1) molecules acting as "brakes" on the immune system. CTLA-4 interacts with CD80/CD86 avoiding the crucial CD28/CD80 and CD28/CD86 based T cell activation, while the PD-1/PD-L1 pathway altered the T-cell metabolism and differentiation leading to reduced T effector cells and increased T regulatory (Treg) and T exhausted cells.

The first evidence of these drugs effects was published in 1996 by Allison et al., and showed the tumor elimination in mice treated with antibodies against CTLA-4 [4]. Although the initial resistance of pharmaceutical companies toward the novel immunotherapies, in 2011 the U.S. Food and Drug Administration (FDA) approved ipilimumab, the first human anti-CTLA-4 antibody for the treatment of patients with metastatic melanoma in which demonstrated an increase of survival of 3.6 months compared with glycoprotein 100 peptide (gp100) vaccine [5]. In these years, novel immunotherapy molecules have been produced and entered the market such as other immune checkpoint modulators blocking the PD-1/PDL-1 pathway (i.e. pembrolizumab and nivolumab targeting the PD-1, avelumab, atezolizumab, and durvalumab acting against the PD-L1) [6], adoptive cell therapy [7] (i.e. chimeric antigen receptor T-cell (CAR T-cell) [8], [9], T-cell receptor (TCR) [10], and tumor-infiltrating lymphocytes (TILs) [11] therapy), therapeutic antibodies, cancer vaccines [12], and immune system modulators.

Recently, also the human leukocyte antigen-G (HLA-G) was considered as an immunocheckpoint molecule due to its primary tolerogenic role, capable of stopping the immune response. All the different HLA-G isoforms generated by the alternative splicing of the primary transcript could exert a negative regulation on immune cells conferring both protection, from the rejection of the fetus and the transplant, and exposure to the tumor development, favoring the immune escape from the antitumor immunity [13].

In this landscape, the investigation of the role of HLA-G could reveal an interesting alternative to fight cancer, discovering novel predictive or prognostic factors in its coding and non-coding regions. In particular, we focused on the impact of *HLA-G* genetics regulation and modulation of the HLA-G expression in tumorigenesis of metastatic colorectal cancer (mCRC). In fact, the mortality rate related to CRC moved to the second place compared to the overall diagnosed cancers. Incidence rate for CRC is also quite high, especially in region with a flourishing socioeconomic development such as Europe, Australia, North America, Eastern Asia, and it is located at the fourth place worldwide [14]. Although the recent treatment regimens using combinations of standard chemotherapies in association with biological drugs prolong the median overall survival of patients with mCRC, a precise and efficacy treatment strategy is still lacking.

The *HLA-G* 3'untraslated region (3'UTR) of the gene has never been explored in mCRC [15] and currently there are no data in the literature regarding the genomic characterization of *HLA-G* polymorphisms and haplotypes in these patients. However, some 3'UTR single nucleotide polymorphisms (SNPs) of the *HLA-G* gene were found by our group to be independently associated with prognosis in non-metastatic stage II-III CRC patients [16]. Several genetic variations have been described in the regulatory *HLA-G* 3'UTR as opposed to the lower variability in the coding regions of classical HLA class I loci. These variants have been described to share a functional impact because could modulate the expression level of the gene, also altering the binding sites of transcription factors or miRNAs [17]–[19]. An interesting aspect to investigate is the genotypic-phenotypic correlation, between germline polymorphisms/haplotypes and plasma levels of

soluble HLA-G in mCRC. Although tissue and/or soluble HLA-G were already observed upregulated in patients with several cancers compared to healthy donors [20]–[24], further investigation about its therapeutic role is needed. All the soluble isoforms, generated through the alternative splicing mechanism, can be detected in body fluids such as plasma, serum, ascites, cerebrospinal fluids exudates, both in physiological conditions and in cancer where their expression can be induced. Monitoring the sHLA-G levels may have an important prognostic/predictive value to early recognized pathological conditions and improve personalization of therapy.

Moreover, the severe toxicities derived from chemotherapy regimens containing cytotoxic drugs such as irinotecan and oxaliplatin, really impact the quality of life of the patients with mCRC and their clinical management. However, the immune system could contribute to modify the toxicity responses of these patients. In light of the immune inhibitory nature of HLA-G, we previously investigated the impact of genetic markers at the *HLA-G* 3'UTR polymorphisms and related haplotypes, on the risk of developing grades 3 and 4 toxicity in patients with non-metastatic CRC treated with an oxaliplatin-based regimen, the FOLFOX4 [25], finding novel predictive markers.

To our knowledge, the *HLA-G* genetic pattern has never been studied in relation to the develop of severe toxicities in patients with mCRC treated with irinotecan-based regimen such as the FOLFIRI. Genetic factors of the MHC have been found to represent a very strong risk factor for idiosyncratic drug reactions (IDRs), such as the HLA-B*57:01 allele in patients treated with abacavir [26], [27]. The molecular mechanism between this specific HLA allele and the hypersensitivity of reverse-transcriptase inhibitor abacavir was investigated with docking simulations analysis and also the crystal structure of this complex is available [28], [29].

A better evaluation of a possible interaction between HLA-G and irinotecan should give novel insight into the mechanism related to these toxicities that could help clinician to better manage the patients.

All these evidences have suggested to us interesting future perspectives in our research focused on several different aspect of HLA-G from the genetics to the protein-drug interaction passing though the evaluation of the modulation of its expression level.

2. AIMS

The primary aim of this project was to explore the role of HLA-G as immune checkpoint molecule in order to suggest novel predictive and prognostic genetic markers related to mCRC treated with FOLFIRI regimen. A prognostic marker is associated with clinical outcome related with time, such as overall survival (OS) or time to progression (TTP), and could be relevant in the clinical management or treatment-choice of patients. On the contrary, a predictive marker is used as indicator of the probable effect to a specific treatment. Moreover, we suggest future strategy to better manage the patients affected by mCRC, to reduce or prevent the toxicity related to the available treatment, based on *HLA-G* genotyping.
In particular, in the retrospective study of the PART1, the presence of new predictive and prognostic genetic markers related to mCRC homogenously treated according to FOLFIRI regimen, was evaluated. PART1 was designed to evaluate the association between genetic polymorphisms and haplotypes in non-coding 3'UTR of *HLA-G* and clinical outcomes such as survival, for its prognostic role, the tumor response, and the first cycle and cumulative toxicity for their predictive effects. Moreover, the possible correlations with clinical outcome and a particular *HLA-G* 3'UTR haplotype reported to be associated with different expression levels of the protein, were studied through the design of two "HLA-G secretor models". Then, the influence of *UGT1A1*28* (rs8175347) polymorphism of each patient, in relation with the HLA-G secretor models, was also investigated for its well-known role in patients with mCRC regarding the response and toxicity outcomes.

In the PART 2, one of the aims was to evaluate the expression of the sHLA-G protein present in plasma of patients affected by mCRC at the baseline level, then treated with FOLFIRI regimen. The characterization of the genetic polymorphisms and haplotypes in non-coding 3'UTR and coding regions (exon 2, 3, and 4) of *HLA-G* was performed. The goal of this PART2 was to investigate the possible correlations between the sHLA-G plasma level and the clinical characteristics of the patients; the germline *HLA-G* polymorphisms found. Finally, an analysis on the relation between irinotecan pharmacokinetic parameters and sHLA-G levels in a subgroup of these patients was realized.

In the last PART 3, an HLA-G structural and *in silico* analysis was conducted to better hypothesize a possible interaction with the irinotecan drug, present in the FOLFIRI regimen, mainly involved in the development of the severe drug-related toxicity. This hypothesis was verified through UV-Vis and fluorescence spectroscopy experiments measuring the energy transfer phenomenon between the two species in a titration assay.

3.INTRODUCTION

Immunogenetics is a field of study that aims to investigate the genetics basis of the immune response. The influence of genetic variations in genes involved in the development of cancer can be very important and is described in numerous types of hereditary cancer [30], [31]. At the germline level, the presence of polymorphisms (minor allele frequency (MAF) ≥1%) or rare genetic variants (MAF <1%) or very rare genetic variants (MAF <0.1%), in immune-related genes, should be considered because the interactions between the immune system and cancer can significantly influence the clinical outcome with respect to predisposition, nature, prognosis and response to treatment in each individual [32]. A polymorphism is defined as a genetic variation in human genome present in at least 1% of the population studied. Due to its allele variant frequency, it is not so deleterious to compromise the life of the bearing individual, therefore was not erased by the evolutionary pressure. The most common form of polymorphism, about 90% of human genome variations, is the single nucleotide polymorphism (SNP), change in a single nucleotide in the genomic DNA sequence that can affects gene function. Other types of polymorphisms are the sequence repeats, variation of number of tandem repeats (VNTR), insertion or deletions (INDELs) of DNA sequences of different length. Polymorphisms may influence the immune related functions at several levels from gene expression/regulation to post-translation modifications: some might act at the epigenetic level changing DNA methylation and then chromatin structure; others may disrupt transcription factor binding site (TFBSs), influencing the expression, or modify the mRNAs splicing patterns (deletion or creation of a new splice site, modification of exon-splicing enhancers and silencers); they could alter the binding site of microRNA (miRNAs), important elements in gene regulation, or cause mRNA instability. Moreover, some polymorphisms may create premature stop codons; exonic polymorphisms can substitute an aminoacid in protein sequence (missense or non-synonymous variation), changing its structure and function. Further, some polymorphisms may change post-translation modification (PMT) sites and consequently influence post-translation modifications [32]. In the last decades, several efforts were done to better characterize the human genome. The Human Genome Project concluded in 2003, contributed to the discovery of genetic polymorphisms and also became a stimulus for other large-scale projects like International HapMap Project and “1000 Genomes Project”. Approximately 25-35% of the estimated nine to ten million SNPs were discovered with these projects. The genotype could guide the prescribing in the growing era of pharmacogenetics. Genetic variability could affect genes whose transduced proteins are critical for the drug response, such as drug-metabolizing enzymes or are involved in the immune system modulation. Around 7% of FDA-approved drugs including chemotherapy drugs and immunosuppressants, and 18% of US outpatient prescriptions are affected by “actionable” inherited pharmacogenes [33]. Moreover, a fine genotyping could be useful in clinic to improve the management of the patients, reducing or avoiding ADEs or drug-related toxicities [34]. Pharmacogenomics has been largely applied in recent years to the personalization of adjuvant treatments for CRC [25], [35], [36].

The immune system and related actors are gaining increasing importance in the field of anti-cancer therapy strategy and several novel immunotherapies were produced to strengthen immune system against cancer. From the concept of Hanahan and Weiberg that proposed that one of the emerging hallmarks of cancer was the capability to evade the immune surveillance [37], we passed to Fouad and Aanel that considered after their revision, the immune modulation as the hallmark of cancer concerning the immune system involvement in the cancer development [38]. The cancer immunoediting phenomenon is the process by which: 1) the immune system eliminates the developing cancers; 2) cancer cells, under immune selection pressure, enter in quiescence or in a slow-cycling state where harbor different mutations; and 3) some cancer clones escape immune surveillance through the emerged new variants that increased resistance to immune pressure. Different mechanisms were used to evade the immune surveillance, like the alteration in

the tumor-antigen recognition (immunoediting), the downregulation of the local immune system, the resistance to cell death, the induction of immunological ignorance and tolerance through immunosuppressive factors secreted by tumor cells, and host immunodeficiency [32]. Expression of immune checkpoint inhibitors such as PD-L1 and CD80, acts to evade immune recognition. An interesting strategy to oppose the cancer activity, different from blocking the PD-1/PD-L1 pathway, should be represented by the modulation of the HLA-G expression. A focus on the human leukocyte antigen-g (HLA-G) due to its tolerogenic role and to its observed increase expression in pathological conditions such as cancer, is an interesting field of research because of its possible involvement in the immune modulation mechanism during the tumorigenesis process. HLA-G can be considered one of the novel immune checkpoint molecules because of its crucial role in the maintenance of self-tolerance and protection from excessive immune activation [39]. In particular, its role in the mCRC disease treated with irinotecan-based regimens has been investigated also because of the preliminary data obtained in patients with non-metastatic CRC [16]. Then, the HLA-G functions were deeper analyzed in this thesis.

3.1. The clinical matter: metastatic colorectal cancer

Colorectal cancer (CRC) is the third most commonly diagnosed cancer and moved to the second place as most common cause of cancer death worldwide from the third of previous years epidemiologic data [14]. Even if in the Europe (EUR), CRC mortality has been declining since 1993 in men and over the whole period in women, it accounts for the second highest number of deaths for 2018 with 177400 deaths [40]. The highest incidence are found in different geographical areas of Europe (Hungary and Norway ranking first among males and females), Australia, Northern America, and Eastern Asia, especially in region with a flourishing socioeconomic development [14]. Migrant studies showed that CRC mortality tends to change faster than for other cancers following lifestyle and diet modification [41]. From a molecular point of view, the development of CRC starts with the accumulation of genetic and epigenetic mutations in epithelial cells from gastrointestinal trait, hyperproliferation of mucosa and subsequent formation of adenoma that evolves into carcinoma and metastasis in about 10 years [42]. CRCs has very heterogeneous causes of disease. Around 70-80% of CRCs are sporadic, due to somatic mutations, while 15-30% are familial without a Mendelian inheritance rule. About 5% are considered with an hereditary/genetic component or linked to specific genetic syndromes such as hereditary non-polyposis colorectal cancer (HNPCC) and adenomatous polyposis syndrome [43]. Several mechanisms are known to be involved in the CRC development: loss of genomic integrity, chromosomal instability (CIN), microsatellite instability (MSI), aberrant DNA methylation and DNA repair [44]. Moreover, signaling pathways usually altered in CRC onset are: the Wnt/APC/β-catenin, contributing to tumor cell proliferation and inhibition of differentiation mainly in sporadic CRC; and the phosphoinositide 3-kinase (PI3K)/AKT/glycogen synthase kinase-3β (GSK-3β), which activates cell growth and inhibits apoptosis in some sporadic and hereditary CRC, transforming growth factor-β (TGF-β)/Smad, NF-kB or mismatch repair genes (MMR) [45]–[47]. Nowadays, the consensus molecular subtype (CMS) classification in four different classes, CMS1 (MSI immune), CMS2 (canonical), CMS3 (metabolic) and CMS4 (mesenchymal), was proposed as gold standard for the CRC classification [48].
The clinical model analyzed in this thesis is represented by metastatic colorectal cancer (mCRC). About 25% of patients with CRC are diagnosed with advanced disease and a further 35% will develop metastases during the course of the disease [43]. Over the past 15 years, the introduction of the new therapeutic strategies determined a progressive improvement of OS. A patient with mCRC typically achieve a OS in a window of 6 to almost 30 months, as a results of the systemic strategy chosen for a “continuum of care” therapy; the median overall survival (mOS) reached the unprecedented value of over 40 months in molecularly selected patients with mCRC, improving the overall treatment plan [43]. Clinical aspects of the CRC depend on the location of the first-tumor site lesion: right and left colon (lesions with distal extension >15 cm from the anal margin), cause hematochezia even if in some case the occult bleeding determines anemia and fatigue, while rectal involvement (lesions with distal extension <15 cm from the anal margin) cause hematochezia, bleeding and tenesmus. Metastases localized mainly in liver but also in lung and are present in almost 20% of patients with CRC at the date of diagnosis. The presence of distant metastases (M stage) is one of the criteria used to the CRCs classification according with local invasion depth (T stage), and lymph node involvement (N stage) [49], [50]. The management of patients with CRC is very complex; therapeutic strategy is generally based on the tumor stage established at the time of diagnosis. Despite the high incidence of this neoplasia, the majority of patients newly diagnosed with CRC have a localized early stage disease that is suitable for curative surgical resection. A suspicion diagnosis of mCRC should always confirmed by adequate radiological imaging and biopsy to evaluate the histology of the primary tumor or metastases [51]. The aims of therapy for mCRC could be: cure (in very selected cases), prolongation of life, palliation of symptoms, improvement of quality of life, delay disease progression and cytoreduction (tumor shrinkage).

3.1.1. Treatment of mCRC

The mCRC treatment strategy lay on the traditional chemotherapy mainly based on FOLFIRI or FOLFOX regimen in association with targeted therapies; the choice of the optimal therapeutic strategy is based on clinical and molecular aspects of the disease, and patient-related characteristics [51]. Therefore, to choose the therapy a clinician should consider as potential prognostic and predictive factors for chemotherapy: 1) clinical aspects such as tumor burden, sidedness of the primary tumor (right or left colon), and the presence of symptomatic or asymptomatic disease [52], [53]; 2) molecular features like the *KRAS*, *NRAS*, and *BRAF* mutational status [54]; 3) performance status of the patient and the presence of comorbidities [55]. A patient can be classified in "fit" or "unfit" and the clinician could decide the treatment to assign more intensive chemotherapy combinations (3-drug cytotoxic regimens such as FOLFOXIRI) with biological drugs [56].
Recently, MMR genes seem relevant to be evaluated because of their role in the mCRC onset [57], [58]. However, in the clinical practice their use is not well defined.
From what we said, at least three different clinical scenarios are possible: patients with "potentially resectable disease" who are treated with upfront cytoreductive chemotherapy that may have a role of conversion therapy; symptomatic patients with compromised quality of life that are treated with palliative first-line therapy to reduce the tumor burden; asymptomatic patients whose disease is "never resectable" and who are treated with sequential drugs with particular attention to the drug toxicities (continuum of care) [43]. Several drugs are actively used for the treatment of mCRC, such as fluoropyrimidines, irinotecan, oxaliplatin, EGFR-inhibitors (cetuximab, panitumumab), antiangiogenic drugs (bevacizumab, aflibercept, ramucirumab (not yet reimbursed in Italy)), regorafenib, trifluridine/tipiracil (TAS-102).
The typical first-line chemotherapy backbone comprises a fluoropyrimidine (intravenous (i.v.) 5-fluorouracil (5-FU) or oral capecitabine (i.e., Xeloda)) used in various combinations and schedules with irinotecan (FOLFIRI) or oxaliplatin (FOLFOX or XELOX) [59]. The combination chemotherapy provides better outcomes in term of tumor response, progression-free survival (PFS) and OS compared to monotherapy. Subsequently, targeted therapy was added to these combinations: for RAS or BRAF mutated, bevacizumab plus FOLFIRI [60], [61], XELOX or FOLFOX-4 [62], showed an improvement in the median duration of survival, and aflibercept in association to FOLFIRI increased the survival benefit [63]; while for *RAS* and *BRAF* wild-type patients, cetuximab was associated among the others with FOLFOX-4 [64], FOLFIRI [65] to achieve a significant increase in overall response rate (ORR); panitumumab can be combined with a first-line chemotherapy such as FOLFIRI in *KRAS* wild type patients [66] to improve patient outcome in patients with some contraindication (reduced organ function, poor performance status, cardiovascular insufficiency). The mutational status of *RAS* and *BRAF* must be considered as predictive factors of cetuximab and panitumumab efficacy, while to date, no specific predictive factors can help in the choice of the most adequate antiangiogenic drugs. The use of bevacizumab (monoclonal antibody against VEGFA) with FOLFOXIRI (folinic acid, 5-FU, oxaliplatin and irinotecan) in patients younger than 75 years, significantly increased response rate (RR), PFS and OS in comparison to FOLFIRI plus bevacizumab [67], [68], but also to FOLFOX6 plus bevacizumab [69], [70]. Generally, in selected patients with asymptomatic primary unresectable metastases, the FOLFOXIRI regiment is used instead of the classical doublet treatments such as FOLFOX, FOLFIRI, XELOX. The maintenance treatment with bevacizumab plus fluoropyrimidines until disease progression could also be a possibility.
The second-line treatments are strategies that should be offered to a patient with good performance status, adequate organ function, and depend on the first-line choice because if an initial chemotherapy has failed, the strategy backbone should be changed. The available treatments in second-line are: bevacizumab, in patients who have not been treated with bevacizumab as first-line chemotherapy or are on progression

after a pretreatment, aflibercept, in combination with FOLFIRI in patients who progressed after a previous oxaliplatin-based therapy, ramucirumab in combination with FOLFIRI in patients previously treated with oxaliplatin and bevacizumab [43]. Some studies evaluate the role of cetuximab and panitumumab in combination with a second-line irinotecan-based chemotherapy demonstrated a significant benefit in terms of RR and PFS, but not in OS [71], [72].
Considering third and subsequent treatment lines, an increasing number of options are available. In particular, in *RAS* and *BRAF* wild-type pre-treated patients with mCRC, cetuximab (alone or with irinotecan) and panitumumab showed prolongation of OS compared with best supportive care (BSC) [73]. Regorafenib, an oral multikinase inhibitor, and TAS-102, an oral combination of the nucleoside analogue trifluridine and a thymidine phosphorylase inhibitor, showed a similar benefit of prolongation of OS in heavily pre-treated patients with mCRC compared with BSC alone [43]. In metastatic disease treatment with all three classical available antiblastic drugs (5-FU/capecitabine, oxaliplatin and irinotecan) used in subsequent lines of therapy, it was showed doubled survival reaching an OS of 20-22 months, compared to chemotherapy with only fluoropyrimidine [74].
The introduction to these therapies of new anti-angiogenic biological (bevacizumab, aflibercept, ramucirumab) and anti-EGFR (cetuximab, panitumumab) drugs, has resulted in a further significant improvement in the clinical efficacy of treatments with a survival that may reach 30 months [75].

3.1.2. FOLFIRI regimen and its related toxicities

Cytotoxic drugs are metabolized through complex mechanisms in which different enzymatic activities are involved. Genetic alterations of the enzymes responsible for the cytotoxic agent's metabolism may determine their reduced functionality resulting in the accumulation of the drug and its active metabolites with possible increase in toxicity. ADEs impact the quality of life of patients leading to dose reductions, delays, and discontinuation of treatment. Among all the mCRC treatments, we focused on the description of FOLFIRI-related toxicities because is the regiment used in the patients with mCRC in our study cohort. FOLFIRI is a combination of three different drugs: irinotecan, 5-FU and folinic acid. Leucovorin, a reduced folate, stabilizes the binding of 5-FU to thymidylate synthase and enhances the inhibition of DNA synthesis. Several studies have shown that genetic variants of enzymes involved in the metabolism of fluoropyrimidines and of irinotecan, are associated with a higher risk of toxicity [76]. The most consistent data in the literature concern dihydropyrimidine dehydrogenase (DPD), the main enzyme of fluoropyrimidine metabolism, and uridine-glucuronosyl-transferase (UGT) involved in the hepatic metabolism of SN-38, an active metabolite of irinotecan [36].

Irinotecan hydrochloride

Irinotecan hydrochloride (CPT-11) is a semisynthetic derivative of the camptothecin that act as antineoplastic enzyme inhibitor of the DNA topoisomerase I (Figure 1). CPT-11 has an molecular weight of 596.678 Da. It prevents re-ligation of the DNA strand by binding to topoisomerase I-DNA complex and interfering with the moving replication fork; it induces replication arrest and determines double-strand breaks in DNA that are not efficiently repaired and causes cell death.

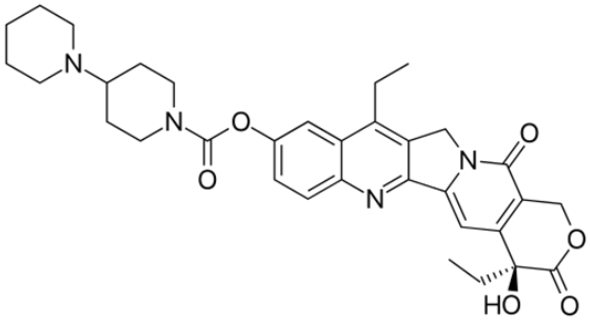

Figure 1. Lewis structure of irinotecan (CPT-11) molecule.

At the end of '90s, the introduction of CPT-11 into the fluoropyrimidine-based regimens for advanced or metastatic CRC was welcomed as an important step forward a better treatment, because it brought to an overall response rate greater than 15%, and it was superior to infusional 5-FU in terms of OS. However, the chances of obtaining significant therapeutic results were partially reduced by the occurrence of severe toxicities, such as diarrhea and neutropenia, that limit the CPT-11 clinical use [77]. Irinotecan-induced diarrhea is categorized as early or late ADE based on the occurrence (within the first 24 hours or more than 24 hours after administration). Other typical chemotherapy toxicities may affect the GI tract due to the high growth rate of epithelial cells that makes them vulnerable to the toxic effects causing anorexia, nausea and vomiting. Analyses of genetic polymorphisms seemed to offer a more reliable and safer approach for the identification of patients at risk to develop severe drug-related toxicities, than pharmacokinetics to optimize CPT-11 dose and schedule. In this view, *UGT1A1* gene was associated with significant changes in disposition of CPT-11 and its metabolites, and consequently with treatment-induced toxicities [77]. In particular, the germline allelic variant *UGT1A1*28* is associated with reduced irinotecan metabolism. Several studies have revealed that patients with this variant are more likely to experience hematological toxicity and diarrhea [36]. The risk of toxicity is dose-dependent and generally not observed at doses <150 mg/m^2 [78], [79]. However, a dose reduction is recommended in subjects homozygous for the *UGT1A1*28* variant. Finally, pharmacogenetic studies have also demonstrated that carriers carrying the wild type *UGT1A1*1* could tolerate higher doses of CPT-11 with better therapeutic outcomes [80]. After i.v. administration, CPT-11 undergoes a two-step metabolic process which first led to the activation of the drug with the production of its active metabolite SN-38 by plasma carboxylesterase isoforms 1 (hCE1) and 2 (hCE2). The hCE1 and hCE2 isoforms remove the bispiperidine moiety from the molecule and subsequently the uridine diphosphate glucuronosil transferase isoform 1A1 (UGT1A1) transform the SN-38 in its inactive glucoronide metabolite. Irinotecan is characterized by a linear pharmacokinetics at doses ranging from 100 up to 450 mg/m^2 or up to 750 mg/m^2 when area under the time concentration/curve (AUC) or maximal plasma concentration (Cmax) parameters are considered after a 1-hr infusion [77]. The CPT-11/SN-38 plasma concentration ratio is approximately 10:1, but it is worth noting that CPT-11 has a negligible cytotoxic activity when compared with the main metabolite SN-38 (100-1000 times more potent than CPT-11). The conjugated SN-38 glucuronide (SN-38G) is mainly excreted through the bile, while renal clearance accounts only to 30% of the total SN-38 clearance. Within the intestinal lumen, CPT-11 and SN-38 are released from glucuronated species by beta-glucuronidase produced by the flora of the gut, and the active metabolite may be further absorbed, or it can exert a cytotoxic effect against intestinal epithelium with the following occurrence of diarrhea [81].

5-Fluorouracil

5-Fluorouracil (5-FU) is a pyrimidine antimetabolite that inhibits the thymidylate synthase enzyme, which is involved in pyrimidine nucleotide synthesis (Figure 2). Fluorouridine triphosphate is incorporated and substituted for uracil during the RNA synthesis causing the broadly cytotoxicity attacking both DNA and RNA synthesis [82].

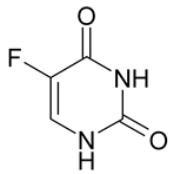

Figure 2. Lewis structure of 5-FU molecule.

The principal side effects 5-FU-related depend on the method and timing of administration. When 5-FU i.v. administration is performed, according to a bolus schedule of five consecutive days' treatment every four to five weeks, neutropenia and stomatitis are the most common toxicities reported. Instead, diarrhea is the more frequent ADE recorded when 5-FU is given according to a weekly i.v. bolus schedule [83]. In regimens where 5-FU is continuously i.v. somministrated, palmar-plantar erythrodysesthesia or hand-foot syndrome are the more common ADEs than hematologic and gastrointestinal (GI) [84]. The DPD is the enzyme, codified by the *DPYD* gene, acting as rate-limiting step of the degradation of pyrimidine base and acts breakdowning the 5-FU. Polymorphisms in this gene could affect the metabolism of drugs and determine the accumulation of toxic metabolites. In particular, it was demonstrated that some polymorphisms such as *DPYD*-rs3918290, *DPYD*-rs55886062 and *DPYD*-rs67376798, were associated with grade ≥3 toxicity [85]. Some common ADEs reported as low-grade toxicities, are ischemia, fatigue, cytopenia and GI side effects such as severe nausea and vomiting, mild loose stools and refractory diarrhea, stomatitis or mucositis [86]. Mucosal damage is observed in many regions of the GI tract and in the oral tract can cause difficulty to chew and swallow, interfering with nutrition. Cardiovascular toxicities are most often associated with 5-FU/Xeloda regimens, causing arrhythmias, myocardial infarction, and sudden cardiac death, but usually are reversible.

3.1.3. Biomarkers for mCRC

In the European Society for Medical Oncology (ESMO) clinical practice guidelines 2017 for mCRC was reported the available biomarkers for the management of mCRC: tumor *KRAS* and *BRAF* mutational status, microsatellite instability (MSI) status, *DPYD* and UDP glucuronosyltransferase 1 family, polypeptide A1 (*UGT1A1*) germline variation status, excision repair cross-complementation group 1 (ERCC1) expression or presence of ERCC1-rs11615 polymorphism, thymidylate synthase (TS) activity [87]. Patients with mCRC with activating mutations in *KRAS* do not receive a benefit from EGFR monoclonal antibody therapy and have a detrimental effect when combined with an oxaliplatin-based cytotoxic backbone (negative predictive marker). A percentage of 8-12% of patients with mCRC are *BRAF*-mutated in the corresponding tumor, and two-thirds of these patient's primary lesions were located on the right colon and associated with an increased incidence of peritoneal and distant lymphnode metastases (negative prognostic marker) [88]. Data about the predictive and prognostic value of MSI status in mCRC are scarce, but some suggest that MSI tumors tend to have lower disease control rates when treated with oxaliplatin-based first-line therapy [89]. DPD metabolic activity is a predictive biomarker of potential toxicity of treatment based on 5-FU and

Capecitabine because deficiencies in DPD, mainly due to *DPYD*2A* polymorphism, cause an alteration in the catabolic process used to eliminate them [90]. Genetic variations within the *UGT1A1* gene have also been associated with drug toxicities; the US FDA recommends a reduction of the starting dose of irinotecan in patients with the *UGT1A1*28/*28* genotype. Moreover, high ERCC1 levels seems to confer oxaliplatin resistance, while ERCC1-rs11615 polymorphism has been associated with a negative predictive marker for oxaliplatin-based therapy in patients with mCRC [91]. The TS activity and CRC sensitivity to 5-FU seem to correlate with *TSER* polymorphisms but should be confirmed in a larger randomized study [92]. The ESMO Guidelines Committee recommended the *DPYD* and *UGT1A1* genetic test before the 5-FU administration and irinotecan treatments, respectively, but this procedure is not yet entered completely in routine clinical practice [51]. Others emerging biomarkers such as *HER2*, *MET*, *KRAS* gene amplification, transforming growth factor-α (TGF-α), amphiregulin and epiregulin expression, *EGFR* mutations and alterations in *HER3*, *PI3KCA* and *PTEN* genes, impact on the response to all classes of target agents, especially EGFR-antibody based-therapies. Primary de novo, and also acquired resistance to anti-EGFR treatment, was driven by mutations in *KRAS*, *NRAS* and *BRAF* and amplification of *HER2* and *MET* [93]. CpG island methylator phenotype (CIMP), an epigenetic alteration, is associated with a worse prognosis in patients with CRC [94]. Further, the inactivation of the *SRBC* gene though the DNA hypermethylation, predicts a shorter PFS, especially in oxaliplatin-treated patients with mCRC for whom metastasectomy was not indicated.

Evidences that also the HLA-G polymorphisms, haplotypes and expression levels could become very interesting predictive and prognostic biomarkers for CRC are reported in the literature and are deeper introduced in the 3.2.5 section [95], [96].

3.2. A focus on HLA-G

3.2.1. HLA-G structure

The HLA-G is a non-classical major histocompatibility complex (MHC) class I molecule. *HLA-G* gene is composed of eight exons and seven introns with a stop codon at exon 6, a quite large 5'upstream regulatory region (5'URR) extending at least 1.4 kb from ATG, and a 3'UTR [97]. The coding exons transduce only the heavy chain of the molecule and are located on chromosome 6, while β2-microglobulin (β2M) is encoded by a separated gene on chromosome 15 [15]. Contrary to classical (class Ia) molecules, the nonclassical HLA-E, -F, -G (class Ib) molecules are characterized by a limited allelic variation, in particular, only 61 alleles of *HLA-G* are described (see http://hla.alleles.org/data/hla-g.html, Release 3.33.0, 11 July 2018). The *HLA-G*01:01* protein allele is the most prevalent in different European populations [98]. A discrepancy in numbering the exons of the *HLA-G* gene is present between Ensembl database, where the *HLA-G* gene have a supplementary exon at the 5'-end, and IMGT/HLA database, where this supplementary exon is absent [99]. We decided to use the current IMGT/HLA nomenclature for the mRNA transcript and protein sequence description.

HLA-G mRNA transcripts may produce seven different isoforms of HLA-G molecule by alternative splicing, four membrane bound (HLA-G1 to -G4), three soluble (HLA-G5 to -G7) generated by the retention of a stop codon after exon 4, all capable of exerting a negative regulation on immune cells [100]. Full-length HLA-G1 is expressed as a cell surface glycoprotein, cross-linked to the membrane with a disulfide-bonded due to the Cys^{42} [101]. Through a proteolytic shedding of the major membrane-bound isoform by metalloproteinases (MMPs), a shed HLA-G1 (sHLA-G1) can be generated. The HLA-G protein sequence has a length of 338 aminoacids (AAs) and a molecular weight of 38.23 kDa (www.bioinformatics.org). The 1-24 AAs codifies for the peptide signal (red, in exon 1), that is lost in the mature protein (full length 314 AAs): according to the official nomenclature of the HLA committee, the first residue is Gly^{25}. Therefore, 1-90 AAs generates the α1 domain (orange, exon 2), 91-182 AAs the α2 domain (yellow, exon 3) and the 183-274 AAs the α3 domain (green, exon 4). Furthermore, an Ig-like type C1 domain is generated by the 185-275 AAs, the 275-284 AAs codify for the connecting peptide (light blue, in exon 5), and the 285-314 AAs for the transmembrane and cytoplasmic domains (black in exon 5 and blue in exon 6, respectively).

All the HLA-G isoforms contains the α1 domain (see Figure 3). The HLA-G1 molecule has a transmembrane heavy chain and shares with the HLA-G5 molecule (its soluble counterpart) the extracellular chain that consists of the α1, α2, and α3 globular domains non-covalently associated with β2M (Figure 4) and encoded by the exons 2, 3 and 4, respectively [102], [103]. The HLA-G2 and HLA-G4 isoforms are composed only by the α1 and α3 domains, the HLA-G4 is lacking of α3 domain, while the HLA-G3 and HLA-G7 are constituted of only the α1 domain linked to two aminoacid residues encoded by intron 2 [104]. HLA-G5 and HLA-G6 retain the intron 4 that encodes the C-terminal of these soluble isoforms: translation is prematurely interrupted by generation of stop codon that prevents the synthesis of the hydrophobic transmembrane domain. The peptide presentation usually depends on the differential interaction of MHC complex with tapasin (TPN) or the transporter associated with antigen processing (TAP) which facilitates efficient loading of short peptides (8-12 amino acids). However, the mechanism used by HLA-G is a TAP-independent and a tissue-specific peptide loading that could be controlled though several cytokines is observed [101]. HLA-G associates with the peptide-loading complex in the endoplasmic reticulum but also binds some signal sequence-derived peptides in a transporter associated with antigen processing (TAP)-independent fashion [101]. HLA-G binds peptides of 9 amino acid residues in length from intracellular proteins with the α1 and α2 domains which defines the peptide-binding cleft at a deeper position compared to other HLA class I

proteins. This feature suggests that the HLA-G2/6 and HLA-G3/7, lacking the α2 domain, are not able to present antigens justifying the peculiarity of the non-classical HLA-G functions.

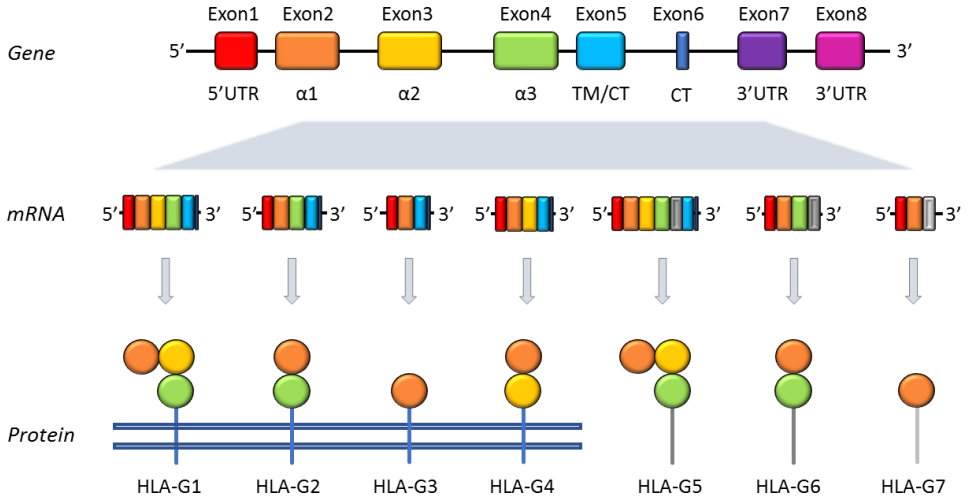

Figure 3. *HLA-G* gene and HLA-G isoforms generated by alternative splicing representation.

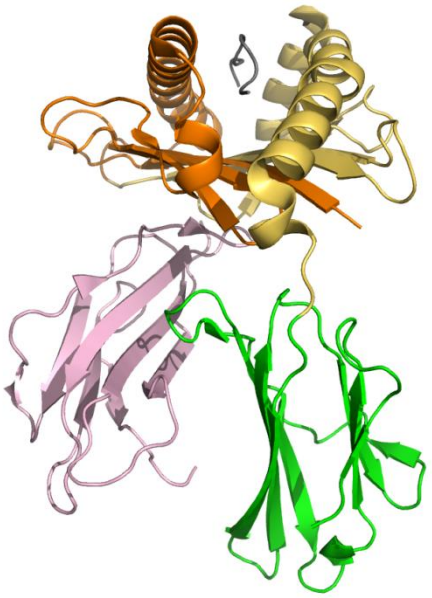

Figure 4. Cartoon representation of the crystal structure of the extracellular HLA-G complex (PDB ID: 1YDP). α1 (orange colour), α2 (yellow colour), and α3 (green colour) globular domains non-covalently associated with β2-microglobulin (pink colour) and antigen peptide (grey colour).

3.2.2. HLA-G functions and interactions

Immune checkpoint molecules as HLA-G are inhibitory agents that physiologically counterbalance the co-stimulatory signals to fine regulate the immune responses. The principal and known functions of HLA-G are related to immune-system tolerance and inhibition and depend on the HLA-G expression [105]. HLA-G exerts three main immunoinhibitory activities: 1) a direct function through blocking effector cells; 2) an indirect function through regulatory cells; 3) other functions with immunoinhibitory consequences [100].

HLA-G directly interacts with the inhibitory receptors leukocyte Ig-like transcript 1 (LIRB1) known also as CD85j/ILT2 (ILT2) and transcript-4 (LIRB2), i.e. CD85d/ILT4 (ILT4), killer Ig-like receptor (KIR) 2 Ig domain and long cytoplasmic tail 4 (KIRD2DL4)/CD158d, CD160 and CD8, favoring the escape from the host's immune surveillance. All these inhibitory receptors are widely expressed on immune system cells. In particular, ILT2 is present on all monocytes, B cells, DCs, myeloid derived suppressive cells, NK cells and T cells [106]. ILT4 is expressed by myeloid lineage cells, dendritic cells (DCs), monocytes, neutrophils [104]. KIR2DL4 is predominately expressed on decidual NK cells [107]. CD160 is expressed by subgroups of $CD8^+$, $CD4^+$, Tγ/δ and $CD56^{dim}$ NK cells, and by activated endothelial cells and intestinal intraepithelial cells. CD8 is characteristic of T cytotoxic cells and some NK cells. In particular, ILT2 and ILT4 are constituted by four tandem immunoglobulin-like extracellular domains, a transmembrane region of 23 aminoacids and three immunoreceptor tyrosine-based inhibitory motifs (ITIMs) in their cytoplasmic tails [108]. They interact with the α3 domain of multiple class I molecules but with a higher affinity to HLA-G. The restricted peptide repertoire of HLA-G suggests a potential role for disulfide-bonded dimerization through the Cys^{42} or Cys^{147} residues that could justify the higher affinity of the ILT2 and ILT4 receptors compared with the other MHC class I molecules. The dissociation constant (Kd) of dimers of β2M-associated HLA-G1 and ILT2 were 0.0067 μM (Kd=3.5 μM for monomers) while of dimers of β2M-associated HLA-G1 and ILT4 were 0.75 μM (Kd=15μM for monomers), then ILT4 binds HLA-G with higher affinity than ILT2 [104]. However, the HLA-G2/4 isoforms of which dimers are difficult to detect, interact only with the ILT4, highlighting a different mechanism of action directed on antigen-presenting cells (APC) and not on lymphocytic effectors compared to the HLA-G1/5 isoforms [104]. The interaction with the ILTs receptors seems to depend on the different capability to dimerize and affinity of the HLA-G isoforms considered. Some of the most significant structural differences of HLA-G are clustered around the antigen-binding cleft and the ILT2/4 binding site. The peptide-binding groove of HLA-G is notably hydrophobic. The presence on HLA-G of Phe^{195} and Tyr^{197} residues, contrary to Ser and His on the classical MHC class I molecules, increases the HLA-G loop hydrophobicity promoting the ILT2 binding with Tyr^{38} and Tyr^{76}. This HLA-G/ILTs interaction causes phosphorylation of ITIMs and recruits protein tyrosine phosphatase Src homology 2 (SH2) domain-containing proteins which initiates the ILT2/4-mediated inhibitory signaling cascade interfering with the $CD8^+$ T cells-mediated inhibition leading to apoptosis [101]. Instead, the presence of Met^{76} and Gln^{79} of the HLA-G α1 domain are critical for the KIR2DL4 recognition. All the HLA-G–receptor interactions represent direct mechanisms to block immune system effectors cells and modulate the immune response.

HLA-G has also an indirect role in the suppression of proliferation in $CD4^+$ T cells and in the induction of $CD4^+CD25^{high}FOXP3^+$ Tregs and Type 1 regulatory T (Tr1) cells interfering with the immune regulation. These regulatory cells are involved in the promotion and maintenance of the "long-term" immune tolerance inhibiting the reactivity of other effectors [105]. In particular, HLA-G could induce the $CD4^+$ and $CD8^+$ T cells to differentiate into regulatory T cells capable of inhibiting the reactivity of other T cells and to lose their capability to respond to antigenic stimulation. The induction of a subset of Tregs, the adaptive allo-specific Tr1 cells, depending on dendritic cells characterized by the production of IL-10 (DC-10), is also crucial for the promotion and maintenance of tolerance [109]. Moreover, activated NK and T cells could transitorily behave as regulatory suppressor cells and inhibit immune responses through HLA-G [100].

HLA-G expression could determine other immunoinhibitory functions including the capability of soluble HLA-G (sHLA-G) to induce apoptosis of NK cells and cytotoxic T cell (CTLs), to inhibit the cytolysis of immune cells, to influence the expression and release of INF-γ by NK cells, to induce the upregulation of inhibitory receptors and tumor-promoting agent such as metalloproteinases (MMPs), the antiangiogenic function of sHLA-G through CD160 binding, and the impair of chemotaxis of different immune effector cells [110].

3.2.3. HLA-G expression mechanisms

The physiological expression of HLA-G is rescrited to few tissues such as trophoblasts, thymic epithelium, cornea, mesenchymal stem cells, nail matrix, pancreatic-β cells, erythroid and endothelial precursors [111]. Mainly, HLA-G is expressed at the maternal-fetal interface and involved in the immune tolerance towards the maternal immune system [112], [113]. However, its expression can be induced in pathological conditions including cancers, transplantation, inflammatory and autoimmune diseases, and viral infections [100], [114]–[116]. Several different mechanisms and regulatory elements are involved in the modulation of the HLA-G expression at transcriptional, post-transcriptional and epigenetic level. DNA methylation of CpG and hypoacetylation of H3 and H4 hystone at the *HLA-G* locus are epigenetic mechanisms that avoid the HLA-G expression, but this could be induced by demethylating agents used in some anticancer therapies [117], [118]. The *HLA-G* transcription is controlled by regulatory elements including machinery of classical HLA class I gene, locus control region (LCR) contained in a 250 bp fragment at 1.2 kb from the starting ATG, CAMP responsive element binding protein 1 (CREB1) factor, interferon regulatory factor 1 (IRF-1), heat shock factor 1 (HSF1), progesterone receptor, Ras responsive element binding protein 1 (RREB-1), GLI-3 factor and long interspersed element-1 (LINE-1) element [119]. Moreover, NF-κB activation not alters HLA-G transcription but its expression on cell surface increasing the proteolytic shedding in tumor cells [120]. At the post-transcriptional levels, HLA-G expression is regulated by mechanisms targeting the 5'URR and 3'UTR that affect mRNA stability and by several specific microRNAs [121], [122]. HLA-G is a non-polymorphic molecules with several genetic variations described in the 5'URR and 3'UTR and a lower variability in the coding regions [17]. However, it is characterized by relevant structural variation that could influence the process of dimerization and the expression depending on the tissue involved, modulating also the biological function of HLA-G [123]. The polymorphisms in the 5'URR can modify methylated CpG sites or alter regulatory elements as RREB-1 binding site, influencing transcriptional activity [106]. Several polymorphic sites are also identified in the *HLA-G* 3'UTR, the most studied region of the gene because of the presence of several miRNAs binding-sites and their involvement in modulation of HLA-G expression. The most studied germline polymorphisms present in the *HLA-G* 3'UTR segment are rs371194629 (+2960 14-base pair (bp) insertion/deletion (INDEL)), rs1707 (+3003 T>C), rs1710 (+3010 C>G), rs17179101 (+3027 C>A), rs17179108 (+3035 C>T), rs1063320 (+3142 G>C), rs9380142 (+3187 A>G), rs1610696 (+3196 C>G), and rs1233331 (+3227 G>A>T). In particular, the presence of 14-bp insertion causes a reduction of HLA-G mRNA and a lower level of sHLA-G in the plasma, but also the formation of a more processed, by the removal of a 92-b fragment, and stable transcript due to the presence of an AU-rich element (ARE) within the 14-b fragment [121]. This higher stability seems not to compensate for the lower HLA-G levels associated with the 14-b insertion [119]. Also the rs9380142-*A* allele at position +3187 influences the HLA-G mRNA stability, decreasing mRNA stability and HLA-G expression [124]. Furthermore, the presence of a Guanine instead of a Cytosine at the +3142 position was demonstrated *in silico* to increase the affinity in the microRNA binding-site, especially with mi-148a, mi-148b, and mi-152 [97], even if Manaster and colleagues showed controversial data that this polymorphism had no effect on the miRNA targeting based on the results of luciferase reporter assay into RKO human colon cancer cells [122]. Nevertheless, a

validation analysis for the functional properties of all these polymorphisms is still necessary. However, it has already been showed that most of 3'UTR polymorphisms are associated with different levels of sHLA-G in the plasma in healthy populations [18]. The HLA-G expression could be influenced by different combinations of specific polymorphisms that generate distinct 3'UTR *HLA-G* haplotypes. In particular, one of the most frequent *HLA-G* UTR-1 haplotype, carries the alleles in brackets (14 bp deletion/+3003 T/+3010 G/+3027 C/+3035 C/+3142 C/+3187 G/+3196 C), has been associated with high HLA-G expression [18]. To date, 61 HLA-G alleles have been officially recognized by the International Immunogenetics Database (IMGT, January 2019, database release 3.35.0) and most of these alleles contain synonymous or intronic substitutions. In addition to the *HLA-G* untranslated regions, the *HLA-G* coding region presents more than 100 polymorphisms and, at the same way, might show affinity for transcription factors that could affect the regulation of the gene expression and of the different HLA-G isoforms generated by alternative splicing of the primary transcript. The possible correlations between *HLA-G* genotype and phenotype were first investigated by Rebmann and colleagues [125]. In particular, patients with *HLA-G*01:04* protein allele had significantly higher sHLA-G levels, while those with *HLA-G*01:03* and *HLA-G*01:05N* protein alleles present lower plasmatic and circulating HLA-G level. The presence of stop codons in the coding region could alter the HLA-G expression limiting the available isoforms and producing those truncated. The HLA-G*01:05N null allele presents a Cytosine deletion leading to a TGA stop signal in codon 189, and produces incomplete formation of the HLA-G1, -G4, and -G5 isoforms and normal expression of HLA-G2, -G3, and -G7 [126]. Similarly, the HLA-G*01:13N allele with a T in the first base of codon 54 that determines the formation of a premature TAG stop codon preventing the production of all HLA-G isoforms and therefore it is probably associated with the absent of expression [127]. Currently, no many data are available in this context and in the literature there are maily studies about the allele frequencies in the different populations [128], [129].
Moreover, enviromental factor as cytokines, hormones, galectin-1, indoleamine 2,3 dioxygenase (IDO) and stress condition such as heat shock, hypoxia, cancer, transplantation and pregnancy could also influences the HLA-G expression [106]. Among the cytokines, the immunosuppressive IL-10 can act at the mRNA and protein level of HLA-G, triggering a vicious circle of immune suppression in cancers [130], the interferon (INF)-γ and INF-β change its expression, and the leukemia inhibitory factor (LIF) stimulates the *HLA-G* promoter [131]. Similarly, the expression of indoleamine 2,3 dioxygenase (IDO) modulate and was modulated by HLA-G.

3.2.4. HLA-G in cancer

The clinical relevance of HLA-G expression in solid cancer was investigated in previous studies and was first demonstrated in the context of melanoma where was observed that HLA-G expression protected tumor cell lines from destruction by cytotoxic cells [132]. HLA-G was detected mainly in primary cancer lesions and its expression was associated with malignant trasformation and not observed in healthy surrounding tissues, suggesting a role in the process of cancer development. Its expression in biopsies and high levels of sHLA-G in plasma are characteristic of solid and hematological tumors of high histological grades and advanced clinical stages [110], [133]–[137]. However, whereas the clinical meaning of HLA-G expression in solid tumors has been widely investigated and correlated generally with a poor prognosis, fewer data are available for hematological malignancies. Specifically, HLA-G over-expression was detected: in advanced ovarian cancer tissues trough immunohistochemistry analysis and was correlated to patients with a worse prognosis [138], [139]; in cervical and endometrial carcinoma lesions [24], [140]; in breast cancer and was associated with high stage and tumor grade [141]–[143]; in non-small cell lung cancer lesions [23], [144]; in colorectal cancer tissue and gastric cancer [134], [145]; in the bladder transitional cell carcinoma [146]; in

primary renal cell carcinoma lesions [147]; in hepatocellular carcinoma and was strongly correlated to advanced disease stage and in pancreatic adenocarcinoma [22], [148], [149]; in esophageal squamous cell carcinoma lesions and was associated with poor prognosis [150]–[152]; in glioblastoma [153]; in nasopharingeal and oral cavity squamous cell carcinoma [154]–[156]; in thyroid carcinoma and was associated with tumor recurrence and poor prognosis [157], [158]; and also in leukemia (both myeloid and lymphocytic) [159]–[162].

Most tumors express HLA-G at different stage of their evolution on their cell surface, but they also release HLA-G as soluble or shedded isoforms in the body fluids such as plasma, serum, ascites, cerebrospinal fluids exudates [163], [164]. Several studies analyzed the expression in these fluids, expecially in serum and plasma samples. In particular, among which: in serum was observed an increased HLA-G level in FAB M4 and M5 subtypes of acute myeloid leukemia, in acute lymphoid leukemia [136], in Waldenstrom macroglobulinemia [165], in melanoma [166] and also in patients with NSCLC and particularly in those with advanced disease stages compared to healthy subjects [167]; in plasma was found a significant increase of sHLA-G levels in patients with chronic lymphatic leukaemia [168], in multiple myeloma compared with healthy donors [169], in gastric and colorectal adenocarcinoma [170], in endometrial carcinoma [171], in non-small cell lung cancer [172], in hepatocellular carcinoma [173], in renal cell carcinoma [147], in metastatic compared to localized neuroblastoma [174]; and in the supernatant of peritoneal ascites of patients with ovarian and breast cancer the sHLA-G levels was also higher [133]. Moreover, *in vitro* study analyzed the secreted sHLA-G5 isoform from melanoma cells that inhibited NK cell cytotoxicity [175] and the role of exosomes with HLA-G, released from cancer cells [176]. Recently, the importance of these HLA-G-bearing extracellular vesicles has been re-discovered for the possible HLA-G role in the intracellular communication that needs further investigations [177].

3.2.5. HLA-G as biomarker in cancer with a focus on CRC

It was noted that the presence of *HLA-G* polymorphisms affects *HLA-G* mRNA translation and protein expression and we just reported that HLA-G expression varies in tumoral conditions, therefore a correlation between the *HLA-G* genetic variations and susceptibility and/or clinical outcome of cancers has already been investigated in some cancers to discover novel prognostic or predictive values. For example, the absence of 14-bp in the *HLA-G* 3'UTR (rs371194629-*Del/Del* genotype vs *Ins/Del* or *Ins/Ins* genotypes) showed a shorter OS in patients with chronic lymphocytic leukemia and diffuse large B-cell lymphoma while those with the rs1233334-*C/C* genotype in the 5'URR had a better overall survival [178], [179]. A recent meta-analysis, showed all possible and published associations between 14-bp polymorphism and the clinical situations/diseases and its findings indicate that the 14-bp variability at the HLA-G 3'UTR segment is not an efficient genetic marker for disease association studies [141]. We must consider that these are complex diseases and this variations is not the only one with a post-transcriptional control of the HLA-G expression; a complete analysis of the *HLA-G* 3'UTR or all the gene remains recommended, since the entire region is under a strong linkage disequilibrium [17]. In particular, the first study showing associations between *HLA-G* 3'UTR polymorphisms and patients with CRC, was a case-control study carried out by our group [180]. The rs371194629-*Ins* allele and rs1610696-*G* allele in the recessive model (G/G vs. C/G+C/C) were significantly more frequent in CRC patients and were associated with increased risk of CRC as the same way of the UTR-2 haplotype. Instead, a "protective" role was observed in the presence of the rs1707-*C* and rs1710-*G* alleles, and the effect was confirmed also in the UTR-4 haplotype. Another important clinical parameter in CRC is the evaluation of grade 3 and 4 (G3-4) toxicities. In particular, rs1610696 and rs371194629 and the related UTR-2 haplotype emerged as predictive markers of G3-4 neutropenia and

neurotoxicity, confirming their emerging role in the CRC outcome and management. Moreover, by our group, the role of these SNPs in the *HLA-G* 3'UTR in patients was investigated in non-metastatic stage II-III CRC, after curative resections and FOLFOX-based chemotherapy [16]. Some SNPs were found to be significantly associated with prognosis in multivariate analysis. In particular, the rs371194629 (dominant model) and rs9380142 (recessive model and *G/G* genotype) SNPs, were associated with a reduced 5-year disease free survival (DFS); the rs9380142 was also significantly associated with a reduced OS in adjusted analysis. These results were confirmed in the *HLA-G* UTR-1 haplotype (high secretor), that contains both rs371194629-*Del* and rs9380142-*G* alleles, found to be significantly associated with a worse DFS and OS in multivariate analysis. Intriguingly, both SNPs have been previously associated with a more stable *HLA-G* mRNA and higher levels of soluble HLA-G. On the contrary, the rs17179108 was significantly linked with a better DFS and it was already observed that subjects with rs17179108-*C/T* genotype had a lower levels of soluble HLA-G compared with the wild type patients due to a more unstable mRNA [18].

Also the variations in the coding region are important to consider because they could produce conformational changes in the molecule that could alter the interaction with cell receptors, the isoform production, the modulation of the immune response, the polymerization and the ability to couple peptides [181]. The occurence of the codon 130*C*-deletion in the *HLA-G* coding sequence was significantly associated with a decreased disease-free and OS in patients with nasopharyngeal carcinoma [182]. Other correlations were found between the G*01:04 allele and progression to high-grade bladder tumor and also between G*01:03 allele and protection against transitional cell carcinoma [183]. To the best of our knowledge, no study analyzed the role of the HLA-G coding region in patients with CRC.

The role of the HLA-G expression in cancer was just reassumed in the previous paragraph, here a focus on the results found both in tissue and in body fluids of patients with CRC was reported. A retrospective study demonstrated in 102 eligible colorectal cancer tissue samples that the expression of HLA-G, detected by immunohistochemistry, is associated with a poor survival compared with its no expression, and revealed a significant predictive value of HLA-G for the prognosis of CRC [145]. Similarly, Zeestraten and colleagues reported an increased expression of HLA-G and HLA-E that was highly correlated with a worst prognosis and survival, and the presence of distant metastasis in people with CRC [184]. Another study evaluated the role of HLA-G expression performing flow cytometry as well as immunohistochemical analysis and compared the results obtained for CRC tissue sections and tumor derived cells lines highlighting that different HLA-G epitopes are expressed in CRC tissue and then the used of several monoclonal antibodies should be also considered to a better interpretation of the results [185]. Özdemir and colleagues suggested to target HLA-G and HLA-E because of their increased expression in CRC tumor sections to eliminate CRC cancer stem cells [186]. The expression and predictive relevance of HLA-G expression in 457 patients with primary CRC was recently investigated by immunohistochemical analysis and showed that patients with higher levels of HLA-G had a significantly worse survival than those with lower levels, and suggested that a stratification based on the HLA-G expression level could be considered an instrument to define a better independent prognostic factor for CRC patients [137]. The problem of the use of different techniques, however, was remarked again in the study of Swets and colleagues where with the immunohistochemistry they were able to detected HLA-G contrary to the used of western blot tecnique [187]. Farjadian and colleagues investigated the results obtained by immunohistochemical analysis in primary cancer lesions and by ELISA in plasma of patients with gastrointestinal cancer declaring that only in the plasma sample the levels of sHLA-G were significantly higher in patients than in healthy controls and that a significant correlation was found between higher sHLA-G level and stage I tumors [170]. However, the results obtained should be confirmed with others analysis. Plasma sHLA-G was evaluated for its role as biomarker for diagnosis of colorectal, gastric, non-small cell lung cancer, and esophageal squamous cell carcinoma [188]. Moreover, in serum and cerebra spinal fluid of patients with multiple sclerosis was demonstrated an

association among higher sHLA-G levels and rs1063320-*C/C*, rs371194629-*Del/Del* genotypes combination [189]. A recent publication, focused on 178 patients diagnosed with CRC, highlighted the importance of the plasma sHLA-G levels as prognostic factor [95]. This analysis showed that plasma sHLA-G level could have not only a prognostic value but also a diagnostic value as precocious biomarker in CRC, because of its increased specificity and sensibility with respect to the wide used carcinoembryonic antigen (CEA) protein [190]. High sHLA-G plasma levels were detected in patients with vascular invasion and in mucinous colorectal carcinoma, and also high HLA-G expression in tumours was associated with poor cancer specific OS in stage II-III, therefore, plasma sHLA-G differential levels could be a useful prognostic marker, and predictive biomarker of response to treatment in progressive CRC stages [96], [191]. To the best of our knowledge, the first distinction between metastatic and non-metastatic cancers based on the HLA-G expression was found in endometrial carcinoma showing a role of HLA-G as clinical marker for tumor metastasis or types of histopathological discrimination, however no data of comparison between metastatic and non-metastatic is available for CRC [192]. Lastly, some experiments in *in vivo* mouse model, confirm that HLA-G expression is associated with tumor metastasis, poor prognosis and counteracts tumor rejection reinforcing the importance to consider HLA-G as a promising target to optimize current cancer immunotherapies [193], [194].

3.2.6. HLA-G-based strategies in cancer immunotherapy and in drug-induced adverse effects

The observation that HLA-G is frequently and specifically expressed on tumor cells with a tolerogenic role, suggests that could be used as target for cancer therapies. However, although there are monoclonal antibodies directed against HLA-G, none is used in clinical cancer protocols. The only study we found using one of these antibodies is the one of Zhang and colleagues that developed a delivery system constituted by nanobubbles of poly(lactic-co-glycolic acid) carrying methotrexate (MTX) to the tumor site thanks to the presence of monoclonal antibody (mAb) anti-HLA-G and showed that they efficiently target HLA-G-positive tumor tissues both *in vitro* and *in vivo* inhibiting the reoccurence of tumors [195]. Another HLA-G-based strategy was the production of HLA-G146-154, a HLA-G-derived peptide, acting to suppress or downregulate the HLA-G by interfering RNA and inducing peptide-specific CTLs cytotoxic activity against HLA-G expressing HLA-A24^{+} renal cell carcinoma [196]. No other study was found in the literature regarding cancer situations. *In vitro* gene delivery strategies was used to efficiently overexpress functional HLA-G on bone marrow-derived mesenchymal stromal cells (MSC) through lentiviral vectors, improving the immunomodulatory function of these cells to treat several immune-based disorders but not cancer yet [197]. Moreover, synthetic HLA-G proteins were produced for therapeutic use in transplantation [198].
Introduction of the new immunotherapy drugs against the immunocheckpoint inhibitors determined an increase in immune-related ADEs (up to 70% of patients after treatment with inhibitors of the PD1 axis) [199]. The principal toxicities described involve the GI tract (diarrhea and enterocolitis as the most important complications), the liver, skin, and endocrine system, causing unspecific symptoms such as fatigue, weakness, nausea, confusion, or headache [200]. An HLA-G-based strategy might be developed also with the aim to manage the idiosyncratic adverse drug reactions (IADRs), type B adverse drug reactions (ADEs) that occur in a small proportion of patients experienced an idiosyncratic drug reaction within the normal therapeutic dose range [201]. They typically have a delayed onset of weeks or more after the primary exposure to a drug, are not related to the therapeutic effects and often are immune-related. IADRs are a major cause of patient morbidity and mortality which may result in drug withdrawal or restricted use, cost to healthcare system, and failure of drugs in development. Recently, genome wide association study

(GWAS) have identified HLA as an important genetic marker for IADRs because of its role as antigen presentation fundamental in the adaptive immune response trigger in a drug-specific way [28], [202]. The underlying mechanism to the IADRs might be the interaction between drugs and specific HLA allele mediated by the adaptive immune system. Traditionally, it was hypothesized that drugs are too small to be presented by MHC molecules because are much smaller than the 8-12 AA peptide of HLA class I and the 9-25 AA of HLA class II molecules. However, three models described how small-molecules might trigger the T-cell reactivity: the hapten concept, the pharmacologic interaction (p-i) concept, and the altered peptide repertoire model concept [203]. In the first model, a reactive metabolite binds covalently to an endogenous protein generating a chemically-modified peptide [204]. The p-i concept, instead, is based on the idea that the drug binds non-covalently to T cell receptor (TCR) or MHC protein in a peptide independent manner to activate T-cells response [205]. In the third mechanism, the offending drug occupies a position in the peptide-binding groove of HLA, changing the chemistry of the binding cleft and the specificity of the peptide for the MHC binding: the peptide is recognized as "foreign" eliciting the T-cells response. Lately, a fourth model was introduced: the altered TCR repertoire model, based on the fact that some drugs directly interact with TCRs changing their conformations [206]. MHC proteins are responsible for the regulation of the immune system and the peptide presented to T-cells could be altered by polymorphisms that affect the shape and electrochemistry of pockets within the peptide-binding groove and impact on interactions with small molecule drugs.

In silico models are important instrument used in docking simulations to investigate possible interactions between specific drugs and HLA molecules and predicts a possible cause of IADRs [29], [207]. Regarding the drug-HLA interactions, a strong association between the *HLA-B*57:01* allele and abacavir, an anti-HIV drug, has led to understand the critical role of T-cells in abacavir hypersensitivity [28]. The important reactions are triggered by the strong binding of abacavir into the antigenic peptide groove of the HLA molecule. Others *HLA* alleles and haplotypes has been found associated with drugs that cause idiosyncratic drug-induced liver and skin injury [208]. In particular, *HLA-B*57:01* allele had been identified as an important biomarker for flucloxancillin-induced liver injury [209]. Another strong HLA-drug association reported was the *HLA-B*15:02* protein with carbamazepine-induced Stevens-Johnson syndrome/toxic epidermal necrolysis (SJS/TEN), a severe cutaneous adverse drug reaction [210]. Carbamazepine may also trigged others hypersensitivity reactions as maculopapular exanthema in the present of *HLA-A*31:01* allele [211], [212]. Morever, the *HLA-B*38:02* and *HLA-DRB1*08:03* alleles are strongly associated with antithyroid drug-induced agranulocytosis. Another association was found between the *HLA-B*58:01* allele and allopurinol hypersensitivity (this drug is a xanthine oxidase inhibitor used in the treatment of gout and hyperuricemia) [213]. Also dapsone and amoxicillin calvulanate induced hypersensitivity expecially in the presence of *HLA-B*13:01*, and *DRB1*15:01, DRB5*01:01; DQB1*06:02, HLA-A*02:01* and *HLA-DQB1*06:02* haplotypes, respectively (for a review see [214]). Several docking techniques and programs were developed in these years and it was reported that through the homology modeling system named HLA-Modeler, is possible to reproduce the 3D structures of the antigenic peptide-binding sites at a level similar to the crystal structure [215]. Moreover, the docking program of ASEDock could predict the bound conformations and free energies for small-molecules ligands to proteins with an accurate precision in the reproducibility of the structure [216]. However, limitations still exist and the use of others approaches is encouraged to support the results found with computational techniques. In conclusion, some HLA-G-based strategies were developed and others could be implement not only for the classical HLA, but also for HLA-G due to its important role as immunocheckpoint inhibitors.

4. PART 1: Genetics characterization

4.1.PATIENTS AND METHODS

4.1.1. Study design and endpoints

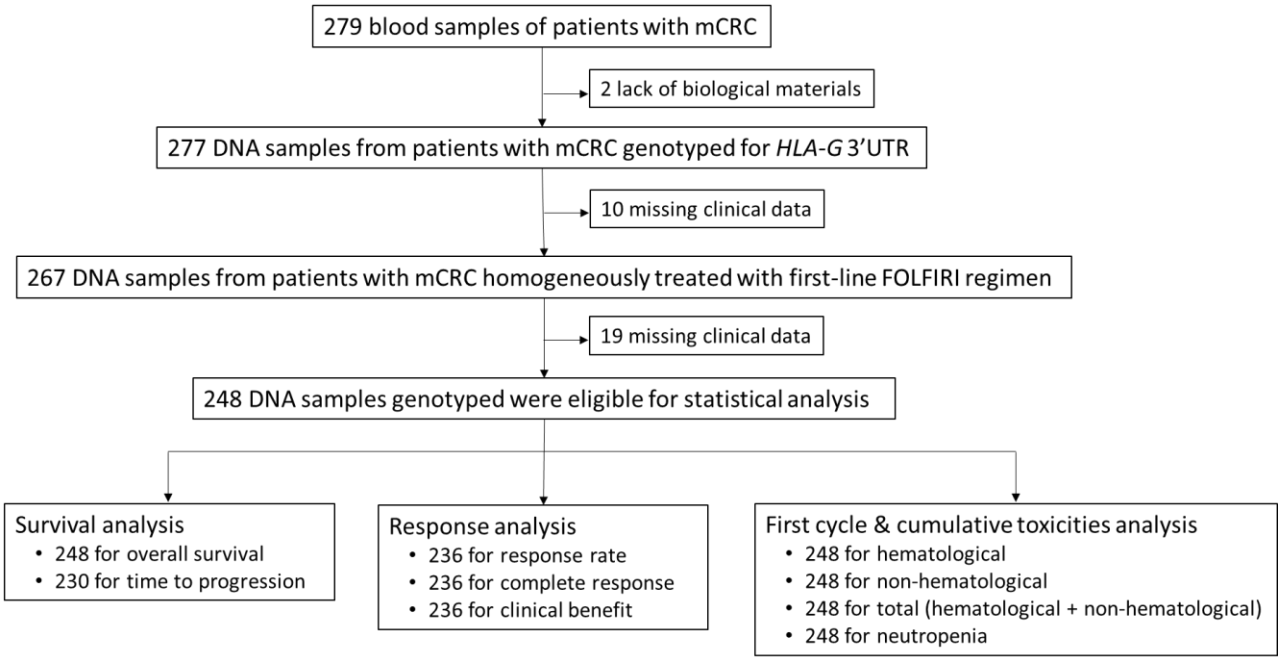

Figure 5. The CONSORT-like diagram.

This retrospective study was based on 279 blood samples from patients with mCRC, stored at -80°C in the existing prospective Biobank in the Experimental and Clinical Pharmacology Unit of National Cancer Institute (IRCCS CRO-Aviano) [78], [80], and the flowchart was described in the CONSORT-like diagram (Figure 5). The follow-up period was between July 2002 and November 2005 [78]. The institutional review board of each participating institution had approved the study protocol, and all patients signed a written informed consent for a genetic analysis before entering the study. A total of 277 available DNA samples were genotyping for *HLA-G* 3'UTR. Of these, 267 were of mCRC patients homogeneously treated with FOLFIRI regimen. We evaluated 248 samples as eligible for statistical analysis because of the presence of complete or partial clinical data availability. About the study cohort considered, the only difference between the previous published study [78] and this retrospective study, is that for 2 patients of the previous eligible cohort (n=250), the *HLA-G* 3'UTR genotyping was not assessable (no availability of whole blood samples), therefore, our study cohort accounts for 248 eligible patients with mCRC. Moreover, we evaluated also the contribution of the *UGT1A1*28* (rs8175347) polymorphism, previously reported to be significantly associated with severe ADE, tumor response and disease progression in the same cohort of patients [78], in combination with the *HLA-G* 3'UTR SNPs here studied. This study was designed to test the association between genetic polymorphisms/haplotypes and clinical outcomes (survival, response, first cycle and cumulative toxicities). In particular, we investigated the association between genotypes/haplotypes with overall survival (OS) and time to progression (TTP) to search novel prognostic

markers. The OS was defined as the time to the first drug administration date to the last date of follow-up or death while the TTP was defined as the time from the first drug administration date to the date of first progression or last follow-up.
Objective tumor response and duration of response were assessed by WHO criteria [217] and response to treatment was evaluated only in patients who had received at least four cycles of chemotherapy [78]. The objective tumor response was defined in three different way: the response rate (RR) (stable disease (SD) + progression disease (PD) vs complete response (CR) + partial response (PR)); the clinical benefit (CB) (PD vs SD+ PR+CR), and the complete response (CR vs SD + PD + PR). The tumor response was measured through computed tomography scans as the objective CR (evaluated with the disappearance of all known disease confirmed at 4 weeks), against all the other parameters: SD (evaluated with neither PR nor PD criteria met), PD (when it was observed an increase of more than 25% of the disease without CR, PR or SD documented before increase disease; new lesion(s) or an increase of more than 25% in 1 lesion), and PR (when it was observed a decrease of ≥50% confirmed at 4 weeks). Toxicity was evaluated according to National Cancer Institute Common Toxicity Criteria [218]. We considered hematological, non-hematological, total and neutropenia toxicities as follows: hematological toxicities were neutropenia, anemia, leukopenia, and thrombocytopenia; non-hematological toxicities were diarrhea, nausea, vomiting, asthenia, alopecia, mucositis, anorexia, and non-neutropenic infections; total toxicities were the sum of hematological and non-hematological toxicities. Finally, neutropenia, diarrhea and gastrointestinal (G.I.) toxicities (grouped in diarrhea, nausea and vomiting) were also evaluated separately because are the most frequent non-hematological toxicities associated with FOLFIRI regiment in mCRC, as also reported in the previous studies by our group [78], [219]. A single cycle of chemotherapy administration was considered enough for evaluation of first cycle toxicity, otherwise the cumulative toxicity was measured. Clinical evaluations were performed blindly with respect to the genetic results, and clinical data were monitored by the previously study sponsor. For each toxicity endpoint, we evaluated the highest grade of toxicity recorded during the treatment to determine association with *HLA-G* polymorphisms and haplotypes.

4.1.2. Patients clinical data and treatment

In this retrospective study, the clinical and treatment data were available from 279 patients diagnosed with mCRC prospectively collected by Toffoli and colleagues for a previous published study [78]. Inclusion criteria in this retrospective study are the same of those previously described [78]. Eligible patients were aged between 18 and 75 years with metastatic colorectal cancer (mCRC) histologically confirmed and no previously treated with chemotherapy for metastatic disease (only adjuvant chemotherapy was allowed, except for irinotecan). Moreover, they had an Eastern Cooperative Oncology Group (ECOG) performance status of 0 to 2; life expectancy more than 3 months; at least one measurable cancer lesion; absolute neutrophil count ≥ 2,000 µl; platelets ≥ 100,000 µl; normal renal function (creatinine clearance > 65 ml/min by Cockcroft formula [220]); ALT and AST less than 2.0x the upper limit of normal and total serum bilirubin less than 1.25x the upper limit of normal. Objective clinical evaluation, blood counts, and hepatic and renal function tests were performed within 48 hours before each cycle. Patients were questioned specifically about nausea and vomiting, mucositis, diarrhea, malaise, and appetite at every cycle. Computed tomography scans of measurable lesions were assessed at baseline and then repeated at least every four cycles. Patients with progressive disease (PD) could leave the study or could continue chemotherapy for two additional cycles according to the physician's decision. Briefly, in the prospective single-arm interinstitutional study a total of 267 patients were treated with irinotecan (CPT-11) in association with 5-fluorouracil (5-FU) and leucovorin (LV) (FOLFIRI regimen: CPT-11 180 mg/m^2 bolus followed by 5-FU 2,400

mg/m^2 continuous infusion during 46 hours plus LV 200 mg/m^2 intravenously for 2 hours on day 1 every 2 weeks) in first-line for metastatic disease. More than 90% of patients received the modified FOLFIRI regiment as described by Tournigand (CPT-11 180 mg/m^2 intravenously for 2 hours on day 1 plus 5-FU 400 mg/m^2 bolus followed by 5-FU 600 mg/m^2 continuous infusion during 22 hours on day 1 and 2 plus LV 200 mg/m^2 on days 1 and 2 every 2 weeks) [221]. Patients were pretreated with atropine 0.5 mg to avoid the cholinergic syndrome frequently seen within the first 24 hours resulting from the inhibition of acetylcholinesterase activity by CPT-11, dexamethasone 8 mg, and granisetron 3 mg or ondansetron 8 mg to control nausea and vomiting. Loperamide 4 mg was used to immediately treated diarrhea at the onset, and then 2 mg every 2 hours were administered until the patients were diarrhea free for at least 12 hours. Chemotherapy was stopped until recovery if neutrophils were ≤ 1,500 µl or in presence of significant, persisting, non-hematological toxicity. In the event of grade 3 or 4 neutropenia, thrombocytopenia, and diarrhea, the irinotecan dose was reduced (from 180 mg/m^2) to 90 to 150 mg/m^2 based on the physician's assessment. Treatment was discontinued in the event of repeated grade 3 to 4 toxicity, despite dose reduction, or because of patient refusal.

4.1.3. DNA extraction, genotyping and frequencies of *HLA-G* polymorphisms

Genomic DNA was extracted from available peripheral whole blood of 277 mCRC patients using the High Pure PCR Template Preparation Kit (Roche Diagnostics GmbH, Mannheim, Germany). The 3'UTR of the *HLA-G* gene was amplified by polymerase chain reaction (PCR) technique using the already published primers [222]: forward HLAG8Fprimer, 5'-TGTGAAACAGCTGCCCTGTGT-3', and reverse HLAG8Rprimer, 5'-GTCTTCCATTTATTTTGTCTCT-3'. The reaction was performed in a final volume of 30 µl with 1.25 mM $MgCl_2$, 0.25mM each deoxynucleotide triphosphates (dNTPs), 5 pmole/primer, about 50-200 ng genomic DNA template, 1X PCR Buffer and 0.5unit of AmpliTaq Gold DNA polymerase (AppliedBiosystems, Foster City, CA, USA). The PCR cycles were as follows: 5 mins of initial denaturation at 94°C, 30 cycles of 45 secs at 95°C, 45 secs at 56°C, 60 secs at 72°C, and the final extension step at 72°C for 7 mins. PCR products (344 bp in presence of deletion and 358 bp for insertion) were checked on 3% agarose gel electrophoresis stained with ethidium bromide. 2.5 µl of PCR reactions were purified using EXOSAP-IT^{TM} (Thermo Fisher Scientific Corporate Group, Waltham, MA, USA) and then 1-2 µl were sequenced (Sanger method) by the use of the Big Dye Terminator kit (Applied Biosystems, Foster City, CA, USA) and an ABI PRISM capillary sequencer with 3.4 µM of the reverse HLAG8R primer to prevent sequence overlaps in heterozygous 14-bp samples. If necessary, a forward sequencing was done. Chromatograms were visualized with Chromas software version 2.01 and all single nucleotide polymorphisms (SNPs), and single nucleotide variants (SNVs) detected were recorded for each study participant to determine their genotypes using the *G*01:01:01:01* genomic sequence reported in the HLA-G specific site (http://hla.alleles.org/data/hla-g.html) as reference (Release 3.33.0, 11 July 2018). In particular, we analyzed the most studied germline polymorphisms present in the *HLA-G* 3'UTR segment: rs371194629 (+2960 14-base pair (bp) INDEL), rs1707 (+3003 T>C), rs1710 (+3010 C>G), rs17179101 (+3027 C>A), rs17179108 (+3035 C>T), rs1063320 (+3142 G>C), rs9380142 (+3187 A>G), rs1610696 (+3196 C>G), and rs1233331 (+3227 G>A>T). Among these 9 polymorphisms, 8 had ≥5% variant allelic frequency (VAF) and were included in the analysis (rs1233331 had VAF <5% and was excluded).
Frequencies of SNPs and SNVs found were calculated and compared to those reported in 1000Genome public database (http://www.internationalgenome.org/1000-genomes-browsers/) for EUR sub-population according to Ensembl GRCh37 access.

4.1.4. Linkage disequilibrium and *HLA-G* haplotypes analyses

The strength of LD between pairs of HLA-G markers with MAF>2% was measured as r^2 using the freely available LDPlotter software (http://www.pharmgat.org/Tools/pbtoldplotform). The statistic $r^2<0.50$ indicates low LD, $0.50\leq r^2<0.80$ moderate high LD, $0.80\leq r^2<1$ high LD and $r^2 = 1$ perfect LD. Haplotype composition, a combination of specific alleles at neighbouring regions that tends to be inherited together, was carried out by using PHASE algorithm version 2.1, which implements a Bayesian statistical method for reconstructing haplotypes from population genotype data [223], [224]. We considered the 7 different haplotypes with the highest worldwide frequencies: UTR-1 (DelTGCCCGC), UTR-2 (InsTCCCGAG), UTR-3 (DelTCCCGAC), UTR-4 (DelCGCCCAC), UTR-5 (InsTCCTGAC), UTR-6/18 (DelTGCCCAC), UTR-7 (InsTCATGAC) (ndr. the sequences in brackets correspond to the alleles of the 8 most abundant polymorphisms reported above) [225].

4.1.5. HLA-G secretor models and UGT1A1-combined models

Patients were also subdivided in classes representing the expression level of the HLA-G protein correlated to the 3'UTR haplotypes composition in the *HLA-G* gene (phenotype-genotype correlation): two HLA-G secretors models were designed, a "2 levels" model and "3 levels" model, according to the findings reported in two already published works [18], [19]. In brief, the 2 levels model [19] divides the patients into "low-HLA-G secretors" when there is presence of *HLA-G* 3'UTR haplotypes composed by *Ins*, *G* and *A* alleles in the rs371194629, rs1063320, and rs9380142 polymorphisms respectively, and in "medium/high-HLA-G secretors" for subjects carrying all the other haplotypes. According to the 3 levels model [18], we considered three groups: "high-HLA-G secretors" represented by subjects with UTR-1 haplotype, "low-HLA-G secretors" (UTR-5 and UTR-7) and "intermediate-HLA-G secretors" (UTR-2/3/4/6).
Moreover, we evaluated also the contribution of the *UGT1A1*28* (rs8175347) polymorphism, previously reported to be significantly associated with severe ADE, tumor response and disease progression in the same cohort of patients [78], in combination with the *HLA-G* 3'UTR SNPs here studied. In particular, we analyzed the co-presence of *HLA-G* 3'UTR haplotypes according to the secretors models and the *UGT1A1* (rs8175347) germline status, already genotyped and analyzed by our group [78], generating new classes of combined haplotype by the PHASE method.

4.1.6. Statistical analysis

All the 248 patients considered eligible for the statistical analysis were investigated for the overall survival, response, first cycle and cumulative toxicities based on their genotype and haplotype. Of these, 230 patients were considered for the time to progression analysis and 236 for the response analysis because of the lack of related-information of 18 and 12 patients, respectively. Univariate (Kaplan-Meier curves with Log-Rank test) and multivariate analyses (Cox proportional hazard models) were used for the survival analysis. Hazard ratio (HR) or odds ratio (OR) and 95% confidence intervals (CI) were reported for each model. Response and toxicities analysis were performed using logistic regression models. The clinical covariates used for model correction were those emerged statistically significant from univariate analysis ($p<0.05$). Dominant, recessive, and log-additive genetic models were considered in the analysis. In the dominant models it was compared the most frequent allele (AA) with the heterozygous plus the recessive allele (Aa+aa), where in the recessive model it is exactly the opposite (aa vs AA+Aa). The log-additive model, instead, could be explain as the risk associated with the progressive addition of a recessive allele

(AA vs Aa vs aa). Consequently, the association analysis was between a given SNP (coded as 0, 1, 2 based on the presence of recessive allele) and a dichotomous trait under study. The significance level was set at $p<0.05$ (two-sided). The most statistically significant genetic model was selected according to the *p*-value corresponding to the likelihood ratio test obtained from a comparison with the null model. Statistical analyses were performed with R software (www.r-project.org).

4.2.RESULTS

4.2.1.Patient characteristics

The main demographic and clinical characteristics of 248 eligible patients with mCRC patients treated with FOLFIRI regiment are shown in Table 1. The median follow-up was 15.37 months (range: 0.7-63.5 months). The majority of the patients with mCRC were staged IV at the time of diagnosis (*n*= 158, 63.7%), received radical surgery (*n*=198, 79.8%), and 48 out of 248 (19.3%) patients had more than 2 metastatic sites at the time of enrollment in the study [78]. We observed a predominance of men (*n*=161, 64.9%) according to the global population incidence [14], [40]. Type of tumor response and its distribution in this study cohort is also reported in Table 1. Most common adverse drug events (ADE) are reported in Table 2.

Table 1. Demographic and clinical characteristics of 248 eligible patients with mCRC.

Characteristic	***n***	**(%)**
Age, yrs		
Median (range)	63.2 (26.3-75.9)	
Gender		
Man	161	64.9
Woman	87	35.1
Primary tumor site		
Right colon	78	31.5
Left colon	99	39.9
Rectum	71	28.6
Radical surgery		
Yes	198	79.8
No	50	20.2
Stage at diagnosis		
I	5	2.0
II	20	8.1
III	65	26.2
IV	158	63.7
Adjuvant (radiotherapy or chemotherapy)		
Yes	87	35.1
No	161	64.9
Number of metastatic sites		
1	107	43.2
2	93	37.5
3	32	12.9
4	9	3.6
5	7	2.8
Follow-up time (months)		
Median (range)	15.37 (0.77-63.50)	
Response		
Complete response	18	7.3
Partial response	84	33.9
Stable disease	66	26.6
Progression disease	68	27.4

Table 2. Most common recorded ADE (G1-2 and G3-4)

	First cycle toxicity (n=248)				Cumulative toxicity (n=248)			
	Grade G1-2		Grade G3-4		Grade G1-2		Grade G3-4	
Adverse events	**n**	**(%)**	**n**	**(%)**	**n**	**(%)**	**n**	**(%)**
Hematological								
Neutropenia	29	11.7	11	4.4	56	22.6	35	14.1
Anemia	32	12.9	1	0.4	56	22.6	3	1.2
Leukopenia	25	10.1	2	0.8	46	18.5	15	6.0
Thrombocytopenia	2	0.8	0	0.0	8	3.2	0	0.0
Fever with concomitant G3-4 neutropenia	0	0.0	2	0.8	4	1.6	3	1.2
Non-hematological								
Diarrhea	62	25.0	5	2.0	93	37.5	20	8.1
Nausea	54	21.8	2	0.8	98	37.5	4	1.6
Vomiting	24	9.6	2	0.8	49	19.8	7	2.8
Asthenia	23	9.2	1	0.4	56	22.6	3	1.2
Alopecia	7	2.8	1	0.4	27	10.9	8	3.2
Mucositis	17	6.8	0	0.0	53	21.4	7	2.8
Anorexia	3	1.2	0	0.0	8	3.2	0	0.0
Infection without concomitant G3-4neutropenia	1	0.4	1	0.4	5	2.0	3	1.2

4.2.2. Germline *HLA-G* 3'UTR polymorphisms, haplotypes and secretor models

From a total of 279 blood samples of mCRC patients, DNA of 277 samples were successfully extracted and evaluated for *HLA-G* 3'UTR genotyping. Among the most studied germline polymorphisms present in this regulatory region, only the rs371194629 (+2960 14-base pair (bp) INDEL), rs1707 (+3003 T>C), rs1710 (+3010 C>G), rs17179101 (+3027 C>A), rs17179108 (+3035 C>T), rs1063320 (+3142 G>C), rs9380142 (+3187 A>G), rs1610696 (+3196 C>G), had a ≥5% variant allelic frequency and were analyzed. All the selected polymorphisms exhibited genotype distributions that conformed to the assumptions of the Hardy-Weinberg equilibrium (HWE) (p>0.05), even if the rs371194629 showed a slight deviation (p=0.05). Moreover, we reported the distributions of alleles and genotypes frequencies with those reported in the 1000Genome Browser for global population, and for European and Italian one. We compared the allele frequency between our eligible population and the European (Table 3). The allele frequencies for the rs1710 (+3010 C>G) SNP differed significantly between the two populations, showing a predominance of the wild type C allele in our mCRC population (p=0.0036). Moreover, a slight increase was observed for the *C/C* and *G/C* genotypes compared to the European population (31.5% and 45.97% vs 20.7% and 50.5%, respectively). The rs1063320 (+3142 G>C) SNP, reported to be in perfect LD, showed the same result as rs1710, with a predominance in the wild type *G* allele in the 248 patients analyzed (p=0.0021). Another significant difference was found for the rs9380142 (+3187 A>G) SNP, with a higher allele frequency of the wild type *A* allele in our mCRC population, compared to the European (p=0.0222). The rs1610696 (+3196 C>G) presented a trend towards higher frequency of the heterozygous *C/G* and lower frequency of the homozygous *C/C* genotypes, compared to the 1000Genome Browser European population (44.35% and 48.39% vs 40.2% and 52.3%). Finally, for the rs371194624 (+2960 14-bp INDEL), a trend towards a higher frequency of the *Ins/Ins* genotype compared to the European population, was observed in our studied population with mCRC.

Table 3. Distributions of alleles and genotypes frequencies identified at *HLA-G* 3'UTR polymorphic sites and comparison with frequencies reported in 1000 Genome Browser: *HLA-G* 3'UTR genotyping was performed in 277 patients genetically evaluated with mCRC of which 248 were eligible for this study.

			1000 Genome Browser			
	Evaluated N_{tot}**=277**	**Eligible** N_{tot}**=248**	**All** N_{tot}**=2504**	**EUR** N_{tot}**=503**	**TSI** N_{tot}**=107**	**Evaluated Eligible 248 vs. EUR**
***HLA-G* 3'UTR SNPs**	**n (%)**	**n (%)**	**n (%)**	**n (%)**	**n (%)**	***p***
+2960 14-bp INDEL (rs371194629)						
Alleles						
Del	324 (58.48)	293 (59.07)	3034 (60.6)	638 (63.4)	131 (61.2)	
Ins	230 (41.52)	203 (40.93)	1974 (39.4)	368 (36.6)	83 (38.8)	
Genotypes						
Del/Del	101 (36.46)	94 (37.90)	920 (36.7)	204 (40.6)	39 (36.4)	
Ins/Del	122 (44.04)	105 (42.34)	1194 (47.7)	230 (45.7)	53 (49.5)	
Ins/Ins	54 (19.50)	49 (19.76)	390 (15.6)	69 (13.7)	15 (14.0)	
+3003 T>C (rs1707)						
Alleles						
T	485 (87.55)	431 (86.90)	4540 (90.7)	851 (84.6)	185 (86.4)	
C	69 (12.45)	65 (13.10)	468 (9.3)	155 (15.4)	29 (13.6)	
Genotypes						
T/T	215 (77.62)	189 (76.21)	2068 (82.6)	367 (73.0)	80 (74.8)	
T/C	55 (19.86)	53 (21.37)	404 (16.1)	117 (23.3)	25 (23.4)	
C/C	7 (2.53)	6 (2.42)	32 (1.3)	19 (3.8)	2 (1.9)	
+3010 C>G (rs1710)						
Alleles						
C	302 (54.51)	268 (54.03)	2994 (59.8)	462 (45.9)	114 (53.3)	0.0036*
G	252 (45.49)	228 (45.97)	2014 (40.2)	544 (54.1)	100 (46.7)	
Genotypes						
C/C	87 (31.41)	77 (31.05)	910 (36.3)	104 (20.7)	30 (28.0)	
G/C	128 (46.21)	114 (45.97)	1174 (46.9)	254 (50.5)	54 (50.5)	
G/G	62 (22.38)	57 (22.98)	420 (16.8)	145 (28.8)	23 (21.5)	
+3027 C>A (rs17179101)						
Alleles						
C	520 (93.86)	464 (93.55)	4569 (91.2)	946 (94.0)	205 (95.8)	
A	34 (6.14)	32 (6.45)	439 (8.8)	60 (6.0)	9 (4.2)	
Genotypes						
C/C	243 (87.75)	216 (87.10)	2105 (84.1)	443 (88.1)	98 (91.6)	
C/A	34 (12.27)	32 (12.90)	359 (14.3)	60 (11.9)	9 (8.4)	
A/A	0 (0.00)	0 (0.00)	40 (1.6)	0 (0.0)	0 (0.0)	
+3035 C>T (rs17179108)						
Alleles						
C	489 (88.27)	437 (88.10)	4281 (85.5)	913 (90.8)	196 (91.6)	
T	65 (11.73)	59 (11.90)	727 (14.5)	93 (9.2)	18 (8.4)	
Genotypes						
C/C	217 (78.34)	194 (78.23)	1838 (73.4)	412 (81.9)	89 (83.2)	
C/T	55 (19.86)	49 (19.76)	605 (24.2)	89 (17.7)	18 (16.8)	
T/T	5 (1.81)	5 (2.01)	61 (2.4)	2 (0.4)	0 (0.0)	

+3142 G>C (rs1063320)						
Alleles						
G	305 (55.05)	270 (54.44)	2995 (59.8)	462 (45.9)	114 (53.3)	0.0021*
C	249 (44.95)	226 (45.56)	2013 (40.2)	544 (54.1)	100 (46.7)	
Genotypes						
G/G	87 (31.41)	77 (31.05)	911 (36.4)	104 (20.7)	30 (28.0)	
G/C	131 (47.29)	116 (46.77)	1173 (46.8)	254 (50.5)	54 (50.5)	
C/C	59 (21.30)	55 (22.18)	420 (16.8)	145 (28.8)	23 (21.5)	
+3187 A>G (rs9380142)						
Alleles						
A	397 (71.66)	355 (71.57)	3715 (74.2)	660 (65.6)	150 (70.1)	0.0222*
G	157 (28.34)	141 (28.43)	1293 (25.8)	346 (34.4)	64 (29.9)	
Genotypes						
A/A	140 (50.54)	126 (50.81)	1390 (55.5)	214 (42.5)	53 (49.5)	
A/G	117 (42.24)	103 (41.53)	935 (37.3)	232 (46.1)	44 (41.1)	
G/G	20 (7.22)	19 (7.66)	179 (7.1)	57 (11.3)	10 (9.3)	
+3196 C>G (rs1610696)						
Alleles						
C	387 (69.86)	350 (70.56)	3755 (75.0)	728 (72.4)	147 (68.7)	
G	167 (30.14)	146 (29.44)	1253 (25.0)	278 (27.6)	67 (31.3)	
Genotypes						
C/C	132 (47.65)	120 (48.39)	1422 (56.8)	263 (52.3)	52 (48.6)	
C/G	123 (44.40)	110 (44.35)	911 (36.4)	202 (40.2)	43 (40.2)	
G/G	22 (7.94)	18 (7.26)	171 (6.8)	38 (7.6)	12 (11.2)	
+3227 G>A (rs1233331)						
Alleles						
G	543 (98.01)	485 (97.78)	4915 (98.1)	970 (96.4)	211 (98.6)	
A	11 (1.99)	11 (2.22)	93 (1.9)	36 (3.6)	3 (1.4)	
Genotypes						
G/G	267 (96.39)	238 (95.97)	2413 (96.4)	468 (93.0)	104 (97.2)	
G/A	9 (3.25)	9 (3.63)	89 (3.6)	34 (6.8)	3 (2.8)	
A/A	1 (0.36)	1 (0.40)	2 (0.1)	1 (0.2)	0(0.0)	

SNPs: Single nucleotide polymorphisms; mCRC: metastatic colorectal cancer; HWE: Hardy-Weinberg equilibrium; N_{tot}: number of individuals; All: global population; EUR: European; TSI: Toscani in Italy.

*calculated with two-sided Fisher's exact test.

A total of 8 different haplotypes were defined by the PHASE algorithm: UTR-1 (DelTGCCCGC), UTR-2 (InsTCCCGAG), UTR-3 (DelTCCCGAC), UTR-4 (DelCGCCCAC), UTR-5 (InsTCCTGAC), UTR-6/18 (DelTGCCCAC), UTR-7 (InsTCATGAC), UTR-13 (DelTCCTGAC). UTR-13 had a frequency of 0.4% (2/496) and was excluded from the analysis. The UTR-6 and UTR-18 differ at only one position in the +3227 G>A (rs1233331) polymorphism (G and A, respectively); this SNP was excluded from the analysis due to its low frequency (<5%) and consequently from the haplotype reconstruction. Therefore, UTR-18 was included in the UTR-6 haplotype in our study (Table 4).

Table 4. *HLA-G* 3' UTR haplotypes: numbers and frequencies identified in 248 eligible patients with mCRC.

***HLA-G* 3'UTR haplotype**	**14-bp**	**+3003**	**+3010**	**+3027**	**+3035**	**+3142**	**+3187**	**+3196**	**Eligible N_{tot}=248 n (%)**
UTR-2									
InsTCCCGAG	**Ins**	T	C	C	C	G	A	**G**	146 (29.5)
Het									109 (44.0)
Hom									18 (7.3)
UTR-1									
DelTGCCCGC	Del	T	**G**	C	C	**C**	**G**	C	140 (28.2)
Het									104 (41.9)
Hom									18 (7.3)
UTR-3									
DelTCCCGAC	Del	T	C	C	C	G	A	C	65 (13.1)
Het									36 (14.5)
Hom									6 (2.4)
UTR-4									
DelCGCCCAC	Del	**C**	**G**	C	C	**C**	A	C	65 (13.1)
Het									53 (21.4)
Hom									6 (2.4)
UTR-7									
InsTCATGAC	**Ins**	T	C	**A**	**T**	G	A	C	32 (6.5)
Het									32 (12.9)
Hom									0 (0.0)
UTR-5									
InsTCCTGAC	**Ins**	T	C	C	**T**	G	A	C	25 (5.0)
Het									21 (8.5)
Hom									2 (0.8)
UTR-6/-18									
DelTGCCCAC	Del	T	**G**	C	C	**C**	A	C	21 (4.2)
Het									15 (6.0)
Hom									2 (0.8)

Het: heterozygous; Hom: homozygous. In bold are highlighted the mutant alleles.

Distributions into classes of the HLA-G secretor models generated by the PHASE method, were described in Table 5. In our study cohort, we observed inside of the different HLA-G expression classes a prevalence of medium-high levels of HLA-G, according to the previously publications in other cancer types [110]. This distribution was more evident in the 2 levels model (3 classes) compared to the 3 levels model (6 classes) where the H/I class was the most frequent (37.5%), followed by the intermediate class (I/I) (34.3%) compared to H/H (7.3%) and S/S (2.0%).

Table 5. HLA-G secretor models: numbers and frequencies observed in 248 mCRC eligible patients.

HLA-G secretor model	Code	Phenotypic effect	Corresponding *HLA-G* 3'UTR haplotypes*	Classes	n (%)
2 levels	M	Medium/High HLA-G expression	UTR-1, UTR-3, UTR-4, UTR-6	M/M	94 (37.9)
	L	Low HLA-G expression	UTR-2, UTR-5, UTR-7	L/M	105 (42.3)
				L/L	49 (19.8)
3 levels	H	High HLA-G expression	UTR-1	H/H	18 (7.3)
	I	Intermediate HLA-G expression	UTR-2, UTR-3, UTR-4, UTR-6	H/I	93 (37.5)
	S	Low HLA-G expression	UTR-5, UTR-7	H/S	11 (4.4)
				I/I	85 (34.3)
				I/S	36 (14.5)
				S/S	5 (2.0)

*Each HLA-G level may be composed by each of the corresponding *HLA-G* 3'UTR haplotype, alone or in combination with one of those listed.

Finally, we also analyzed the co-presence of *HLA-G* 3'UTR haplotypes according to the secretors models and the *UGT1A1* (rs8175347) germline status. In the Table 6, it has been reported the numbers and frequencies of *HLA-G* and *UGT1A1* combined classes, based on their presence in the patients (0, 1, 2 copies). For the HLA-G 2 levels model combined with *UGT1A1* (rs8175347), we observed similar frequencies for *1M and *1L classes (35.5% and 32.9%, respectively) which resulted more represented compared to the *28M and *28L classes (23.6% and 8.0%, respectively). Similarly, for the HLA-G 3 levels model combined with *UGT1A1* (rs8175347), we observed the same trend in the intermediate classes (*1I: 35.3%, *28I: 25.0%) and lower expressing classes (*1S: 10.1%, *28S: 1.4%).

Table 6. HLA-G secretor models and *UGT1A1* rs8175347 combination: numbers and frequencies in 248 eligible patients with mCRC.

				Combo classes n (%) (N_{tot}=248)			
Models	*UGT1A1* alleles	*HLA-G* 3'UTR code*	Combo classes	0 copy n (%)	1 copy n (%)	2 copies n (%)	2N %
UGT1A1	*UGT1A1*1*	M	*1M	115 (46.4)	90 (36.3)	43 (17.3)	35.5
+	*UGT1A1*1*	L	*1L	108 (43.5)	117 (47.2)	23 (9.3)	32.9
HLA-G	*UGT1A1*28*	M	*28M	139 (56.0)	101 (40.7)	8 (3.2)	23.6
2 levels	*UGT1A1*28*	L	*28L	211 (85.0)	34 (13.7)	3 (1.2)	8.0
UGT1A1	*UGT1A1*1*	H	*1H	143 (57.7)	96 (38.7)	9 (3.6)	23.0
+	*UGT1A1*1*	I	*1I	112 (45.2)	97 (39.1)	39 (15.7)	35.3
HLA-G	*UGT1A1*1*	S	*1S	200 (80.6)	46 (18.5)	2 (0.8)	10.1
3 levels	*UGT1A1*28*	H	*28H	221 (89.1)	26 (10.5)	1 (0.4)	5.6
	*UGT1A1*28*	I	*28I	130 (52.4)	112 (45.2)	6 (2.4)	25.0
	*UGT1A1*28*	S	*28S	241 (97.2)	7 (2.8)	0 (0.0)	1.4

*UGT1A1*1*: rs8175347-*(TA)*$_6$ wild type allele for *UGT1A1*; *UGT1A1*28*: rs8175347-*(TA)*$_7$ mutated allele for *UGT1A1*; Combo: combined.

*Each HLA-G level may be composed by each of the corresponding *HLA-G* 3'UTR haplotype, alone or in combination with one of those described.

4.2.3. Linkage disequilibrium patterns

We analyzed the LD between the polymorphisms in the 3'UTR of the HLA-G and we observed perfect linkage disequilibrium (LD) between rs1710 (+3010 C>G) and rs1063320 (+3142 G>C) polymorphisms (r2=1) and a moderate-high LD between 371194629 (+2960 14-bp INDEL) and rs1710 (+3010 C>G), rs1063320 (+3142 G>C), rs1610696 (+3196 C>G) (r2=0.58, 0.58, 0.60, respectively), and between rs17179101 (+3027 C>A) and rs17179108 (+3035 C>T) (r2=0.51) (see Figure 6), according to the previously reported data [25], [226], [227].

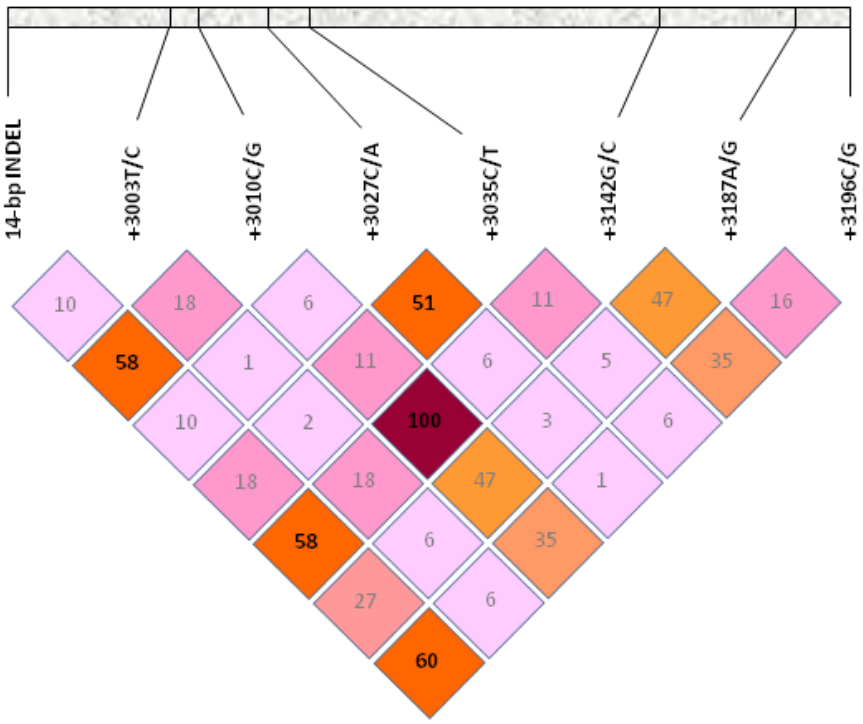

Figure 6. LD patterns at the 3'UTR region of *HLA-G* in 248 patients with mCRC. LD plot generated by LDPlotter shows correlations between all pairs of variants with MAF >2%. The r2 values (x100) for the marker pairs are listed in the corresponding boxes. High pairwise LD (r2) between variants is highlighted in bold.

4.2.4. Survival analysis

The primary endpoint of this study was to test if the *HLA-G* 3'UTR SNPs had a prognostic value as markers for OS and TTP. In the OS analysis the follow-up was cut at 40 months (5 out of 248 patients had an OS>40 months) and for the TTP analysis, follow-up was cut to 30 months (3 out of 248 patients had a TTP>30 months). In univariate analysis clinical variables were tested for their impact in the OS and TTP: only significant variables were used to adjust the results in multivariate analysis.

4.2.4.1. Overall survival

Median OS was 15.37 months (range: 0.77-63.50), and 128 deaths (52%) occurred during this period in the 248 patients evaluated in this study. The univariate and multivariate analysis with clinical variables showed an association only between radical surgery and OS (HR=0.4358, 95% CI: 0.2788-0.6811; *p*=0.0003) (Table 7). Patients who received a radical surgery had a prolonged OS (Figure 7).

Table 7. Multivariate analysis for clinical variables.

Variable	HR	95% CI	*p*-value
Gender	1.0279	0.7007-1.5078	0.8881
Age	0.9969	0.9794-1.0147	0.7315
Primary tumor site rectum	1.1830	0.7274-1.9237	0.4983
Primary tumor site right colon	1.3736	0.8970-2.1033	0.1442
Radical surgery yes	0.4358	0.2788-0.6811	**0.0003**
Stage at diagnosis II	4.4365	0.5096-38.6241	0.1772
Stage at diagnosis III	3.8306	0.4757-30.8456	0.2070
Stage at diagnosis IV	3.1992	0.4347-23.5464	0.2535
Adiuvant yes	0.7576	0.4036-1.4219	0.3875
Number of metastatic sites	1.0092	0.8541-1.1923	0.9147

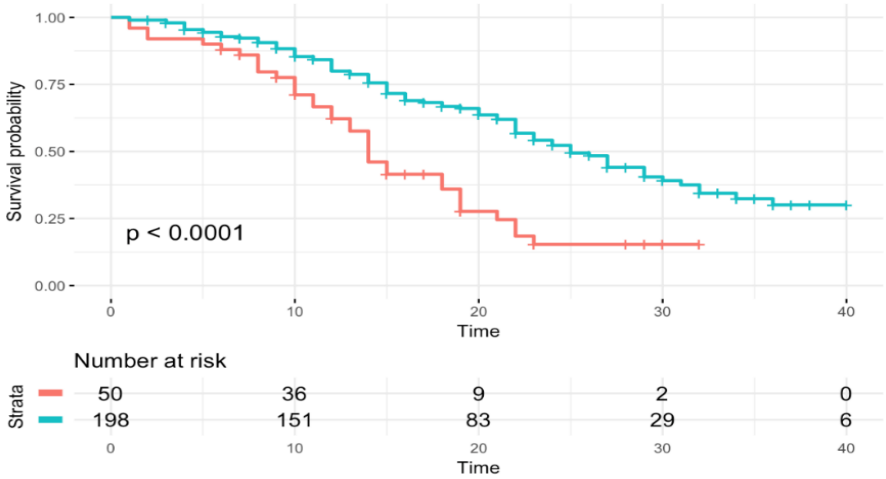

Figure 7. Kaplan-Meier curves according to radical surgery. In red: radical surgery "yes"; in blue: radical surgery "no".

HLA-G 3'UTR polymorphisms/haplotypes and OS

No significant data were found (data not shown).

HLA-G secretor models and OS

None of the *HLA-G* 3'UTR polymorphisms or haplotypes showed a significant association with OS, however, considering the HLA-G secretor 3 levels model, the presence of S/S class was associated with a reduced OS compared to the I/I class (HR=2.91, 95% CI: 1.03-8.18, *p*=0.0432) (Table 8).

Table 8. Significant associations between OS and *HLA-G* secretor models.

HLA-G secretor model	Classes	n (%)	Model	HR*	95% CI*	*p*§
3 levels	H/H	18 (7.3)	S/S vs. I/I	2.91	1.03-8.18	0.0432
	H/I	93 (37.5)				
	H/S	11 (4.4)				
	I/I#	**85 (34.3)**				
	I/S	36 (14.5)				
	S/S	**5 (2.0)**				

H: high level of HLA-G expression in 3 levels model; I: intermediate level of HLA-G expression in 3 levels model; S: low level of HLA-G expression in 3 levels model.

*Estimated from Cox proportional hazard model, adjusted for radical surgery. The HR reported is related to the significant model described in the "Model" column.

§Only the statistically significant data are reported (p < 0.05).

#Reference category.

HLA-G secretor models and *UGT1A1* (rs8175347) combination in OS

We knew that UGT1A1 has a foundamental role in the metabolism of the irinotecan, but this function seems not to cause a positive effect on OS when compared the variant UGT1A1*28 allele subgroup with UGT1A1*1/*1 genotype [78]. We observed an association of the heterozygous *28M combined class with an increased OS compared to the absence of *28M (HR=0.68, 95% CI: 0.47-0.98, *p*=0.0409), while the presence of heterozygous *28S class, compared to the absence of *28S, was associated with a worse OS (Table 9). These associations were shown in the Kaplan-Meier curves (Figure 8, Figure 9). Analyzing the association between HLA-G and UGT1A1, however, we observed a statistical significant effect only in the present of UGT1A1*28 allele. Then, the classes with the presence of lower HLA-G expression are associated with a worse OS both in the HLA-G 2 level and 3 level models.

Table 9. Significant associations between OS and the combination of *HLA-G* 3'UTR haplotypes in secretor models and *UGT1A1* rs8175347 alleles.

UGT1A1* + HLA-G 2 levels	Combo classes	n (%)	Model	HR*	95% CI*	*p*§
*28M	2 copies	8 (3.2)	1 vs. 0	0.68	0.47-0.98	0.0409
	1 copy	**101 (40.7)**				
	0#	**139 (56.1)**				

UGT1A1 + **HLA-G 3 levels**	**Combo classes**	**n (%)**	**Model**	**HR***	**95% CI***	***p*[§]**
*28S	2 copies	0 (0.0)	1 vs. 0	2.53	1.11-5.77	0.027
	1 copy	**7 (2.8)**				
	0[#]	**241 (97.2)**				

L: low level of HLA-G expression in 2 levels model; M: medium/high level of HLA-G expression in 2 levels model; H: high level of HLA-G expression in 3 levels model; I: intermediate level of HLA-G expression in 3 levels model; S: low level of HLA-G expression in 3 levels model; *1: rs8175347-*(TA)*$_6$ wild type allele for *UGT1A1*; *28: rs8175347-*(TA)*$_7$ mutated allele for *UGT1A1;* Combo: combined.
*Estimated from Cox proportional hazard model, adjusted for radical surgery. The HR reported is related to the significant model described in the "Model" column.
§Only the statistically significant data are reported (p < 0.05).
#Reference category.

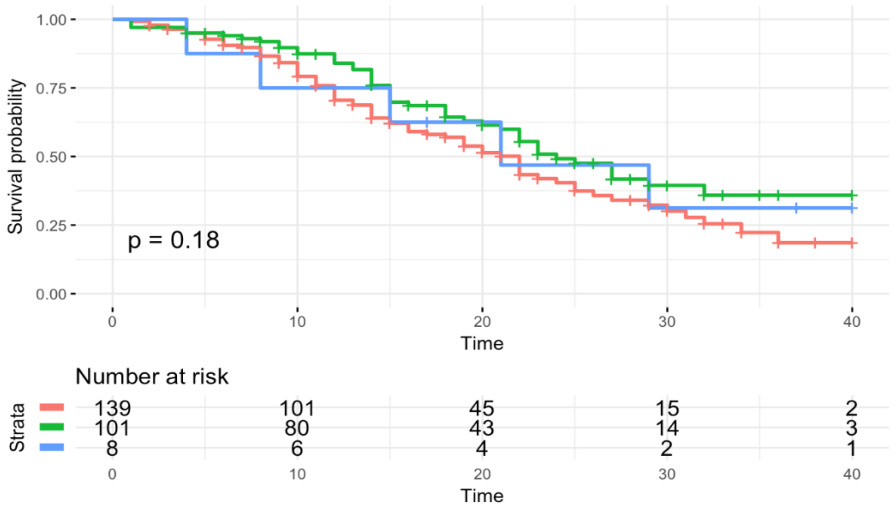

Figure 8. Kaplan-Meier curves for OS. In red: *28M=0; in green: *28M=1; in blue: *28M=2.

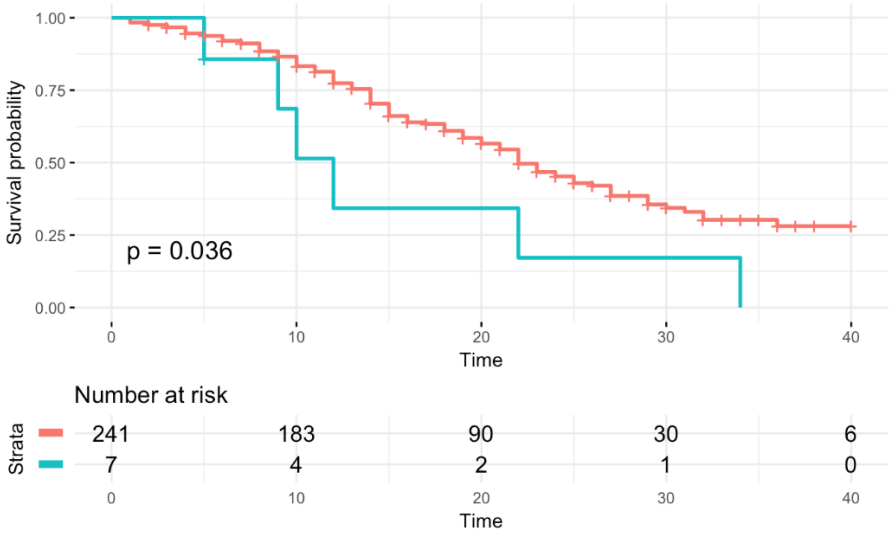

Figure 9. Kaplan-Meier curves for OS. In red: *28S=0; in blue: *28S=1.

In patients with the heterozygous *UGT1A1*1/*28* profile, we identified a trend towards an association between the presence of S/S class with a worse OS compared to I/I and also H/H classes (HR=3.5, 95% CI: 1.02-12.03, *p*=0.0463 and HR=7.65, 95% CI: 1.26-46.53, *p*=0.0271, respectively) (Table 10).

Table 10. Significant associations between OS and HLA-G secretor models in strata of *UGT1A1* rs8175347 genotypes.

***UGT1A1* genotype**	**HLA-G 3 levels classes**	**n (%)**	**Model**	**HR***	**95% CI***	***p*§**
**1/*28*	H/H	8 (7.1)	S/S vs. I/I	3.5	1.02-12.03	0.0463
	H/I	41 (36.3)				
	H/S	5 (4.4)				
	I/I#	**40 (35.4)**				
	I/S	16 (14.2)				
	S/S	**3 (2.6)**				
**1/*28*	H/H#	**8 (7.1)**	S/S vs. H/H	7.65	1.26-46.53	0.0271
	H/I	41 (36.3)				
	H/S	5 (4.4)				
	I/I	40(35.4)				
	I/S	16 (14.2)				
	S/S	**3 (2.6)**				

H: high level of HLA-G expression in 3 levels model; I: intermediate level of HLA-G expression in 3 levels model; S: low level of HLA-G expression in 3 levels model; *1: rs8175347-$(TA)_6$ wild type allele for *UGT1A1*; *28: rs8175347-$(TA)_7$ mutated allele for *UGT1A1*.

*Estimated from Cox proportional hazard model, adjusted for radical surgery. The HR reported is related to the significant model described in the "Model" column.

§Only the statistically significant data are reported ($p < 0.05$).

#Reference category.

4.2.4.2. Time to progression

In this study, 230 patients were eligible for TTP analysis. Median TTP time was 7.43 months (range: 0.73-41.60), and 204 (89%) documented progressions were recorded. The univariate and multivariate analysis with clinical variables showed an association only between the primary tumor site (right colon vs. left colon) in the TTP analysis (HR=1.4885, 95% CI: 1.0547-2.1007; *p*=0.0236) (Table 11). This association was shown in the Kaplan-Meier (Figure 10, Figure 11).

Table 11. Multivariate analysis for clinical variables for TTP.

Variable	**HR**	**95% CI**	***p*-value**
Gender	0.8576	0.6372-1.1544	0.3111
Age	0.9896	0.9751-1.0043	0.1637
Primary tumor site rectum	1.3670	0.9178-2.0360	0.1241
Primary tumor site right colon	**1.4885**	1.0547-2.1007	**0.0236**
Radical surgery yes	0.8953	0.6144-1.3048	0.5650
Stage at diagnosis II	2.2319	0.6936-7.1818	0.1782
Stage at diagnosis III	2.2758	0.7461-6.9414	0.1484
Stage at diagnosis IV	1.8115	0.6528-5.0266	0.2539
Adiuvant yes	0.7469	0.4571-1.2205	0.2442
Number of metastatic sites	1.0283	0.8884-1.1903	0.7084

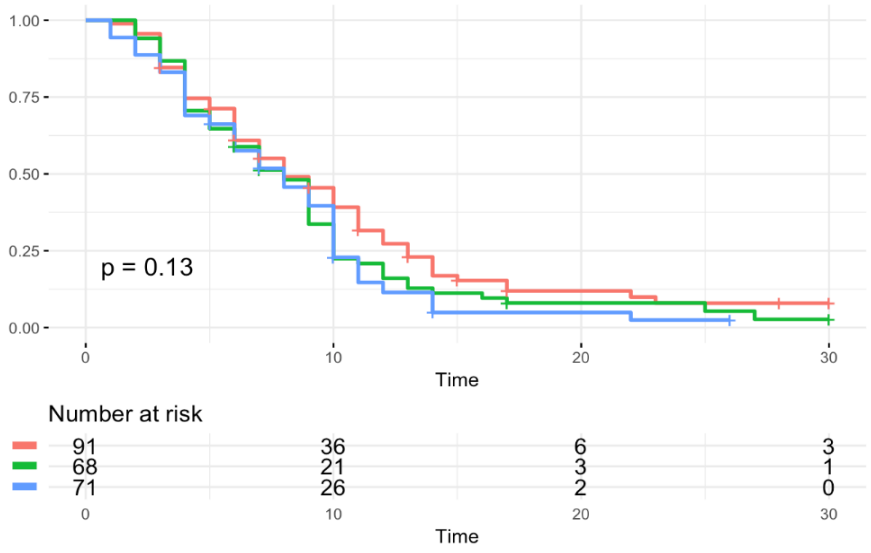

Figure 10. Kaplan-Meier curves for primary tumor site in 248 patients with mCRC. In red: left colon; in green: rectum; in blue: right colon.

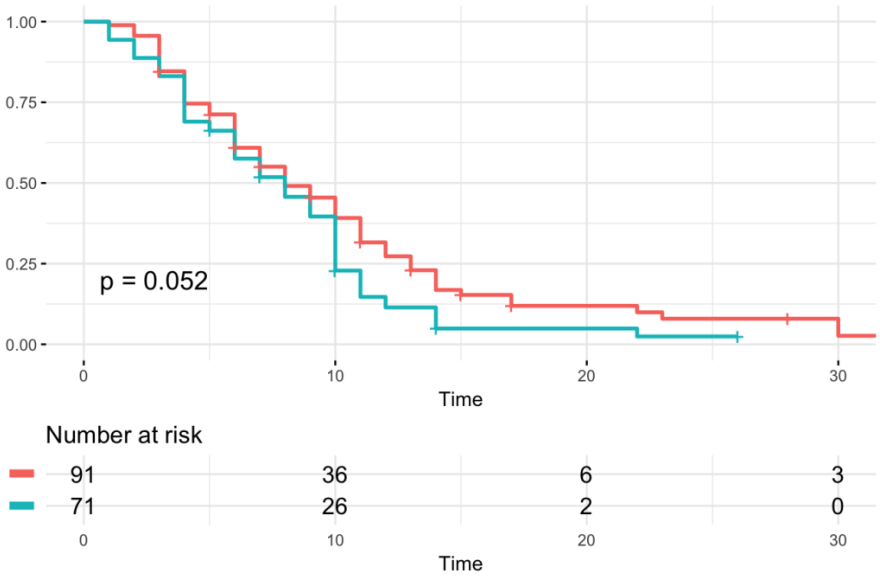

Figure 11. Kaplan-Meier curves for primary tumor site in 248 patients with mCRC. In red: left colon; in blue: right colon.

HLA-G 3'UTR polymorphisms and TTP

No significant data were found (data not shown).

HLA-G secretor models and *UGT1A1* (rs8175347) combination in TTP

In our precedent study, we observed a significant reduction of the risk of progression (TTP longer) for patients with UGT1A1*28/*28 and UGT1A1*28/*1 genotypes compared with the UGT1A1*/*1 genotype [78]. For this reason, the analysis evaluating the combination of *HLA-G* and *UGT1A1* could highlighted a combined effect. We observed that the heterozygous *28M and *28I combined classes were associated with a longer time to progression compared with their absence (HR=0.71, 95% CI: 0.53-0.95; *p*=0.0231 and HR=0.7, 95% CI: 0.53-0.94; *p*=0.0191, respectively), as reported in Table 12. These associations were shown in the Kaplan-Meier curves (Figure 12, Figure 13). A borderline association was found for the combined heterozygous *28H class and a prolonged TTP (HR=0.64, 95% CI: 0.4-1.02; *p*=0.0595) as shown in Figure 14.

Table 12: Significant associations between TTP and the combination of *HLA-G* 3'UTR haplotypes in secretor models and *UGT1A1* rs8175347 alleles.

***UGT1A1* + HLA-G 2 levels**	**Combo classes**	**n (%)**	**Model**	**HR***	**95% CI***	***p*§**
*28M	2 copies	7 (3.0)	1 vs. 0	0.71	0.53-0.95	0.0231
	1 copy	**94 (40.9)**				
	0#	**129 (56.1)**				
***UGT1A1* + HLA-G 3 levels**	**Combo classes**	**n (%)**	**Model**	**HR***	**95% CI***	***p*§**
*28I	2 copies	6 (2.6)	1 vs. 0	0.7	0.53-0.94	0.0191
	1 copy	**101 (43.9)**				
	0#	**123 (53.5)**				

L: low level of HLA-G expression in 2 levels model; M: medium/high level of HLA-G expression in 2 levels model; H: high level of HLA-G expression in 3 levels model; I: intermediate level of HLA-G expression in 3 levels model; S: low level of HLA-G expression in 3 levels model; *1: rs8175347-$(TA)_6$ wild type allele for *UGT1A1*; *28: rs8175347-$(TA)_7$ mutated allele for *UGT1A1;* Combo: combined.
*Estimated from Cox proportional hazard model, adjusted for primary tumor site. The HR reported is related to the significant model described in the "Model" column.
§Only the statistically significant data are reported (p < 0.05).
#Reference category.

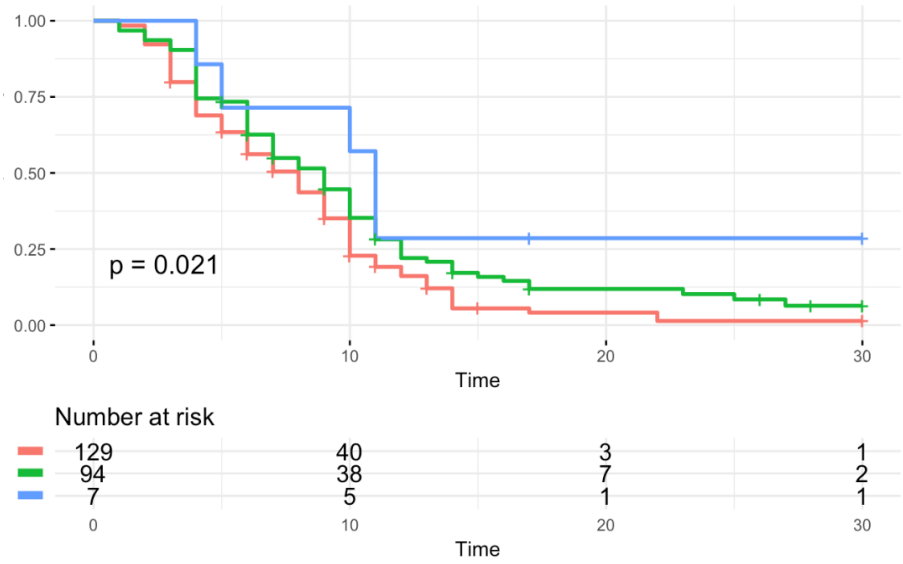

Figure 12.Kaplan-Meier curves for TTP according to combined class *28M. In red: *28M=0; in green: *28M=1; in blue: *28M=2.

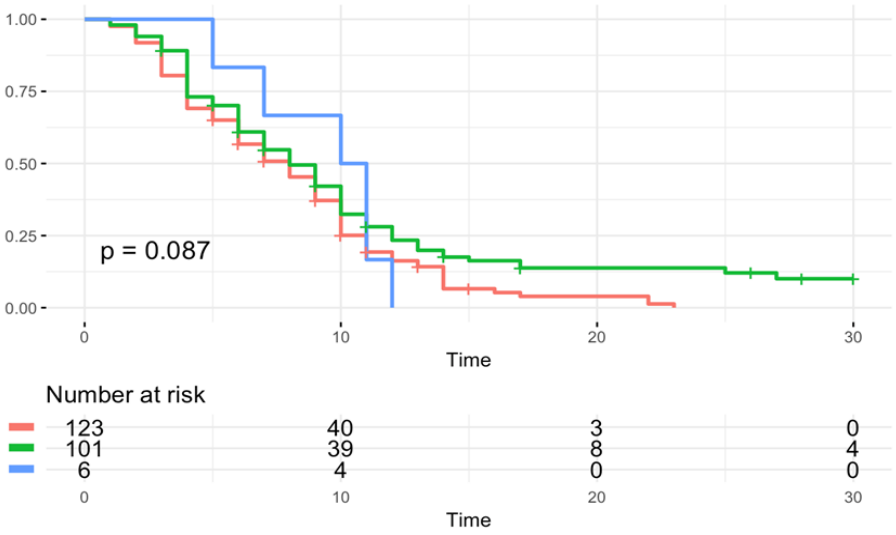

Figure 13. Kaplan-Meier curves for TTP according to combined class *28I. In red: *28I=0; in green: *28I=1; in blue: *28I=2.

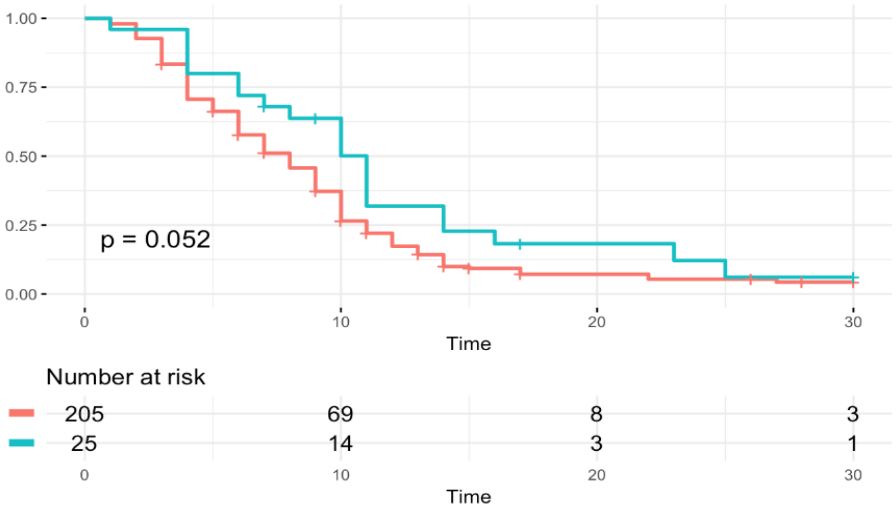

Figure 14. Kaplan-Mayer curves for TTP according to combined class *28H. In red: *28H=0; in blue: *28H=1.

HLA-G secretor models and *UGT1A1* (rs8175347): stratification based on *UGT1A1* genotype in TTP

No significant data were found (data not shown).

4.2.5. Response analysis

The secondary endpoint in PART 1 of this study was the investigation of the relationship between *HLA-G* 3'UTR polymorphisms and tumor response in a set of patients with mCRC after chemotherapy treatment with FOLFIRI regimen. A total of 236 patients were assessable for response analysis as previously described [78]. None of the clinical variables analyzed were significantly associated with the response (Table 13), therefore results were not adjusted for potential confounders.

Table 13. Evaluation of clinical variables with logistic regression models.

Variable	Response Rate (*p*)	Clinical Benefit (*p*)	Complete Response (*p*)
Gender	0.6203	0.4906	0.9619
Age	0.4066	0.2316	0.7904
Primary tumor site	0.8104	0.7419	0.1760
Radical surgery	0.4296	1	1
Stage at diagnosis	0.8916	0.6345	0.6389
Adiuvant	0.4905	1	1
Number of metastatic sites	0.0783	0.1806	0.6354

4.2.5.1. Complete response

Associations between *HLA-G 3'UTR* polymorphisms and CR

We identified significant associations between specific *HLA-G* 3'UTR polymorphisms, haplotypes and phenotypic/secretor models and CR (Table 14). In particular, rs371194629 (+2960 14-bp INDEL), according to the log-additive model, was associated with increased risk of a partial or unfavorable tumor response (SD+PD+PR) (OR=2.34, 95% CI: 1.07-5.12, *p*=0.0453); consequently, the rs371194629-*Del* allele resulted as more frequently associated with a CR. On the contrary, rs9380142 (+3187 A>G), in particular the increase of the *G* alleles for rs9380142 according to the log-additive model, was associated with a favorable tumor response (reduced risk of experiencing SD+PD+PR) (OR=0.33, 95% CI: 0.16-0.69, *p*=0.0031). The rs1710 (+3010 C>G) and rs1063320 (+3142 G>C), in strong LD, were also associated with an increased CR (reduced risk of SD+PD+PR) (OR=0.3, 95% CI:0.14-0.65, *p*=0.001 for both). In particular, these last two polymorphisms were analyzed according to the log-additive model, therefore the increase of the *G* alleles for rs1710, and the *C* allele for rs1063320, were associated with CR. The same correlation with an increased CR was found for the UTR-1 haplotype that includes the rs371194629-*Del*, rs1710-*G*, rs1063320-*C* and rs9380142-*G* alleles: the presence of 2 copies of UTR-1 haplotype was strongly associated with CR (OR=0.1, 95% CI: 0.03-0.41, *p*=0.0013). The genotype-phenotype association was evaluated considering two HLA-G secretor models. We found that the 2 levels model was associated with increased risk of SD+PD+PR) according to the log-additive model. In particular, the progressive increase of L class (corresponding to haplotypes expressing lower levels of HLA-G) was associated with this risk (OR=2.34, 95% CI: 1.07-5.12, *p*=0.0453. These results correlated also with those found for the 3 levels model that showed an association between the classes correlated with higher expression of HLA-G and the CR, especially when 2 copies of H class are present compared to class without H (H/H vs. 0: OR=0.1, 95% CI: 0.03-0.41, *p*=0.0013) (Table 14).

Table 14. Significant associations between CR and *HLA-G* 3'UTR polymorphisms, haplotypes and secretor models.

HLA-G		**CR**	**SD+PD+PR**				
3'UTR SNPs	**Classes**	**n (%)**	**n (%)**	**Model**	**OR***	**95% CI***	***p*§**
+2960 14-bp INDEL	Del/Del#	10 (55.6)	80 (36.7)	log-Additive[1]	2.34	1.07-5.12	0.0453
rs371194629	Ins/Del	8 (44.4)	92 (42.2)				
	Ins/Ins	0 (0.0)	46 (21.1)				
+3010 C>G	C/C#	1 (5.6)	71 (32.6)	log-Additive[2]	0.3	0.14-0.65	0.001
rs1710	G/C	8 (44.4)	104 (47.7)				
	G/G	9 (50.0)	43 (19.7)				
+3142 G>C	G/G#	1 (5.6)	71 (32.6)	log-Additive[3]	0.3	0.14-0.65	0.001
rs1063320	G/C	8 (44.4)	104 (47.7)				
	C/C	9 (50.0)	43 (19.7)				
+3187 A>G	A/A#	5 (27.8)	115 (52.8)	log-Additive[4]	0.33	0.16-0.69	0.0031
rs9380142	A/G	8 (44.4)	91 (41.7)				
	G/G	5 (27.8)	12 (5.5)				
HLA-G		**CR**	**SD+PD+PR**				
haplotypes	**Classes**	**n (%)**	**n (%)**	**Model**	**OR***	**95% CI***	***p*§**
UTR-1	2 copies	5 (27.8)	12 (5.5)	2 vs. 0	0.1	0.03-0.41	0.0013
	1 copy	8 (44.4)	91 (41.7)				
	0#	5 (27.8)	115 (52.8)				
HLA-G secretor		**CR**	**SD+PD+PR**				
model	**Classes**	**n (%)**	**n (%)**	**Model**	**OR***	**95% CI***	***p*§**
2 levels	M/M#	10 (55.6)	80 (36.7)	log-Additive[5]	2.34	1.07-5.12	0.0453
	L/M	8 (44.4)	92 (42.2)				
	L/L	0 (0.0)	46 (21.1)				
3 levels	2 copies	**5 (27.8)**	**12 (5.5)**	2 vs. 0	0.1	0.03-0.41	0.0013
(H class)	1 copy	8 (44.4)	91 (41.7)				
	0#	**5 (27.8)**	**115 (52.8)**				
3 levels	H/H#	**5 (27.8)**	**12 (5.5)**	H/I vs. H/H	4.17	1.17-14.86	0.0278
	H/I	**8 (44.4)**	**80 (36.7)**				
	H/S	0 (0.0)	11 (5.0)				
	I/I	5 (27.8)	75 (34.4)				
	I/S	0 (0.0)	35 (16.1)				
	S/S	0 (0.0)	5 (2.3)				
3 levels	H/H	**5 (27.8)**	**12 (5.5)**	H/H vs. I/I	0.16	0.04-0.64	0.0093
	H/I	8 (44.4)	80 (36.7)				
	H/S	0 (0.0)	11 (5.0)				
	I/I#	**5 (27.8)**	**75 (34.4)**				
	I/S	0 (0.0)	35 (16.1)				
	S/S	0 (0.0)	5 (2.3)				

L: low level of HLA-G expression in 2 levels model; M: medium/high level of HLA-G expression in 2 levels model; H: high level of HLA-G expression in 3 levels model; I: intermediate level of HLA-G expression in 3 levels model; S: low level of HLA-G expression in 3 levels model; CR: complete response; SD: stable disease; PD: progression disease; PR: partial response; log-Additive[1]: Del/Del=0; Ins/Del=1; Ins/Ins=2; log-Additive[2]: C/C=0; C/G=1; G/G=2; log-Additive[3]: G/G=0; C/G=1; C/C=2; log-Additive[4]: A/A=0; A/G=1; G/G=2; log-Additive[5]: M/M=0; L/M=1; L/L=2.

*Estimated from Logistic regression model, not adjusted for any clinical variable. The OR reported is related to the significant model described in the "Model" column.

§Only the statistically significative data are reported ($p < 0.05$).

#Reference category.

HLA-G secretor models and *UGT1A1* (rs8175347) combination in CR

It has been reported that the *UGT1A1*28* allele may influence the response of patients with mCRC [78], therefore we decided to consider if the co-presence with a particular *HLA-G* 3'UTR polymorphism/haplotype could better define the tumor response in these patients. In particular, we analyzed the co-presence of a wild type or mutated *UGT1A1* (rs8175347) allele with a peculiar *HLA-G* 3'UTR haplotype according to the HLA-G secretor models. We observed that the *1M/*1M combination compared to the absence of *1M (2 vs. 0), was associated with a reduced risk of SD+PD+PR (OR=0.24, 95% CI: 0.07-0.81, *p*=0.0209) and the "heterozygous" *1L compared to the absence of *1L (1 vs. 0), was associated with an increased risk of SD+PD+PR (OR=3, 95% CI: 1.03-8.74, *p*=0.044) (Table 15). The same significant trend was confirmed with the HLA-G 3 levels model: the *1H/*1H combination compared to no *1H (2 vs. 0) was associated with a reduced risk of SD+PD+PR (OR=0.09, 95% CI: 0.02-0.46, *p*=0.0037).

Table 15: Significant associations between CR and combination of *HLA-G* 3'UTR haplotypes in secretor models and *UGT1A1* rs8175347 alleles.

UGT1A1+ **HLA-G 2 levels**	**Combo classes**	**CR n (%)**	**SD+PD+PR n (%)**	**Model**	**OR***	**95% CI***	***p*[§]**
*1M	2 copies	7 (38.9)	35 (16.1)	2 vs. 0	0.24	0.07-0.81	0.0209
	1 copy	6 (33.3)	79 (36.2)				
	0[#]	5 (27.8)	104 (47.7)				
*1L	2 copies	0	22 (10.1)	1 vs. 0	3	1.03-8.74	0.044
	1 copy	5 (27.8)	105 (48.2)				
	0[#]	13 (72.2)	91 (41.7)				
UGT1A1+ **HLA-G 3 levels**	**Combo classes**	**CR n (%)**	**SD+PD+PR n (%)**	**Model**	**OR***	**95% CI***	***p*[§]**
*1H	2 copies	3 (16.7)	5 (2.3)	2 vs. 0	0.09	0.02-0.46	0.0037
	1 copy	8 (44.4)	84 (38.5)				
	0[#]	7 (38.9)	129 (59.2)				

L: low level of HLA-G expression in 2 levels model; M: medium/high level of HLA-G expression in 2 levels model; H: high level of HLA-G expression in 3 levels model; *1: rs8175347-$(TA)_6$ wild type allele for *UGT1A1*; *28: rs8175347-$(TA)_7$ mutated allele for *UGT1A1;* CR: complete response; SD: stable disease; PD: progression disease; PR: partial response; Combo: combined.

[*]Estimated from Logistic regression model, not adjusted for any clinical variable. The OR reported is related to the significant model described in the "Model" column.

[§]Only the statistically significant data are reported ($p < 0.05$).

[#]Reference category.

HLA-G secretor models and *UGT1A1* (rs8175347)*:* stratification based on *UGT1A1* genotype in CR

We have subdivided our population in strata by *UGT1A1* (rs8175347) genotype (**1/*1, *1/*28, *28/*28*) and evaluated the different impact on CR of HLA-G secretor models. Only in the **1/*1* stratum was identified a significant association between H/H class with a reduced risk of SD+PD+PR (OR=0.14, 95% CI: 0.02-0.91, *p*=0.0397), compared with I/I class (Table 16). This is mainly a trend towards an association due to the low numbers in each class analyzed.

Table 16. Significant associations between CR and HLA-G secretor models in strata of *UGT1A1* rs8175347 genotypes

***UGT1A1* genotype**	**HLA-G 3 levels classes**	**CR n (%)**	**SD+PD+PR n (%)**	**Model**	**OR***	**95% CI***	***p*§**
*1/*1	H/H	**3 (30.0)**	**5 (5.1)**	H/H vs. I/I	0.14	0.02-0.91	0.0397
	H/I	4 (40.0)	35 (35.7)				
	H/S	0 (0.0)	6 (6.1)				
	I/I#	**3 (30.0)**	**35 (35.7)**				
	I/S	0 (0.0)	15 (15.3)				
	S/S	0 (0.0)	2 (2.1)				

H: high level of HLA-G expression in 3 levels model; I: intermediate level of HLA-G expression in 3 levels model; S: low level of HLA-G expression in 3 levels model; *1: rs8175347-*(TA)*$_6$ wild type allele for *UGT1A1*; CR: complete response; SD: stable disease; PD: progression disease; PR: partial response.

*Estimated from Logistic regression model, not adjusted for any clinical variable. The OR reported is related to the significant model described in the "Model" column.

§Only the statistically significant data are reported (p < 0.05).

#Reference category.

4.2.5.2. Clinical benefit

Association between *HLA-G* 3'UTR polymorphisms and CB

The only significant association was found between the "heterozygous" *HLA-G* UTR-3 haplotype and a reduced risk of SD+PR+CR, i.e CB (OR=0.47, 95% CI: 0.25-0.91, *p*=0.0253) (Table 17). The comparison between the UTR-3/UTR-3 and absence of UTR-3 (2 vs. 0) was not significant (data not shown).

Table 17. Significant associations between CB and *HLA-G* 3'UTR haplotypes.

***HLA-G* haplotypes**	**Classes**	**PD n (%)**	**SD+PR+CR n (%)**	**Model**	**OR***	**95% CI***	***p*§**
UTR-3	2 copies	1 (1.5)	5 (3.0)	1 vs. 0	0.47	0.25-0.91	0.0253
	1 copy	**21 (30.9)**	**29 (17.3)**				
	0#	**46 (67.6)**	**134 (79.8)**				

CR: complete response; SD: stable disease; PD: progression disease; PR: partial response.

*Estimated from Logistic regression model, not adjusted for any clinical variable. The OR reported is related to the significant model described in the "Model" column.

§Only the statistically significant data are reported (p < 0.05).

#Reference category.

HLA-G secretor models and *UGT1A1* (rs8175347) combination in CB

In our previous study, we had analyzed UGT1A1*28/*28 genotype alone and had shown a significantly lower risk of experiencing progression [78]. Considering the combination between the germline setting of the *UGT1A1* (rs8175347) and *HLA-G* 3'UTR polymorphisms, we observed an increased risk of experiencing SD+PD+PR in the presence of *28L haplotype (corresponding to *UGT1A1*28* mutated allele and low expression of HLA-G) according to the dominant model and in present of 1 *28L copy vs. 0 *28L copy (Table 18). Another significant result was the association with a reduced risk of develop CB for *1I/*1I combined class, compared to the absence of *1I (2 vs. 0) (OR=0.41, 95% CI: 0.19-0.92, *p*=0.0305).

Table 18: Significant associations between CB and combination of *HLA-G* 3'UTR haplotypes in secretor models and *UGT1A1* rs8175347 alleles.

***UGT1A1+* HLA-G 2 levels**	**Combo classes**	**PD n (%)**	**SD+PR+CR n (%)**	**Model**	**OR***	**95% CI***	***p*§**
*28L	2 copies	1 (1.5)	2 (1.2)	1 vs. 0	4.52	1.33-15.37	0.0159
	1 copy	**3 (4.4)**	**29 (17.3)**				
	0#	**64 (94.1)**	**137 (81.5)**				
*28L	2 copies	1 (1.5)	2 (1.2)	Dominant[1]	3.62	1.23-10.69	0.0199
	1 copy	**3 (4.4)**	**29 (17.3)**				
	0#	**64 (94.1)**	**137 (81.5)**				
***UGT1A1+* HLA-G 3 levels**	**Combo classes**	**PD n (%)**	**SD+PR+CR n (%)**	**Model**	**OR***	**95% CI***	***p*§**
*1I	2 copies	3 (16.7)	35 (16.1)	2 vs. 0	0.41	0.19-0.92	0.0305
	1 copy	5 (27.8)	85 (39.0)				
	0#	10 (55.6)	98 (44.9)				

L: low level of HLA-G expression in 2 levels model; I: intermediate level of HLA-G expression in 3 levels model; *1:$(TA)_6$ wild type allele for *UGT1A1*; *28: rs8175347-$(TA)_7$ mutated allele for *UGT1A1;* CR: complete response; SD: stable disease; PD: progression disease; PR: partial response; Combo: combined; Dominant[1]: 2 *28L copies + 1 *28L copy vs. 0 *28L copy.

*Estimated from Logistic regression model, not adjusted for any clinical variable. The OR reported is related to the significant model described in the "Model" column.

§Only the statistically significative data are reported ($p < 0.05$).

#Reference category.

***HLA-G* 3'UTR polymorphisms in secretor models and *UGT1A1* (rs81755347): stratification based on *UGT1A1* genotype in CB**

No significant data were found (data not shown).

4.2.5.3. Response rate

No significant associations with RR were found analyzing *HLA-G* 3'UTR polymorphisms/haplotypes/secretor models, and in combination with the *UGT1A1* (rs8175347) alleles and stratifying by *UGT1A1* (rs8175347) genotypes.

4.2.6. Toxicity analysis

Secondary endpoint of this PART 1 was the investigation of the predictive value of the *HLA-G* 3'UTR polymorphisms in the development of severe (G3-4) toxicity after FOLFIRI treatment. All the 248 patients with mCRC were eligible for toxicity analysis. We considered the first cycle and cumulative hematological, non-hematological, total and neutropenia toxicities according to National Cancer Institute Common Toxicity Criteria [218], as reported in the previous study by our group [78]. The possible associations with diarrhea or GI (nausea, vomiting and diarrhea) toxicities were also considered but were finally excluded in the statistical analysis because of the low event numbers (Table 19). No significant covariates emerged for multivariate adjustment of first cycle toxicity analysis, while for cumulative toxicity some clinical variables (gender, primary tumor site and treatment with previous adjuvant therapy) were found to have a significant impact and were used to adjust results in the multivariate analysis (Table 19). However, in our previous study, we had analyzed the relationship between UGT1A1*28 genotype and the same toxicities here considered. It had been reported a significant association between UGT1A1*28 allele and grade 3-4 hematological toxicity during the first cycle of therapy compared with the UGT1A1*/*1 genotype and the same trend was also seen for overall toxicity [78]. On the contrary, the cumulative hematological and non-hematological toxicities had not revealed any clear association with UGT1A1*28 allele. For this reason, the analysis evaluating the combination of *HLA-G* and *UGT1A1* could highlighted a combined effect.

Table 19.Clinical variables analyzed for multivariate adjustment.

	First cycle toxicities				Cumulative toxicities			
Variable	**HT** ***p***	**NHT** ***p***	**TOTAL** ***p***	**NEUTRO** ***p***	**HT** ***p***	**NHT** ***p***	**TOTAL** ***p***	**NEUTRO** ***p***
Gender	0.4236	0.5022	0.1343	0.6786	0.0914	0.0909	**0.0166**	0.2182
Age	0.3491	0.7091	0.5996	0.5680	0.8863	0.0860	0.5899	0.4052
Primary tumor site	0.8831	0.7024	0.9464	0.8433	0.7028	**0.0329**	0.1887	0.5190
Radical surgery	0.4254	1	0.4716	0.3243	1	0.3911	0.7237	0.8003
Stage at diagnosis	0.9638	0.7597	0.8367	0.9688	0.7265	0.8827	0.8228	0.6666
Adjuvant	0.4236	1	0.5881	0.288	0.0914	0.3264	0.0724	**0.0460**
Number of metastatic sites	0.6821	0.6331	0.8236	0.6635	0.5523	0.5153	0.3671	0.3975

HT: hematological toxicity; NHT: non-hematological toxicity; TOTAL: total toxicity; NEUTRO: neutropenia.

4.2.6.1. First cycle toxicity: hematological

***HLA-G* 3'UTR polymorphisms and first cycle hematological toxicity**

No significant data were found (data not shown).

***HLA-G* secretor models and *UGT1A1* (rs8175347) combination in first cycle hematological toxicity**

We observed a trend showing an increase risk of first cycle G3-4 toxicities for patients with *28L/*28L combined class compared to the absence of *28L (OR=58.29, 95% CI: 4.71-721.45, *p*=0.0015). The same trend was seen for the *28S class, where the heterozygous *28S showed an increased risk of association

with first cycle G3-4 toxicities compared with the absence of *28S (OR=9.24, 95% CI: 1.59-53.59, *p*=0.0132) (Table 20).

Table 20. Significant associations between first cycle hematological toxicity and the combination of *HLA-G* 3'UTR haplotypes in secretor models and *UGT1A1* rs8175347 alleles.

***UGT1A1+* HLA-G 2 levels**	**Combo classes**	**G0-2 n (%)**	**G3-4 n (%)**	**Model**	**OR***	**95% CI***	***p*§**
*28L	2 copies	**1 (0.4)**	**2 (16.7)**	2 vs. 0	58.29	4.71-721.45	0.0015
	1 copy	31 (13.1)	3 (0.3)				
	0#	**204 (86.4)**	**7 (58.3)**				
***UGT1A1+* HLA-G 3 levels**	**Combo classes**	**G0-2 n (%)**	**G3-4 n (%)**	**Model**	**OR***	**95% CI***	***p*§**
*28S	2 copies	0 (0.0)	0 (0.0)	1 vs. 0	9.24	1.59-53.59	0.0132
	1 copy	5 (2.1)	2 (16.7)				
	0#	231 (97.9)	10 (83.3)				

L: low level of HLA-G expression in 2 levels model; S: low level of HLA-G expression in 3 levels model; *28: rs8175347-*(TA)*$_7$ mutated allele for *UGT1A1;* G0-2: toxicity grade 0-2; G3-4: toxicity grade 3-4; Combo: combined.

*Estimated from Logistic regression model, not adjusted for any clinical variable. The OR reported is related to the significant model described in the "Model" column.

§Only the statistically significative data are reported (p < 0.05).

#Reference category.

***HLA-G* 3'UTR polymorphisms in secretor models and *UGT1A1* (rs8175347): stratification based on *UGT1A1* genotype in first cycle hematological toxicities**

In the subdivided population based on the *UGT1A1* (rs8175347) genotype, we observed a trend showing how in the population strata homozygous for mutated *UGT1A1**28/*28, the increasing of the L "allele" was associated with an increased risk of first cycle G3-4 hematological toxicities (Table 21).

Table 21. Significant associations between first cycle hematological toxicities and HLA-G secretor models in strata of *UGT1A1* rs8175347 genotypes.

UGT1A1 genotype	**HLA-G 2 levels classes**	**G0-2 n (%)**	**G3-4 n (%)**	**Model**	**OR***	**95% CI***	***p*§**
*28/*28	M/M#	8 (42.1)	0 (0.0)	log-Additive[1]	22.63	1.17-439.49	0.0377
	M/L	10 (52.6)	1 (33.3)				
	L/L	1 (5.3)	2 (66.7)				

L: low level of HLA-G expression in 2 levels model; M: medium/high level of HLA-G expression in 2 levels model; *28: rs8175347-*(TA)*$_7$ mutated allele for *UGT1A1;* G0-2: toxicity grade 0-2; G3-4: toxicity grade 3-4; log-Additive[1]: M/M=0; M/L=1; L/L=2.

*Estimated from Logistic regression model, not adjusted for any clinical variable. The OR reported is related to the significant model described in the "Model" column.

§Only the statistically significative data are reported (p < 0.05).

#Reference category.

4.2.6.2. First cycle toxicity: non-hematological

Associations between *HLA-G* 3'UTR polymorphisms and first cycle non-hematological toxicity

No significant data were found (data not shown).

HLA-G secretor models and *UGT1A1* (rs8175347) combination in first cycle non-hematological toxicity

The results obtained for *28L in the hematological toxicitiy was confirmed also for the non-hematological toxicity where the heterozygous *28L combined class compared to the absence of *28L, showed an increased risk of first cycle non-hematological toxicities (OR=4.56, 95% CI: 1.21-17.09, *p*=0.0246) (Table 22).

Table 22. Significant associations between first cycle non-hematological toxicities and the combination of *HLA-G* 3'UTR haplotypes in secretor models and *UGT1A1* rs8175347 alleles.

***UGT1A1*+ HLA-G 2 levels**	**Combo classes**	**G0-2 n (%)**	**G3-4 n (%)**	**Model**	**OR***	**95% CI***	***p*§**
*28L	2 copies	3 (1.3)	0 (0.0)	1 vs. 0	4.56	1.21-17.09	0.0246
	1 copy	**30 (12.6)**	**4 (40.0)**				
	0#	**205 (86.1)**	**6 (60.0)**				

L: low level of HLA-G expression in 2 levels model; *28: rs8175347-*(TA)*$_7$ mutated allele for *UGT1A1;* G0-2: toxicity grade 0-2; G3-4: toxicity grade 3-4; Combo: combined.

*Estimated from Logistic regression model, not adjusted for any clinical variable. The OR reported is related to the significant model described in the "Model" column.

§Only the statistically significative data are reported ($p < 0.05$).

#Reference category.

HLA-G secretor models and *UGT1A1* (rs8175347): stratification based on *UGT1A1* genotype in first cycle non-hematological toxicity

No significant data were found (data not shown).

4.2.6.3. First cycle toxicity: total

Associations between *HLA-G* 3'UTR polymorphisms and first cycle total toxicity

Regarding the correlation between *HLA-G* 3'UTR polymorphisms/haplotypes and first cycle total toxicities, some interesting results emerged (Table 23). The *HLA-G* rs9380142 (+3187 A>G) polymorphism was associated with a reduced risk of G3-4 first cycle total toxicity (OR=0.38, 95% CI: 0.14-1.02, *p*=0.0447), according to the dominant model; this means that the rs9380142-*G* allele was associated with a reduced risk of development of G3-4 first cycle total toxicity. The presence of *HLA-G* UTR-1 haplotype in one copy compared to its absence (1 vs. 0) was associated with a reduced risk of G3-4 first cycle total toxicity (OR=0.22, 95% CI: 0.06-0.78, *p*=0.0192); the same result was confirmed in the HLA-G 3 levels model with the H class. We also identified an association with reduced risk of G3-4 first cycle total toxicity in the presence of HLA-G 3 levels H/I class, compared to the H/H class (OR=0.17, 95% CI: 0.03-0.9, *p*=0.0378).

Table 23. Significant associations between first cycle total toxicity and *HLA-G* 3'UTR polymorphisms, haplotypes and secretor models.

***HLA-G* 3'UTR SNPs**	**Classes**	**G0-2 n (%)**	**G3-4 n (%)**	**Model**	**OR***	**95% CI***	***p*§**
+3187 A>G	A/A#	**111 (48.9)**	**15 (71.4)**	Dominant[1]	0.38	0.14-1.02	0.0447
rs9380142	A/G+G/G	**116 (51.1)**	**6 (28.6)**				

***HLA-G* Haplotype**	**Classes**	**G0-2 n (%)**	**G3-4 n (%)**	**Model**	**OR***	**95% CI***	***p*§**
UTR-1	2 copies	15 (6.6)	3 (14.3)	1 vs. 0	0.22	0.06-0.78	0.0192
	1 copy	**101 (44.5)**	**3 (14.3)**				
	0#	**111 (48.9)**	**15 (71.4)**				
HLA-G secretor model	**Classes**	**G0-2 n (%)**	**G3-4 n (%)**	**Model**	**OR***	**95% CI***	***p*§**
3 levels (H class)	2 copies	15 (6.6)	3 (14.3)	1 vs. 0	0.22	0.06-0.78	0.0192
	1 copy	**101 (44.5)**	**3 (14.3)**				
	0#	**111 (48.9)**	**15 (71.4)**				
3 levels	H/H#	**15 (6.6)**	**3 (14.3)**	H/I vs. H/H	0.17	0.03-0.9	0.0378
	H/I	**90 (39.6)**	**3 (14.3)**				
	H/S	11 (4.8)	0 (0.0)				
	I/I	77 (33.9)	8 (38.1)				
	I/S	29 (12.8)	7 (33.3)				
	S/S	5 (2.2)	0 (0.0)				

H: high level of HLA-G expression in 3 levels model; I: intermediate level of HLA-G expression in 3 levels model; S: low level of HLA-G expression in 3 levels model; G0-2: toxicity grade 0-2; G3-4: toxicity grade 3-4; Dominant[1]: (GG+AG) vs AA.

*Estimated from Logistic regression model, not adjusted for any clinical variable. The OR reported is related to the significant model described in the "Model" column.

§Only the statistically significative data are reported ($p < 0.05$).

#Reference category.

HLA-G secretor models and *UGT1A1* (rs8175347) combination in first cycle total toxicity

The present of the combined *28L class showed an increased risk of G3-4 first cycle total toxicity, both in the 2 copies and 1 copy vs. none models (OR=30.46, 95% CI: 2.59-358.4, *p*=0.0066 and OR=3.26, 95% CI: 1.15-9.28, *p*=0.0265, respectively) (Table 24). The same trend was identified in the *28S class where the co-presence of mutated *UGT1A1*28* allele and low-level class of HLA-G 3 level model was associated in heterozygous to an increased risk of G3-4 first cycle total toxicity compared to absence of *28S (OR=9.29, 95% CI: 1.93-44.76, *p*=0.0055). On the contrary, the heterozygous *1H class was associated with a reduced risk of G3-4 first cycle total toxicity (OR=0.26, 95% CI: 0.07-0.9, *p*=0.0343).

Table 24. Significant associations between first cycle total toxicity and the combination of *HLA-G* 3'UTR haplotypes in secretor models and *UGT1A1* rs8175347 alleles.

***UGT1A1+* HLA-G 2 levels**	**Combo classes**	**G0-2 n (%)**	**G3-4 n (%)**	**Model**	**OR***	**95% CI***	***p*§**
*28L	2 copies	**1 (0.4)**	**2 (9.5)**	2 vs. 0	30.46	2.59-358.4	0.0066
	1 copy	28 (12.3)	6 (28.6)				
	0#	**198 (87.2)**	**13 (61.9)**				
*28L	2 copies	1 (0.4)	2 (9.5)	1 vs. 0	3.26	1.15-9.28	0.0265
	1 copy	**28 (12.3)**	**6 (28.6)**				
	0#	**198 (87.2)**	**13 (61.9)**				
***UGT1A1+* HLA-G 3 levels**	**Combo classes**	**G0-2 n (%)**	**G3-4 n (%)**	**Model**	**OR***	**95% CI***	***p*§**
*1H	2 copies	7 (3.1)	2 (9.5)	1 vs. 0	0.26	0.07-0.9	0.0343
	1 copy	93 (40.1)	3 (14.3)				
	0#	127 (55.9)	16 (76.2)				

*28S	2 copies	0	0	1 vs. 0	9.29	1.93-44.76	0.0055
	1 copy	4 (1.8)	3 (14.3)				
	0[#]	223 (98.2)	18 (85.7)				

L: low level of HLA-G expression in 2 levels model; H: high level of HLA-G expression in 3 levels model; S: low level of HLA-G expression in 3 levels model; *1: rs8175347-*(TA)*$_6$ wild type allele for *UGT1A1*; *28: rs8175347-*(TA)*$_7$ mutated allele for *UGT1A1*; G0-2: toxicity grade 0-2; G3-4: toxicity grade 3-4; Combo: combined.

[*]Estimated from Logistic regression model, not adjusted for any clinical variable. The OR reported is related to the significant model described in the "Model" column.

[§]Only the statistically significative data are reported (p < 0.05).

[#]Reference category.

HLA-G secretor models and *UGT1A1* (rs8175347): stratification based on *UGT1A1* genotype in first cycle total toxicity

No significant data were found (data not shown).

4.2.6.4. First cycle toxicity: neutropenia

Associations between *HLA-G* 3'UTR polymorphisms and first cycle neutropenia

No significant data were found (data not shown).

HLA-G secretor models and *UGT1A1* (rs8175347) combination in first cycle neutropenia

A trend towards an association with an increased risk of G3-4 neutropenia was found for the homozygous *28L/*28L and for the heterozygous *28S combined classes compared to absence of *28L and *28S respectively (OR=68.33, 95% CI: 5.42-861.26; *p*=0.0011 and OR=10.31, 95% CI: 1.76-60.53, *p*=0.0098) (Table 25).

Table 25. Significant associations between first cycle neutropenia and the combination of HLA-G secretor model and *UGT1A1* rs8175347 alleles.

***UGT1A1* + HLA-G 2 levels**	**Combo classes**	**G0-2 n (%)**	**G3-4 n (%)**	**Model**	**OR***	**95% CI***	***p*[§]**
*28L	2 copies	**1 (0.4)**	**2 (18.2)**	2 vs. 0	68.33	5.42-861.26	0.0011
	1 copy	31 (13.1)	3 (27.3)				
	0[#]	**205 (86.5)**	**6 (54.5)**				
***UGT1A1* + HLA-G 3 levels**	**Combo classes**	**G0-2 n (%)**	**G3-4 n (%)**	**Model**	**OR***	**95% CI***	***p*[§]**
*28S	2 copies	0 (0.0)	0 (0.0)	1 vs. 0	10.31	1.76-60.53	0.0098
	1 copy	5 (2.1)	2 (18.2)				
	0[#]	232 (97.9)	9 (81.8)				

L: low level of HLA-G expression in 2 levels model; S: low level of HLA-G expression in 3 levels model; *28: rs8175347-*(TA)*$_7$ mutated allele for *UGT1A1*; G0-2: toxicity grade 0-2; G3-4: toxicity grade 3-4; Combo: combined.

[*]Estimated from Logistic regression model, not adjusted for any clinical variable. The OR reported is related to the significant model described in the "Model" column.

[§]Only the statistically significative data are reported (p < 0.05).

[#]Reference category.

HLA-G secretor models and *UGT1A1* (rs8175347): stratification based on *UGT1A1* genotype in first cycle neutropenia

In our patients carried for homozygous mutated *UGT1A1*28/*28* genotype, an association between HLA-G 2 levels model, according to log-additive model, and an increased risk of G3-4 first cycle neutropenia, was observed (OR=22.63, 95% CI: 1.17-439.49, *p*=0.0377) (Table 26).

Table 26. Significant associations between first cycle neutropenia and HLA-G secretor models in strata of *UGT1A1* rs8175347 genotypes.

***UGT1A1* genotype**	**HLA-G 2 levels model**	**G0-2 n (%)**	**G3-4 n (%)**	**Model**	**OR***	**95% CI***	***p*§**
*28/*28	M/M#	8 (42.1)	0 (0.0)	log-Additive[1]	22.63	1.17-439.49	0.0377
	M/L	10 (52.6)	1 (33.3)				
	L/L	1 (5.3)	2 (66.7)				

L: low level of HLA-G expression in 2 levels model; M: medium/high level of HLA-G expression in 2 levels model; *28: rs8175347-*(TA)*$_7$ mutated allele for *UGT1A1;* G0-2: toxicity grade 0-2; G3-4: toxicity grade 3-4; log-Additive[1]: M/M=0, M/L=1, L/L=2.
*Estimated from Logistic regression model, not adjusted for any clinical variable. The OR reported is related to the significant model described in the "Model" column.
§Only the statistically significative data are reported (p < 0.05).
#Reference category.

4.2.6.5. Cumulative toxicity: hematological

Association between *HLA-G* 3'UTR polymorphisms and cumulative hematological toxicity

All significant associations found were reported in Table 27. The *HLA-G* 3'UTR rs1707 (+3003 T>C) genotype, according to the log-additive model, was associated with an increased risk of G3-4 cumulative hematological toxicity (OR=2.02, 95% CI: 1.09-3.74, *p*=0.0315). The *HLA-G* 3'UTR rs1710 (+3010 C>G) and *HLA-G* 3'UTR rs1063320 (+3142 G>C) genotypes, according to the recessive model and in perfect LD (r^2=1), were also associated with an increased risk of G3-4 cumulative hematological toxicity (OR=2.52, 5% CI: 1.2-5.33, *p*=0.0182). Moreover, the *HLA-G* UTR-4 haplotype, in 2 copies, was associated with an increased risk of G3-4 cumulative hematological toxicity compared to the absence of UTR-4 (2 vs. 0) (OR=6.87, 95% CI. 1.31-36.03, *p*=0.0226).

Table 27. Significant associations between cumulative hematological toxicity and *HLA-G* 3'UTR polymorphisms and haplotypes.

***HLA-G* 3'UTR SNPs**	**Classes**	**G0-2 n (%)**	**G3-4 n (%)**	**Model**	**OR***	**95% CI***	***p*§**
+3003 T>C	T/T#	165 (78.2)	24 (64.9)	log-Additive[1]	2.02	1.09-3.74	0.0315
rs1707	T/C	43 (20.4)	10 (27.0)				
	C/C	3 (1.4)	3 (8.1)				
+3010 C>G	G/G	41 (19.4)	14 (37.8)	Recessive[2]	2.52	1.2-5.33	0.0182
rs1710	(C/C +C/G)#	170 (80.6)	23 (62.2)				
+3142 G>C	C/C	41 (19.4)	14 (37.8)	Recessive[3]	2.52	1.2-5.33	0.0182
rs1063320	(G/G+G/C)#	170 (80.6)	23 (62.2)				

HLA-G Haplotype	Classes	G0-2 n (%)	G3-4 n (%)	Model	OR*	95% CI*	*p*§
UTR-4	2 copies	3 (1.4)	3 (8.1)	2 vs. 0	6.87	1.31-36.03	0.0226
	1 copy	43 (20.4)	10 (27.0)				
	0#	165 (78.2)	24 (64.9)				

G0-2: toxicity grade 0-2; G3-4: toxicity grade 3-4; log-Additive[1]: T/T=0, T/C=1, C/C=2; Recessive[2]: G/G vs. (C/C+C/G); Recessive[3]: C/C vs. (G/G+G/C).

*Estimated from Logistic regression model, not adjusted for any clinical variable. The OR reported is related to the significant model described in the "Model" column.

§Only the statistically significant data are reported (p < 0.05).

#Reference category.

HLA-G secretor models and *UGT1A1*28* combination in cumulative hematological toxicity

A trend of association of *28L/*28L combined class with the increased risk of G3-4 cumulative hematological toxicity compared to the absence of *28L were observed (OR=12.07, 95% CI: 1.06-137.25, *p*=0.0447) (Table 28). The heterozygous *28I combined class was also associated with an increased risk of G3-4 cumulative hematological toxicity (OR=2.14, 95% CI: 1.04-4.4, *p*=0.0381).

Table 28. Significant associations between cumulative hematological toxicity and the combination of HLA-G secretor models and *UGT1A1* rs8175347 alleles.

UGT1A1+ HLA-G 2 levels	Combo classes	G0-2 n (%)	G3-4 n (%)	Model	OR*	95% CI*	*p*§
*28L	2 copies	**1 (0.5)**	**2 (5.4)**	2 vs. 0	12.07	1.06-137.25	0.0447
	1 copy	29 (13.7)	5 (13.5)				
	0#	**181 (85.8)**	**30 (81.1)**				
***UGT1A1+* HLA-G 3 levels**	**Combo classes**	**G0-2 n (%)**	**G3-4 n (%)**	**Model**	**OR***	**95% CI***	***p*§**
*28I	2 copies	6 (2.8)	0 (0.0)	1 vs. 0	2.14	1.04-4.4	0.0381
	1 copy	**89 (42.2)**	**23 (62.2)**				
	0#	**116 (55.0)**	**14 (37.8)**				

L: low level of HLA-G expression in 2 levels model; I: intermediate level of HLA-G expression in 3 levels model; *28: rs8175347-$(TA)_7$ mutated allele for *UGT1A1;* G0-2: toxicity grade 0-2; G3-4: toxicity grade 3-4; Combo: combined.

*Estimated from Logistic regression model, not adjusted for any clinical variable. The OR reported is related to the significant model described in the "Model" column.

§Only the statistically significant data are reported (p < 0.05).

#Reference category.

HLA-G secretor models and *UGT1A1* (rs8175347): stratification based on *UGT1A1* genotype in cumulative hematological toxicity

No significant data were found (data not shown).

4.2.6.6. Cumulative toxicity: non-hematological

Associations between *HLA-G* 3'UTR polymorphisms and cumulative non-hematological toxicity

The presence of the heterozygous S class was associated with an increased risk of G3-4 cumulative non-hematological toxicity as reported in the Table 29.

Table 29. Significant associations between cumulative non-hematological toxicity and HLA-G secretor models.

HLA-G secretor model	Classes	G0-2 n (%)	G3-4 n (%)	Model	OR*	95% CI*	$p^{§}$
3 levels	2 copies	5 (2.4)	0	1 vs. 0	2.52	1.16-5.48	0.0196
(S class)	1 copy	**34 (16.5)**	**13 (31.0)**				
	0[#]	**167 (81.1)**	**29 (69.0)**				

S: low level of HLA-G expression in 3 levels model; G0-2: toxicity grade 0-2; G3-4: toxicity grade 3-4.

*Estimated from Logistic regression model, adjusted for primary tumor site. The OR reported is related to the significant model described in the "Model" column.

§Only the statistically significative data are reported ($p < 0.05$).

#Reference category.

HLA-G secretor models and *UGT1A1* (rs8175347) combination in cumulative non-hematological toxicity

We observed that the heterozygous *1S combined class was associated with an increased risk of G3-4 cumulative non-hematological toxicity compared to the absence of *1S (Table 30).

Table 30. Significant associations between cumulative non-hematological toxicity and the combination of HLA-G secretor models and *UGT1A1* rs8175347 alleles.

UGT1A1+ HLA-G 3 levels	Combo classes	G0-2 n (%)	G3-4 n (%)	Model	OR*	95% CI*	$p^{§}$
*1S	2 copies	2 (1.0)	0 (0.0)	1 vs. 0	2.34	1.06-5.18	0.0356
	1 copy	**34 (16.5)**	**12 (28.6)**				
	0[#]	**170 (82.5)**	**30 (71.4)**				

S: low level of HLA-G expression in 3 levels model; *1: rs8175347-*(TA)*$_6$ wild type allele for *UGT1A1*; G0-2: toxicity grade 0-2; G3-4: toxicity grade 3-4; Combo: combined.

*Estimated from Logistic regression model, adjusted for primary tumor site. The OR reported is related to the significant model described in the "Model" column.

§Only the statistically significative data are reported ($p < 0.05$).

#Reference category.

HLA-G secretor models and *UGT1A1* (rs8175347): stratification based on *UGT1A1* genotype in cumulative non-hematological toxicity

In the homozygous strata of wild type *UGT1A1*1/*1* patients, it was identified an association between the HLA-G 2 level model, according to the log-additive model, and an increased risk of G3-4 cumulative non-hematological toxicity (Table 31).

Table 31: Significant associations between cumulative non-hematological toxicity and HLA-G secretor models in strata of *UGT1A1* rs8175347 genotypes.

***UGT1A1* genotype**	**HLA-G 2 levels**	**G0-2 n (%)**	**G3-4 n (%)**	**Model**	**OR***	**95% CI***	***p*§**
*1/*1	M/M	39 (41.1)	4 (22.2)	log-Additive[1]	2.09	1.04-4.19	0.0333
	M/L	40 (42.1)	7 (38.9)				
	L/L	16 (16.8)	7 (38.9)				

L: low level of HLA-G expression in 2 levels model; M: medium/high level of HLA-G expression in 2 levels model; *1: rs8175347-*(TA)*$_6$ wild type allele for *UGT1A1;* G0-2: toxicity grade 0-2; G3-4: toxicity grade 3-4; log-Additive[1]: M/M=0, M/L=1, L/L=2.

*Estimated from Logistic regression model, adjusted for primary tumor site. The OR reported is related to the significant model described in the "Model" column.

§Only the statistically significative data are reported (p < 0.05).

#Reference category.

4.2.6.7. Cumulative toxicity: total

Association between *HLA-G* 3'UTR polymorphisms and cumulative total toxicity

The significant associations found were reported in Table 32. The *HLA-G* 3'UTR rs9380142 (+3187 A>G) genotype according to the dominant model was associated with a reduced risk of G3-4 cumulative total toxicity (OR=0.5, 95% CI: 0.27-0.89, *p*=0.0181). The *HLA-G* UTR-1 haplotype, when present in single copy, showed an association with reduced risk of G3-4 cumulative total toxicity compared to the absence of UTR-1 (OR=0.44, 95% CI: 0.23-0.82, *p*=0.0102). On the contrary, the UTR-6 in single copy (1 vs. 0), was associated with an increased risk of G3-4 cumulative total toxicity (OR=3.34, 95% CI: 1.24-9.02, *p*=0.0173). Moreover, we observed a reduced risk of G3-4 cumulative total toxicity in presence of HLA-G 3 levels model with the heterozygous H class (OR=0.44, 95% CI: 0.23-0.82, *p*=0.0102)(Table 32).

Table 32: Significant associations between cumulative total toxicity and *HLA-G* 3'UTR polymorphisms, haplotypes and secretor models.

***HLA-G* 3'UTR SNPs**	**Classes**	**G0-2 n (%)**	**G3-4 n (%)**	**Model**	**OR***	**95% CI***	***p*§**
+3187 A>G	A/A#	85 (47.0)	41 (61.2)	Dominant[1]	0.5	0.27-0.89	0.0181
rs9380142	A/G+G/G	96 (53.0)	26 (38.8)				
***HLA-G* haplotype**	**Classes**	**G0-2 n (%)**	**G3-4 n (%)**	**Model**	**OR***	**95% CI***	***p*§**
UTR-1	2 copies	12 (6.6)	6 (9.0)	1 vs. 0	0.44	0.23-0.82	0.0102
	1 copy	84 (46.4)	20 (29.9)				
	0#	85 (47.0)	41 (61.1)				
UTR-6	2 copies	1 (0.5)	1 (1.5)	1 vs. 0	3.34	1.24-9.02	0.0173
	1 copy	9 (5.0)	9 (13.4)				
	0#	171 (94.5)	57 (85.1)				
HLA-G secretor model	**Classes**	**G0-2 n (%)**	**G3-4 n (%)**	**Model**	**OR***	**95% CI***	***p*§**
3 levels	2 copies	12 (6.6)	6 (9.0)	1 vs. 0	0.44	0.23-0.82	0.0102
(H class)	1 copy	84 (46.4)	20 (29.9)				
	0#	85 (47.0)	41 (61.1)				

H: high level of HLA-G expression in 3 levels model; G0-2: toxicity grade 0-2; G3-4: toxicity grade 3-4; Dominant[1]: (A/G+G/G) vs. A/A.
*Estimated from Logistic regression model, adjusted for gender. The OR reported is related to the significant model described in the "Model" column.
§Only the statistically significative data are reported (p < 0.05).
#Reference category.

HLA-G secretor models and *UGT1A1* (rs8175347) combination in cumulative total toxicity

A significant association was found for heterozygous *1H combined class (1 vs. 0) with a reduced risk of G3-4 cumulative total toxicity (OR=0.45, 95% CI: 0.24-0.85, *p*=0.0145) and for the heterozygous *28I combined class (1 vs. 0) with an increased risk of G3-4 cumulative total toxicity (OR=1.96, 95% CI: 1.09-3.54, *p*=0.0247) (Table 33).

Table 33: Significant associations between cumulative total toxicity and combination of HLA-G secretor models and *UGT1A1* rs8175347 alleles.

***UGT1A1* + HLA-G 3 levels**	**Combo classes**	**G0-2 n (%)**	**G3-4 n (%)**	**Model**	**OR***	**95% CI***	***p*§**
*1H	2 copies	5 (2.8)	4 (6.0)	1 vs. 0	0.45	0.24-0.85	0.0145
	1 copy	**78 (43.1)**	**18 (26.8)**				
	0#	**98 (54.1)**	**45 (67.2)**				
*28I	2 copies	4 (2.2)	2 (3.0)	1 vs. 0	1.96	1.09-3.54	0.0247
	1 copy	**75 (41.4)**	**37 (55.2)**				
	0#	**102 (56.4)**	**28 (41.8)**				

H: high level of HLA-G expression in 3 levels model; I: intermediate level of HLA-G expression in 3 levels model; *1: rs8175347-*(TA)*$_6$ wild type allele for *UGT1A1;* *28: rs8175347-*(TA)*$_7$ mutated allele for *UGT1A1;* G0-2: toxicity grade 0-2; G3-4: toxicity grade 3-4; Combo: combined.
*Estimated from Logistic regression model, adjusted for gender. The OR reported is related to the significant model described in the "Model" column.
§Only the statistically significative data are reported (p < 0.05).
#Reference category.

HLA-G secretor models and *UGT1A1* (rs8175347): stratification based on *UGT1A1* genotype in cumulative total toxicity

No significant data were found (data not shown).

4.2.6.8. Cumulative toxicity: neutropenia

Association between *HLA-G* 3'UTR polymorphisms and cumulative neutropenia

The same polymorphisms of those emerged in the cumulative hematological toxicity emerged in the analysis focused on cumulative neutropenia (Table 34). In particular, *HLA-G* 3'UTR rs1707 (+3003 T>C) genotype, according to the log-additive model, was associated with an increased risk of G3-4 cumulative neutropenia (OR=2.02, 95% CI: 1.06-3.84, *p*=0.0386); the *HLA-G* 3'UTR rs1710 (+3010 C>G) and the *HLA-G* 3'UTR rs1063320 (+3142 G>C) genotypes, according to the recessive model, were both also associated with an increased risk of G3-4 cumulative neutropenia (OR=2.53, 95% CI: 1.16-5.5, *p*=0.023). The *HLA-G* UTR-4

haplotype, in 2 copies, was associated with an increased risk of G3-4 cumulative neutropenia (OR=9.06, 95% CI: 1.66-49.58, *p*=0.011).

Table 34: Significant associations between cumulative neutropenia and *HLA-G* 3'UTR polymorphisms and haplotypes.

***HLA-G* 3'UTR SNPs**	**Classes**	**G0-2 n (%)**	**G3-4 n (%)**	**Model**	**OR***	**95% CI***	***p*§**
+3003 T>C	T/T#	166 (77.9)	23 (65.7)	log-Additive[1]	2.02	1.06-3.84	0.0386
rs1707	T/C	44 (20.7)	9 (25.7)				
	C/C	3 (1.4)	3 (8.6)				
+3010 C>G	G/G	42 (19.7)	13 (37.1)	Recessive[2]	2.53	1.16-5.5	0.023
rs1710	C/C+C/G#	171 (80.3)	22 (62.9)				
+3142 G>C	C/C	42 (19.7)	13 (37.1)	Recessive[3]	2.53	1.16-5.5	0.023
rs1063320	G/G+ G/C#	171 (80.3)	22 (62.9)				
***HLA-G* haplotype**	**Classes**	**G0-2 n (%)**	**G3-4 n (%)**	**Model**	**OR***	**95% CI***	***p*§**
UTR-4	2 copies	3 (1.4)	3 (8.6)	2 vs. 0	9.06	1.66-49.58	0.011
	1 copy	44 (20.7)	9 (25.7)				
	0#	166 (77.9)	23 (65.7)				

G0-2: toxicity grade 0-2; G3-4: toxicity grade 3-4; log-Additive[1]: T/T=0, T/C=1, C/C=2; Recessive[2]: G/G vs. (C/C+C/G); Recessive[3]: C/C vs. (G/G+G/C).

*Estimated from Logistic regression model, adjusted for adiuvant. The OR reported is related to the significant model described in the "Model" column.

§Only the statistically significative data are reported (p < 0.05).

#Reference category.

HLA-G secretor models and *UGT1A1* (rs8175347) combination in cumulative neutropenia

We observed a trend of association between the*28L combined class when presented in 2 copies with an increased risk of G3-4 cumulative neutropenia compared to its absence of *28L (Table 35).

Table 35: Significant associations between cumulative neutropenia and combination of *HLA-G* 3'UTR haplotypes in secretor models and *UGT1A1* rs8175347 alleles.

***UGT1A1* + HLA-G 2 levels**	**Combo classes**	**G0-2 n (%)**	**G3-4 n (%)**	**Model**	**OR***	**95% CI***	***p*§**
*28L	2 copies	**1 (0.5)**	**2 (5.7)**	2 vs. 0	14.25	1.2-169.07	0.0353
	1 copy	29 (13.6)	5 (14.3)				
	0#	**183 (85.9)**	**28 (80.0)**				

L: low level of HLA-G expression in 2 levels model; *28: rs8175347-$(TA)_7$ mutated allele for *UGT1A1;* G0-2: toxicity grade 0-2; G3-4: toxicity grade 3-4; Combo: combined.

*Estimated from Logistic regression model, adjusted for adiuvant. The OR reported is related to the significant model described in the "Model" column.

§Only the statistically significative data are reported (p < 0.05).

#Reference category.

HLA-G secretor models and *UGT1A1* (rs8175347): stratification based on *UGT1A1* genotype in cumulative neutropenia

No significant data were found (data not shown).

4.3.DISCUSSION

The influence of germline polymorphisms of *HLA-G* in the 3'UTR region in determine the clinical outcome of patients with CRC, was first investigated by our group [15]. In particular, some 3'UTR polymorphisms of the *HLA-G* gene were found to be independently associated with prognosis of patients with non-metastatic stage II-III CRC who received primary surgery and chemotherapy treatment based on fluoropyrimidine [226]. In this PART 1, we present results based on 248 patients with metastatic CRC, homogeneously treated with FOLFIRI regimen. We characterized *HLA-G* 3'UTR genotypes and haplotypes and we defined the *HLA-G* haplotypes with a class corresponding to the HLA-G expression levels according to two new defined HLA-G secretor models. Moreover, genotyping for the *UGT1A1* (rs8175347) was considered because of the FDA recommendation of reduction of the starting dose of irinotecan for patients with *UGT1A1*28/*28* (rs8175347-$(TA)_7/(TA)_7$) genotype and our previous reported results [78]. The co-presence of peculiar HLA-G characteristics and a *UGT1A1*28* polymorphims seems to determine a variation in the clinical important outcome we analyzed. The *HLA-G* genetic and phenotypic features, never considered in this set of patients, could assume an important role as novel prognostic and predictive immune biomarkers. We performed statistical analysis to correlate the HLA-G genetic and phenotypic results with clinical parameters such as survival, tumor response and first cycle and cumulative toxicity.

4.3.1. Survival analysis

Despite of the recent increase of the median overall survival due to the introduction of the new targeted therapies, the CRC is still at the first positions for mortality. Then, the discovery of novel prognostic biomarker for OS or TTP could contribute to prolong the patients' survival. The slight deviation from the HWE noticed for the rs371194629-*Ins* allele, shows the same trend of abundance of this mutated allele observed in the previous case-control study in patients with CRC [180]. First of all, we investigated the role of *HLA-G* 3'UTR polymorphisms and haplotypes in both OS and TTP, however, we did not find any significant association. The results differ from those emerged in patients with non-metastatic stage II-III CRC treated with FOLFOX regimen [226]. In non-metastatic patients the rs17179108-*C/C,* the rs9380142-*G/G* and the rs371194629-*Del/Del* genotypes and the UTR-1 haplotype in homozygous state, were associated with a reduced disease free-survival (DFS), defined as the time from date of surgery to date of clinically detectable recurrence (local, regional or distant), death from any cause, or last follow-up evaluation. Moreover, the homozygous mutant rs9380142-*G/G* genotype and the homozygous UTR-1 haplotype, that includes the rs9380142-*G/G,* were associated with reduced OS in non-metastatic CRC patients. The first important observation was that our previously reported HLA-G-related prognostic biomarker for non-metastatic CRC were not confirmed in patient with the metastatic disease, highlighting that they could considered prognostic biomarkers only for early stage CRC [226]. To our knowledge, these results are the only one that face of the prognostic value of 3' UTR *HLA-G* polymorphisms and CRC [16]. Late-stage CRC disease in which the cancer has metastasized to distant site, still have a 5-year survival rate of about 12-15%, while for the early-stage CRC it is about 90% [228], [229]. The poor survival in stage IV CRC maybe conceal the effect we observed in patients with a better prognosis, indicating that other factors are involved in its determination. It was previously reported that the UTR-1 haplotype was associated with higher levels of soluble HLA-G [18]. Then, the genetic variations influence the HLA-G expression and several efforts have been made to define whether it can be associated with tolerance and/or others complications [230]. The levels of the HLA-G increase in pathological conditions having an important role in the modulation of the immune system response against cancer [20]–[22], [24], [150]. To our knowledge, the

evaluation of the possible correlation between particular HLA-G expression level classes, determined on the base of the *HLA-G* genotype, could give us important information about the survival. We had supposed that higher levels of sHLA-G contribute to the tolerance of the malignant cancer reducing the OS in patients with mCRC as we observed in early-stage CRC. However, analyzing only the contribution of *HLA-G* 3'UTR polymorphisms in the HLA-G secretor models, we found an association between the presence of the low-level S/S class and a significant reduced OS compared to the intermediate I/I class (Table 8), according to the 3 levels secretor model. Moreover, this result was supported by the finding of a borderline association between the S class in 2 copies, representing the low secretor UTR-5 and UTR-7 haplotypes, and a worse OS. In the same way, these results regarding the HLA-G secretor models differ from those already published for the non-metastatic CRC. There no data in the literature that could explain this evident inconsistency. However, we observed that the treatment used for the two different population analyzed, are different. In fact, the adiuvant chemotherapy received in the early-stage CRC was based on fluoropyrimidine alone or with oxaliplatin, while the therapeutic regimen of the mCRC was FOLFIRI treatment, irinotecan-based. The possible implication of the HLA-G levels and the presence of irinotecan in the mCRC treatment for the patient's prognosis was considered in the PART 3 of this thesis.

In the previous work from our group on the same cohort of patients, carriers for the *UGT1A1*28/*28* genotype had a significant prolonged TTP while were not associated with an increased OS compared to carriers for the *UGT1A1*1/*1* and *UGT1A1*1/*28* genotypes [78]. The only significant result from the *HLA-G* 3'UTR-*UGT1A1* combined analysis using the 3 levels secretor model, was identified in carriers for at least a *UGT1A1*28* allele and at least one S class haplotype (UTR-5 or UTR-7 (Table 9). Intriguingly, the allelic count for *UGT1A1*28* was inversely and significantly (Chi-square test, $p<0.0001$) correlated in the sub-groups analyzed: 79% in the *28S, and 30% in the reference group (no *28S combination), without carriers for the *UGT1A1*1/*1* genotype in the *28S sub-group with worse OS.

To better define the influence of *UGT1A1*28* allele, we also analyzed the impact of HLA-G secretor models according to the *UGT1A1* genotype, finding significant results only when the HLA-G 3 levels model was applied into the heterozygous *UGT1A1*1/*28* strata were we identified an association of HLA-G low levels S/S class with a significant reduced OS, compared to the intermediate level I/I class. We did not observe the same effect in the others homozygous *UGT1A1* strata, probably cause to the low or null number of patients in the S/S and I/I classes due to the strong stratification. Further, in the same *UGT1A1* heterozygous strata, comparing patients carrying for the low S/S and the HLA-G high levels H/H class, a significant association was still found, even if, these results should be carefully considered due to the low number of patients analyzed. In fact, one of the limits of the used of these two secretor models is the low number of patients per class due to the retrospective nature of this study, particulary evident in the HLA-G 3 levels model that subdivides in 6 classes of HLA-G expression levels the eligible samples.

Concerning the HLA-G 2 levels secretor models, we observed that it perfectly stratified patients according to the *HLA-G* 3'UTR rs371194629, therefore the survival results in terms of TTP and OS, were superimposable to those calculated for rs371194629. However, this secretor model increases the numerosity of HLA-G low levels class within the presence of UTR-2 haplotype and includes the intermediate haplotypes (UTR-3, UTR-4, UTR-6 according to the 3 levels model) into the medium-high level. The only significant result from the *HLA-G* 3'UTR-*UGT1A1* combined analysis using the 2 levels secretor model, was identified in carriers for at least a *UGT1A1*28* allele and at least one M haplotype (UTR-1 or UTR-3 or UTR-4 or UTR-6): the *28M combo class (0/101 patients were *UGT1A1*/*1*) that shared a prolonged OS (Table 9). Again, the allelic count for *UGT1A1*28* was inversely and significantly (Chi-square test, $p<0.0001$) correlated in the sub-groups analyzed: 55% in the *28M, and 10% in the reference group (no *28M combination), without carriers for the *UGT1A1*1/*1* genotype in the *28M sub-group with better OS.

In the previous study, from which our patient's cohort derived, the *UGT1A1* genotyping showed that carriers for the *UGT1A1*28/*28* genotype were associated, but not significantly, with a better OS compared to carriers for the *UGT1A1*1/*1* and *UGT1A1*1/*28* genotypes [76]. Otherwise, the TTP analysis showed a significant survival advantage for patients carrying of the *UGT1A1*28/*28* with respect to the wild type and heterozygous patients with mCRC [76]. The influence of *UGT1A1*, mostly in TTP, should be taken in account in the combined classes here analyzed and identified as prognostic potential markers.

Significant results of TTP were obtained only in the analysis of both HLA-G secretor models combined with *UGT1A1*28* polymorphism and were concordant with those found in OS. Further, the trend with a prolonged TTP and a high HLA-G secretor class, was also confirmed when 2 copies of the M class were coupled in patients homozygous for *UGT1A1*28* allele, even if, not significantly.

In the HLA-G 2 levels secretor model, carriers for the heterozygous *28M combo class shared a prolonged TTP (Table 12). Again, the allelic count for *UGT1A1*28* was inversely and significantly (Chi-square test, $p<0.0001$) correlated in the sub-groups analyzed: 56% in the *28M, and 10% in the reference group (no *28M combination), without (0/94) carriers for the *UGT1A1*1/*1* wild type genotype in the *28M sub-group with better TTP.

In the HLA-G 3 levels secretor model, carriers for the heterozygous *28I combo class shared a prolonged TTP (Table 12). Again, the allelic count for *UGT1A1*28* was inversely and significantly (Chi-square test, $p<0.0001$) correlated in the sub-groups analyzed: 56% in the *28I, and 7% in the reference group (no *28I combination), without (0/101) carriers for the *UGT1A1*1/*1* wild-type genotype in the *28I sub-group with better TTP.

Finally, considering only the contribute of *HLA-G* 3'UTR polymorphisms, a single prognostic marker of worse OS emerged and was related to the more complex HLA-G 3 levels secretor model. The negative impact on OS survival defined by the homozygous UTR-5/UTR-7 (S/S) low levels class in this model, seems to be independent from the effect exerted by *UGT1A1*28 allele, as confirmed in the "strata" analysis. Moreover, considering the combined heterozygous *28S class, and the preferential distribution of the mutated *UGT1A1*28* allele in this sub-group, the strength of HLA-G low secretor 3 levels model was further evidenced. Conversely, when the HLA-G 2 levels secretor model was applied, the protective results in the heterozygous *28M combined class were probably influenced by the presence of the *UGT1A1*28 allele that significantly segregated with the M related high/intermediate levels haplotypes (UTR-1 or UTR-3 or UTR-4 or UTR-6). In TTP, candidate markers emerged only considering the contribution of both HLA-G and *UGT1A1*28* polymorphism, with a similar trend. In the HLA-G 3 levels secretor model analyzed, the combined heterozygous *28I class, and the preferential distribution of the mutated *UGT1A1*28* allele in this sub-group, stratified again the patients with an unfavorable outcome. According to the HLA-G 2 levels secretor model, the protective results in the heterozygous *28M combined class were probably influenced by the presence of the *UGT1A1*28 allele that significantly segregated with the M related high/intermediate levels haplotypes: considering that none of the *HLA-G* 3'UTR polymorphism/haplotypes was found to be significantly associated with TTP, it is probable that the favorable prognostic effect is due to the strength of *UGT1A1*28* allele. Overall, the results highlight an association with a "high" secretor HLA-G with a favorable prognosis, and "low" secretor HLA-G, with a poor and unfavorable prognosis.

However, the more accurate and restrict HLA-G 3 levels classification (6 classes) needs a higher numerosity for each class to reach the statistical significance and solid results.

Our findings highlight that the *HLA-G* 3'UTR polymorphisms categorized in secretor models corresponding to complex haplotype blocks, should be considered as prognostic markers of OS and TTP. Anyway, their validation in a bigger and prospective cohort is required to confirm the prognostic value here found.

Table 36: Summary of the survival results.

	OS		TTP	
	Good (longer OS)	***Poor (shorter OS)***	***Good (longer TTP)***	***Poor (shorter TTP)***
SNP	-	-	-	-
Haplotype	-	-	-	-
HLA-G secretor model	-	S/S vs. I/I	-	-
***UGT1A1* + HLA-G secretor models**	*28M (1 vs. 0 copy)	*28S (1 vs 0 copy)	*28M (1 vs. 0 copy) *28I (1 vs. 0 copy)	-
In stratum UGT1A11/*28* genotype, HLA-G secretor models**	-	S/S vs. I/I S/S vs. H/H	-	-

4.3.2. Response analysis

Evaluation of the effects of *HLA-G* 3'UTR polymorphisms on tumor response was the secondary endpoint of this first part of the study. The tumor shrinkage is an important goal to reach for the treatment of mCRC. Therefore, we evaluated the associations between *HLA-G* 3'UTR polymorphisms, haplotypes, secretor models and complete response. The presence of this insertion/deletions polymorphism in the *HLA-G* 3'UTR region has been widely reported as causing mRNA instability and then controlling post-transcriptional regulation. In particular, the rs371194629-*Ins* allele confers better resistance to mRNA degradation [121] and has been associated with higher TNM stage and early relapse risk in non-metastatic CRC [231]. Moreover, the homozygous rs371194629-*Ins* genotype corresponds to lower plasma levels of HLA-G protein [232], [233]. In our study we observed that the rs371194629-*Ins* allele is associated with an increased risk of partial or unfavorable tumor response, according to the log-additive model; then, the rs371194629-*Del* allele resulted more frequently associated with a CR. In general, the presence of the 14-bp nucleotide sequence has been associated with lower HLA-G levels both membrane-bound and soluble, even if there are controversial results [125], [189], [234]. Moreover, according to the log-additive model, also the increase in the mutant *G* alleles for rs1710 and rs9380142, and in the *C* allele for rs1063320, all polymorphisms in linkage disequilibrium, were associated with CR. Although no specific regulation mechanism has been described in the literature, these polymorphic sites might influence microRNA binding site. In particular, the miR-148a and miR-152, binding a sequence involving the nucleotide variation at position +3142 (rs1063320), downregulate HLA-G expression irrespective of the alleles [122], while others studies reported that the presence of guanine at this position increases the binding affinity of miRNAs leading to mRNA degradation and translation suppression [97], [235]. Regarding the rs9380142, the presence of the A allele lead to a decreased HLA-G expression due to the increased numbers of adenines in the near AU-rich motif that mediates mRNA degradation [119]. Then, the alleles we found associated with CR, suggest the presence of an higher HLA-G expression. The rs371194629-*Del* and rs1710-*G* alleles were found to share a "protective" role in non-metastatic CRC [180]. These data suggest that rs371194629 and rs1710 could be predictive biomarker in both the setting of the disease. The same correlation with an increased CR was found in presence of two *HLA-G* UTR-1 copies that includes the rs371194629-*Del,* rs1710-*G*, rs1063320-*C,* and rs9380142-*G* alleles. The presence of one UTR-3 copy was associated with the reduced risk of CB and then was associated with PD. In the *HLA-G* UTR-3 haplotype is present the rs1710-*C*, rs1063320-*G,* and rs9380142-*A* wild type alleles, not associated with the CR, and the rs371194629-*Del* associated with CR. In the literature, a less intensely downregulation of the gene expression in *in vitro* gene

reporter test was observed in presence of the rs371194629-*Del* allele (UTR-1, UTR-3, UTR-4 and UTR-18) compared with the rs371194629-*Ins* allele (UTR-2, UTR-5 and UTR-7) [225]. However, our findings showed that rs371194629-*Del* allele might not be involved alone in the prediction of response, in particular regarding the results for CB. The protective role of the UTR-1 haplotype in the tumor response was confirmed in both HLA-G secretor models. The *HLA-G* UTR-1 haplotype corresponds to high secretor in both models, especially in the 3 levels model, and the *HLA-G* UTR-3 haplotype corresponds to a high or intermediate HLA-G secretor class based on the 2 or 3 levels model, respectively. From the 2 levels secretor model, the progressive increase of L "low" HLA-G level class was associated with a worse tumor response. The increase of L class corresponds to a reduction of the *HLA-G* UTR-1 and UTR-3 and also the UTR-4 and UTR-6 haplotypes. All these four haplotypes in the M class share the rs371194629-*Del*, rs17179101-*C*, rs17179108-*C* and rs1610696-*C* wild type alleles. The positive association of UTR-1 with CR, was confirmed in the 3 levels secretor model where this haplotype constitutes the high secretor class alone. In particular, comparing the intermediate H/I and I/I classes with the H/H class, an association with a worse CR was highlighted, confirmed the results above mentioned. However, we observed that we obtained the same significant results for CR studying the 2 levels and the rs371194629 and the 3 levels model and the UTR-1 haplotype (see Table 14). These findings suggest that genotypes and haplotypes correlated with low levels of sHLA-G might determine poor clinical response and then overall prognosis in mCRC. In general, it has been reported that high levels of soluble HLA-G correlated with poor prognosis in solid tumors, while it was observed the opposite for hematological tumors [106]. Moreover, it has already been noted in other disease that HLA-G expression could be different in metastatic and non-metastatic carcinoma [110], [192].

An evaluation of the possible influence of the rs8175347 polymorphism in *UGT1A1* gene in the presence of a particular set of *HLA-G* polymorphisms grouped in the secretor level models was also performed. In the previous published study [78], the *UGT1A1*28/*28* genotype was associated with a better tumor response (CR+PR) and a favorable CB (reduced PD) in mCRC. Nevertheless a following meta-analysis reported that the *UGT1A1* genotype globally not impact upon ORR, but in the discussion was specified that in the studies with smaller variance the OR estimate tends to indicate a higher ORR for the *28/*28 group [236]. The most interesting results, were observed for the *1M and *1L combined classes, where the presence of HLA-G M or L secretor class in patients carrier for the *UGT1A1*1/*1* wild-type allele, causes a different predictive value for the CR. Patients with two copies of *1M class were associated with a better CR, while the heterozygous *1L was associated with a worse tumor response (CR). These results suggest that *HLA-G* 3'UTR haplotypes, associated to higher HLA-G levels, may predictive of a CR in patients with mCRC, regardless of *UGT1A1*28* allele. Patients with *UGT1A1*1/*1* genotype, known to be associated with a worse response compared to *UGT1A1*28/*28* genotype carriers, could have a complete response with a high HLA-G secretor level model. Moreover, considering only the wild type patients for rs8175347 (*UGT1A1*1/*1* stratum) according to the 3 levels secretor model, the trend was the same: in particular, patients with the high HLA-G secretor level (H/H class), were associated with CR compared to the I/I class, and the same trend was observed in the *UGT1A1*1/*28* stratum, even if not significant (not shown).

From the results concerning the response after the FOLFIRI treatment, rs371194629-*Del,* rs1710-*G*, rs9380142-*G*, and rs1063320-*C* alleles, included in the UTR-1 haplotype and consequently in the M and H classes of the HLA-G high secretor models, resulted more frequently associated with a CR. The higher HLA-G expression is probably due to a mechanism of less intensely post-transcriptional downregulation, already demostrated [225]. Finally, we also highlighted the predictive role of the HLA-G high secretor level for CR, regardless of the rs8175347 polymorphisms in the *UGT1A1* gene.

Table 37: Summary of the response results.

	CR		CB (PD vs. SD+PR+CR)	
	Good (CR)	***Poor (SD+PD+PR)***	***Good (SD+PR+CR)***	***Poor (PD)***
SNP (genetic model)	rs1710-*G* (log-add) rs1063320-*C* (log-add) rs9380142-*G* (log-add)	rs371194629-*Ins* (log-add)	-	-
Haplotype (model)	UTR-1 (2 vs. 0 copy)	-	UTR-3 (1 vs. 0 copy)	
HLA-G secretor model	H class (2 vs. 0 copy) H/H vs. I/I	L class (log-add) H/I vs. H/H	-	-
***UGT1A1* + HLA-G secretor models**	*1M (2 vs. 0 copy) *1H (2 vs. 0 copy)	*1L (1 vs. 0 copy)	*28L (1 vs. 0 copy)	*28L/*28L (dom) *1I (2 vs. 0 copy)
In stratum UGT1A11/*1* genotype, HLA-G secretor models**		H/H vs. I/I	-	-

4.3.3. Toxicity analysis

Patients with cancer are reluctant in the treatment choice often due to the treatment-related toxicities that affect their quality of life. The risk of developing severe G3-4 toxicity was evaluated considering the first-cycle (first cycle toxicities) and total amount of planned FOLFIRI chemotherapy cycles (cumulative toxicities). In particular, the main and severe toxicities that occur in patients treated with irinotecan-containing regimen such as FOLFIRI regimen, are diarrhea, neutropenia and other toxicities of the GI tract that limit the clinical use of the treatment (nausea, vomiting). Significant associations with specific *HLA-G* 3'UTR polymorphisms were identified in the total toxicity, and in neutropenia and hematological cumulative toxicities. The overall toxicities evaluation derives from the sum of the hematological and non-hematological events. In particular, rs9380142-*G* allele, according to the dominant model, was associated with a reduced risk of G3-4 total toxicity, both at the first cycle and in the cumulative option. Moreover, patients heterozygous for the *HLA-G* UTR-1 haplotype, that includes the rs9380142-*G* allele, were associated with reduced risk of G3-4 total toxicity confirming the value of single SNP. Remembering that the *A* allele instead of *G* allele at position +3187 lead to a decreased HLA-G expression, we could suppose that higher level of HLA-G we had previously demonstrated to be correlated with better survival and response, maybe are linked with the association found with a reduced risk of G3-4 toxicity. On the contrary, the *HLA-G* heterozygous UTR-6 haplotype characterized by the presence of rs9380142-*A* allele, was associated with an increased risk of G3-4 total cumulative toxicity.

Patients with heterozygous UTR-6 haplotype share the contribution of rs1710-*G* and rs1063320-*C* alleles that were associated with increased risk of G3-4 cumulative neutropenia. The rs1707-*C*, rs1710-*G* and rs1063320-*C* alleles were found to be predictive of G3-4 cumulative neutropenia according to the log-additive model for rs1707 and to the recessive model for the other two SNPs (rs1710 and rs1063320). Patients homozygous for *HLA-G* UTR-4 haplotype, that includes the rs1707-*C*, rs1710-*G* and rs1063320-*C* mutated alleles, were associated with increased risk of G3-4 neutropenia confirming the value of single SNPs. Similarly, these polymorphisms and the UTR-4 haplotype were identified as predictors of G3-4

hematological toxicity, in the cumulative toxicity analysis. These findings are explained by the main contribution of neutropenia among the hematological toxicities. To our knowledge, genetic markers predictive of G3-4 toxicity in patients with mCRC treated with irinotecan-containing regimens have not been reported for *HLA-G* 3'UTR polymorphisms. Comparing these results with those obtained in our published retrospective study in patients with non-metastatic CRC treated with folinic acid/5-FU/oxaliplatin (FOLFOX4) [237], the prognostic biomarkers identified were not the same, highlighting a value related to the characteristics of the population and treatment analyzed. The predictive value of rs1610696-*G/G* and rs371194629-*Ins/Ins* genotypes for increased risk of G3-4 neutropenia and neurotoxicity in the patients treated with FOLFOX4, and of UTR-2 haplotype for neurotoxicity, suggests us a specific mechanism drug-related that should be better investigate.

The others important toxicities affecting patients with mCRC treated with FOLFIRI regimen such as diarrhea and those affecting GI tract were grouped into the non-hematological toxicity. Evaluation of diarrhea and GI tract toxicities (diarrhea, nausea and vomiting) as single classes was not realized in association with the *HLA-G* 3'UTR polymorphisms due to the low number of the events. Moreover, the only HLA-G-related association with increased risk of G3-4 cumulative non-hematological toxicity that emerged from the analysis was the heterozygous S class of the HLA-G 3 level secretor model.

In the HLA-G 3 level secretor model the H class was associated with a reduced risk of G3-4 total toxicity both first cycle and cumulative.

Regarding the HLA-G secretor models and *UGT1A1*28* combination, the results required a more careful analysis. In fact, previous published results have attempted to identify novel predictive genetic biomarkers to better personalize irinotecan-based therapy [238]–[240]. Currently, the *UGT1A1*28* polymorphism is the only one pharmacogenetics marker adopted as predictive marker by international guidelines for adjusting the irinotecan dose [241]. *UGT1A1*28/*28* had showed an association with the G3-4 hematological and total toxicity only after the first cycle [78]. The evaluation of the L and M classes of the *HLA-G* 2 level secretor model suggests a trend towards an association with an increased risk of G3-4 first cycle toxicities (hematological, non-hematological, total and neutropenia) for the L class, corresponding to the HLA-G low secretor level, and an opposite trend for HLA-G medium-high secretor level (M class) (data not shown). The same trends were observed in the *HLA-G* 3 level secretor model, excepted for the non-hematological and total toxicities. However, the distribution of the *UGT1A1*28* polymorphism in these classes vary and then, due to the known effect of association with increased risk of G3-4, interpretation of these preliminary data become more difficult. We could only make some considerations based on the results obtained for *HLA-G* 3'UTR UTR-4 haplotype in the cumulative hematological and neutropenia toxicities. In the two secretor models, UTR-4 belongs to the M and I classes. Therefore, combining the *UGT1A1*28* and HLA-G secretor levels, we observed in the *28M and *28L class a trend towards a reduced risk of G3-4 with an OR higher for the *28M class compared to the *28L class. Maybe this means that the presence of the M class with the UTR-4, tends to shift the low risk of G3-4 hematological toxicity for patients with *UGT1A1*1/*1* towards a higher risk of G3-4 toxicity. Another result to consider, was obtained in the stratification based on *UGT1A1*28* genotype where in the *UGT1A1*1/*1* stratum was identifed an association with an increased risk of G3-4 cumulative non-hematological toxicity, according to the log-additive model. The peculiarity was that no statistically relevant significance was found in the two classes of the HLA-G 2 level secretor model.

Summarized, novel possible predictive biomarkers were identified for patients with mCRC treated with FOLFIRI regimen. The common factor of all the associations we found was the fact that lower HLA-G expression has been correlated with an increase in the G3-4 toxicity, while higher HLA-G expression with a reduction, as we had observed in the survival and response analysis. The only explanation we could hypotized was an interaction between the HLA-G with the drugs used in the treatment, in particular with

irinotecan. In fact, an aspecific protein-drug interaction could alter the levels of available soluble HLA-G and also of the irinotecan in blood. In the PART 3 we decided to deepen this hypothesis.

Table 38: Summary of the first cycle and cumulative toxicity results.

	Cumulative toxicity			
	G3-4 Hematological & Neutropenia	**G3-4 Non-hematological**	**G3-4 Total**	
	Increase	**Increase**	**Reduction**	**Increase**
SNP allele or genotype (genetic model)	**- rs1707-C (log-add)** **- rs1710-G/G (rec)** **- rs1063320-C/C (rec)**			**- rs9380142-*A/A* (dom)**
Haplotype (model)	**- UTR-4 (2 vs. 0 copy)**		**- UTR-1 (1 vs. 0 copy)**	**- UTR-6 (1 vs. 0 copy)**
HLA-G secretor model	**- *28L (2 vs. 0 copy)** **- *28I (1 vs. 0 copy)**[1]	**- S class (1 vs. 0 copy)**	**- H class (1 vs. 0 copy)** - H/I vs. H/H	
***UGT1A1* + HLA-G secretor models**	- *28L (2 vs. 0 copy) - *28S (1 vs. 0 copy)	**- *1S (1 vs. 0 copy)** - *28L (1 vs. 0 copy)	**- *1H (1 vs. 0 copy)**	**- *28I (1 vs. 0 copy)** - *28L (2 vs. 0 copy) - *28L (1 vs. 0 copy) -*28S (1 vs. 0 copy)
In stratum UGT1A11/*1* genotype, HLA-G secretor models**		**- L class (log-add)**		
In stratum UGT1A128/*28* genotype, HLA-G secretor models**	- L class (log-add)			

[1]association not present for neutropenia toxicity.

In text not italiac neither bold the association regarding the first cycle toxicity, in bold the association regarding the cumulative toxicity; in bold and grey the association regarding both the first cycle and cumulative toxicity.

5. PART 2: Soluble characterization

5.1. PATIENTS AND METHODS

5.1.1. Patients clinical data and treatment

Clinical data, genomic DNA from blood samples and plasma samples were obtained from patients diagnosed with mCRC prospectively collected by Toffoli and colleagues for two previous published studies [78], [80]. Inclusion criteria for our retrospective study was defined by plasma sample availability: 25 samples derived from the first study [78] and 15 from the genotype-driven phase I study published in 2010 [80]. Patient's characteristics and treatment were already described in the PART 1, 4.1.1 section for the first study. About the second study [80], patients were selected for *UGT1A1*1/*1* and *UGT1A1*1/*28* genotypes and treated with FOLFIRI regiment: irinotecan was administered at doses higher than the standard 180 mg/m^2 starting from 215 mg/m^2 with a 20% increase every-2-weeks dose until a dose of 370 mg/m^2 for *UGT1A1*1/*28* patients and of 420 mg/m^2 for *UGT1A1*1/*1* patients, with unchanged dose of infusional FU/LV. FOLFIRI was discontinued because of disease progression, intolerable side effects, patient refusal, or physician assessment. According to the eligible criteria of the study [80], patients who carried the *UGT1A1*28/*28* were excluded.

5.1.2. Study design and endpoints

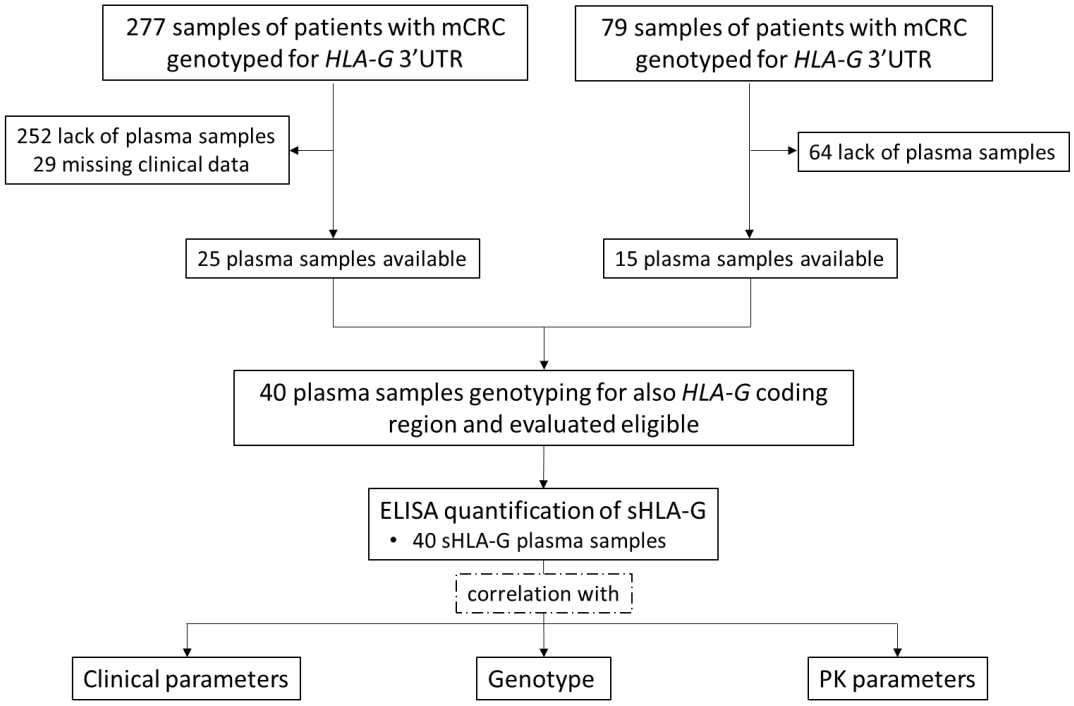

Figure 15: CONSORT-like diagram. PK: pharmacokinetics.

Our exploratory retrospective study aims to evaluate the plasma levels of soluble HLA-G (sHLA-G) and the correlations between these levels and: clinical parameters, the genomic non-coding setting in the *HLA-G* 3'UTR region, and the genomic coding setting in *HLA-G.* The flowchart was described in the CONSORT-like diagram (Figure 15). We focused on a subgroup of patients affected by mCRC treated with FOLFIRI regiment of which plasma samples were available. The institutional review board of each participating institution had approved both the study protocols, and all patients signed a written informed consent for a genetic analysis before entering the studies. All the 40 eligible samples were investigated for the polymorphisms and haplotypes in the *HLA-G* coding and non-coding 3'UTR regions.

5.1.3. Genotyping assays of genomic coding and non-coding regions of *HLA-G*

From the existing Biobank of blood samples stored at Clinical and Experimental Pharmacology Unit of National Cancer Institute (IRCCS CRO-Aviano), 279 mCRC patients of the prospective study and 77 of the phase 1 study were considered. For all the samples genomic DNA was extracted from peripheral whole blood and the *HLA-G* 3'UTR was sequenced using Sanger method with the protocol described in PART 1, 4.1.3 section. For the 40 selected samples, exons 2, 3 and 4 of the *HLA-G* gene were amplified by PCR technique using the already published primers [242]: forward HLAG-Ex2F primer, 5'-GGGTCGGGCGGGTCTCAA-3', and reverse HLAG-Ex2R primer, 5'-TCCGTGGGGCATGGAGGT-3'for the *HLA-G* exon 2; forward HLAG-Ex3F primer, 5'-CCCAGACCCTCTACCTGGGAGA-3', and reverse HLAG-Ex3R primer, 5'-CTCTCCTTGTGCTAGGCCAGGCTGAGAGG-3'for the *HLA-G*exon 3; and finally forward HLAG-Ex4F primer, 5'-CCATGAGAGATGCAAAGTGCT-3', and reverse HLAG-Ex4R primer, 5'-TGCTTTCCCTAACAGACATGAT-3'for the *HLA-G*exon 4. The reaction was performed in a final volume of 30 µl with 1.25 mM $MgCl_2$, 0.25 mM dNTPs, 0.5 µM each primer, about 50-200 ng genomic DNA template, 1X PCR Buffer and 0.17 unit of AmpliTaq Gold DNA polymerase (AppliedBiosystems, Foster City, CA, USA). Only for the amplification of *HLA-G* exon 2 and 4 was added DMSO 5%. The PCR cycles for the exon 2 were as follows: 2 mins of initial denaturation at 94°C, 30 cycles of 30 secs at 94°C, 30 secs at 62°C, 60 secs at 72°C, and the final extension step at 72°C for 10 mins; for the exon 3, the PCR cycles were of 2 mins of initial denaturation at 94°C, 35 cycles of 60 secs at 94°C, 90 secs at 65°C, 120 secs at 72°C, and the final extension step at 72°C for 10 mins; and finally for the exon 4 were: 2 mins of initial denaturation at 94°C, 30 cycles of 30 secs at 94°C, 30 secs at 55°C, 60 secs at 72°C, and the final extension step at 72°C for 10 mins. The protocol used for the purification of PCR products and sequencing are the same described for the *HLA-G* 3'UTR (see PART 1, 4.1.3 section).
To determine the *HLA-G* genotype in the coding regions, we analyzed the entire 2-4 exonic regions and related intronic boundaries. Genetic data were annotated considering nomenclature reported for *HLA-G* coding region in the public HLA official website (http://hla.alleles.org/data/hla-g.html).

5.1.4. Linkage disequilibrium and *HLA-G* haplotype analysis

HLA-G coding haplotypes were reconstructed through alignment of our sequences with the *G*01:01:01:01* genomic sequence reported in the HLA-G official website (http://hla.alleles.org/data/hla-g.html). The PHASE algorithm version 2.1 was also used to determine the *HLA-G* coding haplotypes frequencies using the nomenclature reported in the official HLA-G public website (http://hla.alleles.org/data/hla-g.html). The strength of LD between pairs of HLA-G markers with MAF>2% was measured as r^2 using the freely available LDPlotter software (http://www.pharmgat.org/Tools/pbtoldplotform). The statistic $r^2<0.50$ indicates low LD, $0.50\leq r^2<0.80$ moderate high LD, $0.80\leq r^2<1$ high LD and $r^2 = 1$ perfect LD.

5.1.5. Soluble *HLA-G* ELISA protocol

We analyzed for sHLA-G the 40 available EDTA plasma samples using the sHLA-G ELISA kit (BioVendor – Laboratorní medicína a.s. and EXBIO Praha a.s.), a sandwich enzyme immunoassay for the quantitative measurement of sHLA-G. This ELISA assay detects both shedded HLA-G1 and soluble HLA-G5. The primary monoclonal anti-sHLA-G antibody used recognizes the specific soluble isoforms of the HLA-G protein, whereas the secondary antibody is the monoclonal anti-human β2-microglobulin antibody labelled with horseradish peroxidase (HRP). The remaining HRP conjugate can react with the substrate solution (TMB) and the reaction is stopped by the addition of acidic solution. The absorbance of the resulting yellow product is measured by Infinite F200 PRO (TECAN, Männedorf, Switzerland) set to 450 nm with the reference wavelength set to 630 nm and it is proportional to the concentration of sHLA-G (U/ml). All samples were assayed in duplicate, and the total sHLA-G levels were determined from a five-point standard curve, using dilutions of calibrator (7.81, 15.63, 31.25, 62.5, and 125 U/ml) purchased by the kit as a standard reagent. Results were expressed as U/ml.

5.1.6. Statistical analysis

The plasma levels of sHLA-G were evaluated in relation to clinical variables, patient's genotype and pharmacokinetic (PK) parameters: comparison between two groups was performed using the non-parametric Mann-Whitney test and between three or more groups with the Krustal-Wallis test. Analyses were done with GraphPad Prism 3.02 (GraphPad Software). All the statistical analyses were two-tailed and a $p<0.05$ was considered statistically significant.

5.2.RESULTS

5.2.1.Patient characteristics

The subgroup of 40 patients with mCRC patients treated with FOLFIRI regimen, was described in Table 39. The plasma sample availability determined the numerosity of this exploratory study. The median follow-up was 15.42 months (range: 1.15-31.48 months). Most of the patients with mCRC were staged III-IV at the time of diagnosis (*n*=35, 87.5%), received radical surgery (*n*=26, 65.0%), and 8 out of 40 (20.0%) patients had more than 2 metastatic sites at the time of enrollment in the study. In this subgroup, we observed some differences in the distribution of demographic and clinical characteristics with respect to those reported for the 248 patients analyzed in the PART 1, with predominance of women (*n*=25, 62.5%) and of rectum and right colon as primary tumor site (*n*=18, 45.0% and *n*=14, 35.0%, respectively).

Table 39. Demographic and clinical characteristics of 40 eligible patients with mCRC.

Characteristic	***n***	**(%)**
Age, yrs		
Median (range)	57.94 (26.5-75.5)	
Gender		
Man	15	37.5
Woman	25	62.5
Primary tumor site		
Right colon	14	35.0
Left colon	8	20.0
Rectum	18	45.0
Radical surgery		
Yes	26	65.0
No	14	35.0
Stage at diagnosis		
I-II	5	12.5
III-IV	35	87.5
Adjuvant (radiotherapy or chemotherapy)		
Yes	14	35.0
No	26	65.0
Number of metastatic sites		
1	17	45.9
2	12	32.4
3	2	5.4
4	5	13.5
5	1	2.7
Follow-up time (months)		
Median (range)	15.42 (1.15-31.48)	

5.2.2. Germline *HLA-G* 3'UTR polymorphisms, haplotypes, secretor models and coding regions

Forty patients were considered for the genotype-phenotype analysis of this PART 2. They derived from the previous analyzed patients (see PART 1) and other 79 patients with mCRC homogeneously treated with FOLFIRI regimen (data not shown), and also from these latter, the DNA was successfully extracted and genotyping for *HLA-G* 3'UTR polymorphisms considering for the analysis the same *HLA-G* 3'UTR

polymorphisms and haplotypes analyzed in PART 1 (Table 40). Also for these selected 40 patients the polymorphisms respect the Hardy-Weinberg equilibrium (HWE) (p>0.05), even if the rs371194629 had showed a slight deviation (*p*=0.05). As expected from the frequencies in the European population, the UTR-2 and UTR-1 were the most abundant *HLA-G* 3'UTR haplotypes.

Table 40. Distributions of alleles, haplotypes and genotypes frequencies identified at *HLA-G* 3'UTR polymorphic sites: *HLA-G* 3'UTR genotyping was performed in 277 + 79 patients genetically evaluated with mCRC of which 40 were eligible for this study and reported in this table.

***HLA-G* 3'UTR SNPs**		**Eligible N_{tot}=40, n (%)**	***HLA-G* haplotypes**		**Eligible N_{tot}=40, n (%)**
+2960 14-bp INDEL	**Alleles**		**UTR-2**	**Haplotypes**	
(rs371194629)	Del	45 (56.3)		**Ins**TCCCGA**G**	27 (33.75)
	Ins	35 (43.7)			
	Genotypes			**Genotypes**	
	Del/Del	14 (35.0)		Het	17 (42.5)
	Ins/Del	17 (42.5)		Hom	5 (12.5)
	Ins/Ins	9 (22.5)			
+3003 T>C	**Alleles**		**UTR-1**	**Haplotype**	
(rs1707)	T	75 (93.8)		Del**T**GCC**CG**C	26 (32.5)
	C	5 (6.2)			
	Genotypes			**Genotypes**	
	T/T	36 (90.0)		Het	12 (30.0)
	T/C	3 (7.5)		Hom	7 (17.5)
	C/C	1 (2.5)			
+3010 C>G	**Alleles**		**UTR-3**	**Haplotype**	
(rs1710)	C	48 (60.0)		DelTCCCGAC	12 (15.0)
	G	32 (40.0)			
	Genotypes			**Genotypes**	
	C/C	17 (42.5)		Het	10 (25.0)
	G/C	14 (35.0)		Hom	1 (2.5)
	G/G	9 (22.5)			
+3027 C>A	**Alleles**		**UTR-7**	**Haplotype**	
(rs17179101)	C	73 (91.3)		**Ins**TC**AT**GAC	7 (8.75)
	A	7 (8.7)			
	Genotypes			**Genotypes**	
	C/C	34 (85.0)		Het	5 (12.5)
	C/A	5 (12.5)		Hom	1 (2.5)
	A/A	1 (2.5)			
+3035 C>T	**Alleles**		**UTR-4**	**Haplotype**	
(rs17179108)	C	71 (88.8)		Del**CG**CC**C**AC	5 (6.25)
	T	9 (11.2)			
	Genotypes			**Genotypes**	
	C/C	32 (80.0)		Het	3 (7.5)
	C/T	7 (17.5)		Hom	1 (2.5)
	T/T	1 (2.5)			
+3142 G>C	**Alleles**		**UTR-6/-18**	**Haplotype**	
(rs1063320)	G	45 (60.0)		Del**T**GCC**C**AC	2 (2.5)
	C	32 (40.0)			
	Genotypes			**Genotypes**	
	G/G	17 (42.5)		Het	2 (5.0)
	G/C	14 (35.0)		Hom	0 (0.0)
	C/C	9 (22.5)			
+3187 A>G	**Alleles**		**UTR-5**	**Haplotype**	
(rs9380142)	A	55 (68.8)		**Ins**TCC**T**GAC	1 (1.25)
	G	25 (31.2)			

	Genotypes			**Genotypes**	
	A/A	22 (55.0)		Het	1 (2.5)
	A/G	11 (27.5)		Hom	0 (0.0)
	G/G	7 (17.5)			
+3196 C>G	**Alleles**		**UTR-13**	**Haplotype**	
(rs1610696)	C	53 (66.3)		DelTCC**T**GAC	1 (1.25)
	G	27 (33.7)		**Genotypes**	
	Genotypes			Het	1 (2.5)
	C/C	18 (45.0)		Hom	0 (0.0)
	C/G	17 (42.5)			
	G/G	5 (12.5)			

SNPs: Single nucleotide polymorphisms; N_{tot}: number of individuals.Het: heterozygous; Hom: homozygous. In bold are highlighted the mutant alleles.

We analyzed the entire 2-4 exonic regions and related intronic boundaries. In the Table 41, we reported all the sequence of *HLA-G* coding alleles of which the polymorphisms with synonymous or non-synonymous variations we found. From our genetic analysis, 5 different proteins (*G*01:01, G*01:03, G*01:04, G*01:05N, G*01:06* coding alleles) were codified by the *HLA-G* gene in population affected by mCRC here analyzed (Table 41).

Table 41. *HLA-G* coding alleles corresponding to a specific HLA-G protein defined by the 10 *HLA-G* germline variations found in coding regions genotyped in mCRC patients.

		All the variations in the 40 patients' coding allele									
HLA-G coding alleles	**HLA-G Alleles**	***c.163 A>T***	***c.177 G>A***	***c.243 G>A>C***	***c.351 C>T***	***c.393 A>T***	***c.400 C>A***	***c.460 C>Del***	***c.636 C>T***	***c.845 C>T***	***c.873 G>A***
G*01:01	G*01:01:01:01	A	G	G	C	A	C	C	C	C	G
	G*01:01:01:02	A	G	G	C	A	C	C	C	C	G
	G*01:01:01:03	A	G	G	C	A	C	C	C	C	G
	G*01:01:01:04	A	G	G	C	A	C	C	C	C	G
	G*01:01:01:05	A	G	G	C	A	C	C	C	C	G
	G*01:01:01:06	A	G	G	C	A	C	C	C	C	G
	G*01:01:01:07	A	G	G	C	A	C	C	C	C	G
	G*01:01:01:08	A	G	G	C	A	C	C	C	C	G
	G*01:01:02:01	A	G	A	T	A	C	C	C	C	G
	G*01:01:02:02	A	G	A	T	A	C	C	C	C	G
	G*01:01:03:01	A	G	A	C	T	C	C	C	C	G
	G*01:01:03:02	A	G	A	C	T	C	C	C	C	G
	G*01:01:03:03	A	G	A	C	T	C	C	C	C	G
	G*01:01:04	A	G	G	C	A	C	C	*	*	*
	G*01:01:05	A	G	G	C	T	C	C	C	C	G
	G*01:01:06	A	G	G	C	A	C	C	T	C	G
	G*01:01:07	A	G	A	T	T	C	C	C	C	G
	G*01:01:08	A	G	A	C	A	C	C	C	C	G
	G*01:01:09	A	G	G	C	A	C	C	C	C	G
	G*01:01:11	A	G	G	C	A	C	C	C	C	G
	G*01:01:12	A	G	A	T	A	C	C	C	C	G
	G*01:01:13	A	G	A	T	A	C	C	*	*	*
	G*01:01:14	A	A	A	T	A	C	C	C	C	G
	G*01:01:15	A	G	G	C	A	C	C	C	C	G
	G*01:01:16	A	G	A	T	A	C	C	C	C	G
	G*01:01:17	A	G	A	T	A	C	C	C	C	G

	G*01:01:18	A	G	A	T	A	C	C	C	C	G
	G*01:01:19	A	G	A	T	A	C	C	C	C	G
	G*01:01:20	A	G	G	T	A	C	C	C	C	G
	G*01:01:21	A	G	G	C	A	C	C	C	C	G
	G*01:01:22	A	G	A	T	A	C	C	C	C	G
G*01:03	G*01:03:01:01	**T**	G	G	C	A	C	C	T	C	G
	G*01:03:01:02	**T**	G	G	C	A	C	C	T	C	G
G*01:04	G*01:04:01	A	G	A	C	A	**A**	C	C	C	G
	G*01:04:02	A	G	C	C	A	**A**	C	T	C	G
	G*01:04:03	A	G	G	C	A	**A**	C	C	C	G
	G*01:04:04	A	G	A	C	A	**A**	C	C	C	A
	G*01:04:05	A	G	A	C	A	**A**	C	C	C	G
G*01:05N	G*01:05N	A	G	A	T	A	C	**.**	C	C	G
G*01:06	G*01:06	A	G	A	T	A	C	C	C	**T**	G

In bold are highlighted the nucleotide variants corresponding to a non-synonymous change.

We specified the type of variations occurred in the protein that was determined by the DNA substitutions in the coding sequences (HLA-G exon 2, 3 and 4) (Table 42).

Table 42. All HLA-G coding SNPs with a synonymous or non-synonymous/missense variants detected in the 40 patients with mCRC analyzed.

HLA-G exon	rs code	HLA-G coding SNPs	HLA-G protein variations	Type of variant
Exon 2	rs41551813	c.163 A>T	p.Thr31Ser	Non-synonymous/missense
	rs72558174	c.177 G>A	p.Arg35Arg	Synonymous
	rs1130355	c.243 G>A>C	p.Pro57Pro	Synonymous
Exon 3	rs1130356	c.351 C>T	p.His93His	Synonymous
	rs3873252	c.393 A>T	p.Gly107Gly	Synonymous
	rs12722477	c.400 C>A	p.Leu110Ile	Non-synonymous/missense
	rs41557518*	c.460 DelC	p.Leu130fs*189	Frameshift with formation of stop codon
Exon 4	rs41562616	c.636 C>T	p.His188His	Synonymous
	rs12722482	c.845 C>T	p.Thr258Met	Non-synonymous/missense
	rs17875406	c.873 G>A	p.Pro267Pro	Synonymous

All the rs refer toEnsembl GRCh37. *In the HLA site was reported c.460 DelC, however the frameshift variant in Ensembl GRCh37 (rs41557518) was reported as c.459 DelC corresponded to the p.Asp130fs*189.

Through the UCFS Chimera software, we verified the non-synonymous variations in the HLA-G protein (Figure 16). The p.Thr31Ser variation was located in one of the β-sheets of the α1 domain, the p.Leu110Ile and p.Leu130Cys variations were in distinct β-sheet strands in the α2 domain, the two domains responsible of the formation of the HLA-G pocket. All side chains of these missense variations are structurally external to the antigen-binding cleft of HLA-G. The missense variation p.Thr258Met, localized in the exon 4, is not involved in the genesis of the pocket.

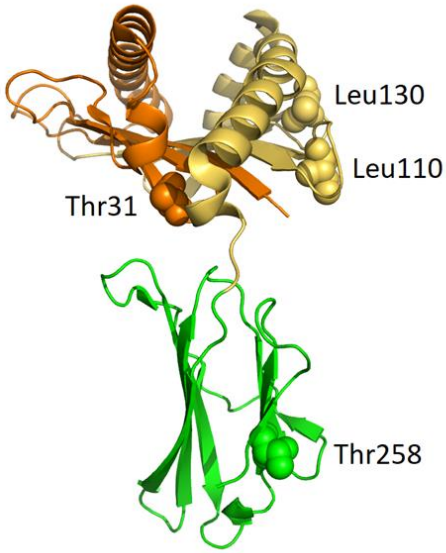

Figure 16. Three-dimensional (3D) cartoon representation of HLA-G mature protein (1-274 AAs). Side chains of residues involved in a missense change and found to be mutated in 40 patients with mCRC: Thr31Ser (α1 domain, orange), Leu110Ile (α2 domain, yellow), Leu130Cys (α2 domain, yellow), and Thr258Met (α3 domain, green) in secondary structures, are highlighted. Picture was created using UCSF Chimera software (PDB ID: 1YDP).

Moreover, distribution of the *HLA-G* polymorphisms and haplotypes in coding regions considered was described in Table 43. The rs41551813 SNP was detected in only one patient (coding allele *G*01:03*). Similarly, also the *G*01:05N* null allele and the synonymous c.636 C>T (rs41562616) change (coding allele *G*01:01*), were found in only two distinct patients, respectively. The synonymous change for rs72558174 SNP was detected in only two distinct patients. We observed that in the forty patients with mCRC, rs1130355 SNP alleles frequencies had an inverse distribution of wild type (*A*=58.8%) and mutated (*G*=41.2%) compared to EUR population (*G*=57%, *A*=43%), even if in the TSI sub-population the frequencies were slightly different (*G*=50%, *A*=50%) (data reported in the 1000Genome Browser). However, the differential distribution in these alleles, did not reach a significant statistical value when we compared our small population with TSI.

Table 43. Distribution of alleles, haplotypes and genotypes frequencies identified at *HLA-G* coding regions that defines the 5 *HLA-G* coding alleles identified in 40 patients with mCRC.

***HLA-G* coding SNPs**		**Eligible N_{tot}=40, n (%)**	***HLA-G* coding alleles**		**Eligible N_{tot}=40, n (%)**
Exon 2			G*01:01	**Alleles**	
c.163 A>T	**Alleles**			G*01:01	62 (77.5)
(rs41551813)	A	79 (98.8)			
	T	1 (1.2)		**Genotypes**	
	Genotypes			Het	12 (30.0)
	A/A	39 (97.5)		Hom	25 (62.5)
	A/T	1 (2.5)	G*01:04	**Alleles**	
	T/T	0 (0.0)			

c.177 G>A	**Alleles**	
(rs72558174)	G	78 (97.5)
	A	2 (2.5)
	Genotypes	
	G/G	38 (95.0)
	G/A	2 (5.0)
	A/A	0 (0.0)
c.243 A>G	**Alleles**	
(rs1130355)	A	47 (58.8)
	G	33 (41.2)
	Genotypes	
	A/A	16 (40.0)
	A/G	15 (37.5)
	G/G	9 (22.5)
Exon 3		
c.351 C>T	**Alleles**	
(rs1130356)	C	53 (66.3)
	T	27 (33.7)
	Genotypes	
	C/C	18 (45.0)
	C/T	17 (42.5)
	T/T	5 (12.5)
c.393 A>T	**Alleles**	
(rs3873252)	A	73 (91.3)
	T	7 (8.7)
	Genotypes	
	A/A	34 (85.0)
	A/T	5 (12.5)
	T/T	1 (2.5)
c.400 C>A	**Alleles**	
(rs12722477)	C	67 (83.8)
	A	13 (16.2)
	Genotypes	
	C/C	29 (72.5)
	C/A	9 (22.5)
	A/A	2 (5.0)
c.460 DelC	**Alleles**	
(rs41557518)	C	79 (98.8)
	Del	1 (1.2)
	Genotypes	
	C/C	39 (97.5)
	Del/C	1 (2.5)
	Del/Del	0 (0.0)
Exon 4		
c.636 C>T	**Alleles**	
(rs41562616)	C	79 (98.8)
	T	1 (1.2)
	Genotypes	

	G*01:04	13 (16.3)
	Genotypes	
	Het	9 (22.5)
	Hom	2 (5.0)
G*01:06	**Alleles**	
	G*01:06	3 (3.7)
	Genotypes	
	Het	3 (7.5)
	Hom	0 (0.0)
G*01:03	**Alleles**	
	G*01:03	1 (1.25)
	Genotypes	
	Het	1 (2.5)
	Hom	0 (0.0)
G*01:05N	**Alleles**	
	G*01:05N	1 (1.25)
	Genotypes	
	Het	1 (2.5)
	Hom	0 (0.0)

	C/C	39 (97.5)
	C/T	1 (2.5)
	T/T	0 (0.0)
c.845 C>T	**Alleles**	
(rs12722482)	C	77 (96.3)
	T	3 (3.7)
	Genotypes	
	C/C	37 (92.5)
	C/T	3 (7.5)
	T/T	0 (0.0)
c.873 G>A	**Alleles**	
(rs17875406)	G	76 (95.0)
	A	4 (5.0)
	Genotypes	
	G/G	36 (90.0)
	G/A	4 (10.0)
	A/A	0 (0.0)

SNPs: Single nucleotide polymorphisms; N_{tot}: number of individuals. Het: heterozygous; Hom: homozygous.

We chose to analyze also the distribution into classes of the HLA-G 2 level and 3 level secretor models generated by the PHASE method. As aspected, classes of the HLA-G 3 level secretor model showed low numerosity. In our study cohort, we observed a prevalence of heterozygous low/medium-high levels in the HLA-G 2 level secretor model, and a prevalence of homozygous intermediate level class in the HLA-G 3 levels model (Table 44).

Table 44. HLA-G secretor models: numbers and frequencies observed in 40 eligible patients with mCRC.

HLA-G secretor model	**Code**	**Phenotypic effect**	**Corresponding *HLA-G* 3'UTR haplotypes***	**Classes**	**n (%)**
2 levels	M	Medium/High HLA-G expression	UTR-1, UTR-3, UTR-4, UTR-6	M/M	14 (35.0)
	L	Low HLA-G expression	UTR-2, UTR-5, UTR-7	L/M	17 (42.5)
				L/L	9 (22.5)
3 levels	H	High HLA-G expression	UTR-1	H/H	7 (17.5)
	I	Intermediate HLA-G expression	UTR-2, UTR-3, UTR-4, UTR-6	H/I	11 (27.5)
	S	Low HLA-G expression	UTR-5, UTR-7	H/S	0 (0.0)
				I/I	15 (37.5)
				I/S	6 (15.0)
				S/S	1 (2.5)

*Each HLA-G level may be composed by each of the corresponding *HLA-G* 3'UTR haplotype, alone or in combination with one of those listed.

5.2.3. Linkage disequilibrium patterns

We realized the LD analysis between all the polymorphisms found in the 40 patients, considering both in *HLA-G* 3'UTR and in coding region (Figure 17). We had observed perfect LD (r^2=1) between rs1710 (+3010 C>G) and rs1063320 (+3142 G>C) polymorphisms also in this subgroup of 40 patients, between rs3873252 (c.393 A>T) and rs17179101 (+3027 C>A), and between rs1130356 (c.351 C>T) and rs1610696 (+3196 C>G).

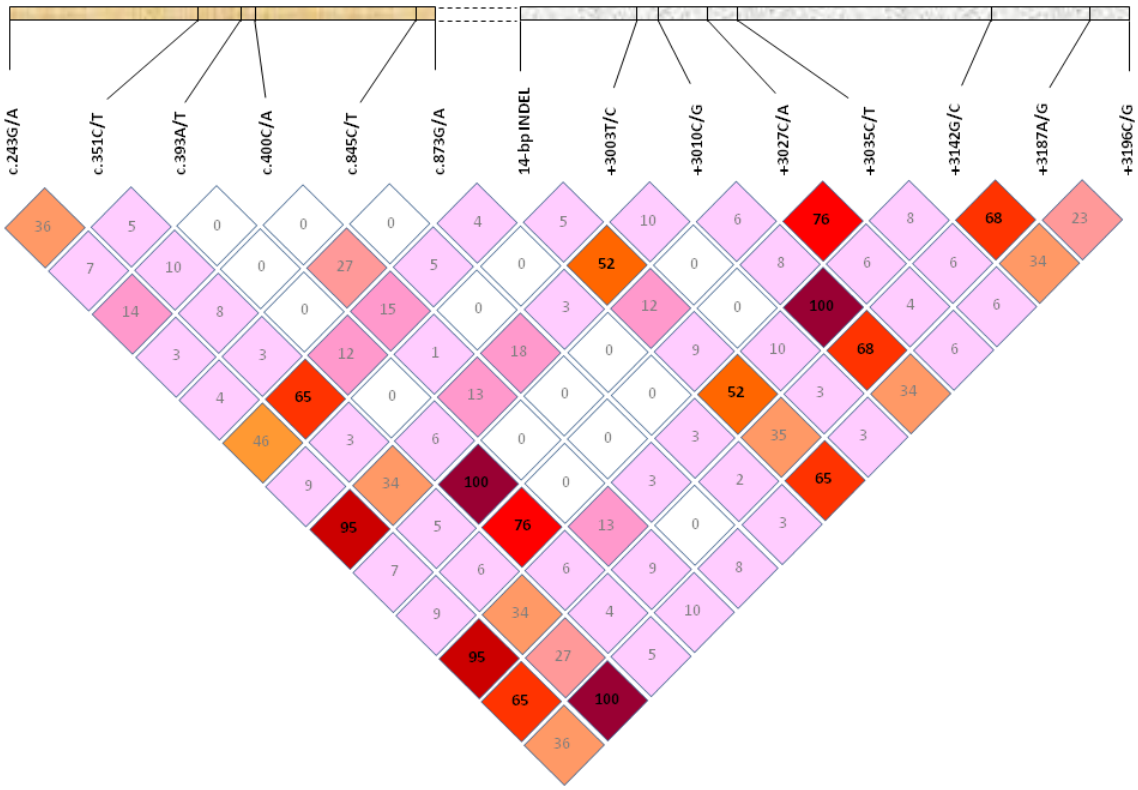

Figure 17. LD patterns at the 3'UTR region and coding region of HLA-G in 40 patients with mCRC. LD plot generated by LDPlotter shows correlations between all pairs of variants with MAF >2%. The r^2 values (x100) for the marker pairs are listed in the corresponding boxes. High pairwise LD (r^2) between variants is highlighted in bold.

Moreover, we also analyzed the LD between the coding and non-coding regions of *HLA-G* only for the 25 patients we had the pharmacokinetic parameters and no substantial difference we found (data not shown).

5.2.4. Soluble HLA-G plasmatic levels and clinical parameters

An ELISA test was realized to measure the sHLA-G in the 40 plasma samples. The five-point standard curve used to determine the total sHLA-G levels was realized thanks to a standard reagent at the concentrations reported in the Table 45. The standard curve and R^2 value are reported in Figure 18.

Table 45. ELISA assay: mean of duplicate measurement of absorbance.

Calibrator (U/ml)	A450-630 nm (a.u.)
7.81	0.11
15.63	0.20
31.25	0.36
62.5	0.81
125	1.60

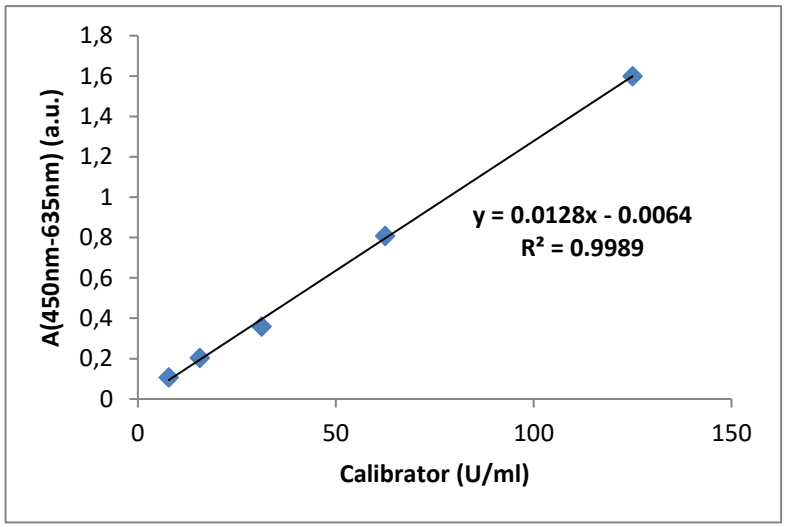

Figure 18. Standard curve

In the 40 patients with mCRC analyzed, the median sHLA-G was 116.44 U/ml (range: 12.8-1552.7). We stratified the median sHLA-G plasma levels measured with the ELISA assay, according to the clinicopathological parameters in patients with mCRC (Table 46).

Table 46. Association of sHLA-G plasma expression with clinicopathological parameters in patients with mCRC.

Variables	N.	sHLA-G median (range, U/ml)	*p**
mCRC	40	116.4 (12.8-1552.7)	
Gender			
Female	25	121.4 (43.1-1552.7)	0.1464
Male	15	91.5 (12.8-254.5)	
Age (61.6 years)			
≤61.6	20	117.9 (43.1-1552.7)	0.8604
>61.6	20	116.4 (12.8-806.0)	
Primary tumor site			
Colon	22	110.04 (12.8-199.7)	0.5054[a]
Rectum	18	123.6 (43.1-1552.7)	
Left colon	8	113.8 (12.8-199.7)	0.8112
Right colon	14	94.2 (44.4-806.0)	
Radical surgery			
Yes	26	125.8 (48.4-806.0)	0.1605
No	14	89.3 (12.8-806.0)	
Number of metastatic sites			
1	17	84.1 (43.1-711.4)	**0.0214**
>1	20	150.1 (12.8-1552.7)	
Adjuvant CT			
No	26	102.0 (12.8-1552.7)	0.6000
Yes	14	123.6 (61.8-366.5)	

*Mann-Whitney-test; [a]: *p*= 0.759 with Kruskal-Wallis test comparing rectum, left colon and right colon.

No significant association was observed between median sHLA-G plasma levels and the clinical characteristics except for the number of metastatic site. Significantly lower sHLA-G level (*p*=0.0214) was

observed in patients with 1 metastatic site (median 84.1 U/ml) compared with patients with more than 1 metastatic site that were associated with an higher sHLA-G level (median 150.1 U/ml); this result was also confirmed (p=0.0486) when patients were stratified according to 1 metastatic site, 2 metastatic sites and more than 2 (>2) extra-regional lesions (Figure 19).

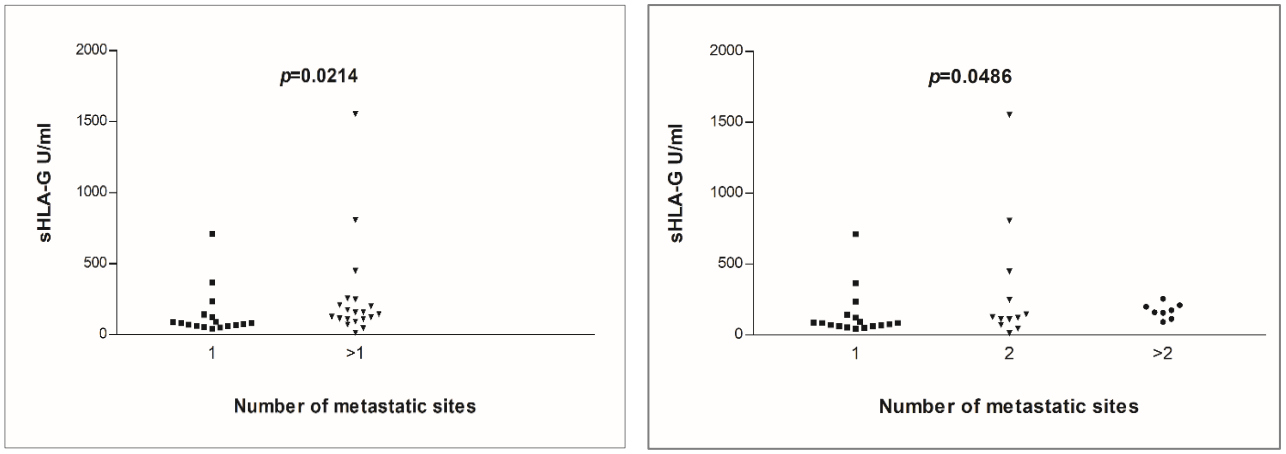

Figure 19. On the left: statistical comparison between 1 metastatic site and >1 metastatic site based on Mann-Whitney test; on the right: statistical comparison between 1, 2 and >2 metastatic site based on Krustal Wallis test.

5.2.5. Phenotypic-genetic associations

5.2.5.1. sHLA-G plasmatic levels and *HLA-G* 3'UTR polymorphisms

When we analyzed the possible association between sHLA-G plasmatic levels and HLA-G 3'UTR genotypes we did not find significant results. The sHLA-G median values according to the rs371194629 polymorphism were: 109.8 U/ml (range: 44.4-1552.7) for patients with the *Del/Del* genotype, 116.4 U/ml (range: 12.8-711.4) in heterozygous *Ins/Del* and 91.5 U/ml (range: 43.2-806.0) in patients that were double mutated with the *Ins/Ins* genotype. In the recessive model, the (*Del/Del*+*Ins/Del*) patient's group had a sHLA-G median value of 120.4 U/ml (range: 12.8-1552.7), but the difference with the mutated subgroup was not statistically significant. The highest sHLA-G value (1552.7 U/ml) was identified in a patient carrier of the rs371194629-*Del/Del* genotype and UTR-1/UTR-1 haplotype, both related as high HLA-G secretors. Scatter plots showing sHLA-G distribution in patients according to *HLA-G* rs371194629 genotypes, were represented (Figure 20).

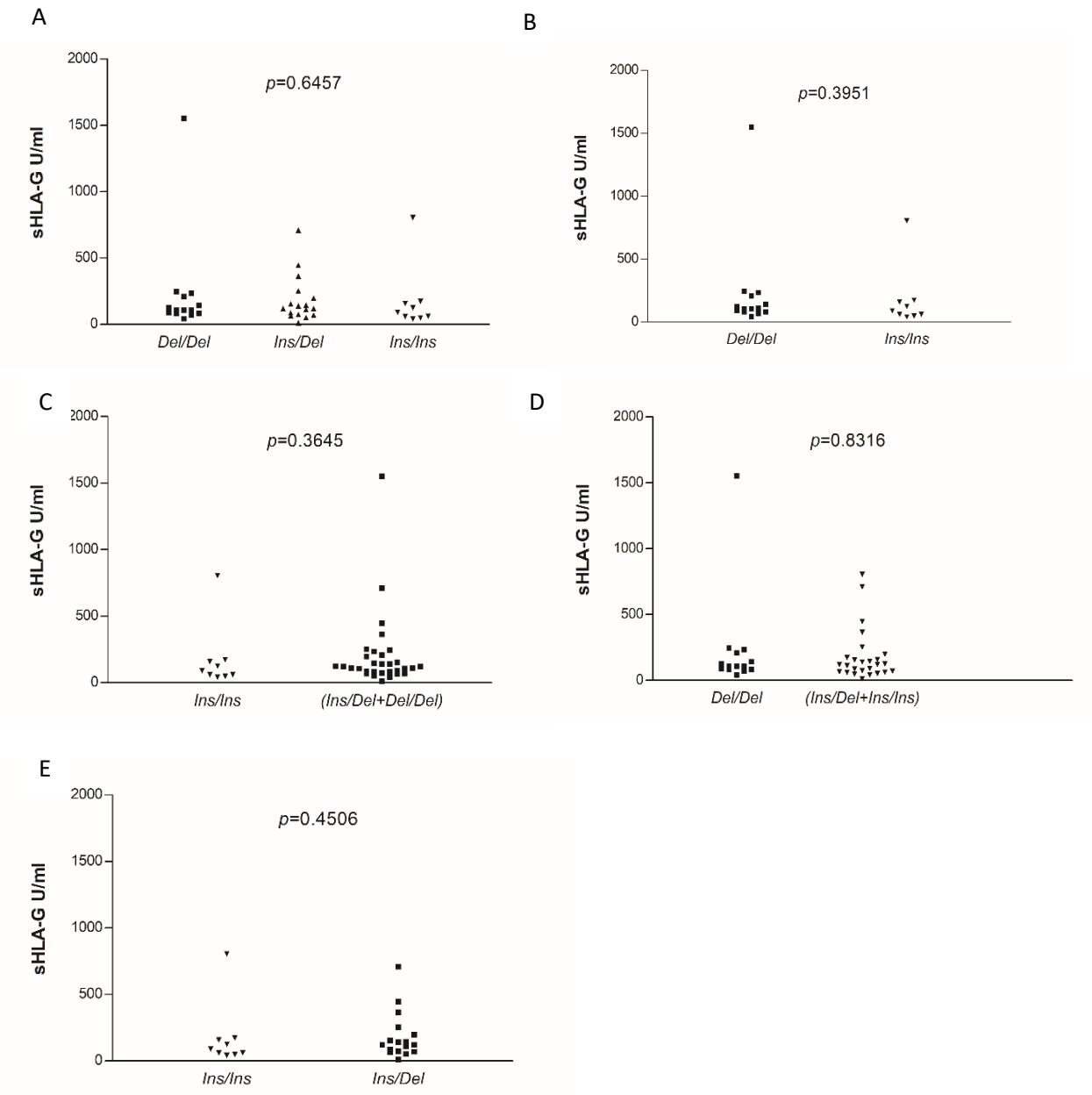

Figure 20. the 5 scatter plots represented the genetic models and single analyses of the INDEL polymorphism of the HLA-G 3'UTR region. A: all the three genotypes of rs371194629 (+2960 14-bp INDEL); B: high vs low levels; C: recessive model rs371194629 (+2960 14-bp INDEL); D: dominant model rs371194629 (+2960 14-bp INDEL); E: Ins/Ins vs Ins/Del.

5.2.5.2. sHLA-G plasmatic levels and *HLA-G* 3'UTR haplotypes

No significant data were found. Scatter plots showing sHLA-G distribution in patients according to *HLA*-G UTR1 and UTR-2 haplotype combinations were represented (Figure 21).

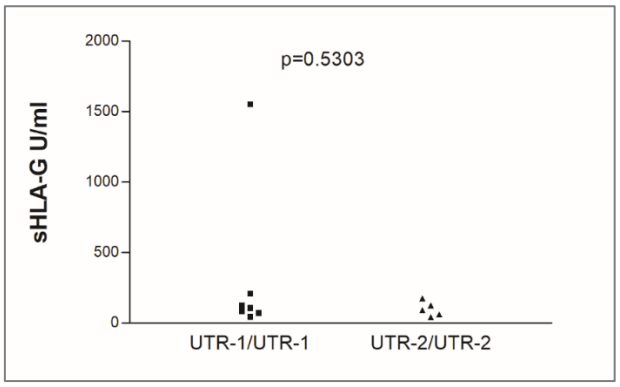

Figure 21. Correlation between sHLA-G levels and the most abundant homozygous *HLA-G* 3'UTR haplotypes.

5.2.5.3. sHLA-G plasmatic levels and HLA-G 2 levels secretor model

No significant data were found analyzing the associations between sHLA-G plasmatic levels and HLA-G 2 level secretor model (p=0.6457). To evaluate the HLA-G 2 level secretor model is equivalent to analyze the +2960 14-bp INDEL (rs371194629) polymorphisms. The M/M class corresponds to the rs371194629-*Del/Del*, the L/M class to the rs371194629-*Ins/Del* and the L/L class to the rs371194629-*Ins/Ins* genotypes (Figure 22).

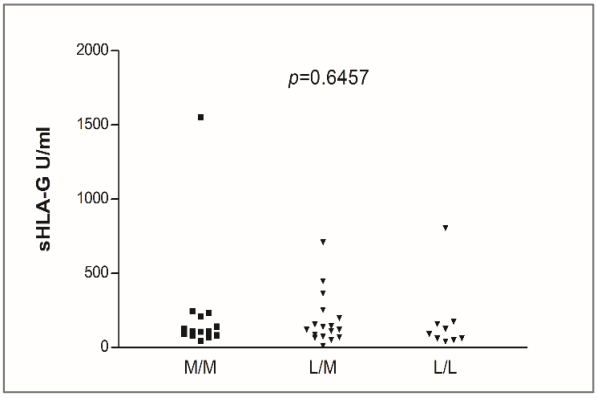

Figure 22. Correlation between sHLA-G levels and HLA-G 2 level secretor model.

5.2.5.4. sHLA-G plasmatic levels and HLA-G 3 levels secretor model

No significant data were found for the correlation between sHLA-G plasmatic levels and HLA-G 3 levels secretor model (Figure 23).

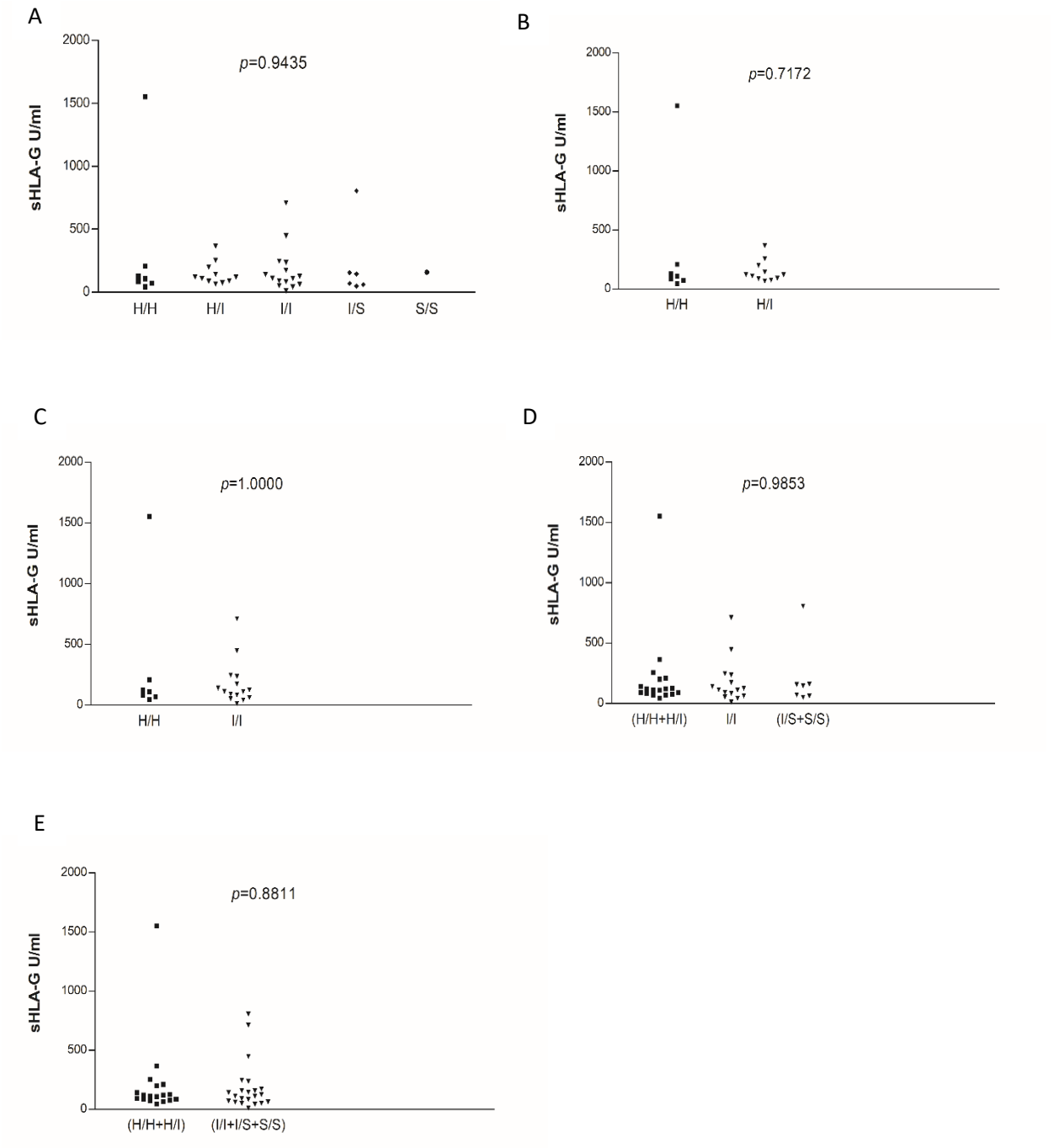

Figure 23. Correlation between sHLA-G levels and classes of HLA-G 3 level secretor model. A: all the six classes of the HLA-G 3 level secretor model; B: H/H vs. H/I; C: H/H vs. I/I; D: H/H+H/I vs. I/I vs. I/S+S/S; E: H/H vs. I/I+I/S+S/S.

5.2.5.5. sHLA-G plasmatic levels and *HLA-G* coding region SNPs

For this analysis, the coding *HLA-G* SNPs (rs12722482 rs1130355, rs1130356, rs3873252 and rs17875406) were considered. The other less frequent SNPs identified, rs72558174, rs41551813, rs41557518, rs41562616, were excluded from this analysis.

No significant data were found for the correlation between the sHLA-G plasmatic levels and HLA-G coding regions SNPs. Intriguingly, we observed a trend of the distribution sHLA-G concentration comparing rs1130355 genotype (p=0.8192): wild type rs1130355-*A/A* genotype (median value 135.4 U/ml; range: 12.8-806.0), heterozygous rs1130355-*A/G* genotype (120.4 U/ml; range: 48.5-448.2) and mutated rs1130355-*G/G* genotype (median value 90.6 U/ml; range: 44.4-1552.7) genotypes. However, the distribution of sHLA-G was not significantly different also comparing with the recessive model groups (p=0.5383): *G/G* vs (*G/A+A/A*) (median value 121.4; range: 12.8-806.0). Scatter plots for the *HLA-G* rs12722477 and rs1130355 SNPs, were represented (Figure 24).

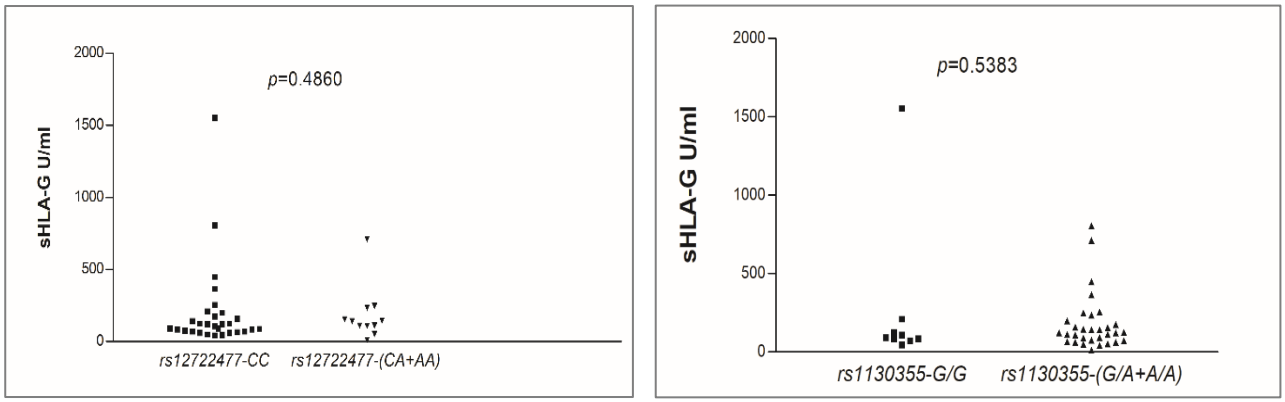

Figure 24. Correlation between sHLA-G plasmatic levels and rs12722477 and rs1130355 according to the dominant (on the left) and the recessive (on the right) model.

5.2.6. sHLA-G plasmatic levels and CPT-11 pharmacokinetic parameters

Based on the possible role of the sHLA-G plasmatic level in patients with mCRC treated with first-line FOLFIRI regimen, we decided to investigate the possible correlation between sHLA-G levels and pharmacokinetic parameters (PK), the latter values obtained by our previous work (in particular, CPT-11 AUC, SN38 AUC (active form of irinotecan), SN38G (inactive glucuronidation form of irinotecan), GR (glucuronidation ratio) and BI (biliary index)) [78]. The BI was defined as the product of the irinotecan AUC and the ratio of the SN38 AUC over the SN38G AUC [78].

For this analysis, patients were stratified in two groups: above (sHLA-G High) and equal or below (sHLA-G Low) the median sHLA-G value (86.4 U/ml). Patients with sHLA-G High were associated with lower levels of plasmatic CPT-11 (p=0.0216) (Figure 25). Similarly, a significant inverse correlation was found between sHLA-G and BI (p=0.0181): patients with sHLA-G High were associated with lower levels of BI (median value 3.6 U/ml; range: 1.4-7.7) and patients with sHLA-G Low with increased levels of BI (median value 7.6 U/ml; range: 3.2-12.1) (Figure 25). Further, we observed that *Del* allele was prevalent in patients in the sHLA-G High subgroup (*Del*=69%, *Ins*=31%), while *Ins* allele was the most abundant in patients in the sHLA-G Low (*Del*=46%, *Ins*=54%); the difference in the distribution of *Del* and *Ins* alleles and related genotypes comparing the subgroups, was not statistically significant (*Del* vs *Ins*, p=0.1513; *Del/Del* vs *Ins/Del* vs *Ins/Ins*, p=0.2399).

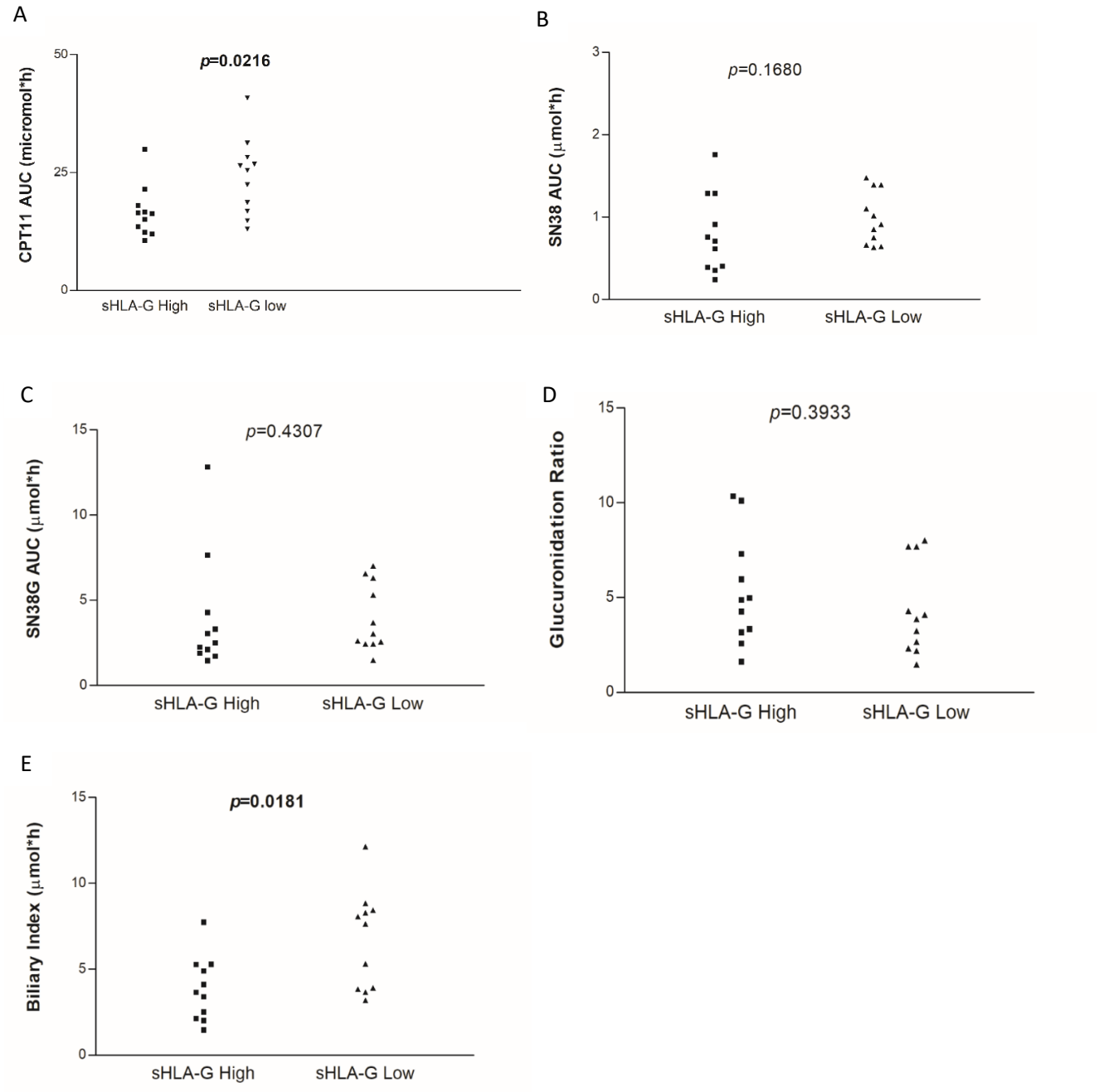

Figure 25. Distribution of PK parameters according to plasma sHLA-G High or Low group. A: CPT-11 AUC; B: SN38 AUC; C: SN38G AUC; D: glucuronidation ratio; E: biliary index.

5.3.DISCUSSION

The soluble HLA-G expression has been investigated for its possible role as clinical biomarker in cancer [110]. The possibility to measure and to monitor the level of soluble HLA-G in body fluids such as plasma, serum, ascites, cerebrospinal fluids exudates, assumed increased importance in the last years, to early find variations in sHLA-G levels as a tool for a precocious diagnosis of the cancer disease and to monitor the appearance of tumor recurrence. In this PART 2, we analyzed a subgroup of 40 patients with mCRC treated with first-line FOLFIRI regimen. To the best of our knowledge, this is the first study in which the *HLA-G* coding exons (exon 2, 3, 4) for the mature protein were genotyped. From the sequencing of the HLA-G coding regions we observed the distributions of mutant and wild type alleles that correspond to 5 different HLA-G proteins: *G*01:01, G*01:03, G*01:04, G*01:05N* and *G*01:06*. In particular, the *G*01:01* was the coding allele more frequent (77.5%) and it was characterized by the wild type alleles for all the HLA-G coding polymorphisms corresponding to missense variations. The rs12722477-*A* allele (c.400 C>A) defines the *HLA-G*01:04* coding allele (the second allele more frequent with a percentuage of 16.3%) and determine a missense variation at the codon 110 (p.Leu110Ile) located on a connecting β-sheet strand (S6 β-sheet) in the α2 domain. However, the two amino acids are both no polar hydrophobic residues and then the molecule structure should not undergo drastic changes, even if, an additional methyl group in the side chain of Ile residue may significantly perturb DNA-protein interactions [243]. The rs41551813-*T* allele (c.163 A>T) characterizes the *HLA-G*01:03* coding allele that presents a missense variation at codon 31 (p.Thr31Ser), located in the S3 β-sheet, on the bottom of the HLA-G pocket. This amino acid change should not determine an important variation in the HLA-G local structure because of the chemical nature remains apolar and without a charge. However, we had hyphotesized that this SNP could have a role in the binding between *HLA-G* and *HIF-1* (Hipoxia inducible factor-1) genes in hypoxia conditions, typical of a tumoral environment [244]. The rs41557518-*Del* allele (c.460 C>Del) causes a frameshift with the downstream formation of a stop codon that characterizes the truncated protein *HLA-G*01:05N* [245]. This deletion is present in the short loop (Glu128-Arg131) between the S7 and S8 β-sheets. Finally, the rs12722482-*T* allele is peculiar for the *HLA-G*01:06* coding allele and corresponds to the missense variation at the codon 258, from a polar to a no polar residue (p.Thr258Met), located in the α3 domain of the HLA-G. This change determines the introduction of the sulfur on the protein structure in the S15 β-sheet. Modulation of the HLA-G expression is observed in many pathological conditions such as cancer, inflammatory and autoimmune diseases, viral infections, transplantations [246]. Several studies have suggested an HLA-G genotype-phenotype association such as *HLA-G*01:04* has been associated with high HLA-G secretion levels while *HLA-G*01:03* and *HLA-G*01:05N* with low levels [115], [247]. In healthy subjects, low sHLA-G levels were associated with genotypes carrying the *G*01:01:03:01/02/03* or *G*01:20* (corresponding to the old nomenclature *G*01013*) and *G*01:05N* (corresponding to *G*0105N*) alleles, whereas high sHLA-G levels were preferentially associated with *G*01:04:01/04/05* or *G*01:07*, *G*01:15, G*01:19* (corresponding to the old nomenclature *G*01041*) carriers [125].

The overall distribution of sHLA-G is not a Gaussian distribution. The median sHLA-G level was 116.44 U/ml. No significant associations were found between sHLA-G levels and *HLA-G* 3'UTR polymorphisms, haplotypes, secretor models, coding polymorphisms, and coding alleles.

Our preliminary data from the genotype-phenotype analysis, may suggest that the mechanisms involved in the regulation of the individual sHLA-G level are not based on the genetics of HLA-G. However, the small sample size compared to the number of genetic variables to analyze, may explain the unreached significance. In fact, although we did not find significant associations, we observed that the distribution of sHLA-G in the functional *HLA-G* rs371194629 genotypes, the most reported to impact the HLA-G expression among the *HLA-G* 3'UTR polymorphisms, confirmed the literature data [119]. Moreover, this data was also

confirmed, even if not significantly, analyzing the CPT11/BI-PK parameters in subgroups of patients stratified according to the median sHLA-G value: patients grouped in sHLA-G High had a prevalence in *Del* allele, and those in sHLA-G Low had a prevalence in *Ins* allele. Of note, another interesting result was found for the coding *HLA-G* rs1130355 SNP that corresponds to a synonymous change. Although the mutated rs1130355-*G* allele was associated, but not significantly, with low sHLA-G levels, a trend was observed suggesting a possible mechanism in the alteration of the miRNA binding sites, influencing *HLA-G* transcription efficiency, that for the first time is here hyphothesized. Furthermore, the unexpected prevalence of the rs1130355-*A* allele, reported as the wild type allele in EUR subpopulation, may suggest an association with the CRC disease, that should be explored and confirmed in a larger cohort of patients.

The only significant association with a clinical parameter was found analyzing the number of metastatic sites. Higher sHLA-G levels were associated with more than 1 metastatic site. The correlation between expression of sHLA-G and tumor metastasis seems to indicate that HLA-G could facilitate tumor immune escape, invasiveness and disease progression, as supported by literature data that associated increased sHLA-G with a worse prognosis [145], [184]. It was reported that HLA-G positive primary CRC tumors are associated with a shorter survival [96], [145], [248] and associated liver metastasis [249]. High levels of the peripheral sHLA-G, also showed a negative prognostic value in CRC [95].

Nonetheless, a recent study for the first time showed that high sHLA-G levels were associated with a favorable outcome in stage III patients with CRC, in particular with a significantly longer liver metastasis free survival, even if the chemotherapy treatment was not specified [96].

The results found in PART 1 revealed an association of *HLA-G* 3'UTR polymorhisms related to a "high" secretor HLA-G, with a favorable prognosis. The results in PART 2 highlighted the association of high sHLA-G levels with an increased (>1) number of metastases in 37 metastatic patients with CRC. Although this result is concordant with the aspected increase in the tumor escape phenomenon [39], we also observed that patients with high sHLA-G levels were associated with diminished irinotecan and BI levels, suggesting an improved efficacy of the drug metabolism and glucuronidation and detoxification pathways in these patients. It is possible that in presence of tumor metastasis the increased expression of sHLA-G may represent an indirect mechanism of protection against cancer. About that, the interesting result for the PK parameters could suggest a possible mechanism of irinotecan captured by the HLA-G molecule that alters the interaction of HLA-G with its target receptors, blocking the induction of tolerogenic effect that facilitates tumor immune escape and then determines a favorable prognosis.

6. PART 3: HLA-G/irinotecan interaction

6.1. MATERIALS AND METHODS

6.1.1. Recombinant HLA-G structure determination: *in silico* analysis

6.1.1.1. Molecular visualization programs

Before running a docking analysis, a visualization and evaluation of the structure of the protein of interest is necessary. Biophysical techniques, such as X-ray crystallography or less frequent NMR spectroscopy, are normally used to determine the experimental protein structures. The coordinates of the structural models of proteins are usually deposited in the most comprehensive resource, the RCSB PDB Protein Data Bank (http://www.rcsb.org). Then, the structural models taken from the PDB database are generally the starting points for docking analysis. These coordinates, opportunely adapted, can be used as input for specific docking program. The crystal structure of HLA-G (PDB ID: 1YDP) [101], determined through a molecular replacement method using the HLA-E structure as search model, was visualized with Visual Molecular Dynamics (VMD) 1.9.3 version, and PyMOL software, molecular visualization programs for displaying, animating, and analyzing large biomolecular systems using 3D graphics and built-in scripting. The docking results obtained were analyzed with UCSF Chimera program, another highly extensible program for interactive visualization and analysis of molecular structures and related data: with it were also generated all the high-quality images here presented.

6.1.1.2. Molecular docking

The task of docking strategy is to predict the structure (pose), with the minimum free energy of the global system, of the resulting complex between a protein and a ligand such as small molecules or peptides. This method was used to decide whether irinotecan possesses a good steric and chemical complementarily to the HLA-G protein. The correct selection of the preferred ligand orientation to form a stable complex is fundamental to predict the strength of non-covalent association or binding affinity. Crucial is the choice of scoring functions that better fit to the contest. It is preferable to choose one that it has been parameterized and tested for the same protein family and ligand class under study [250]. Several different docking software have been developed and are now available [251]. SwissDock, the online docking web server of the Swiss Institute of Bioinformatics is one of these platforms with an easily-accessible web portal, realized with the aim of extending the use of protein-small molecule docking software [252]. In particular, it allows the download of manually assembled protein structures and the preparation of original PDB files through *ad hoc* scripts and the online visualization of the predicted binding modes. Moreover, a more detailed analysis can be performed using the supported integration with the UCSF Chimera molecular viewer. Flexible docking option was used, where the internal torsions of the ligand changes, while the protein remains fixed. SwissDock is based on the docking software EADock DSS [253]. This software generated a huge number of poses and estimated their CHARMM energies on a grid considering the solvent effect and ranking the most favorable clusters. We used the SwissDock web server to perform EADdock-based docking simulations between HLA-G protein (PDB ID: 1YDP) and irinotecan (ZINC1612996 code). The virtual screening-based, ready-to-dock, 3D structure of irinotecan was available from ZINC Database (download at http://zinc.docking.org) and then CHARMM topology, parameters and coordinates are derived automatically from the Merck Molecular Force Field (MMFF) [252]. In order to simplify the molecular

docking, the protein model was restricted to the α1 and α2 domains containing the antigen-binding cleft of HLA-G (180 residues from Ser2 to Arg181 of 1YDP PDB coordinates). Beta-2-microglobulin protein, histone 2a peptide and water molecules were omitted from the starting model of the protein. The structural model of the α1 and α2 domain of HLA-G was completed adding the hydrogen atoms by PyMOL software. Eight different possible orientations of irinotecan into the antigen-binding cleft were investigated (two opposite orientations with 4 different rotations around to the main axis of the molecule differing by 90°). In order to remove the conformation rigidity of a single protein model, these 8 structural models of the complex (assembled by Coot graphical program) were relaxed by REFMAC5 software [254] and were used as representative models of a possible adaptation of the antigen-binding cleft to the ligand binding. These protein models were used as input for the molecular docking of HLA-G/irinotecan complexes. In the SwissDock three different docking protocols can be selected: very fast, fast and accurate. We selected the accurate type because irinotecan has a limited conformational freedom with few rotable bonds.

6.1.2. Chemical and reagents

CPT-11 hydrocloride (Sigma-Aldrich, Saint Louis, MO, USA), with a molecular weight of 677.18 g/mol, was dissolved in dimethylsulfoxide (DMSO) solution at a stock concentration of 1.6 mM and then diluited with milliQ water to the different work concentrations used in the experiment. CPT-11 hydrocloride were kept at 4 °C and fresh solution were prepared every 3-6 hours to consider the hydrolysis of the lactone.
Bovine serum albumin (BSA) was purchased from Sigma-Aldrich (Saint Louis, MO, USA) and its molecular weight was assumed to be 66 kDa to calculate the molar concentrations. The stock solution was prepared at a concentration of 1 mg/ml (corresponding to 15 µM) in milliQ water. After reconstitution, for short term storage the protein was kept at 4°C, while for long period it was storage at -20°C.
We purchased two batch of the recombinant glycosylated HLA-G protein expressed in mammalian cells with N-terminal with a 6xHis-tagged (MyBioSource, San Diego, CA, USA). It was lyophilized from a 0.2 µm filtered 10 mM Tris-HCl, 1 mM EDTA, pH 8.0. The HLA-G aminoacidic sequence, related to the mature HLA-G protein of 314 AAs, is described below:

```
        10         20         30         40         50         60
GSHSMRYFSA AVSRPGRGEP RFIAMGYVDD TQFVRFDSDS ACPRMEPRAP WVEQEGPEYW
        70         80         90        100        110        120
EEETRNTKAH AQTDRMNLQT LRGYYNQSEA SSHTLQWMIG CDLGSDGRLL RGYEQYAYDG
       130        140        150        160        170        180
KDYLALNEDL RSWTAADTAA QISKRKCEAA NVAEQRRAYL EGTCVEWLHR YLENGKEMLQ
       190        200        210        220        230        240
RADPPKTHVT HHPVFDYEAT LRCWALGFYP AEIILTWQRD GEDQTQDVEL VETRPAGDGT
       250        260        270        280        290        300
FQKWAAVVVP SGEEQRYTCH VQHEGLPEPL MLRWKQSSLP TIPIMGIVAG LVVLAAVVTG
       310
AAVAAVLWRK KSSD
```

The HLA-G was reconstituted in 100 µl of deionized sterile water (milliQ) to a final concentration of 1 mg/ml, corresponding to a concentration of 25 µM, according to the datasheet instructions. The predicted band size of the HLA-G expressed and declared in the datasheet was 39.6 kDa; in the first batch purchased the band size observed by SDS-PAGE quantitative densitometry by Coomassie Blue Staining was of 40 kDa,

while in the second batch was of 42 kDa (Figure 26). The purity declared was ≥90%. After reconstitution, for short term storage the protein was kept at 4°C, while for long period it was frozen at -20°C.

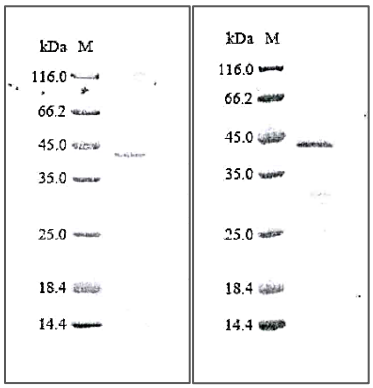

Figure 26. SDS-PAGE and quantitative densitometry by Coomassie Blue Staining declared for the HLA-G. On the left, the first batch HLA-G; on the right, the second batch HLA-G.

6.1.3. Bradford protein assay

We used the Bradford protein assay to measure the protein concentration, in particular for the HLA-G (MyBioSource, San Diego, CA, USA) and bovine serum albumin (BSA) (Sigma-Aldrich, Saint Louis, MO, USA). The binding of the basic amino acid residues of the proteins (Arginine, Lysine and Histidine) to Coomassie dye under acidic conditions, results in a color change from brown to blue. Sample preparation provided the addition of 800 µl of milliQ water, 1 µl of HLA-G 25 µM or 1µl of BSA at 0.01, 0.05, 0.1, 1, 5, 10, 25, 50, 100, 150 and 200 µM and 200 µl of Bradford 5X in 10 mm plastic cuvette. The absorbance of the blank, HLA-G and BSA samples was read in the NanoDrop 2000C UV-Vis Spectrophotometer (ThermoFisher Scientific, Waltham, MA, USA) at 595 nm using a 5-point BSA calibration curve from 0 to 20 mg/ml. The absorbance measurements at 595 nm are strictly linear with protein concentration according to the Beer-Lambert law [255] (Figure 27).

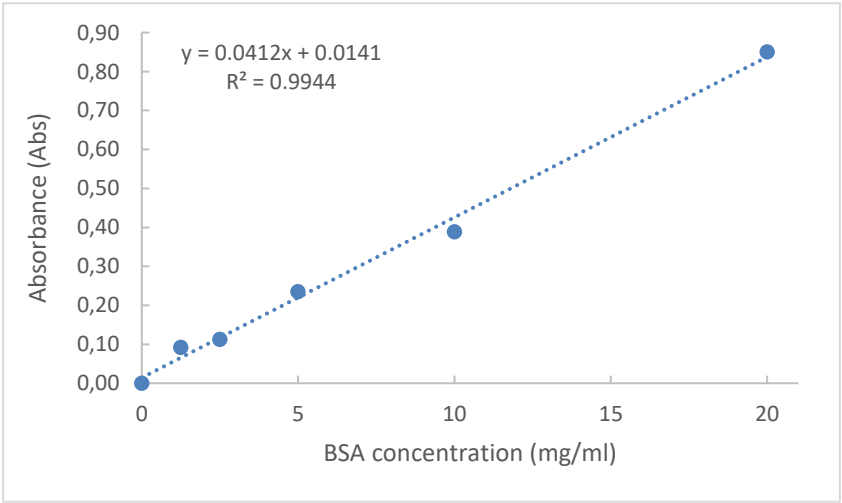

Figure 27. BSA calibration curve.

6.1.4. UV-Vis absorption experiments

The UV-Vis spectrophotometry is based on capability of the molecules to absorb UV-rays proportional to their concentration. The instrument measures the ratio between the intensity of the light transmitted and that incident, in particular between two beams having the same wavelength and placed on the same optical path. The UV-Vis absorption spectra were recorded over a wavelength range of 200-800 nm. Absorption spectra of blank, HLA-G, BSA and CPT-11 were recorded on the Cary Eclipse UV-Vis Spectrophotometer (Agilent Technologies, Santa Clara, CA, USA) and were the mean of 4 spectra acquisitions. The final concentrations analyzed were 0.01, 0.05 and 0.1 µM for HLA-G, and 0.01, 0.05, 0.1, 1, 5, 10, 25, 50 µM for CPT-11. The BSA was also analyzed at several concentrations. Quartz cuvettes with a pathlength of 10 mm and a volume of 700 µl (Hellma-Analytics, Müllheim, Germany) were used. We calculated the molar concentration though the Beer-Lambert law [255] using the theoretical extinction coefficient obtained from the primary sequence of the proteins calculated with the ExPASy ProtParam tool (http://www.expasy.org/tools/pi_tool.html).

6.1.5. Fluorescence measurements

The fluorescence spectroscopy is used to investigate several chemical aspects such as the interaction between molecules that produces a fluorescence quencing. Among the very interesting characteristics of this technique, there are the high sensibility (pg or fg), high selectivity, a linearity of the response in a wide concentration range higher than the one available for UV-Vis spectrophotometry, and a great versatility due to the possibility to change the excitation and emission wavelengths. To note is that the signal produced by a spectrofluorimeter is not an absolute measure; it depends on the sample characteristics and other instrumental aspects. It measures only the intensity of the emission light at a desired wavelength using an arbitrary scale and the signal is expressed in arbitrary unit (a.u.). The intensity of the emission light could be correlated with the sample concentration not according to the Beer-Lambert law as in the spectrophotometry and only the ratio between measurements done in the same experimental conditions are significant. Moreover, attention should be paid to impurities presents in the sample or in the solvent and to Rayleigh-Tyndall and Raman scattering. The Rayleigh-Tyndall bands are due to the light diffusion of solvent molecules or air/wall interfaces of the cell and wall/solution (Rayleigh scattering) or to the presence of small particles in suspension (Tyndall scattering). The light diffusion is inversely proportional to the absorbance at that λ_{exc} and could be attenuated changing the λ_{exc} or λ_{em} and/or reduced the slits. Raman bands, instead, derived from the interaction between photons of the excitation radiation and the vibrational levels of the solvent molecules, to which they yield a constant fraction of their energy. The maximum of this band is strictly correlated with the λ_{exc}, then it moved with the λ_{exc} changing. Moreover, without a cutoff filter, in a spectrum we can observed also the harmonics bands at integer multiples (m) of the λ_{exc} used, with a lower intensity increasing m.

Emission spectra give important information about the nature and energy of the excited species. To record an emission spectrum, the sample is excited at a fix wavelength (λ_{exc}) based on awavelength absorption of the analyzing sample, and then the intensity variation of the luminescence with the emission wavelength (λ_{em}) is measured.

To record an excitation spectrum, instead, the λ_{em} is fixed and it is generally chosen on the maximum of sample emission, and then the λ_{exc} varies on the range of sample absorbance. It is proportional to the absorbance spectrum only in particular conditions and after opportune corrections. However, the shape of both is the same. The measurement of an excitation spectrum could be used to quantify the presence of

energy transfer phenomena to a molecule (A) we evaluated the emission from another molecule (B) present in the solution. In fact, without energy transfer, the excitation spectrum is proportional to the absorbance spectrum of A, while in the present of energy transfer, the excitation spectrum is proportional to the sum of the absorbance spectrum of A and B (efficiency of 100%).

The fluorescence spectroscopy allows us to establish if a molecule is a good fluorophore. The emission-quantum yield, defined as the ratio between the number of emitted photons and those absorbed, is independent of the λ_{exc} and could give an information compared with a standard fluorophore at the same experimental conditions.

Fluorescence emission spectrum of blank, HLA-G, BSA and CPT-11 was measured using the Cary Eclipse Fluorescence Spectrophotometer (Agilent Technologies, Santa Clara, CA, USA). The concentrations analyzed were 0.01, 0.05 and 0.1 µM for HLA-G, 1µM for BSA and 1, 5, 10, 15, 20 µM for CPT-11. Emission spectrum of the molecules was recorded from 280 to 550 nm and from 365 to 550 nm at an excitation wavelength of 275/280 and 360 nm, respectively. The emission and excitation slits widths were set at 2.5 and 5 nm with a photomultiplier tube (PMT) detector voltage of 800 and 500 Volts, respectively. Quartz cuvettes composed exclusively of silicon dioxide (SiO2) with a pathlength of 3 mm x 3 mm and a volume of 45 µl (Hellma-Analytics, Müllheim, Germany) were used. Using the same instrument setting, we performed a titration assay measuring the fluorescence emission spectra of HLA-G and CPT-11, adding to a fix concentration (0.1 µM) of HLA-G a certain amount of CPT-11 each time, changing concentrations of CPT-11 in solution to 1, 5, 10, 15, 20 µM during the interaction. We investigated the interaction between BSA protein and CPT-11 at the same conditions.

6.2.RESULTS

6.2.1. Structural analysis

6.2.1.1. Recombinant HLA-G structure

The recombinant HLA-G protein is composed by 314 residues and has a molecular weight of 35606.92 Da. The protein has a theoretical isoelectric point (pI), the pH at which it carries no net electrical charge or is electrically neutral in the statistical mean, of 5.45. The amino acid (AA) number and frequency in the HLA-G protein sequence was determined with the ExPASy ProtParam tool (http://www.expasy.org/tools/pi_tool.html) as well as other chemical characteristics such as the extinction coefficient at 280 nm measured in water of 76235 $M^{-1}cm^{-1}$, the aliphatic index of 69.33 and the grand average of hydropathicity of -0.561. The total number of aromatic residues present in the sequence of HLA-G is 31 (Table 47).

Table 47. Amino acid composition of HLA-G protein sequence.

Classification	Amino acid	Code	Letter	n	%
No polar	Alanine	Ala	A	32	10.2
	Glycine	Gly	G	22	7.0
	Isoleucine	Ile	I	8	2.5
	Leucine	Leu	L	24	7.6
	Methionine	Met	M	8	2.5
	Valine	Val	V	21	6.7
Polar, no charged	Asparagine	Asn	N	6	1.9
	Cysteine	Cys	C	6	1.9
	Glutamine	Gln	Q	17	5.4
	Proline	Pro	P	16	5.1
	Serine	Ser	S	17	5.4
	Threonine	Thr	T	19	6.1
Aromatic	Phenylalanine	Phe	F	7	2.2
	Tyrosine	Tyr	Y	14	4.5
	Tryptophan	Trp	W	10	3.2
Basic	Arginine	Arg	R	24	7.6
	Histidine	His	H	9	2.9
	Lysine	Lys	K	10	3.2
Acidic	Aspartic acid	Asp	D	18	5.7
	Glutamic acid	Glu	E	26	8.3

The aminoacids were divided by chemical characteristic of R-group.

Localization of the aromatic residues (in red) in the AA primary sequence of the recombinant HLA-G:

```
         10         20         30         40         50         60
GSHSMRYSA AVSRPGRGEP RFIAMGYVDD TQFVRFDSDS ACPRMEPRAP WVEQEGPEYW
         70         80         90        100        110        120
EEETRNTKAH AQTDRMNLQT LRGYYNQSEA SSHTLQWMIG CDLGSDGRLL RGYEQYAYDG
        130        140        150        160        170        180
KDYLALNEDL RSWTAADTAA QISKRKCEAA NVAEQRRAYL EGTCVEWLHR YLENGKEMLQ
        190        200        210        220        230        240
RADPPKTHVT HHPVFDYEAT LRCWALGFYP AEIILTWQRD GEDQTQDVEL VETRPAGDGT
        250        260        270        280        290        300
FQKWAAVVVP SGEEQRYTCH VQHEGLPEPL MLRWKQSSLP TIPIMGIVAG LVVLAAVVTG
        310
AAVAAVLWRK KSSD
```

6.2.1.2. *In silico* analysis of HLA-G protein

The most interesting aspect to highlight was not only to analyze the primary sequence of the mature protein but also the secondary, and mainly the tertiary and quaternary structures by 3D visualization of the AAs localization, to investigate the chemistry of the possible HLA-G interactions.

Five crystal structures of the HLA-G complex are available on RCSB Protein Data Bank (PDB) Database (http://www.rcsb.org) with the lower resolution of 3.2-Å, for the crystal structure of disulfide-linked HLA-G dimer (PDB ID: 2D31) [256]. The crystal structure of the HLA-G complex with the PDB ID=1YDP obtained at 1.9-Å resolution was chosen for our analysis [101]. HLA-G is a heterodimeric protein composed by a heavy chain (HLA-G) and a light chain (β2M) that binds with non-covalent interactions with HLA-G. The crystal structure presents also the histone 2a peptide (Figure 28, in purple), crystalized together with the HLA-G protein. The histone 2a peptide is positioned in the antigen-binding cleft of the protein. The HLA-G is composed by 4 α-helices and several β-sheet strands. The α1-α2 domains of HLA-G structure hosts the antigen-binding cleft while the α3 domain is involved in the heterodimeric interaction with the β2-microglobulin chain. Considering the aromaticity of the irinotecan molecule, the HLA-G structure was analyzed highlighting in red the aromatic AAs that are potentially involved in π–π stacking interactions (Figure 28). In particular, the helices 2-3 constitute the groove for the molecular recognition of the antigen. In this site there are several aromatic residues (tryptophan, tyrosine and phenylalanine) that could participate to the specific molecular recognition of aromatic molecules such as irinotecan. The analysis of the HLA-G complex also revealed the presence of a central “hole” between HLA-G and β2-microglobulin surfaces particularly rich of aromatic residues (Figure 28).

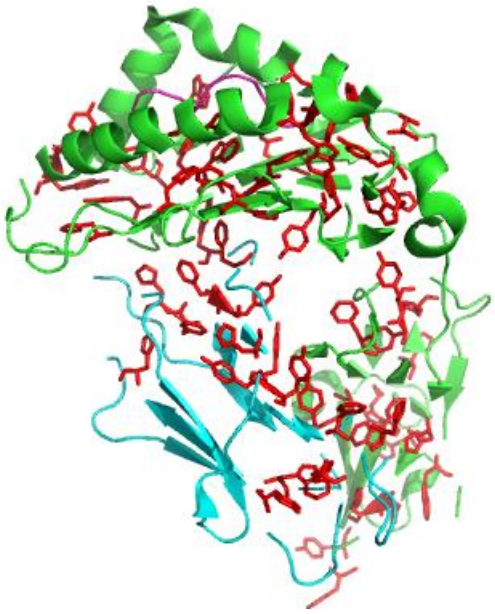

Figure 28. Cartoon representation of the crystal structure of HLA-G complex (PDB ID: 1YDP). The heterodimer consists of a heavy chain, the HLA-G protein (in green) and a light chain, the β2-microglobulin (light blue). The histone 2a peptide in the antigen-binding cleft of HLA-G, is also represented (purple). The aromatic residues are evidenced in red.

We decided to investigate, as priority, the antigen-binding cleft of HLA-G as possible binding site for irinotecan since its specific morphology and presence of aromatic residues. For that reason, we considered the α1-α2 domains of HLA-G alone, without the β2-microglobulin, to set the model and evaluated our hypothesis.

The possible irinotecan conformations adopted within the binding sites of the 8 structural models of α1-α2 domains of HLA-G (representative of the antigen-binding cleft flexibility) were investigated by SwissDock platform. The software produced more than 2000 different possible complexes. These structural models were clustered and ordered in term of binding score and analyzed with Chimera program. Within the cluster of conformations having the best binding score we have chosen, as representative of ligand/protein interactions, the HLA-G/irinotecan complex that showed a less deformed binder. The full fitness was 11004.30 Kcal/mol with a Gibbs free energy of about -9.8 kcal/mol.

In the following pictures, obtained with Chimera program, we reproduce this selected model of HLA-G/irinotecan complex obtained with the molecular docking, evidencing the chemical characteristics that allow the specific intermolecular interactions between protein and ligand.

The three orthogonal views of the structural model of the complex depicted in Figure 29, with cartoon representation of the protein and stick representation of irinotecan molecule, show the irinotecan molecule well inserted in the antigen-binding cleft of HLA-G. The irinotecan molecule adopts an extended conformation.

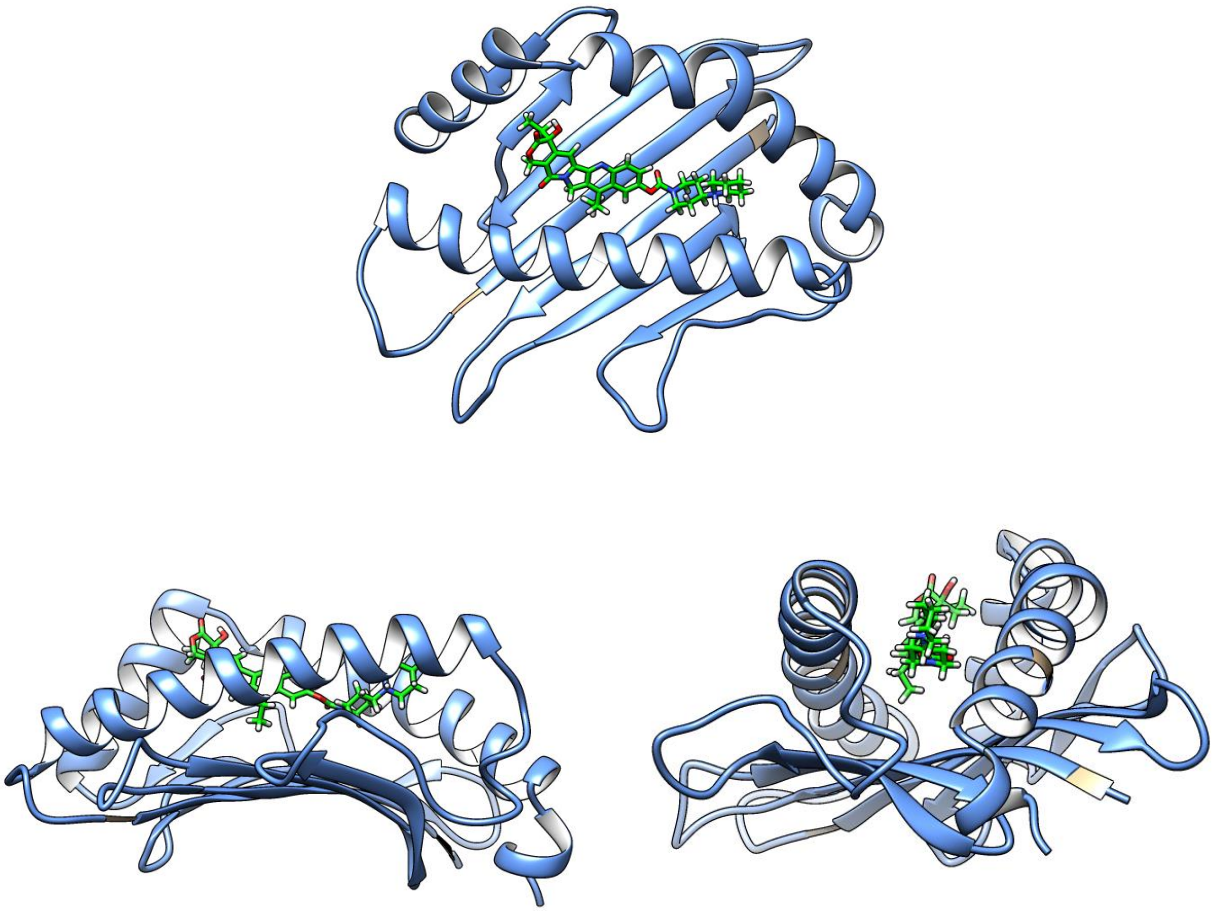

Figure 29. Three orthogonal views of the structural model of the complex HLA-G/irinotecan.

The size and shape of the irinotecan molecule fits remarkably well into the antigen-binding cleft of HLA-G (Figure 30). The whole cleft is filled by the guest molecule. Furthermore, irinotecan is predominantly a hydrophobic molecule and correspondingly the groove of HLA-G is notably hydrophobic (see Figure 30).

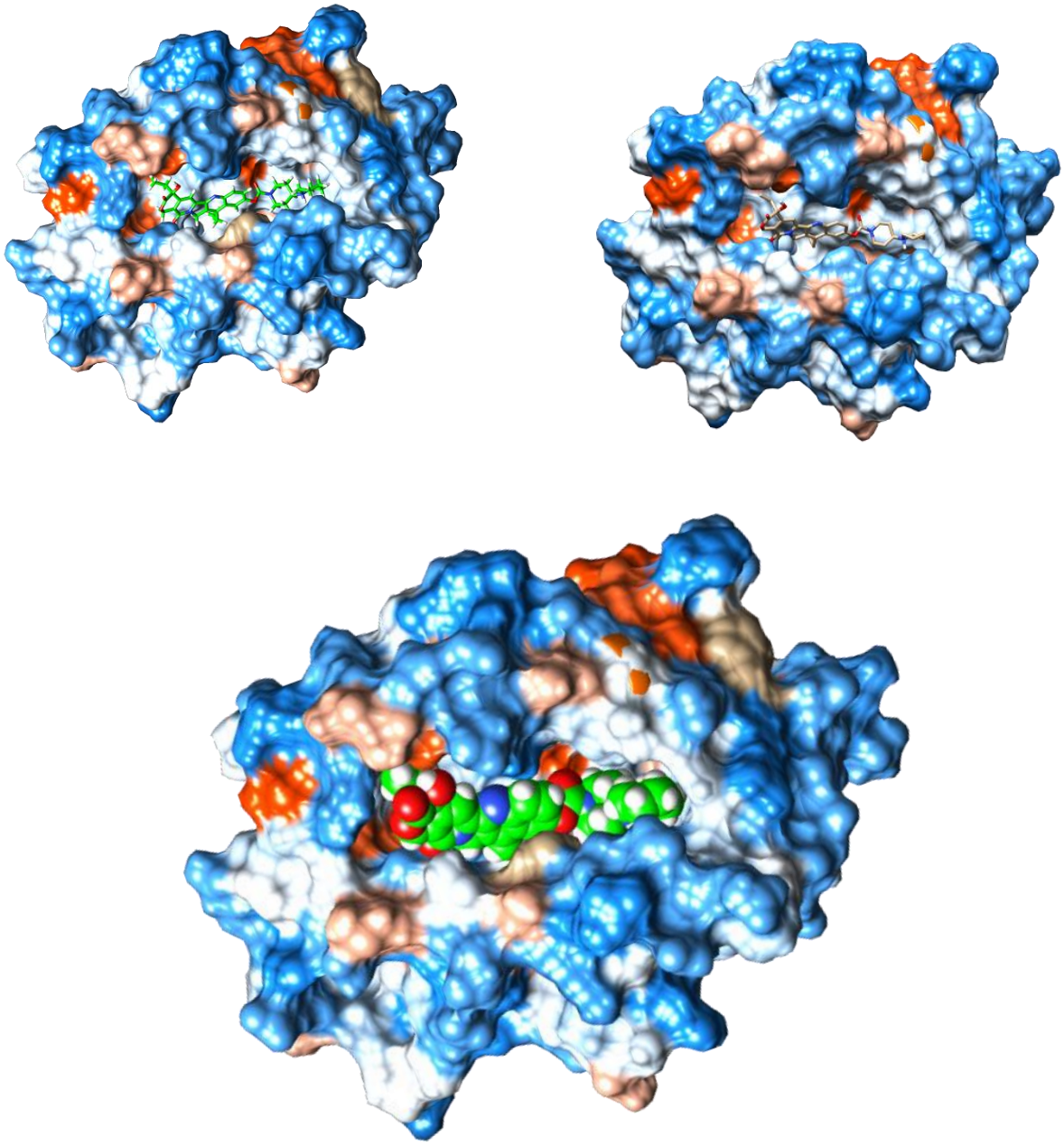

Figure 30. Hydrophobicity surface of the HLA-G/irinotecan complex shows amino acid hydrophobicity in the Kyte-Doolittle scale with colors ranging from dodger blue for the most hydrophilic to white at 0.0 to orange red for the most hydrophobic.

We observed the well-known HLA-G groove formed by the α-helics as flanks and β-sheets as bottom that harbors the irinotecan. The isoelectric point of the protein calculated from the amino acid sequence is 5.45. This means that HLA-G is a slightly acid protein. The number of acidic residues (Asp + Glu = 44 residues) exceeds the number of basic residues (Lys + Arg = 34 residues). HLA-G has also 9 His residues. These residues, if protonated, can partially reduce the total negative charge of the protein. Also the β-2-microglobuline is a slightly acid protein (theoretical IP = 6.05) with 15 positively charges AAs and 13 negatively charges AAs. Then, the overall HLA-G heterodimeric complex should be negatively charged at neutral pH. Analogue situation is observed if we restrict the analysis to α1-α2 domains of HLA-G protein, such as in this theoretical study. In this case, the total number of negatively charged residues (Asp + Glu =

31) is not balanced by the total number of positively charged residues (Arg + Lys =25). The irinotecan molecule is monoprotonated at neutral pH. Then, a negatively charged protein is an ideal counterpart for formation of a complex. In particular, it should be considered that long-range electrostatic interactions are the driving forces for molecular recognition. The possible charge complementarity of HLA-G/irinotecan complex has been confirmed by the electrostatic potential of the protein and irinotecan surfaces calculated using Amber ff14SB force field. In Figure 31 are depicted the electrostatic potential of the surfaces of protein and irinotecan, according to Coulomb's law. The negative part of the protein (red color) is located close to the N-terminal, while on the other side of the crevice a positive surface (blue color) of protein is present. The molecular docking model shows that the irinotecan molecule is oriented in such a way that its positive terminal part containing the protonated N-amino group (blue color) interacts with the negative part of the cleft (red color). On the other side, a slightly negatively charged surface of irinotecan (pale red) lies close to the positive part of the protein (blue). Then from a preliminary analysis, the final model seems to point out a precise interaction into the cleft. However, a much more detailed analysis should be carried on identify possible binding residues on the HLA-G groove and irinotecan. Moreover, the type of interaction involved should be better defined.

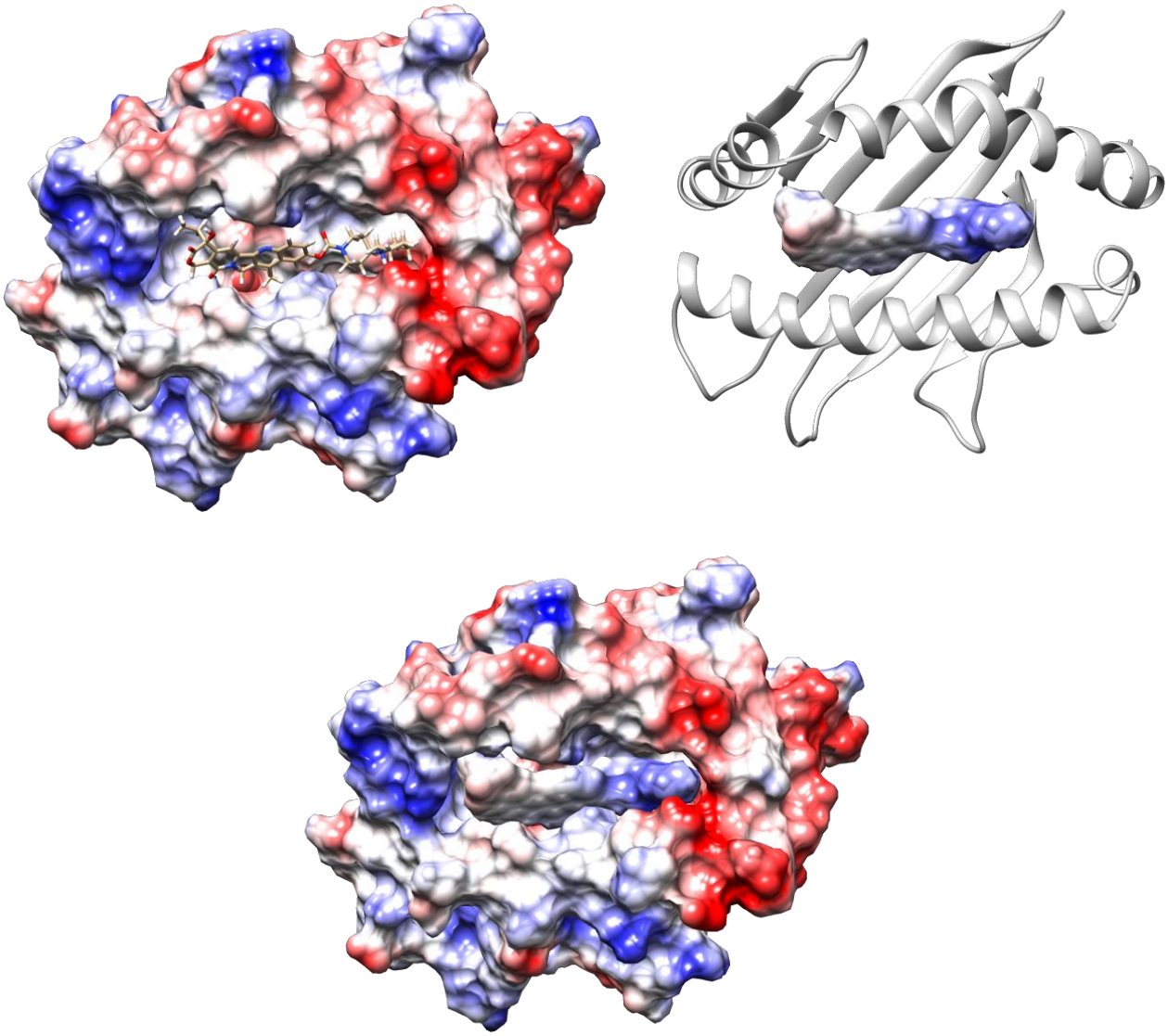

Figure 31. Electrostatic potential of the surfaces of protein and irinotecan according to Coulomb's law.

In order to identify the key residues involved in the formation of the complex, the aromatic residues and all the residues involved in the HLA-G/irinotecan interaction on the HLA-G groove are evidenced in Figure 32. Five Trp residues and 11 Tyr were identified into the groove. Among these, only Trp^{97}, Trp^{167} and Tyr^{7}, Tyr^{59}, Tyr^{116}, Tyr^{159}, Tyr^{171}, could interact with irinotecan. Moreover, the Thr^{163}, Glu^{63}, Asn^{66}, His^{70}, Thr^{73}, Ser^{9}, Ile^{99}, Arg^{153}, Val^{152}, Ala^{150}, Cys^{147}, Lys^{146}, Asn^{77} contribute to stabilize the interaction.

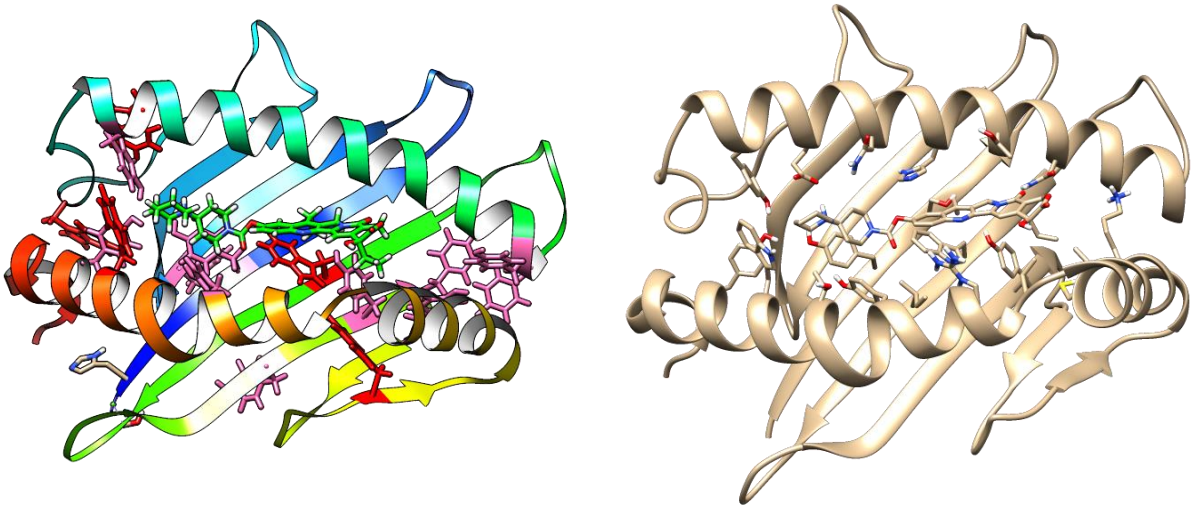

Figure 32. On the left, all the aromatic residues present in the groove: Trp (in red) and Tyr (in purple). On the right all the residues involved in the interaction.

6.2.2. Spectrophotometric analysis

An experimental analysis to better explore the possible interaction between HLA-G and irinotecan is the fluorescence spectroscopy. We based on the known absorption and fluorescent characteristics of irinotecan (CPT-11) to investigate the production of a fluorescence quenching or fluorescence enhancement in the presence of the HLA-G. First of all, we decide to set the method using the BSA protein. In fact, it is widely used as a model protein to study the binding interaction of drug with protein because of the 76% structural similarity with the serum albumin, the most abundant protein in blood plasma with a high affinity for drugs and metabolites [257]. BSA is a protein constituted by three homologous α-helics domains and with two major binding sites. Two Tryptophan residues, Trp^{134} and Trp^{213}, are located on the surface of the protein and within the hydrophobic binding pocket [258]. These hydrophobic characteristics are similar to those described before for HLA-G.

6.2.2.1. BSA protein: Bradford assay

The Bradford assay was realized to measure the BSA concentration using both the Nanodrop and Agilent instruments based on the calibration curve. We observed that the absorbance measurements have similar sensitivity (Figure 33 and Figure 34).

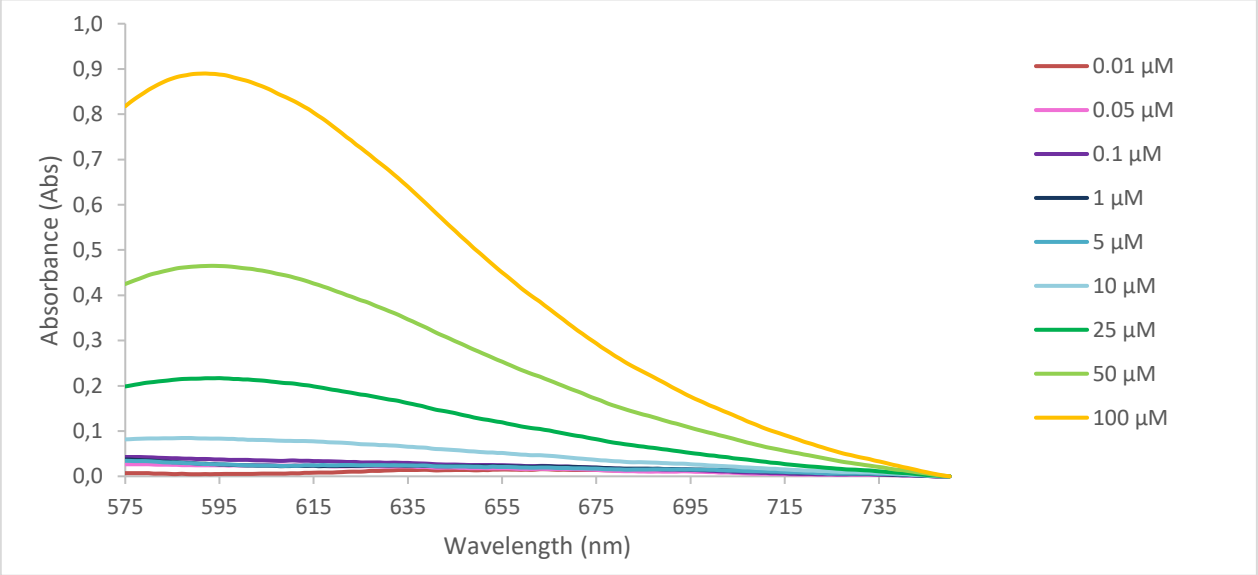

Figure 33. Absorbance spectrum of 2 µl BSA in milliQ water obtained from the mean of 3 spectra acquisitions though the Bradford assay. Calibrator curve used: y=0.0412x + 0.0141; R^2=0.9944 (Figure 27). Nanodrop.

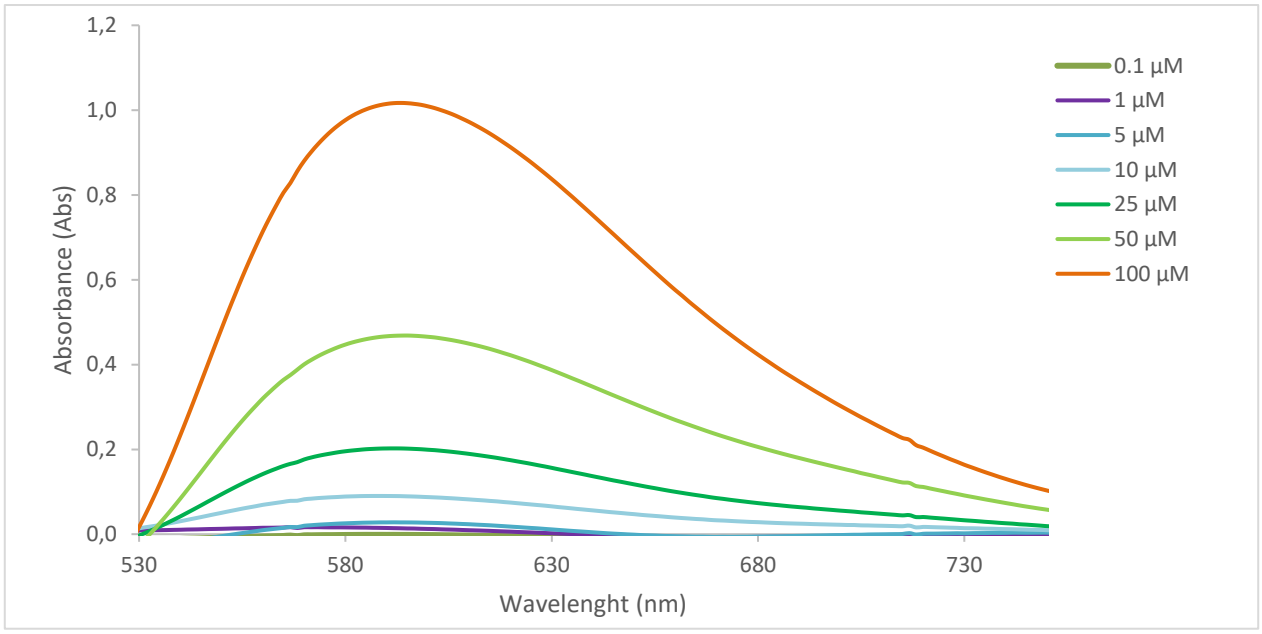

Figure 34. Absorbance spectrum of BSA in milliQ water mean of 3 spectra acquisitions though the Bradford assay. Agilent.

6.2.2.2. BSA protein: absorption spectra

UV-Vis absorption spectra of BSA show a weak band around 280 nm that belong to the π ->π* transition of the aromatic amino acids such as Trp, Tyr, and Phe, and a strong band below 240 nm (around 210 nm) due to the peptide bonds of BSA backbone (Figure 35 and Figure 36). The law that allows to linear couple the absorbance of a molecule with its concentration is the Beer-Lambert law (A=εcl), where A is the molecule absorbance, ε is the wavelength-dependent molar absorptivity coefficient ($M^{-1}cm^{-1}$), c is the sample concentration in solution ($molL^{-1}$), l represents the optical path of the cuvette expressed in cm.

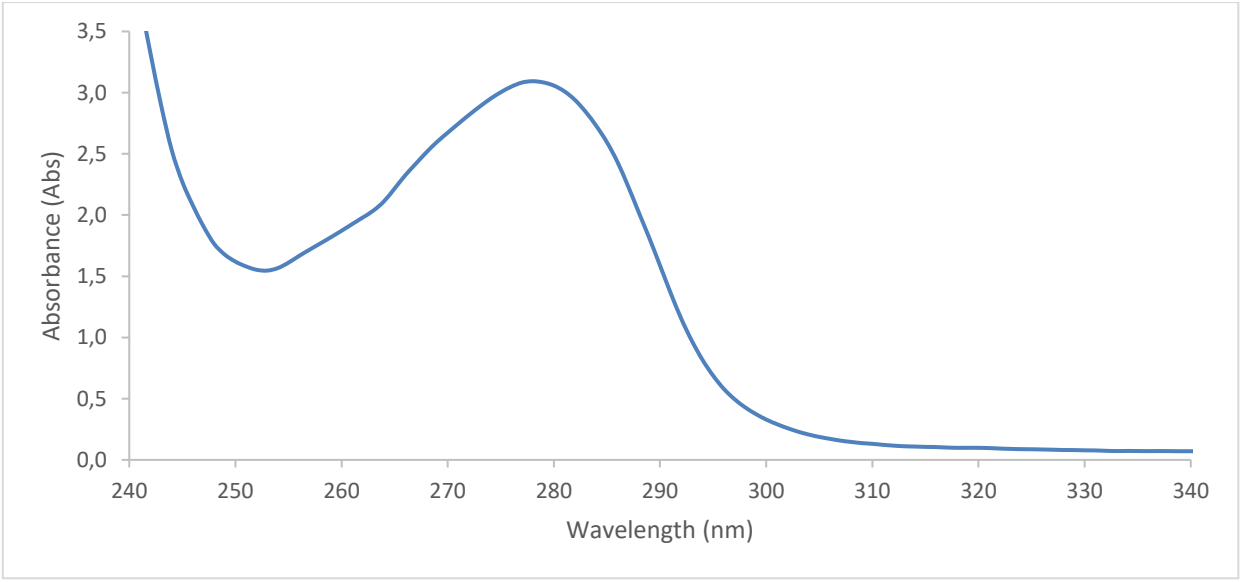

Figure 35.Absorbance spectrum of 1µl BSA in milliQ water and at concentration of 25 µM obtained from the mean of 6 spectra acquisitions though the absorbance spectrum at λ_{exc}=280 nm. Nanodrop.

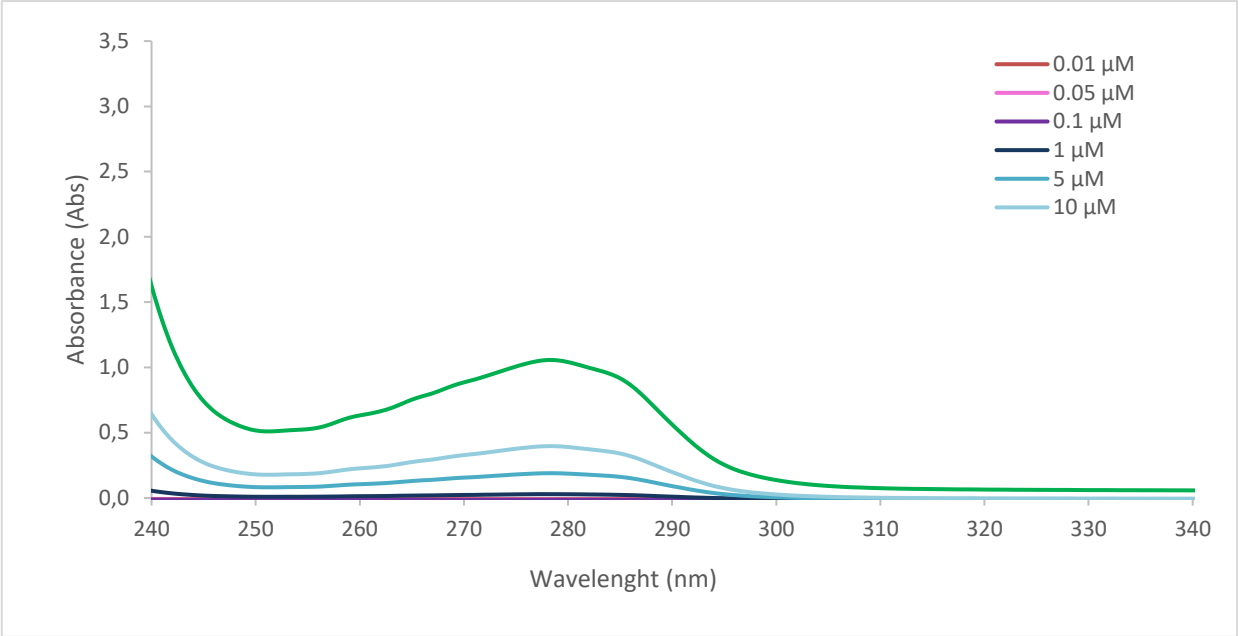

Figure 36. Absorbance spectra of BSA at λ_{exc}=280 nm. Agilent.

6.2.2.3. BSA protein: fluorescence spectra

BSA containing tryptophan, tyrosine and phenylalanine residues, has endogenous fluorescence. The main contributor to the intrinsic fluorescence of BSA was the tryptophan (Trp) that is the most sensitive residue to the changes in the micro-environment, and then widely used as an endogenous fluorescence probe to research the protein-drug interaction. BSA showed a fluorescence emission peak at 344 nm when excited at 280 nm (Figure 37).

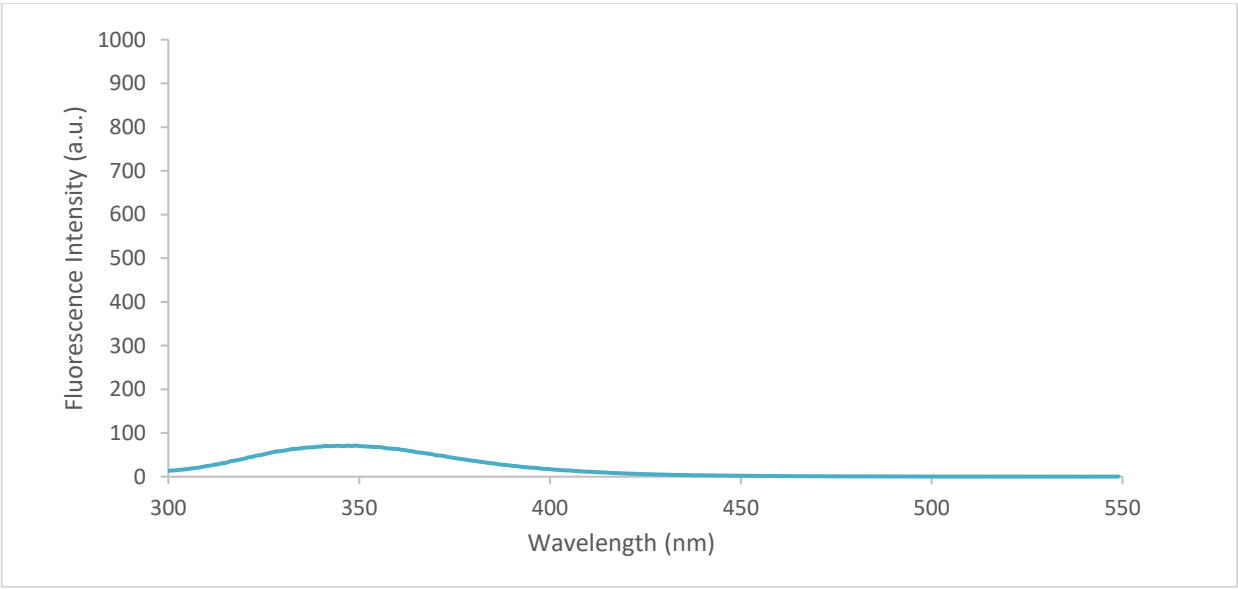

Figure 37. Fluorescence spectrum of 1µM BSA at λ_{exc}=280 nm, width of excitation and emission slits=5 nm, PMT=600 V. Agilent.

6.2.2.4. Irinotecan: absorption spectra

UV-Vis absorption spectra of irinotecan exhibit a shape and a band around 370 nm as reported in the literature [259], [260] (Figure 38). The intensity of the absorbance was proportional to the irinotecan concentrations used.

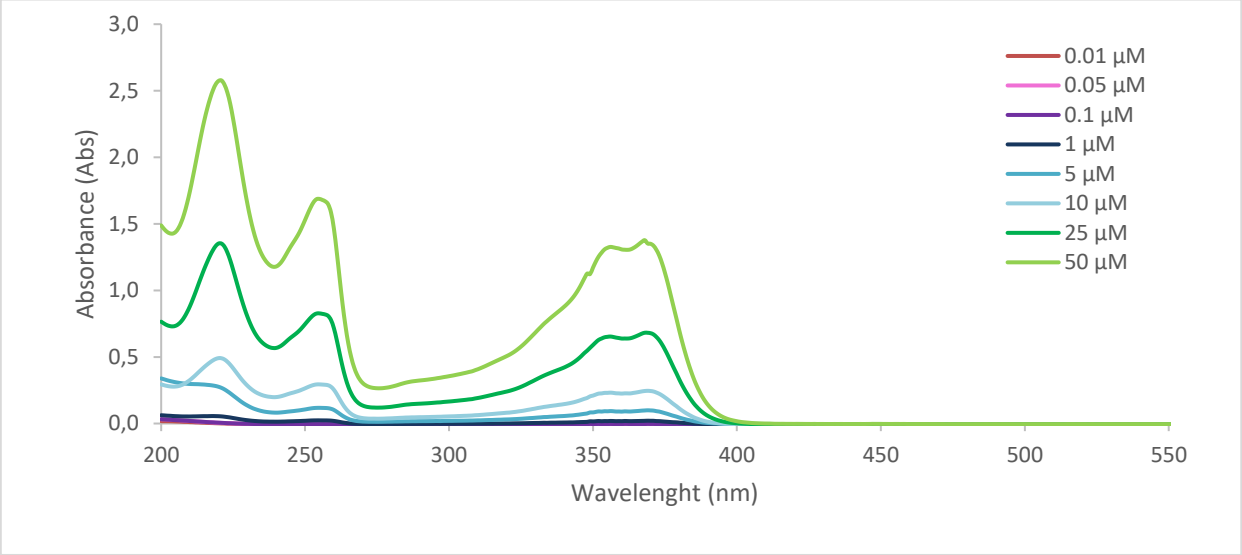

Figure 38. Absorbance spectra of CPT-11 in milliQ water, subtraction of blank. Agilent.

6.2.2.5. Irinotecan: fluorescence spectra

The recorded fluorescence emission spectra of irinotecan showed a peak at around 430 nm when excited at 280 (Figure 39) and 360 nm (Figure 40), as reported in the literature [259], [261]. The intensity of the fluorescence varies with the drug concentration.

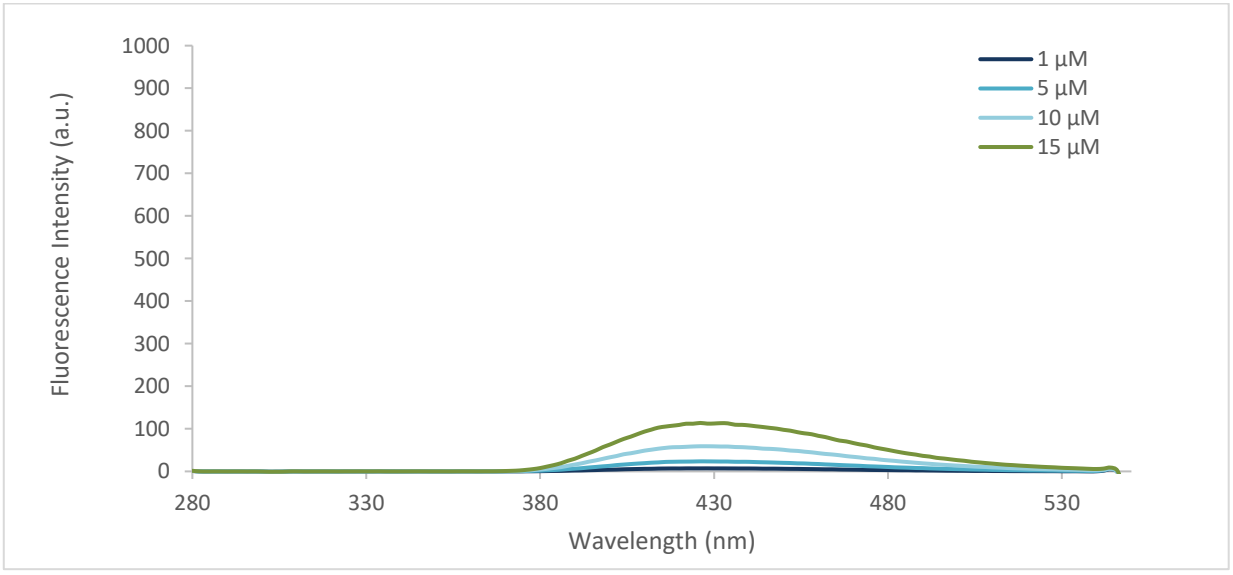

Figure 39. Emission spectra of CPT-11 at λ_{exc}=280 nm, width of excitation and emission slits=5 nm, PMT=500 V, subtraction of blank. Agilent.

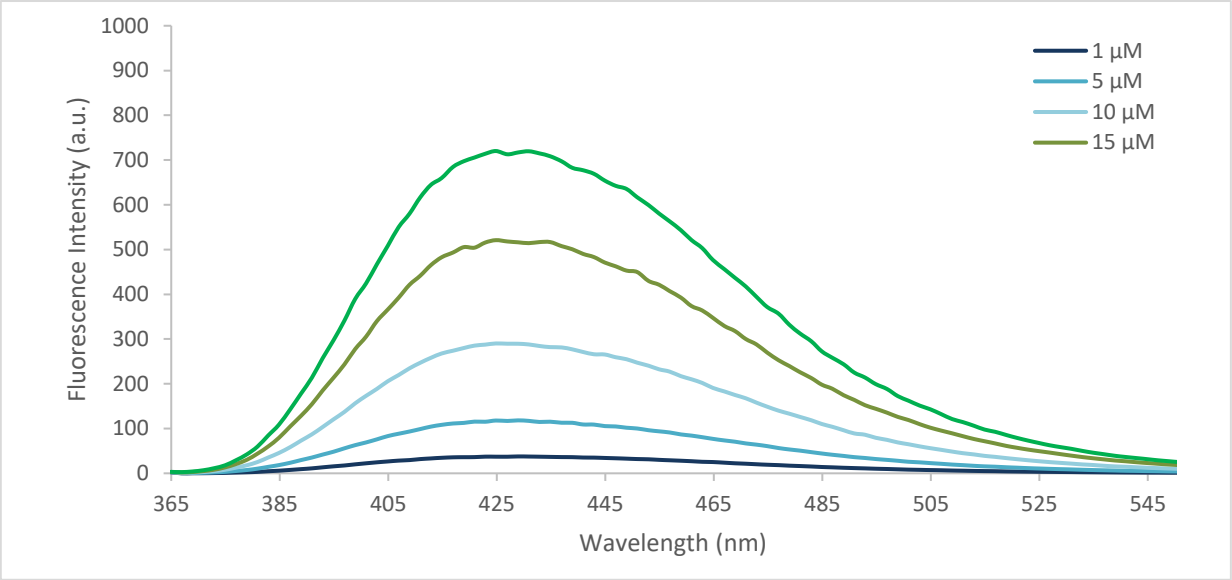

Figure 40. Emission spectra of CPT-11 λ_{exc}=360 nm, width of excitation and emission slits=5 nm, PMT=500 V. Each spectrum was the mean of 4 scans. Agilent.

6.2.2.6. Absorbance spectrum of CPT-11 and emission spectrum of BSA

The emission spectrum of BSA and the absorbance spectrum of irinotecan overlap, suggesting the possibility for a FRET phenomenum occurence (Figure 41).

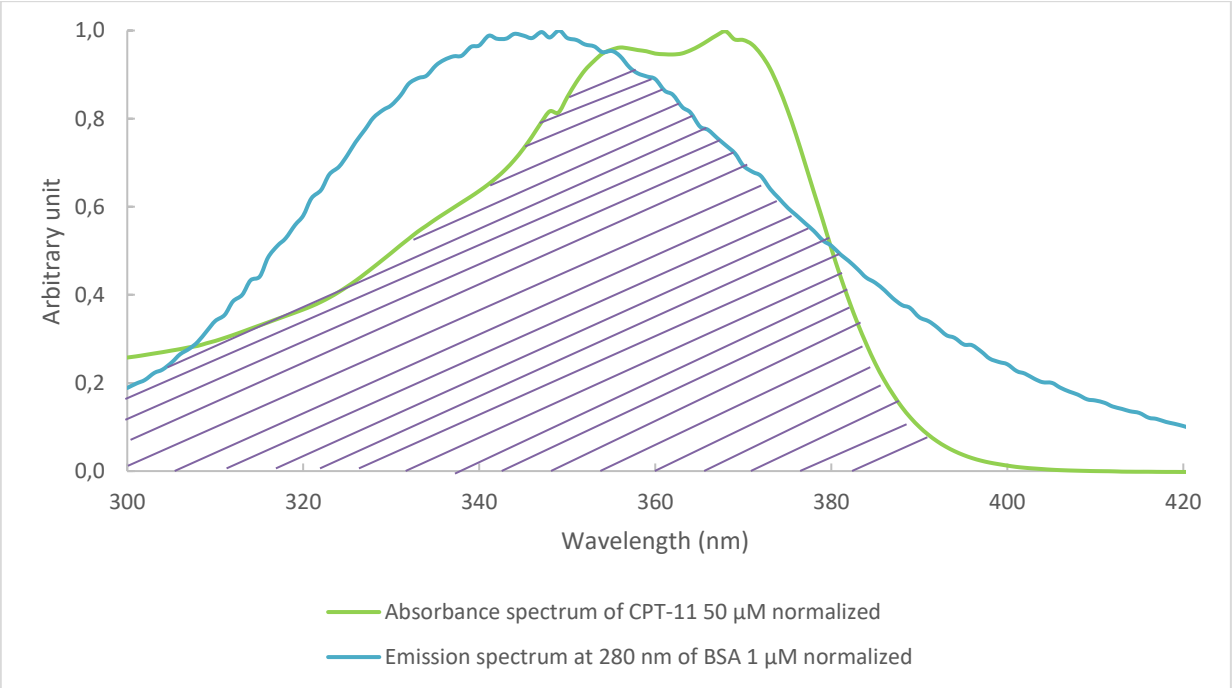

Figure 41. Absorbance spectrum of CPT-11 50 µM plus emission spectrum of BSA 1 µM λ_{exc}=280 nm, width of excitation and emission slits=5 nm, PMT=600 V.

6.2.2.7. BSA and irinotecan: titration assay

Fluorescence spectroscopy technique allows measuring the binding mechanisms and constants between BSA and irinotecan. The fluorescence emission spectra of BSA 1 µM and 10 µM with various concentrations of irinotecan were shown in Figure 42 and Figure 44, respectively. The fluorescence intensity of BSA decreased with the gradually increasing concentrations of irinotecan, indicating that the binding of irinotecan to BSA quenched the intrinsic fluorescence of the protein. Moreover, none visible red shift at the maximum λ_{em} was observed in the binding process despite of the 7 nm red shift observed for the binding of astilbin, a flavonoid molecule, and BSA [257]. The measurement of the excitation spectrum of the BSA-irinotecan complex at the end of the titration should guarantee the control of the status of the active lactone of the irinotecan to avoid the hydrolyzation which determined the opening of the ring and the inactivation to carboxylate (Figure 43).

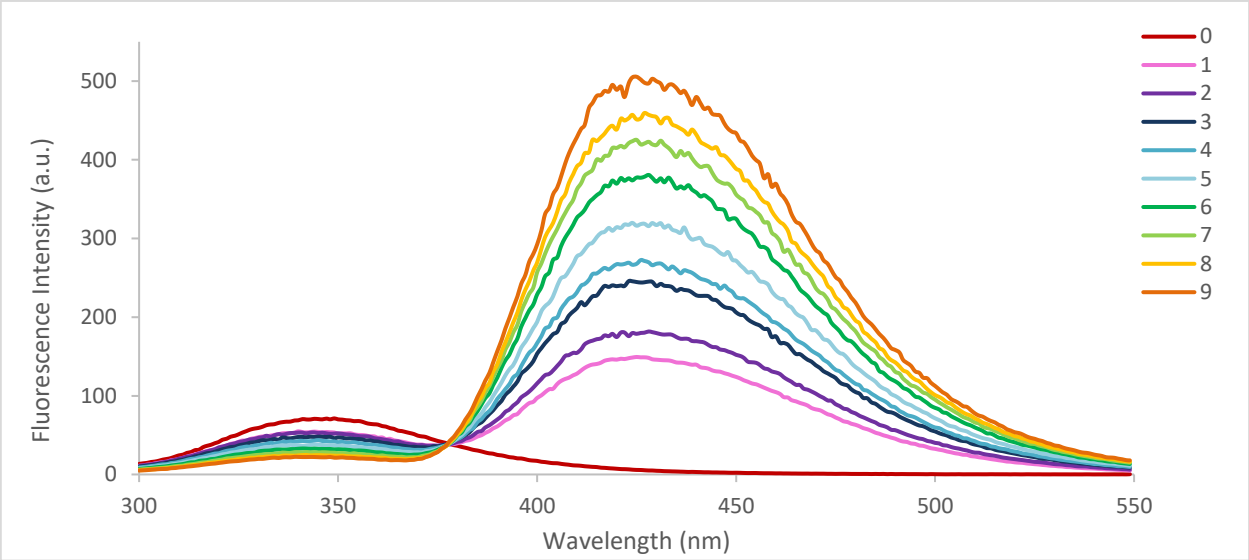

Figure 42. Titration curve BSA 1µM (0) with addition of CPT-11 at different concentrations at λ_{exc}=280 nm, width of excitation and emission slits=5 nm, PMT=600 V. Agilent.

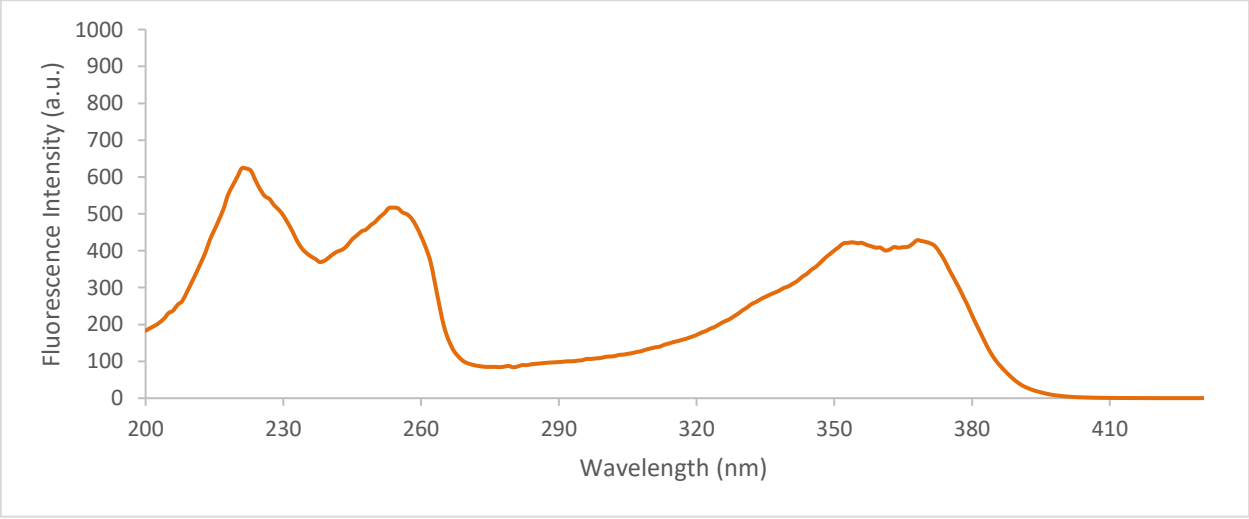

Figure 43. Excitation spectrum of the solution with BSA+CPT-11 post titration assay. λ_{em}=440 nm, width of excitation and emission slits=5 nm, PMT=500 V. The spectrum was the mean of 4 scans. Agilent.

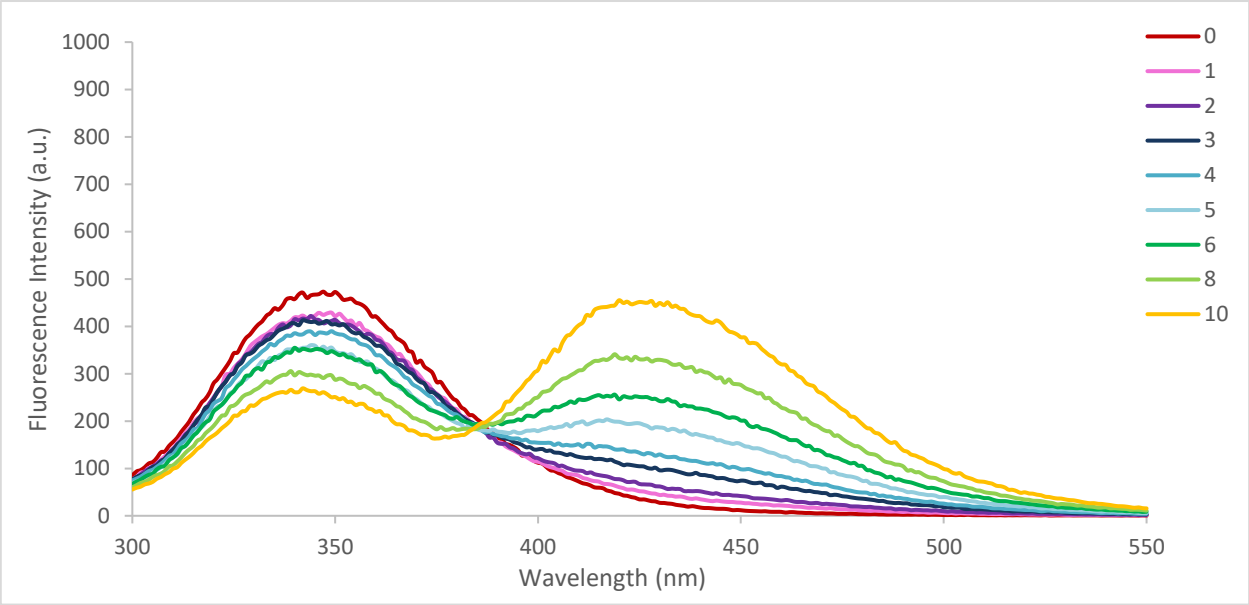

Figure 44. Titration curve BSA 10µM (0) with addition of CPT-11 at different concentrations at λ_{exc}=280 nm, width of excitation and emission slits=2.5 nm, PMT=800 V, subtraction of blank. Agilent.

6.2.2.8. Fitting of BSA and irinotecan: titration assay data

The addition of irinotecan to a solution of BSA induces quenching of the protein fluorescence. No significant changes in spectral shape were observed (Figure 39 and Figure 41). This behavior indicates that the chemical microenvironment of tryptophan and tyrosine in the BSA protein does not change appreciably. The quenching ratios of the BSA fluorescence, Fo/F (where Fo and F are the emission intensities in the absence and in the presence of irinotecan, respectively), were plotted against the irinotecan concentration (Stern−Volmer plot, Figure 45).

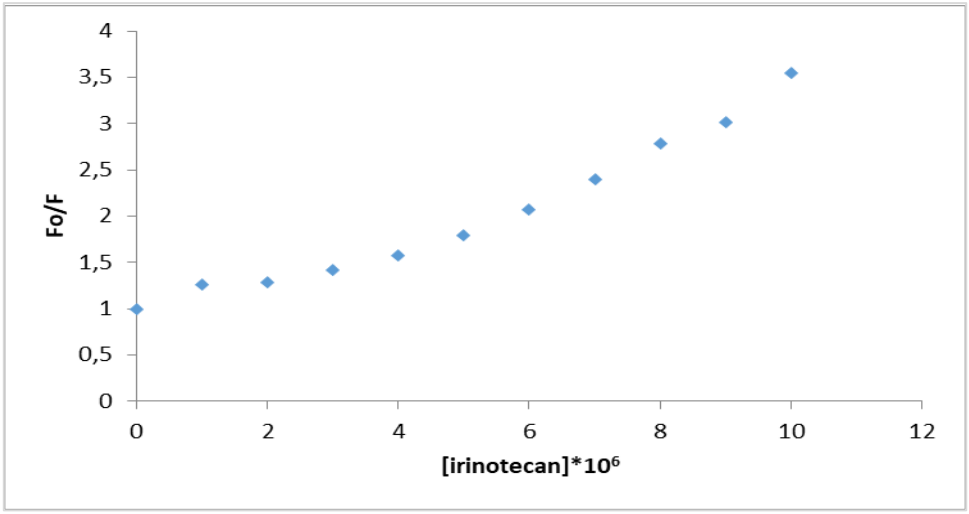

Figure 45. Stern−Volmer plot for the quenching of BSA (1×10^{-6} M) fluorescence by irinotecan (λ_{exc}=280 nm and λ_{obs}=345 nm).

The plot shows an upward curvature suggesting the coexistence of both static and dynamic quenching. Considering the behavior already observed for BSA fluorescence in the presence of Flindersine (a natural

occurring photochromic compound) [262], the upward curvature is determined by a significant contribution of dynamic quenching at higher ligand concentration (more than 4 x 10^{-6} M).

Static quenching is described by the general chemical equation (1). In this equation, n represents the number of irinotecan molecules associated with each protein.

$$(1) \qquad BSA + n\ irinotecan \leftrightarrows BSA(irinotecan)_n$$

With $K_a=[BSA(irinotecan)_n]/([BSA][irinotecan]^n)$.

From the quenching of the fluorescence of BSA by irinotecan, the association constant (K_a) and stoichiometry (n) of the binding process can be determined, according to the Scatchard equation (2):

$$(2)\ \log(Fo/F-1) = \log K_a + n\log[irinotecan]$$

The graphical representation of this treatment is reported in Figure 46.

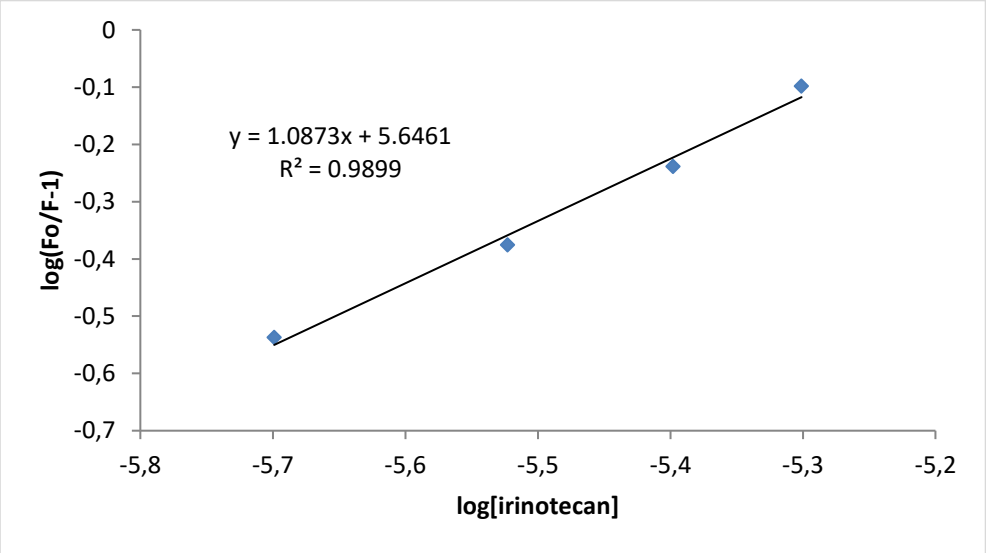

Figure 46. Application of equation $\log(Fo/F-1) = \log K_a + n\log[irinotecan]$ to the fluorescence quenching data of BSA by irinotecan.

From the linear regression of data reported in Figure 46, n indicates that just one molecules of irinotecan binds BSA protein. Because the quenching behavior of irinotecan is analogue to that observed by Flindersine [262], it can be envisaged that for analogy it involves the hydrophobic pocket of domain II where Trp^{212} is located. The value of $\log K_a$ of 5.6 (K_a=4.4 x 10^5 M^{-1}) obtained in the linear regression fitting is in agreement with the K_a value of 4.73 × 10^5 M^{-1} reported in literature for BSA-irinotecan complex [263].

The binding of irinotecan was investigated also through the enhancement of fluorescence of irinotecan through fluorescence resonance energy transfer (FRET) from BSA tryptophan.

Considering the formation of BSA-irinotecan complex (3):

$$(3)\ BSA + irinotecan \leftrightarrows BSA\text{-}irinotecan$$

The fluorescence emission at 430 nm was considered as sum of the fluorescence of bound and free irinotecan (4).

(4) F= F(irinotecan) + F(BSA-irinotecan)

The fluorescence of each species can be safely considered linearly dependent on concentration at high dilution of the reagents. Then F(irinotecan)= F^{o}_{f}[irinotecan] and F(BSA-irinotecan)=F^{o}_{b}[BSA-irinotecan]. Where F^{o}_{f} is the fluorescence of free Irinotecan at 1 x 10^{-6} M (17 a.u.) measured experimentally and F^{o}_{b} is the fluorescence of complex BSA-irinotecan (bound irinotecan) which is a fitting parameter along with the Ka.

The results of the non-linear fitting, carried out by EXCELL solver and solverAID macro, are reported in Figure 47.

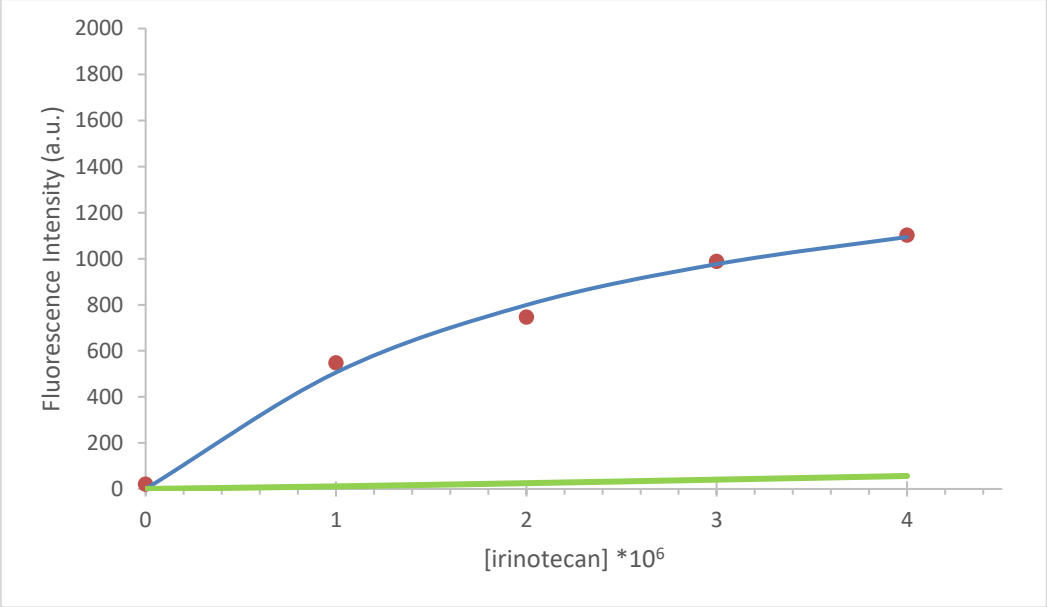

Figure 47. Fit (blue curve) to the experimental fluorescence data (red points) at 430 nm of BSA (1x10^{-6} M) irinotecan titration. Green curve represents the contribution to the fluorescence of free irinotecan.

The Figure 47 shows the good fitting of fluorescence data. The green curve evidences that the contribution of the free irinotecan plays a minor role. The fitting afforded a K_a value of 8±3x10^{5} M^{-1} and a F^{o}_{b} of 1430±166. The K_a obtained by the non-linear fit is in agreement with the value obtained by the linear regression analysis reported in Figure 46.

Analogue non-linear fit was performed also for the quenching signal collected at 345 nm.

In this case the fluorescence signal was calculated as contribution of F (BSA) + F (BSA-irinotecan).

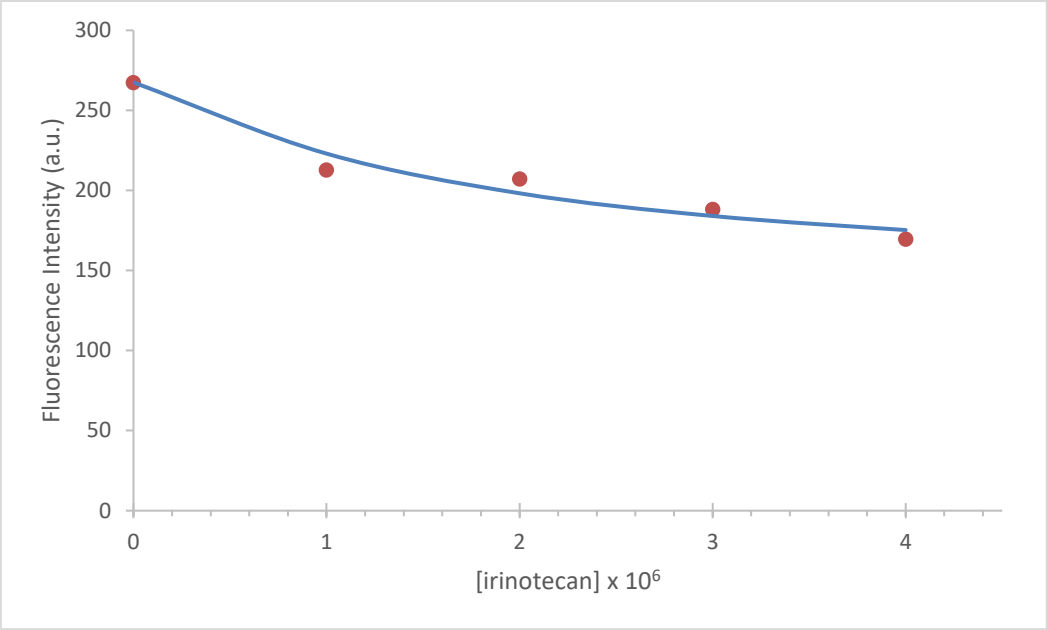

Figure 48. Fit (blue curve) to the experimental fluorescence data (red points) at 345 nm of BSA ($1x10^{-6}$ M) irinotecan titration.

The Figure 48 shows the good fitting of fluorescence quenching data. The fitting afforded a K_a value of $8\pm7x10^5$ M^{-1} and a F^o_b of 142±34. The K_a obtained following the quenching of fluorescence signal of protein by irinotecan is in perfect agreement with the K_a obtained following the fluorescence signal of irinotecan.

6.2.2.9. HLA-G and irinotecan: fluorescence spectra (First batch of protein)

To confirm *in silico* results of a possible strong interaction between HLA-G and irinotecan, we investigated by spectrophotometric analysis the presence, as in the case of BSA-irinotecan, of fluorescence resonance energy transfer (FRET) signal between HLA-G and CPT-11. The first experiment was performed mixing HLA-G protein with irinotecan drug in equimolar ratio. In Figure 49 was reported the emission spectra recorded from 300 to 550 nm (λ_{em}) range with an excitation wavelength (λ_{exc}) of 280 nm, with an excitation and emission slit widths of 5 nm, and a PMT of 700 V, of: 1) irinotecan in milliQ water and 2) in PBS at a concentration of 1µM 3) HLA-G at a concentration of 0.1 µM in milliQ water and 4) the spectra of co-presence of HLA-G 1 µM and irinotecan 1 µM (ratio 1:1) in milliQ water. PBS solution was chosen because more similar to real conditions, however, the ionic strength not determined any appreciable variation in the shape and intensity of the irinotecan emission spectra. We observed an important variation of the fluorescence intensity when the HLA-G and irinotecan where in the same solution in a 1:1 ratio and they could interact, compared to the signals produced by the HLA-G and irinotecan alone.

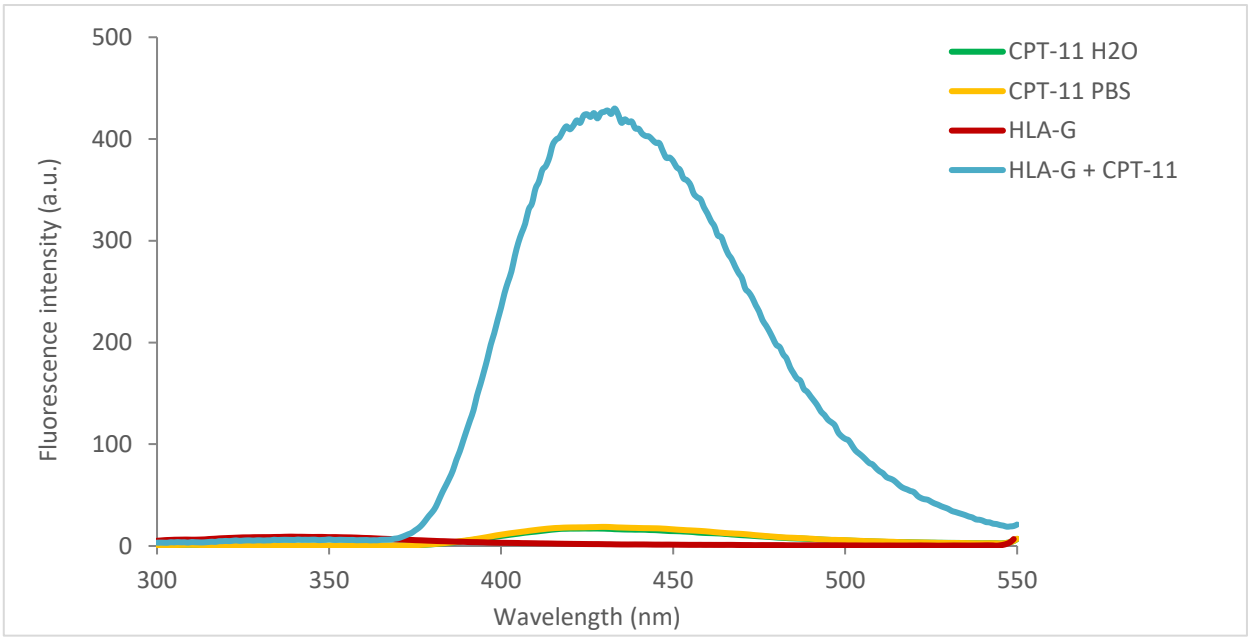

Figure 49. Emission spectra of irinotecan and HLA-G. Agilent.

The absorbance spectrum of CPT-11 was represented with the emission spectrum of HLA-G, to show their superimposition and then the possibility of FRET phenomenon occurrence (Figure 50).

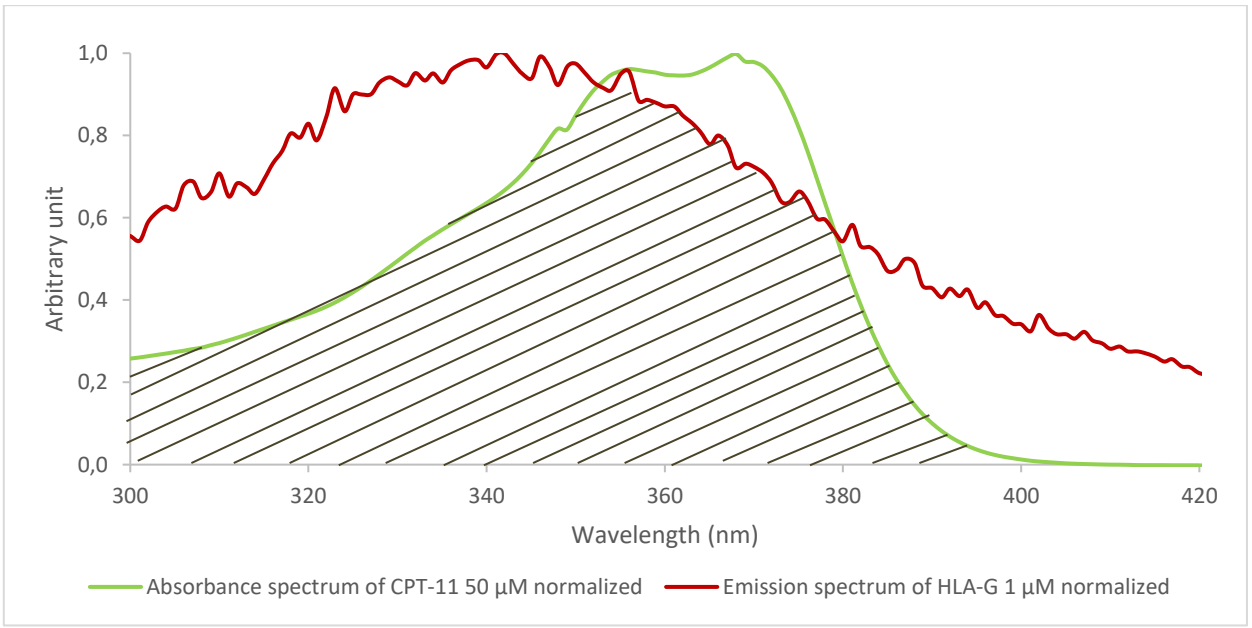

Figure 50. Absorbance spectrum of CPT-11 50 µM plus emission spectrum of HLA-G 1 µM λ_{exc}=280 nm, widths of excitation and emission slits= 5nm, PMT=700 V.

6.2.2.10. HLA-G and irinotecan: titration assay (First batch of protein)

We evaluated the effect of interaction between the HLA-G and irinotecan with a titration assay. The emission spectra were recorded from 300 to 550 nm (λ_{em}) range with an excitation wavelength (λ_{exc}) of 280 nm, with an excitation and emission slit widths of 5 nm, and a PMT of 700 V. The fluorescence emission spectra of 1 µM HLA-G with various concentrations of irinotecan progressively were added to a final concentration of 0.1, 0.2, 0.4, 0.8, 1, 2, 4, 8 µM. The fluorescence intensity of the HLA-G gradually decreased with the increasing concentrations of irinotecan, indicating that the binding of irinotecan to HLA-G quenched the intrinsic fluorescence of HLA-G (Figure 51).

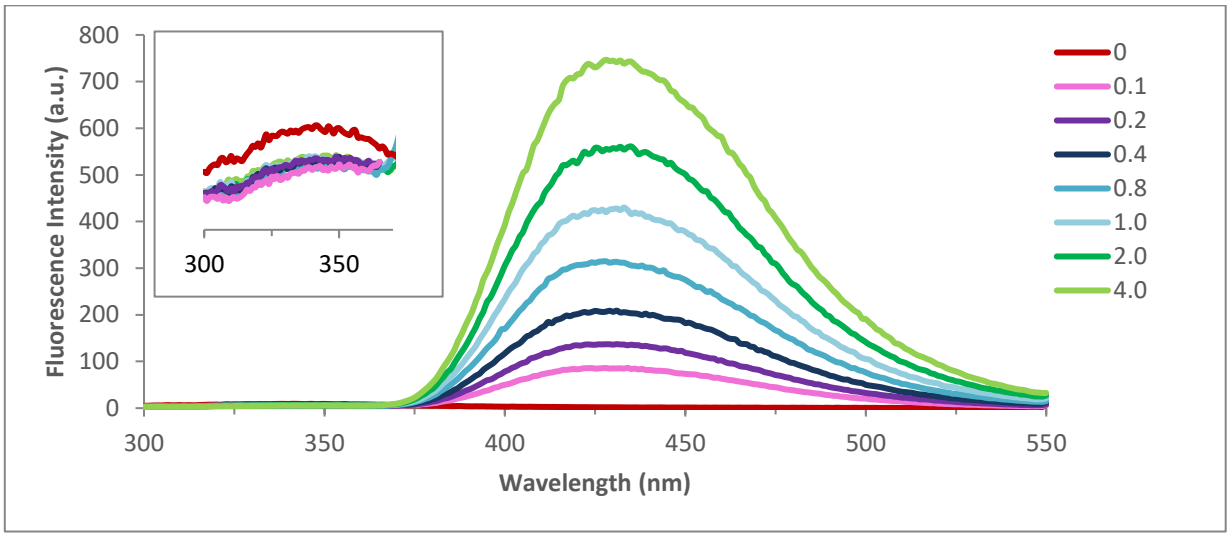

Figure 51. Titration curve HLA-G 1 µM adding different concentration of CPT-11 at λ_{exc}=280 nm, width of excitation and emission slits=5 nm, PMT=700 V. Each spectrum was the mean of 4 scans. Agilent.

6.2.2.11. HLA-G: absorbance spectrum (First batch of protein)

We checked the absorbance spectrum of the recombinant HLA-G to verify the stability of the batch after almost a month, and we not observed anymore the presence of the HLA-G peak of absorbance (Figure 52). Therefore, the purchased of a new batch was necessary to confirm the results found.

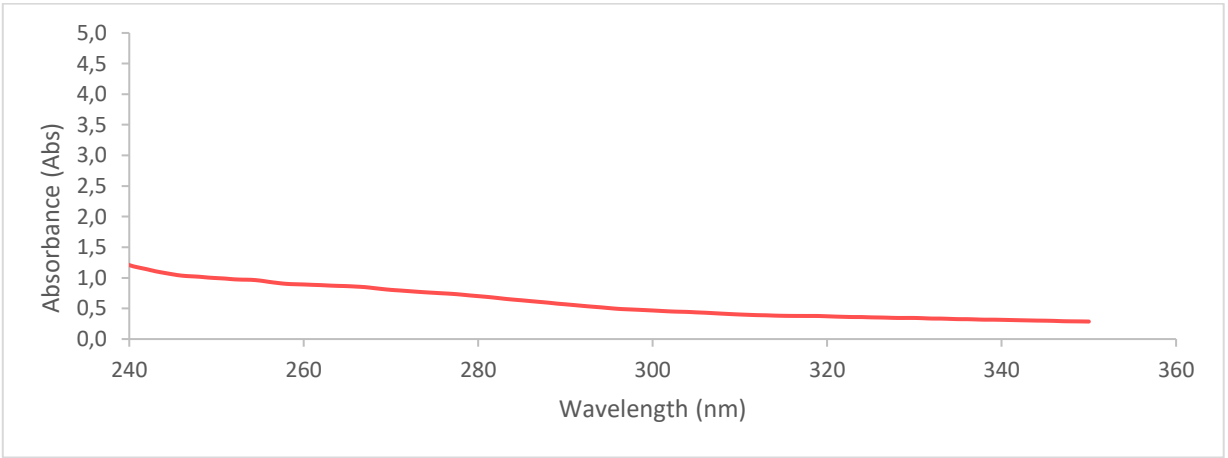

Figure 52. Absorbance spectrum of HLA-G in milliQ water at concentration of 25µM obtained from the mean of 2 scans. Nanodrop.

6.2.2.12. Fitting of HLA-G (first batch) and irinotecan: titration assay data

The addition of irinotecan to a solution of HLA-G (first batch) induces a quenching of the protein fluorescence and a significant increase of the fluorescence signal of irinotecan (Figure 49). The behavior is similar to the experiment with BSA. The main difference is related to the weak fluorescence signal of tryptophan (Figure 51). For that reason, we have used only the fluorescence emission at 430 nm of irinotecan for the fitting analysis. Using an analogue model and treatment of data (5):

(5) HLA-G + irinotecan $\rightleftarrows$ HLA-G-irinotecan

The fluorescence emission at 430 nm was considered as sum of the fluorescence of bound and free irinotecan (6).

(6) F= F(irinotecan) + F(HLA-G-irinotecan)

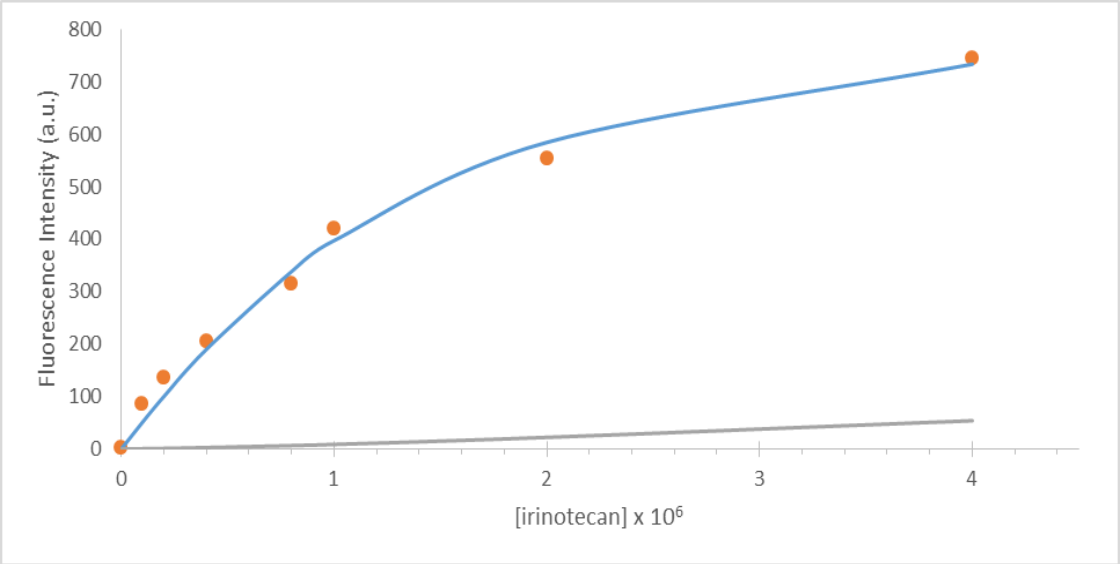

Figure 53. Fit (blue curve) to the experimental fluorescence data (red points) at 430 nm of HLA-G ($1x10^{-6}$ M) irinotecan titration. Grey curve represents the contribution to the fluorescence of free irinotecan.

The Figure 53 shows the good fitting of fluorescence data. The grey curve evidences that the contribution of the free irinotecan molecule plays a minor role. The fitting afforded a K_a value of $1.9\pm6x10^6$ M^{-1} and a F^o_b of 800±60.

The low fluorescence signal at 340 nm suggests that the concentration of the commercial protein should be significantly lower. We have than repeated the simulation using a nominal protein concentration of $0.2x10^{-6}$ M.

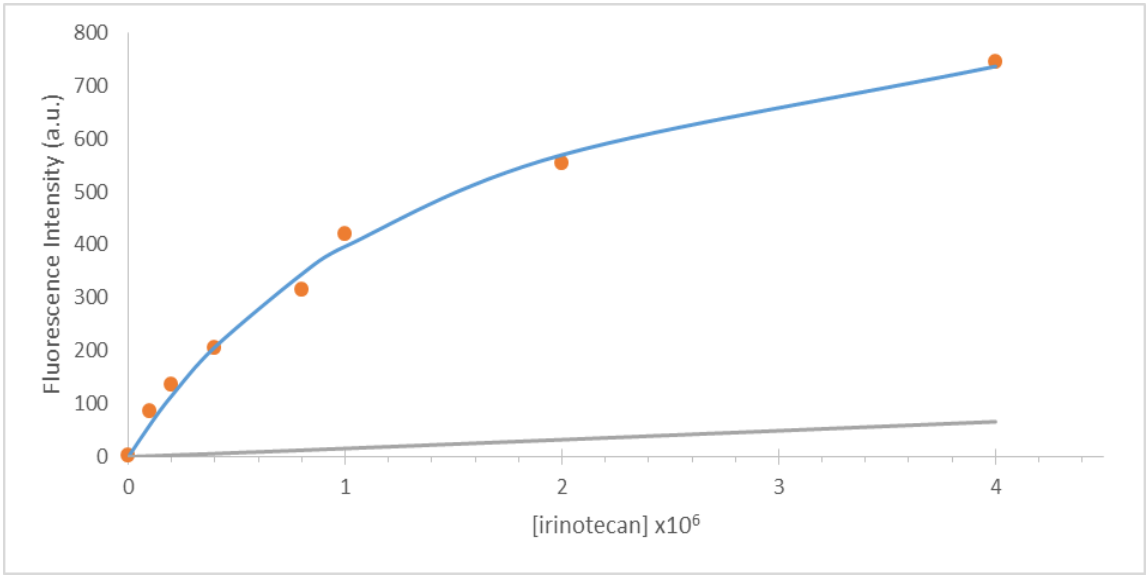

Figure 54. Fit (blue curve) to the experimental fluorescence data (red points) at 430 nm of HLA-G ($0.2x10^{-6}$ M) irinotecan titration. Grey curve represents the contribution to the fluorescence of free irinotecan molecule.

The Figure 54 confirms the good fitting of fluorescence data even at lower concentration of the protein. The fitting afforded a K_a value of $0.8±1x10^6$ M^{-1} and a F^o_b of 4400±200.

6.2.2.13. Recombinant HLA-G protein: Bradford assay (Second batch)

The Bradford assay was also realized to measure the HLA-G concentration using the Nanodrop instrument and the BSA calibration curve. We observed an absorbance peak at a higher intensity of that expected (Figure 55). After a month, we observed a reduction of the intensity of the absorbance peak maybe due to the precipitation or degradation of the protein in solution (Figure 56).

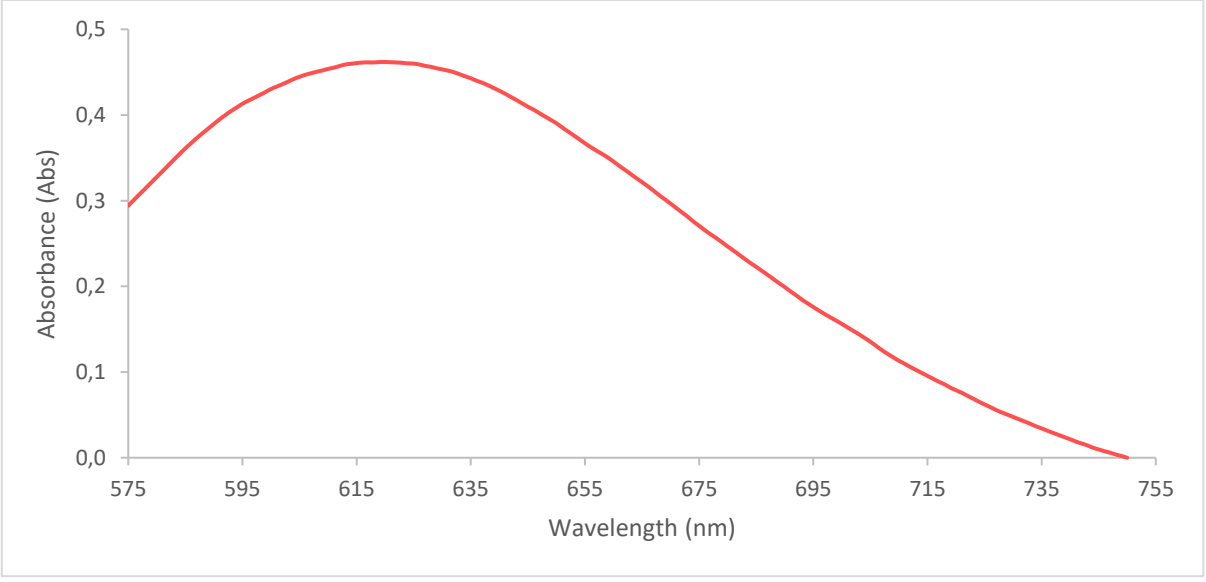

Figure 55.Absorbance spectrum of HLA-G in milliQ water at concentration of 25 µM obtained from the mean of 9 scans though the Bradford assay. Nanodrop.

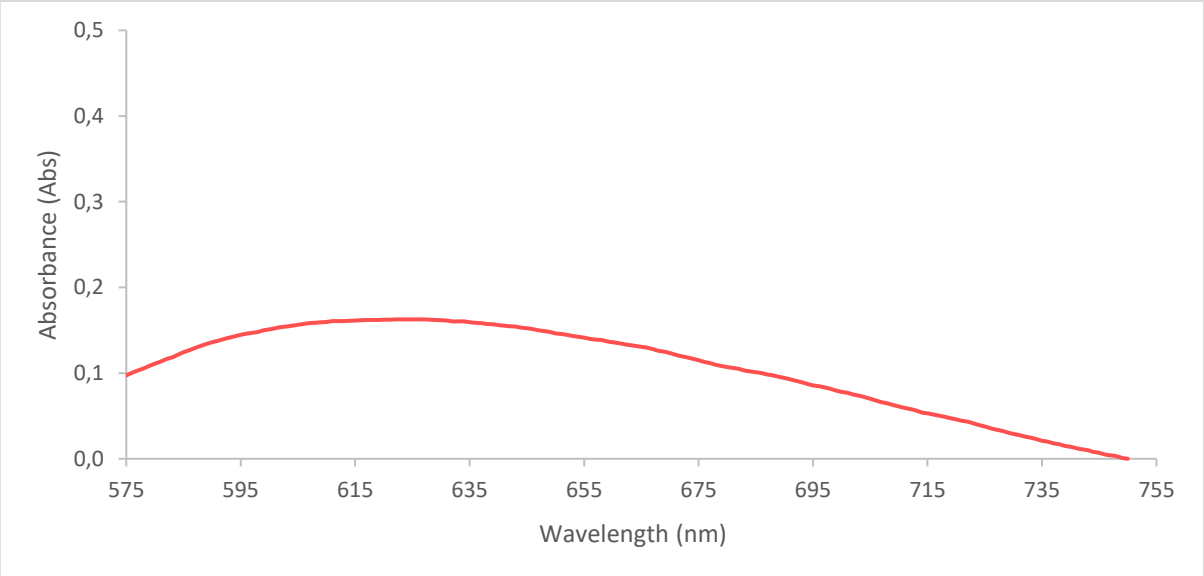

Figure 56. Absorbance spectrum of 2 µl HLA-G in milliQ water at concentration of 25 µM obtained from the mean of 3 scans though the Bradford assay. Nanodrop.

6.2.2.14. Recombinant HLA-G protein: absorption spectra (Second batch)

The extinction coefficient for the HLA-G ($\varepsilon_{HLA\text{-}G}$) estimated using the ExPASy tool is 76235 $M^{-1}cm^{-1}$ at 280 nm. Regarding the spectrophotometric determination of the concentration of a protein in solution, it was exploited the absorbing properties at 280 nm of the aromatic groups of the side chains of Trp and Tyr. The absorbance measurements at 595 nm are strictly linear with protein concentration according to the Beer-Lambert law [255]. We measured the HLA-G absorbance spectra both in milliQ water and PBS solution at different concentrations (Figure 57, Figure 58, Figure 59, and Figure 60). As for the BSA protein, we observed a weak band around 275 nm belongs to the π -> π* transition of the aromatic amino acids such as Trp, Tyr and Phe, while a strong band around 230 nm reflects the presence of the peptide bonds in the HLA-G backbone. We observed a blue shift of the major peak from 280 nm to 275 nm compared to the BSA spectrum.

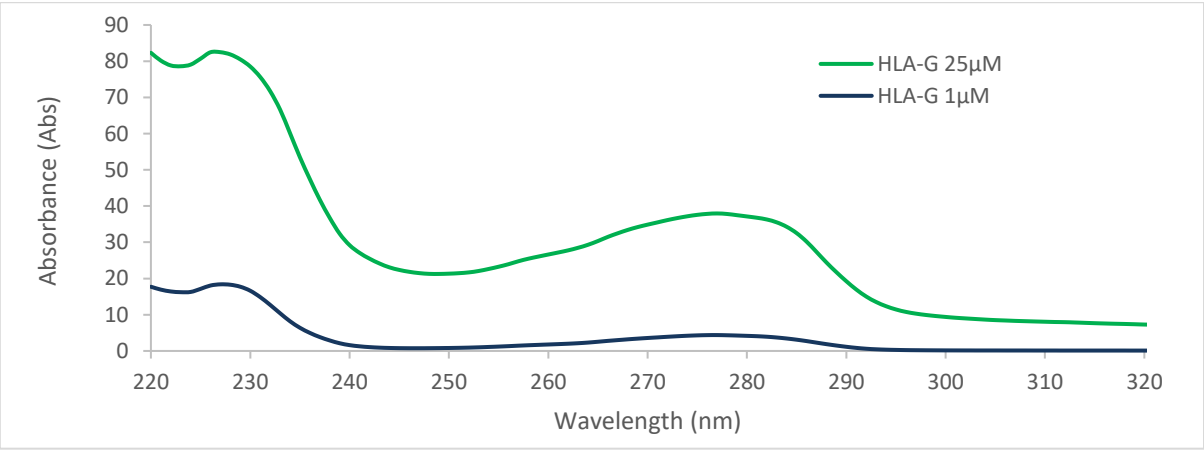

Figure 57. Absorbance spectrum of 1µl HLA-G in milliQ water at concentration of 25 µM and 1 µM obtained from the mean of 3 scans. Nanodrop.

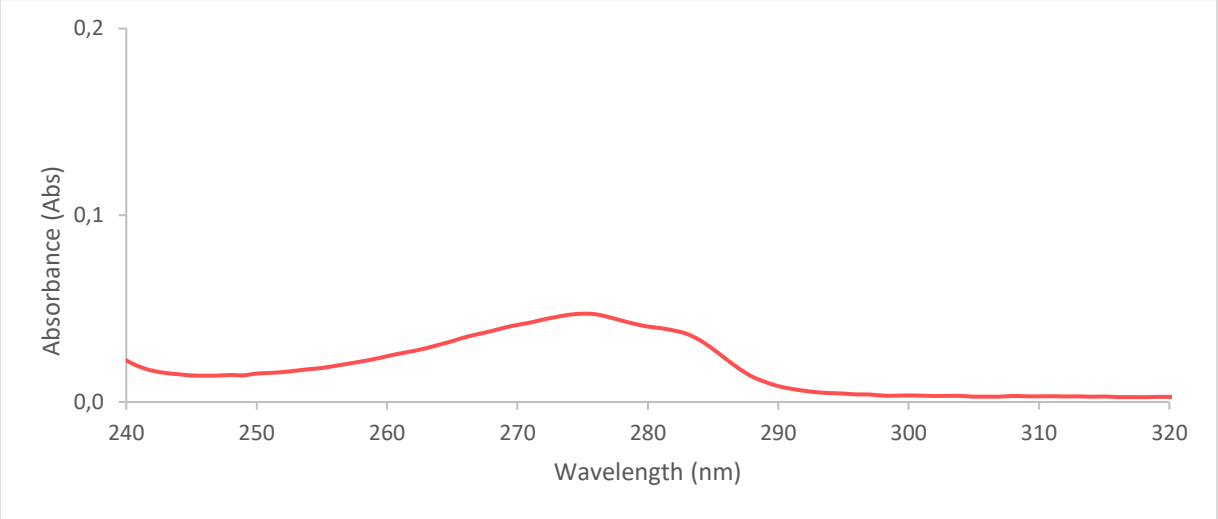

Figure 58. Absorbance spectrum of recombinant HLA-G in PBS at concentration of 0.01 µM obtained from the median of 4 scans. Agilent.

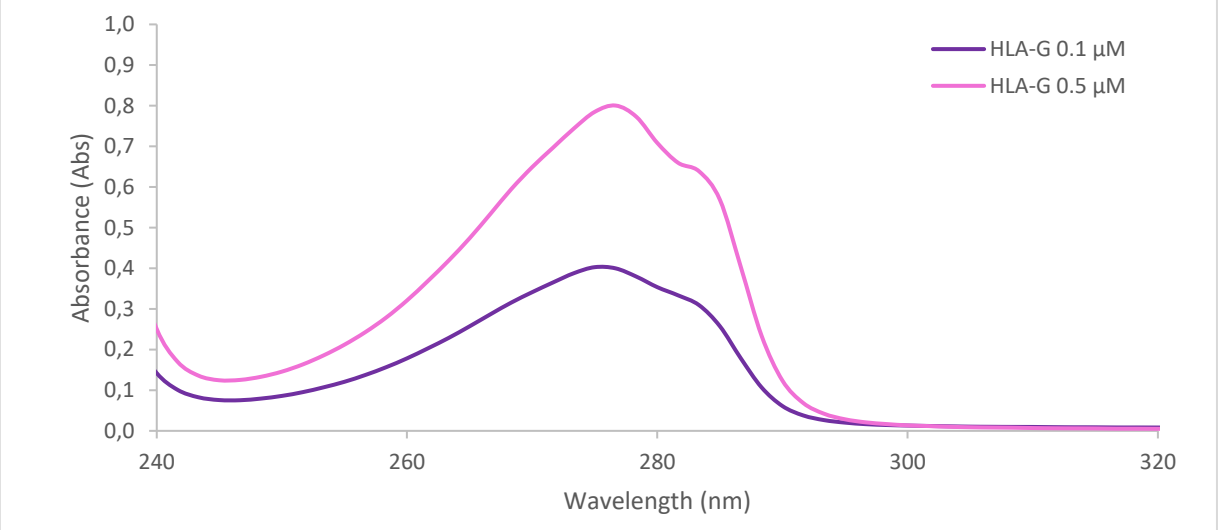

Figure 59. Absorbance spectrum of HLA-G in milliQ water obtained from the mean of 4 scans. Agilent.

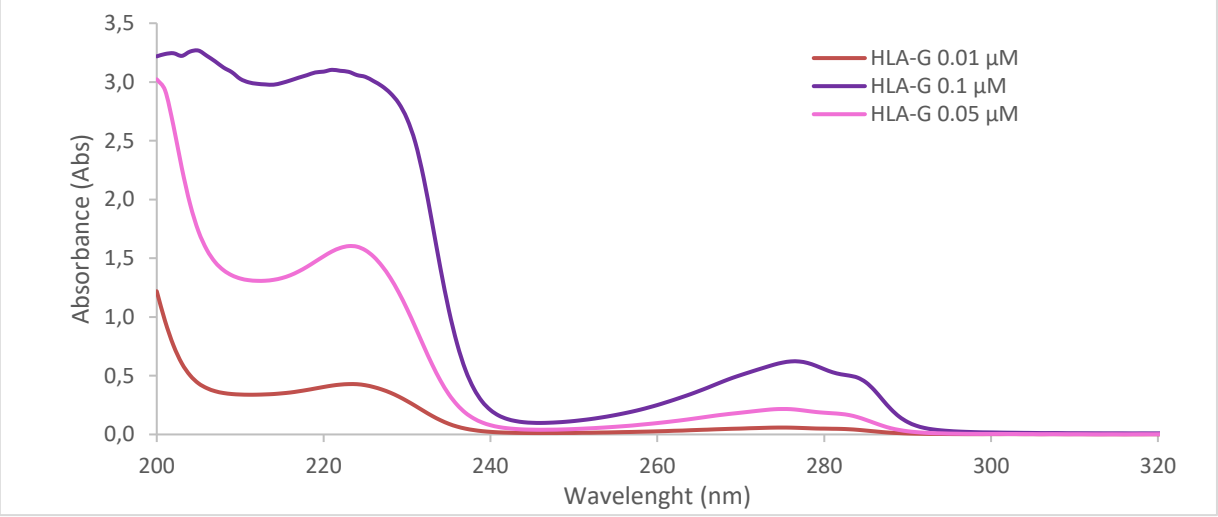

Figure 60. Absorbance spectra of HLA-G in milliQ water. Each spectrum was the mean of 4 scans. Agilent.

6.2.2.15. Recombinant HLA-G protein: fluorescence spectra (Second batch)

HLA-G contains tryptophan, tyrosine and phenylalanine residues, and has endogenous fluorescence like BSA protein. The main contributor to the intrinsic fluorescence of HLA-G was the tryptophan that is the most sensitive residue to the changes in the micro-environment. HLA-G showed a fluorescence emission peak around 300 nm when excited at 275 nm both in milliQ water and in PBS (Figure 61 and Figure 62). Also in this case, we observed a significant blue shift of the emission peak compared to those observed on the BSA spectrum, from 344 nm for BSA to 300 nm for HLA-G. We recorded also the excitation spectra of HLA-G in PBS and the shape obtained was the same observed for the absorbance spectra (Figure 63).

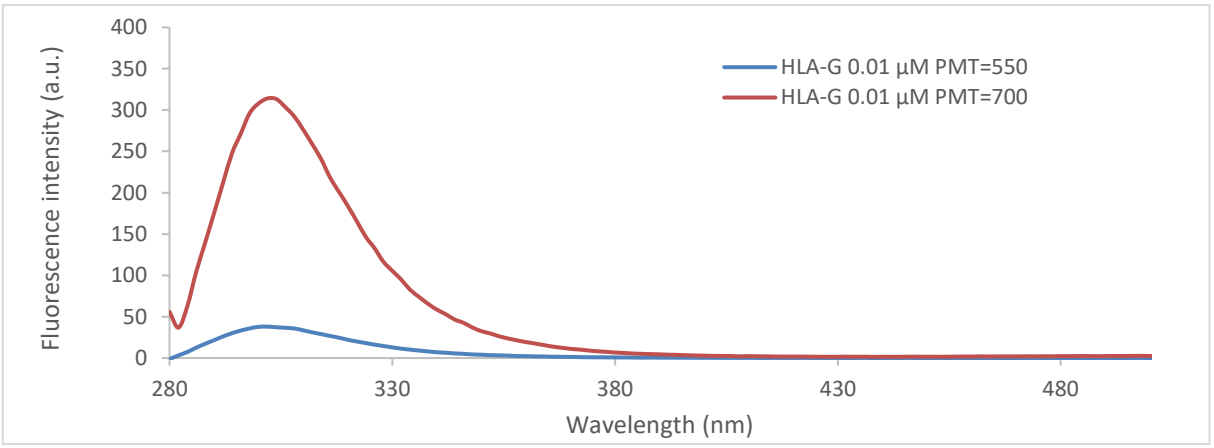

Figure 61. Emission spectra of HLA-G 0.01µM in PBS at λ_{exc}=275 nm, width of excitation and emission slits=5 nm, PMT as indicated in the legend, subctraction of blank.

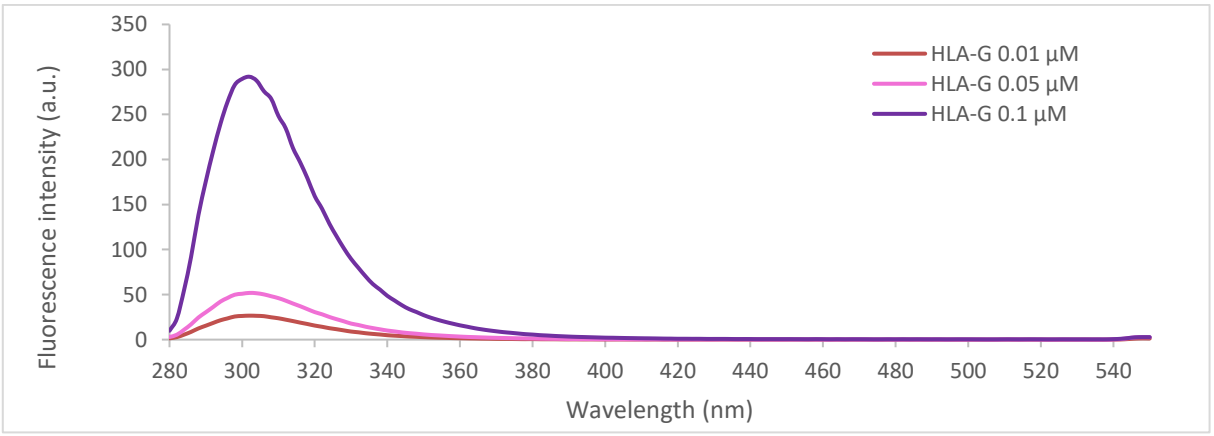

Figure 62. Emission spectra of HLA-G in milliQ water. λ_{exc}=275 nm, width of excitation and emission slits=5 nm, PMT=500 V.

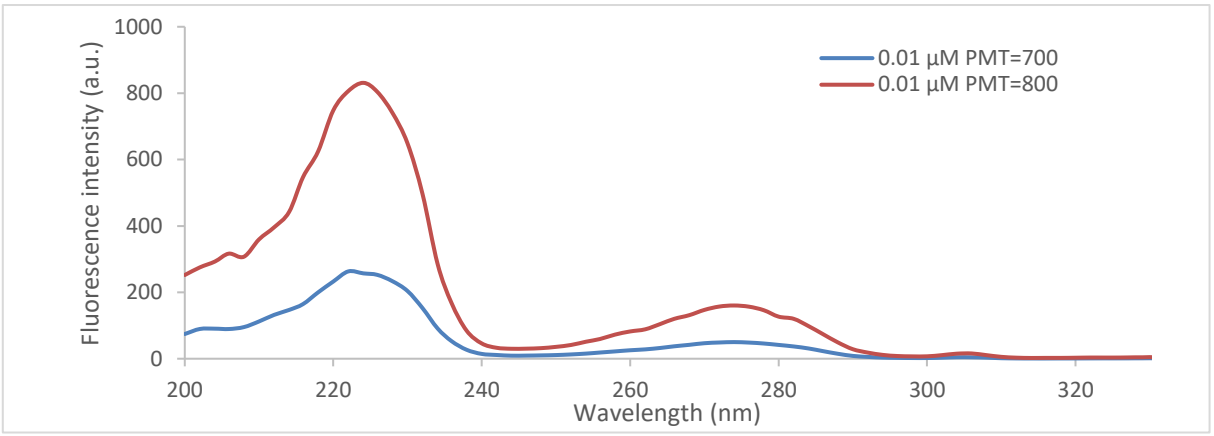

Figure 63. Excitation spectra of HLA-G 0.01 µM in PBS, λ_{em}=340 nm, width of excitation and emission slits=5 nm.

6.2.2.16. HLA-G and irinotecan: titration assay (Second batch of protein)

We also repeated the evaluated effect of interaction between the HLA-G and irinotecan with a titration assay using the second batch of HLA-G. The emission spectra were recorded from 280 to 550 nm (λ_{em}) range with an excitation wavelength (λ_{exc}) of 275 nm, with an excitation and emission slit widths of 5 nm, and a PMT of 500 V. The fluorescence emission spectra of 0.05 µM HLA-G (teorical measurement) with various concentrations of irinotecan progressively were added to a final concentration of 1, 5, 10, 15, 20 µM (Figure 64). The fluorescence intensity of the HLA-G gradually decreased with the increasing concentrations of irinotecan, indicating that the binding of irinotecan to HLA-G quenched the intrinsic fluorescence of HLA-G.

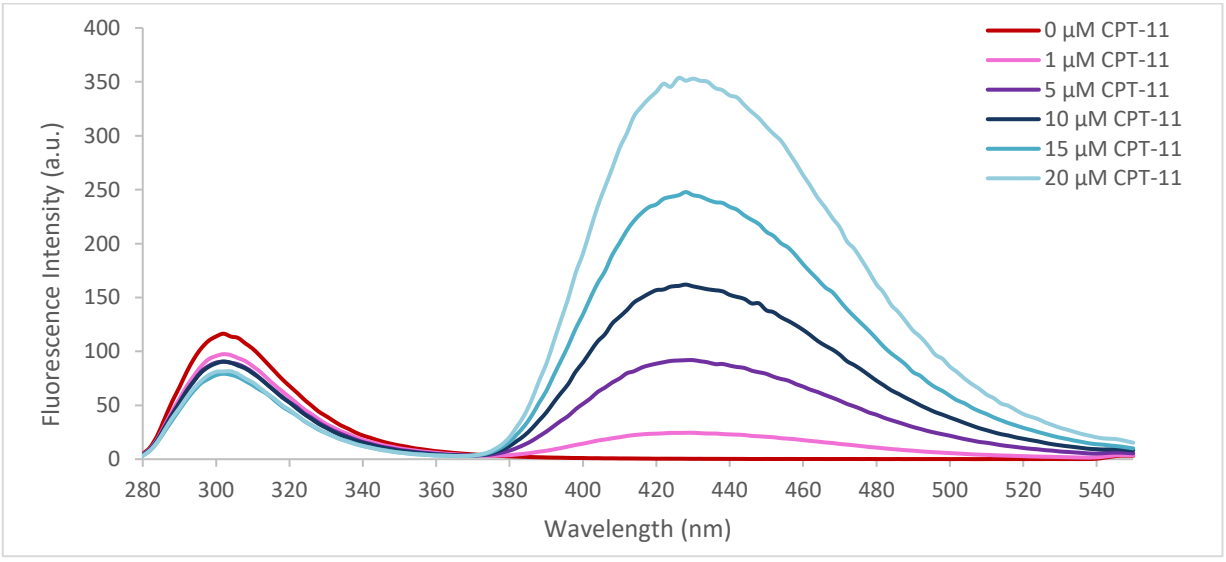

Figure 64. Titration curve of HLA-G 0.05 µM adding different concentration of CPT-11 at λ_{exc}=275 nm, width of excitation and emission slits=5 nm, PMT=500 V, subtraction of blank. Agilent.

The emission spectra of another titration assay, recorded from 280 to 550 nm (λ_{em}) range with an excitation wavelength (λexc_c) of 275 nm, with an excitation and emission slit widths of 5 nm, and a PMT of 550 V, were obtained for 0.5 µM HLA-G (teorical measurement) with various concentrations of irinotecan

progressively added (Figure 65). The fluorescence intensity of the HLA-G gradually decreased with the increasing concentrations of irinotecan, indicating that the binding of irinotecan to HLA-G quenched the intrinsic fluorescence of HLA-G. Moreover, the measurement of the excitation spectrum of the HLA-G-irinotecan complex at the end of the titration should guarantee the control of the status of the active lactone of the irinotecan to avoid thehydrolyzation which determined the opening of the ring and the inactivation to carboxylate (Figure 66).

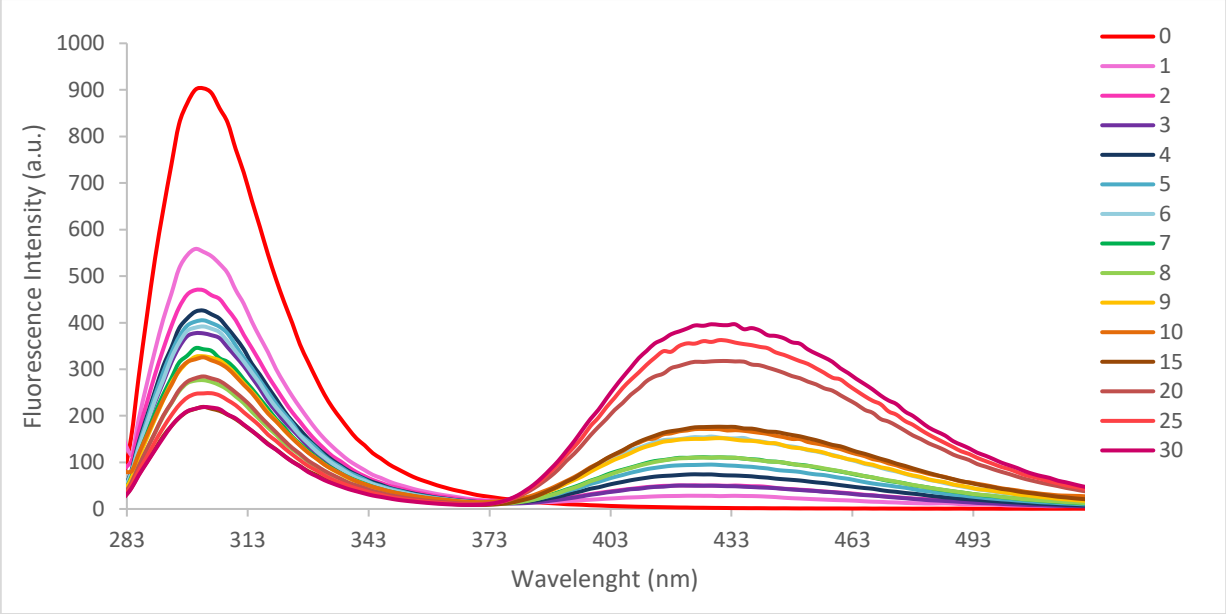

Figure 65. Titration curve of HLA-G 0.5 µM adding CPT-11 at different concentrations at λ_{exc}=275 nm, width of excitation and emission slits=5 nm, PMT=550 V, subtraction of blank. Agilent.

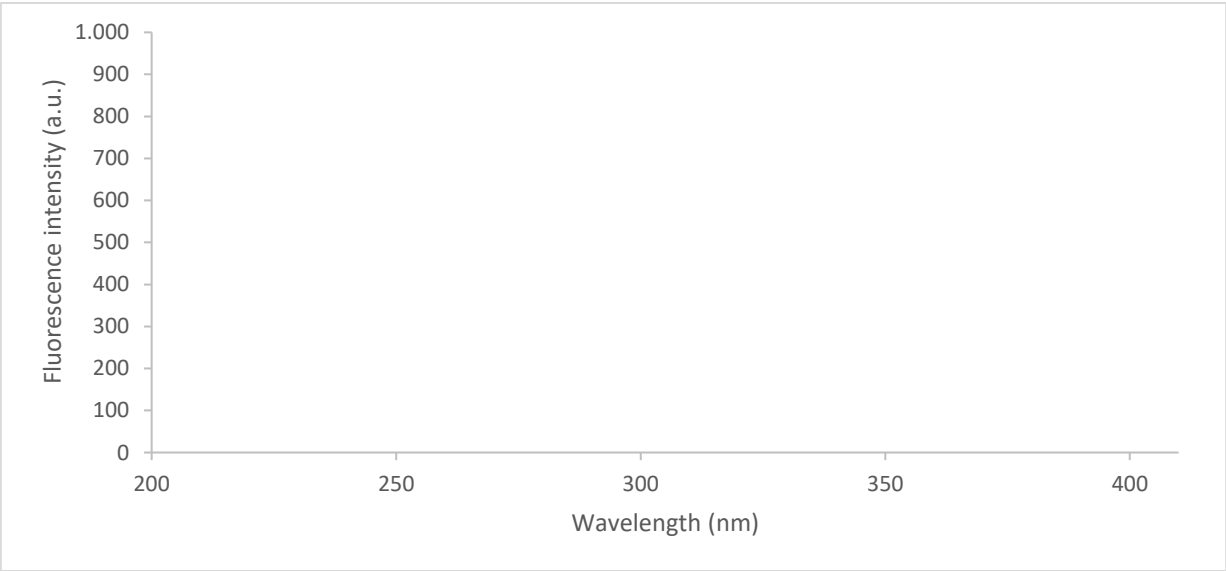

Figure 66. Emission spectrum of the solution with HLAG 0.5 µM and CPT-11 post titration assay of Figure 65 at λ_{em}=440 nm, width of excitation and emission slits=5 nm, PMT=550 V. Mean of 4 scans. Agilent.

We decided to modify the pH to the solution of the final titration solution adding NaOH to verify if the emission fluorescence spectra we had observed to shift towards the blue wavelengths, compared to those of BSA, is due to the presence of other species present in the solution or is a intrinsic characteristic of the HLA-G molecule. We observed a red shift only for the peak at 430 nm to 450 nm, characteristic of the irinotecan molecule, probably due to deprotonation of ammonium group (Figure 67). The peak at 300 nm, corresponding to the HLA-G emission, it does not undergo variations in the wavelength and shape.

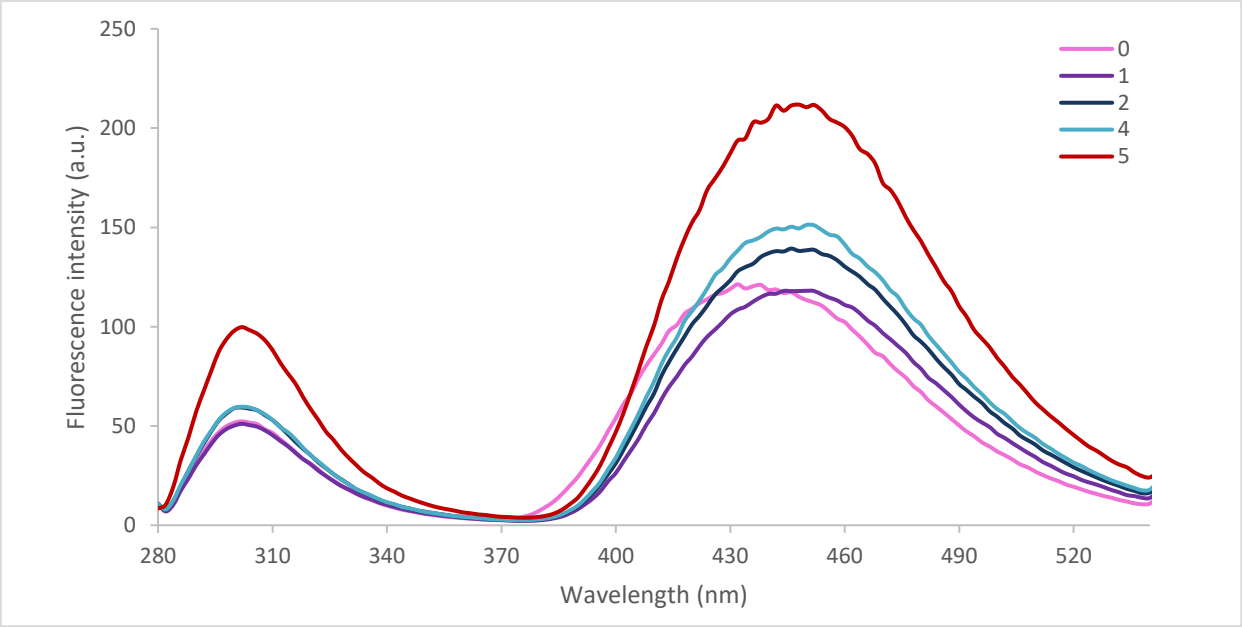

Figure 67. Emission spectra of final titration solution reported in Figure 65, adding NaOH 0.1 µM at λ_{exc}=275 nm, slits= 5 nm, PMT=550 V. Agilent.

We analyzed the solution of the second batch of HLA-G with liquid chromatography-mass spectrometry (HPLC-MS) to elucidate the composition of the solution and to identify unknown compounds within this second batch of HLA-G.

In order to pick some relevant information, solution was directly infused at 10 µL/min in an electrospray source of an API 4000 Hybrid Q-Trap (Sciex) instrument. MS spectra either in positive and negative mode revealed to be useless since the strong presence of contaminants. Therefore, an LC-MS experiment was set-up by using a C8 column (3 µm, 2x100 mm) kept at 50°C and flowed at 250 µL/min by an aqueous solution of trifluoroacetic acid (TFA, 0.05%) and acetonitrile (ACN) with an extended gradient from 20 to 98% ACN in 13 minutes.

Chromatographic process was able to provide a better separation of the components including a considerable presence of polyethylene glycol (PEG) molecules (see in Figure 68, the spots spaced by 44 Thomsons (Th) and with slightly decreasing lipofilicity) and other detergents. Unfortunately, the signal of the HLA-G expected to appear as multiply-charged ions looked to be still hidden by the surrounding matrix components (Figure 68). A viable solution should have been a pretreatment where the targeted protein, likely present at a significantly low concentration, have been trapped either on-line or off-line in order to reduce the strong impact of the surrounding components (detergents and other disturbing components). This pretreatment should be accompanied by a miniaturized LC (like nanoLC) for improving the instrumental sensitivity.

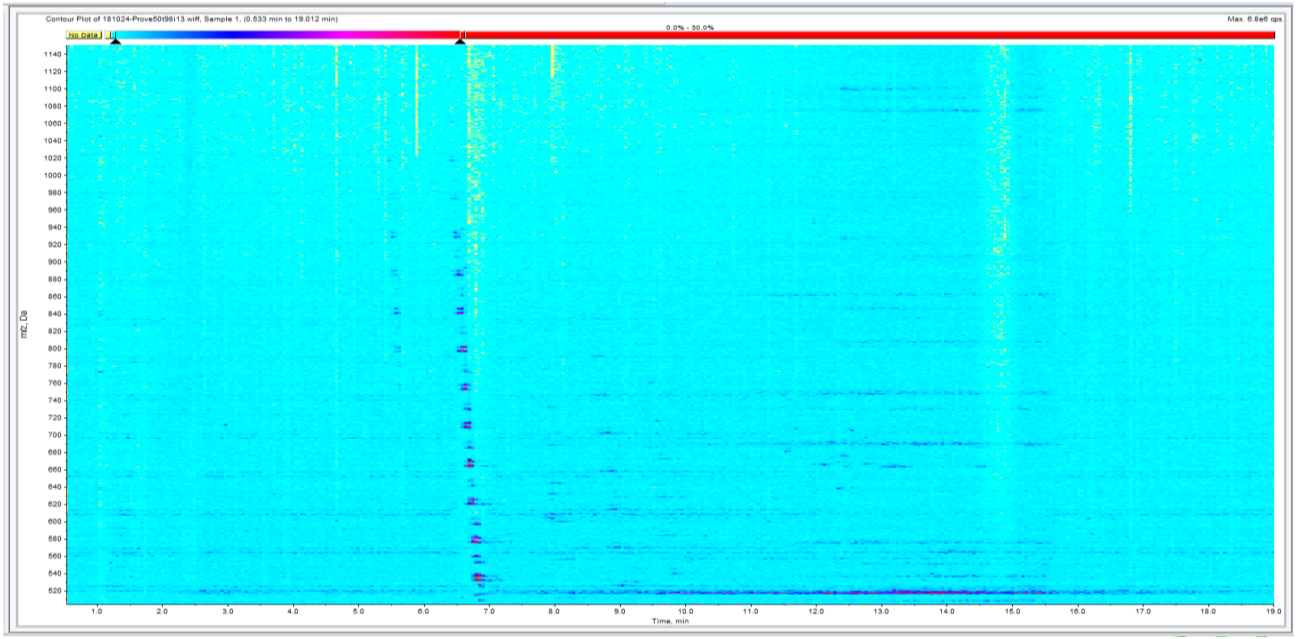

Figure 68. Contour Map of HLA-G second batch solution. HPLC-LC-MS.

In conclusion, the measurements of the HLA-G absorbance for this second batch were quite high compared to what was expected using the extinction coefficient for the HLA-G ($\varepsilon_{HLA\text{-}G}$=76235 $M^{-1}cm^{-1}$) estimated using the ExPASy tool at 280 nm in water. The absorbance spectra do not show any difference in the shape; however, the absorbance peak slightly shifts from 280 nm to 275 nm, which is mainly the absorbance wavelength of Tyr. Generally, proteins containing Trp residues show analogue absorbance spectra with a major peak at 280 nm. This could explain the differences between the spectra obtained for HLA-G and those of the BSA. Moreover, we observed a blue shift in fluorescence emission spectra. The major fluorescence emission peak observed when excited at 275 nm was at 300 nm. The BSA fluorescence emission peak exciting at 280 nm, was recorded at 344 nm [257]. For these reasons, we suspected some problem with this second batch. In fact, the shape of the first batch, despite of the low intensity of the signals, was similar to those of BSA. Since we suspected the presence of a large number of aromatic molecules containing phenolic moieties (such as tyrosine), we modified the pH of the solution adding NaOH, to induce the conversion of the contaminants from phenols to phenolate. We observed a red shift involving the fluorescence peak of irinotecan at 430 nm, but no significant change in the fluorescence peak at 300 nm. Finally, from the mass spectrometry analyses, we observed a considerable presence of polyethylene-glycol molecules and other detergents that hide the signal of the HLA-G.

6.3. DISCUSSION

Irinotecan is an antineoplastic drug present in several regimens used in combination with other chemotherapeutics to treat mCRC, and was also investigated for the clinical application in gastric, lung, pancreas, cervix, and ovarian cancer [264]–[269]. It acts as antineoplastic enzyme inhibitor of the DNA topoisomerase I, interfering with the replication of the DNA and inducing double-breaks in the strand, causing cell death [270], [271]. The therapeutic efficacy was partially reduced by the occurrence of severe toxicities, such as diarrhea and neutropenia, that limits the irinotecan clinical use [272]. Irinotecan has a relatively poor oral bioavailability, then it is mainly i.v. administered with a prevalent distribution in liver and bile to be metabolized. Research of potential genetic biomarkers to reduce or control irinotecan toxicity profile, was mainly focused in metabolic enzymes: in fact, to date, *UGT1A1*28* polymorphism is the only validated pharmacogenetic biomarker recommended by FDA guidelines for a reduction of the starting dose of irinotecan [273]. However, the novel findings of this project described in PART 1 and 2 may suggest that also an immunocheckpoint molecule like the HLA-G, could be an interesting candidate biomarker for irinotecan-containing regimens. In PART 3, for the first time, the physical interaction between the HLA-G molecule and the irinotecan has been deepened.

The visualized HLA-G crystal structure, shows the HLA-typical pocket for the molecular recognition of the antigen, that is constituted by several aromatic residues which could participate to the specific molecular recognition of aromatic molecules such as irinotecan. The analysis of HLA-G structure and the results obtained in PART 1 and 2, suggest a possible interaction between HLA-G and irinotecan. In fact, the irinotecan could be involved in specific hydrophobic interactions in the groove created by the α1-α2 domain and the β-sheet floor of HLA-G, as it happens with the peptide antigen presentation in the classical MHC complex. Moreover, we observed a central aromatic “hole” that could be a further possible site of interaction of an aromatic ligand. Through the docking analyses, it was highlighted that irinotecan is well inserted in the cleft of HLA-G. We calculated the electrostatic potential both for HLA-G and irinotecan, and confirmed a charge complementary of the complex. The docking models support this hypothesis showing a specific orientation of the irinotecan into the groove, with its positive terminal part containing the protonated N-amino group close to the negative part of the HLA-G cleft. Moreover, the molecular docking demostrated that the predominant binding force during the formation of the irinotecan/HLA-G complex was represented by the hydrophobic interactions. The negative Gibbs free energy (ΔG^0= -9.8 kcal/mol), calculated with molecular docking software, supports the spontaneous nature of the HLA-G/irinotecan binding process. It should be stressed that this is just a preliminary analysis of theoretical data. The hypothesis of strong interaction between HLA-G and irinotecan should be confirmed with experimental *in vitro* assay. However, this theoretical finding shows an interesting possibility that could have implications also from a therapeutic point of view. This preliminary data of an hypothetical interaction, was also confirmed through fluorescence spectroscopic experiments: the results obtained in this thesis were compared with those reported in the literature for the bovine serum albumine (BSA) [263]. BSA carries different metabolites and drugs and bind them with high affinity but in a reversible manner [274]. For this reason, it is considered as a model for studying drug-protein interaction *in vitro* [257], [275], [276].

In the BSA/irinotecan titration assay, the fluorescence of BSA decreased with the gradually increasing concentration of irinotecan, according to a quencing of the intrinsic fluorescence of the protein. The isosbestic point is present at 386 nm, indicating that there is an equilibrium state between BSA and irinotecan. These remarkable hypochromic effects without a red shift and the value of the equilibrium binding constant (K_a value of $8x10^5$ M^{-1}), indicate the existance of a strong interaction between irinotecan and BSA molecule with a stoichiometry of the binding process of 1:1, in perfect agreement with the known

function of BSA as transporter of different metabolites and drugs in blood plasma and with the previously published results [262].

In the HLA-G/irinotecan titration assay, we observed the basic condition of the medium from the recorded spectra, then we hypothesized the possible residual presence of TrisHCl in the recombinant HLA-G solution. In fact, irinotecan presents a constitutive fluorescence properties with slight differences based on the pH of the medium (emission fluorescence of the irinotecan is quite constant under physiological conditions (pH range: 4-8)) and its excitation spectra, showed three maxima located at 222, 255 and 368 nm according to the literature data [277]. A decrease in the HLA-G fluorescence with the gradually increased concentrations of irinotecan, was observed, suggesting a similar interaction of those reported for the BSA/irinotecan complex. We investigated the possibility of fluorescence resonance energy transfer (FRET) according to these assumptions: 1) HLA-G (first batch) emits fluorescence at 344 nm even if with a weaker intensity compared to BSA spectrum; 2) the emission spectrum of HLA-G and the absorption spectrum of irinotecan, overlap; and 3) if the two molecules are close enough, as in the case of an interaction, the FRET could occur. The binding constant, evaluated in the micromolar range, was similar to the results found for BSA. The presence of a single binding site for irinotecan in the HLA-G structure, supposed from the docking analysis, is in agreement with the results obtained from fluorescence spectroscopy. Unfortunately, these results were not confirmed analyzing another batch of the protein: we found some problems in the absorbance and emission spectra of this new sample. Therefore, this second batch of HLA-G was analyzed with liquid chromatography-mass spectrometry (HPLC-MS) to elucidate the composition of the solution and to identify unknown compounds. Chromatographic process showed a considerable presence of polyethylene glycol (PEG) molecules and other detergents.

It should be highlighted that these results of HLA-G/irinotecan interaction should be confirmed with others techniques such as circular dichroism (CD), for monitoring conformational alterations in protein folding. Moreover, pH, temperature, ionic strength parameters and folding process should be better evaluated. A potential effect of irinotecan in the secondary and terziary structure of the HLA-G should be also evaluated. Moreover, the HLA-G interaction with SN38, the active metabolite of irinotecan, could be investigated. Even if we did not measure the effect of SN38, it is well known that the absorption and fluorescence properties of the two molecules are almost the same under physiological condition used in our experiments, as reported in the literature [278]. However, the lost of the piperidinopiperidine group from the CPT-11 to form the SN38 molecule could change the affinity of interaction with HLA-G.

In conclusion, to the best our knowledge, this is the first study that investigated the possible interactions between irinotecan and HLA-G. The docking results combined with those obtained from the fluorescence spectroscopy, also compared with those of BSA, strongly suggest an interesting novel interaction between irinotecan and HLA-G, that could have revelant implications also from a therapeutical point of view.

7. CONCLUSIONS

The research of novel immunotherapy strategy has been widespread during these years. Several immunotherapeutic drugs entered the market as effective strategies in the struggle against cancer. They are mainly monoclonal antibodies or small molecules targeting the immunocheckpoints inhibitors such as PD-L1, PD-1, CTLA-4. In this contest, the HLA-G could become a very interesting molecule to be investigated because of its immunotolerance function and its role in the tumor immune escape mechanism.

In this thesis, a global focus of the potential clinical implications of HLA-G was realized, considering the genetics pattern and soluble expression levels of the HLA-G molecule in relation with clinical outcome and pharmacokinetics data in patients with mCRC treated with FOLFIRI regimen and the HLA-G-irinotecan interactions.

Some *HLA-G* polymorphisms, haplotypes or secretor classes resulted interestly associated with favorable clinical outcomes. However, these potential predictive and prognostic values should be confirmed using a huge validation cohort, composed of patients with mCRC with the same clinical characteristics and homogeneously treated with FOLFIRI as first-line treatment. A pre-planned stratification of the patients based on their *HLA-G* secretor classes and *UGT1A1*28* polymorphisms, could be a very important improvement not only in the interpretation and analysis of the data but also in the clinical applications. Moreover, further *in vitro* functional studies should be done to confirm the role of these preliminary biomarkers. The results of PART 1 seem to disagree with the tolerogenic role of HLA-G towards several type of cancer. However, the evaluation of the soluble plasmatic HLA-G level could give a response to these apparently conflicting results. In fact, the inverse relation between sHLA-G and irinotecan AUC levels observed in PART 2, suggests an irinotecan-based effect of sHLA-G in patient with mCRC. Although the result found for sHLA-G and metastasis number is concordant with the expected increase in the tumor escape phenomenon, we also observed that patients with high sHLA-G levels were associated with diminished irinotecan and BI levels, suggesting an improved efficacy of the drug metabolism and glucuronidation and detoxification pathways in these patients. It is possible that in presence of tumor metastasis the increased expression of sHLA-G may represent an indirect mechanism of protection against cancer. About that, the interesting result for the PK parameters could suggest a possible mechanism of irinotecan captured by the HLA-G molecule that alters the interaction of HLA-G with its target receptors, blocking the induction of tolerogenic effect that facilitates tumor immune escape and then determines a favorable prognosis.

The main limits of these analyses are the small sample size, especially for the sHLA-G evaluation, and the selection bias that could be present using a retrospective collection of available tumour specimens for the plasma analysis. Finally, the structural and docking analysis highlighted an almost perfect fit of irinotecan into the pocket constitute by the HLA-G α1-α2 domains. This physical interaction was also suggested by the gradually reduction of the fluorescence intensity of HLA-G with increasing of the irinotecan concentrations observed in the fluorescence spectra of the titration assays recorded with spectroscopy analysis. These preliminary data seem to confirm the docking evidences, but the hypothesis of strong interaction between HLA-G and irinotecan should be confirmed with others *in vitro* experimental assays. Irinotecan could assume a relevant role as immunotherapeutic drug due to its new role as HLA-G inhibitor. In fact, the hypothesized aspecific binding to the HLA-G pocket, could inhibit the binding with HLA-G receptors on the other immune system cells avoiding the tolerogenic effect of HLA-G and allow a quicker immune response against cancer.

Novel efficacious cancer drugs, deeper knowledge of disease molecular characteristics, application of personalized patient-centred strategies and evolution of multidisciplinary teams are needed to

continuously ameliorate the treatment of patients and also the presence of polymorphisms in immune-related genes cannot be ignored. In fact, the interaction between the immune system and cancer can significantly influence the clinical management of the patients with respect to predisposition, nature, prognosis and response to treatment of each individual. In conclusion, the overall findings for HLA-G molecule in mCRC, treated with first-line FOLFIRI regimen, represent an additional value for the management of patients that could have implications also from a therapeutic point of view in the field of precision medicine.

8. ACKNOWLEDGMENTS

I want to acknowledge my supervisor, Prof. Gabriele Grassi and co-supervisors, Prof. Mario Grassi and Prof. Silvano Geremia of the University of Trieste and Dr. Giuseppe Toffoli and Dr. Marica Garziera of the National Cancer Institute (CRO) IRCCS of Aviano (PN). In particular, I want to gratefully acknowledge Dr. Marica Garziera for her constant support in all the course of the realization and revision of this thesis and also the Prof. Silvano Geremia and Prof. Fioretta Asaro for their precious and very helpful support, expecially for PART 3. I want also to acknowledge my Ph.D. coordinators Prof. Barbara Milani and Mauro Stener of the University of Trieste, Prof. Alessandro Scarso and Prof. Maurizio Selva of the Ca' Foscari University in Venice for their support during these years. I want also to thank Dr. Chiara Romualdi and Dr. Marica Garziera for the statistical analysis, Bruno Casetta for providing mass-spectrometry technical assistance, and all the colleagues of the Experimental and Clinical Pharmacology research group of the National Cancer Institute (CRO) IRCCS of Aviano (PN) for their suggestions and advices.

9. REFERENCES

[1] K. Van der Jeught, H.-C. Xu, Y.-J. Li, X.-B. Lu, and G. Ji, "Drug resistance and new therapies in colorectal cancer," *World J. Gastroenterol.*, vol. 24, no. 34, pp. 3834–3848, Sep. 2018.

[2] D. S. Dizon *et al.*, "Clinical Cancer Advances 2016: Annual Report on Progress Against Cancer From the American Society of Clinical Oncology," *J. Clin. Oncol. Off. J. Am. Soc. Clin. Oncol.*, vol. 34, no. 9, pp. 987–1011, Mar. 2016.

[3] J. Couzin-Frankel, "Cancer Immunotherapy," *Science*, vol. 342, no. 6165, pp. 1432–1433, Dec. 2013.

[4] D. R. Leach, M. F. Krummel, and J. P. Allison, "Enhancement of antitumor immunity by CTLA-4 blockade," *Science*, vol. 271, no. 5256, pp. 1734–1736, Mar. 1996.

[5] F. S. Hodi *et al.*, "Improved survival with ipilimumab in patients with metastatic melanoma," *N. Engl. J. Med.*, vol. 363, no. 8, pp. 711–723, Aug. 2010.

[6] J. Sheng *et al.*, "Clinical Pharmacology Considerations for the Development of Immune Checkpoint Inhibitors," *J. Clin. Pharmacol.*, vol. 57 Suppl 10, pp. S26–S42, Oct. 2017.

[7] E. N. Baruch, A. L. Berg, M. J. Besser, J. Schachter, and G. Markel, "Adoptive T cell therapy: An overview of obstacles and opportunities," *Cancer*, vol. 123, no. S11, pp. 2154–2162, Jun. 2017.

[8] H. J. Jackson, S. Rafiq, and R. J. Brentjens, "Driving CAR T-cells forward," *Nat. Rev. Clin. Oncol.*, vol. 13, no. 6, pp. 370–383, 2016.

[9] M. B. Geyer and R. J. Brentjens, "Review: Current clinical applications of chimeric antigen receptor (CAR) modified T cells," *Cytotherapy*, vol. 18, no. 11, pp. 1393–1409, 2016.

[10] E. S. Hickman, M. E. Lomax, and B. K. Jakobsen, "Antigen Selection for Enhanced Affinity T-Cell Receptor-Based Cancer Therapies," *J. Biomol. Screen.*, vol. 21, no. 8, pp. 769–785, Sep. 2016.

[11] H. Shen *et al.*, "Enhancing cancer immunotherapy through nanotechnology-mediated tumor infiltration and activation of immune cells," *Semin. Immunol.*, vol. 34, pp. 114–122, Dec. 2017.

[12] O. J. Finn, "The dawn of vaccines for cancer prevention," *Nat. Rev. Immunol.*, Dec. 2017.

[13] R. D. Schreiber, L. J. Old, and M. J. Smyth, "Cancer immunoediting: integrating immunity's roles in cancer suppression and promotion," *Science*, vol. 331, no. 6024, pp. 1565–1570, 2011.

[14] F. Bray, J. Ferlay, I. Soerjomataram, R. L. Siegel, L. A. Torre, and A. Jemal, "Global cancer statistics 2018: GLOBOCAN estimates of incidence and mortality worldwide for 36 cancers in 185 countries," *CA. Cancer J. Clin.*, Sep. 2018.

[15] M. Garziera and G. Toffoli, "Inhibition of host immune response in colorectal cancer: Human leukocyte antigen-G and beyond," *World J. Gastroenterol. WJG*, vol. 20, no. 14, pp. 3778–3794, Apr. 2014.

[16] M. Garziera *et al.*, "HLA-G 3'UTR Polymorphisms Impact the Prognosis of Stage II-III CRC Patients in Fluoropyrimidine-Based Treatment," *PLoS ONE*, vol. 10, no. 12, Dec. 2015.

[17] E. C. Castelli, C. T. Mendes-Junior, L. C. Veiga-Castelli, M. Roger, P. Moreau, and E. A. Donadi, "A comprehensive study of polymorphic sites along the HLA-G gene: implication for gene regulation and evolution," *Mol. Biol. Evol.*, vol. 28, no. 11, pp. 3069–3086, Nov. 2011.

[18] G. Martelli-Palomino *et al.*, "Polymorphic Sites at the 3' Untranslated Region of the HLA-G Gene Are Associated with Differential hla-g Soluble Levels in the Brazilian and French Population," *PLoS ONE*, vol. 8, no. 10, Oct. 2013.

[19] F. M. B. Zambra, V. Biolchi, C. C. S. de Cerqueira, I. S. Brum, E. C. Castelli, and J. a. B. Chies, "Immunogenetics of prostate cancer and benign hyperplasia--the potential use of an HLA-G variant as a tag SNP for prostate cancer risk," *HLA*, vol. 87, no. 2, pp. 79–88, Feb. 2016.

[20] V. Rebmann, J. Regel, D. Stolke, and H. Grosse-Wilde, "Secretion of sHLA-G molecules in malignancies," *Semin. Cancer Biol.*, vol. 13, no. 5, pp. 371–377, Oct. 2003.

[21] A. Lin *et al.*, "HLA-G expression in human ovarian carcinoma counteracts NK cell function," *Ann. Oncol. Off. J. Eur. Soc. Med. Oncol.*, vol. 18, no. 11, pp. 1804–1809, Nov. 2007.

[22] Y. Wang, Z. Ye, X.-Q. Meng, and S.-S. Zheng, "Expression of HLA-G in patients with hepatocellular carcinoma," *Hepatobiliary Pancreat. Dis. Int. HBPD INT*, vol. 10, no. 2, pp. 158–163, Apr. 2011.

[23] A. Lin *et al.*, "Clinical relevance and functional implications for human leucocyte antigen-g expression in non-small-cell lung cancer," *J. Cell. Mol. Med.*, vol. 14, no. 9, pp. 2318–2329, Sep. 2010.

[24] N. Zheng *et al.*, "Up-regulation of HLA-G expression in cervical premalignant and malignant lesions," *Tissue Antigens*, vol. 77, no. 3, pp. 218–224, Mar. 2011.
[25] M. Garziera *et al.*, "HLA-G 3'UTR Polymorphisms Predict Drug-Induced G3-4 Toxicity Related to Folinic Acid/5-Fluorouracil/Oxaliplatin (FOLFOX4) Chemotherapy in Non-Metastatic Colorectal Cancer," *Int. J. Mol. Sci.*, vol. 18, no. 7, Jun. 2017.
[26] S. Mallal *et al.*, "Association between presence of HLA-B*5701, HLA-DR7, and HLA-DQ3 and hypersensitivity to HIV-1 reverse-transcriptase inhibitor abacavir," *Lancet Lond. Engl.*, vol. 359, no. 9308, pp. 727–732, Mar. 2002.
[27] C.-Y. Cheng, S.-C. Su, C.-H. Chen, W.-L. Chen, S.-T. Deng, and W.-H. Chung, "HLA Associations and Clinical Implications in T-Cell Mediated Drug Hypersensitivity Reactions: An Updated Review," *J. Immunol. Res.*, vol. 2014, 2014.
[28] P. T. Illing *et al.*, "Immune self-reactivity triggered by drug-modified HLA-peptide repertoire," *Nature*, vol. 486, no. 7404, pp. 554–558, Jun. 2012.
[29] N. Hirayama, "Docking simulations between drugs and HLA molecules associated with idiosyncratic drug toxicity," *Drug Metab. Pharmacokinet.*, vol. 32, no. 1, pp. 31–39, Feb. 2017.
[30] M. Dominguez-Valentin *et al.*, "Identification of genetic variants for clinical management of familial colorectal tumors," *BMC Med. Genet.*, vol. 19, Feb. 2018.
[31] B. A. Talseth-Palmer and R. J. Scott, "Genetic Variation and its Role in Malignancy," *Int. J. Biomed. Sci. IJBS*, vol. 7, no. 3, pp. 158–171, Sep. 2011.
[32] A. Hirbod-Mobarakeh, A. A. Amirzargar, B. Nikbin, M. H. Nicknam, A. Kutikhin, and N. Rezaei, "Immunogenetics of Cancer," in *Cancer Immunology: A Translational Medicine Context*, N. Rezaei, Ed. Berlin, Heidelberg: Springer Berlin Heidelberg, 2015, pp. 295–341.
[33] M. V. Relling and W. E. Evans, "Pharmacogenomics in the clinic," *Nature*, vol. 526, no. 7573, pp. 343–350, Oct. 2015.
[34] J. Abbasi, "Getting Pharmacogenomics Into the Clinic," *JAMA*, vol. 316, no. 15, pp. 1533–1535, Oct. 2016.
[35] G. Cavaletti, P. Alberti, and P. Marmiroli, "Chemotherapy-induced peripheral neurotoxicity in the era of pharmacogenomics," *Lancet Oncol.*, vol. 12, no. 12, pp. 1151–1161, Nov. 2011.
[36] L. Cortejoso and L. A. López-Fernández, "Pharmacogenetic markers of toxicity for chemotherapy in colorectal cancer patients," *Pharmacogenomics*, vol. 13, no. 10, pp. 1173–1191, Jul. 2012.
[37] D. Hanahan and R. A. Weinberg, "Hallmarks of Cancer: The Next Generation," *Cell*, vol. 144, no. 5, pp. 646–674, Mar. 2011.
[38] Y. A. Fouad and C. Aanei, "Revisiting the hallmarks of cancer," *Am. J. Cancer Res.*, vol. 7, no. 5, pp. 1016–1036, May 2017.
[39] J. A. Rodríguez, "HLA-mediated tumor escape mechanisms that may impair immunotherapy clinical outcomes via T-cell activation (Review)," *Oncol. Lett.*, vol. 14, no. 4, pp. 4415–4427, Oct. 2017.
[40] M. Malvezzi *et al.*, "European cancer mortality predictions for the year 2018 with focus on colorectal cancer," *Ann. Oncol. Off. J. Eur. Soc. Med. Oncol.*, vol. 29, no. 4, pp. 1016–1022, Apr. 2018.
[41] L. N. Kolonel and L. R. Wilkens, *Migrant Studies*. Oxford University Press, 2006.
[42] M. De Rosa *et al.*, "Genetics, diagnosis and management of colorectal cancer (Review)," *Oncol. Rep.*, vol. 34, no. 3, pp. 1087–1096, Sep. 2015.
[43] L. Salvatore *et al.*, "Management of metastatic colorectal cancer patients: guidelines of the Italian Medical Oncology Association (AIOM)," *ESMO Open*, vol. 2, no. 1, Apr. 2017.
[44] D. Colussi, G. Brandi, F. Bazzoli, and L. Ricciardiello, "Molecular Pathways Involved in Colorectal Cancer: Implications for Disease Behavior and Prevention," *Int. J. Mol. Sci.*, vol. 14, no. 8, pp. 16365–16385, Aug. 2013.
[45] L. Novellasdemunt, P. Antas, and V. S. W. Li, "Targeting Wnt signaling in colorectal cancer. A Review in the Theme: Cell Signaling: Proteins, Pathways and Mechanisms," *Am. J. Physiol. - Cell Physiol.*, vol. 309, no. 8, pp. C511–C521, Oct. 2015.
[46] J. A. McCubrey *et al.*, "GSK-3 as potential target for therapeutic intervention in cancer," *Oncotarget*, vol. 5, no. 10, pp. 2881–2911, May 2014.

[47] A. Tiwari, S. Saraf, A. Verma, P. K. Panda, and S. K. Jain, “Novel targeting approaches and signaling pathways of colorectal cancer: An insight,” *World J. Gastroenterol.*, vol. 24, no. 39, pp. 4428–4435, Oct. 2018.

[48] J. Guinney *et al.*, “The consensus molecular subtypes of colorectal cancer,” *Nat. Med.*, vol. 21, no. 11, pp. 1350–1356, Nov. 2015.

[49] S. B. Edge and C. C. Compton, “The American Joint Committee on Cancer: the 7th edition of the AJCC cancer staging manual and the future of TNM,” *Ann. Surg. Oncol.*, vol. 17, no. 6, pp. 1471–1474, Jun. 2010.

[50] J. Shia, D. S. Klimstra, P. Bagci, O. Basturk, and N. V. Adsay, “TNM staging of colorectal carcinoma: issues and caveats,” *Semin. Diagn. Pathol.*, vol. 29, no. 3, pp. 142–153, Aug. 2012.

[51] E. Van Cutsem *et al.*, “ESMO consensus guidelines for the management of patients with metastatic colorectal cancer,” *Ann. Oncol. Off. J. Eur. Soc. Med. Oncol.*, vol. 27, no. 8, pp. 1386–1422, 2016.

[52] K. W. J. Loh *et al.*, “Tumor burden (TB) as a prognostic indicator in patients with metastatic colorectal cancer (mCRC).,” *J. Clin. Oncol.*, vol. 32, no. 3_suppl, pp. 572–572, Jan. 2014.

[53] S. C. Kamran *et al.*, “Primary tumor sidedness is an independent prognostic marker for survival in metastatic colorectal cancer: Results from a large retrospective cohort with mutational analysis,” *Cancer Med.*, vol. 7, no. 7, pp. 2934–2942, May 2018.

[54] R. Dienstmann, L. Vermeulen, J. Guinney, S. Kopetz, S. Tejpar, and J. Tabernero, “Consensus molecular subtypes and the evolution of precision medicine in colorectal cancer,” *Nat. Rev. Cancer*, vol. 17, no. 2, pp. 79–92, 2017.

[55] A. Vogel, R. D. Hofheinz, S. Kubicka, and D. Arnold, “Treatment decisions in metastatic colorectal cancer – Beyond first and second line combination therapies,” *Cancer Treat. Rev.*, vol. 59, pp. 54–60, Sep. 2017.

[56] R. I. López *et al.*, “Consensus on management of metastatic colorectal cancer in Central America and the Caribbean: San José, Costa Rica, August 2016,” *ESMO Open*, vol. 3, no. 3, p. e000315, 2018.

[57] A. Kalyan, S. Kircher, H. Shah, M. Mulcahy, and A. Benson, “Updates on immunotherapy for colorectal cancer,” *J. Gastrointest. Oncol.*, vol. 9, no. 1, pp. 160–169, Feb. 2018.

[58] C.-M. Zhang *et al.*, “Role of Deficient Mismatch Repair in the Personalized Management of Colorectal Cancer,” *Int. J. Environ. Res. Public. Health*, vol. 13, no. 9, Sep. 2016.

[59] E. Van Cutsem, A. Cervantes, B. Nordlinger, D. Arnold, and ESMO Guidelines Working Group, “Metastatic colorectal cancer: ESMO Clinical Practice Guidelines for diagnosis, treatment and follow-up,” *Ann. Oncol. Off. J. Eur. Soc. Med. Oncol.*, vol. 25 Suppl 3, pp. iii1-9, Sep. 2014.

[60] F. F. Kabbinavar, J. Hambleton, R. D. Mass, H. I. Hurwitz, E. Bergsland, and S. Sarkar, “Combined analysis of efficacy: the addition of bevacizumab to fluorouracil/leucovorin improves survival for patients with metastatic colorectal cancer,” *J. Clin. Oncol. Off. J. Am. Soc. Clin. Oncol.*, vol. 23, no. 16, pp. 3706–3712, Jun. 2005.

[61] E. Van Cutsem *et al.*, “Safety and efficacy of first-line bevacizumab with FOLFOX, XELOX, FOLFIRI and fluoropyrimidines in metastatic colorectal cancer: the BEAT study,” *Ann. Oncol. Off. J. Eur. Soc. Med. Oncol.*, vol. 20, no. 11, pp. 1842–1847, Nov. 2009.

[62] L. B. Saltz *et al.*, “Bevacizumab in combination with oxaliplatin-based chemotherapy as first-line therapy in metastatic colorectal cancer: a randomized phase III study,” *J. Clin. Oncol. Off. J. Am. Soc. Clin. Oncol.*, vol. 26, no. 12, pp. 2013–2019, Apr. 2008.

[63] E. Van Cutsem *et al.*, “Addition of aflibercept to fluorouracil, leucovorin, and irinotecan improves survival in a phase III randomized trial in patients with metastatic colorectal cancer previously treated with an oxaliplatin-based regimen,” *J. Clin. Oncol. Off. J. Am. Soc. Clin. Oncol.*, vol. 30, no. 28, pp. 3499–3506, Oct. 2012.

[64] C. Bokemeyer *et al.*, “Fluorouracil, leucovorin, and oxaliplatin with and without cetuximab in the first-line treatment of metastatic colorectal cancer,” *J. Clin. Oncol. Off. J. Am. Soc. Clin. Oncol.*, vol. 27, no. 5, pp. 663–671, Feb. 2009.

[65] A. P. Venook *et al.*, “CALGB/SWOG 80405: Phase III trial of irinotecan/5-FU/leucovorin (FOLFIRI) or oxaliplatin/5-FU/leucovorin (mFOLFOX6) with bevacizumab (BV) or cetuximab (CET) for patients (pts) with KRAS wild-type (wt) untreated metastatic adenocarcinoma of the colon or rectum (MCRC).,” *J. Clin. Oncol.*, vol. 32, no. 15_suppl, pp. LBA3-LBA3, May 2014.

[66] C.-H. Köhne *et al.*, "First-line panitumumab plus irinotecan/5-fluorouracil/leucovorin treatment in patients with metastatic colorectal cancer," *J. Cancer Res. Clin. Oncol.*, vol. 138, no. 1, pp. 65–72, Jan. 2012.

[67] G. Tomasello, F. Petrelli, M. Ghidini, A. Russo, R. Passalacqua, and S. Barni, "FOLFOXIRI Plus Bevacizumab as Conversion Therapy for Patients With Initially Unresectable Metastatic Colorectal Cancer: A Systematic Review and Pooled Analysis," *JAMA Oncol.*, vol. 3, no. 7, p. e170278, Jul. 2017.

[68] L. Shui, Y.-S. Wu, H. Lin, P. Shui, Q. Sun, and X. Chen, "Triplet Chemotherapy (FOLFOXIRI) Plus Bevacizumab Versus Doublet Chemotherapy (FOLFOX/FOLFIRI) Plus Bevacizumab in Conversion Therapy for Metastatic Colorectal Cancer: a Meta-Analysis," *Cell. Physiol. Biochem.*, vol. 48, no. 5, pp. 1870–1881, 2018.

[69] C. Cremolini *et al.*, "FOLFOXIRI plus bevacizumab versus FOLFIRI plus bevacizumab as first-line treatment of patients with metastatic colorectal cancer: updated overall survival and molecular subgroup analyses of the open-label, phase 3 TRIBE study," *Lancet Oncol.*, vol. 16, no. 13, pp. 1306–1315, Oct. 2015.

[70] T. Gruenberger *et al.*, "Bevacizumab plus mFOLFOX-6 or FOLFOXIRI in patients with initially unresectable liver metastases from colorectal cancer: the OLIVIA multinational randomised phase II trial," *Ann. Oncol. Off. J. Eur. Soc. Med. Oncol.*, vol. 26, no. 4, pp. 702–708, Apr. 2015.

[71] H. Hasegawa *et al.*, "Efficacy of Second-Line Bevacizumab-Containing Chemotherapy for Patients with Metastatic Colorectal Cancer following First-Line Treatment with an Anti-Epidermal Growth Factor Receptor Antibody," *Oncology*, vol. 92, no. 4, pp. 205–212, 2017.

[72] M. Peeters *et al.*, "Analysis of KRAS/NRAS Mutations in a Phase III Study of Panitumumab with FOLFIRI Compared with FOLFIRI Alone as Second-line Treatment for Metastatic Colorectal Cancer," *Clin. Cancer Res. Off. J. Am. Assoc. Cancer Res.*, vol. 21, no. 24, pp. 5469–5479, 15 2015.

[73] R. M. Goldberg *et al.*, "Optimising the use of cetuximab in the continuum of care for patients with metastatic colorectal cancer," *ESMO Open*, vol. 3, no. 4, p. e000353, May 2018.

[74] A. S. Lucas, B. H. O'Neil, and R. M. Goldberg, "A Decade of Advances in Cytotoxic Chemotherapy for Metastatic Colorectal Cancer," *Clin. Colorectal Cancer*, vol. 10, no. 4, pp. 238–244, Dec. 2011.

[75] J. W. Holch, I. Ricard, S. Stintzing, D. P. Modest, and V. Heinemann, "The relevance of primary tumour location in patients with metastatic colorectal cancer: A meta-analysis of first-line clinical trials," *Eur. J. Cancer Oxf. Engl. 1990*, vol. 70, pp. 87–98, 2017.

[76] E. Cecchin, E. De Mattia, F. Ecca, and G. Toffoli, "Host genetic profiling to increase drug safety in colorectal cancer from discovery to implementation," *Drug Resist. Updat. Rev. Comment. Antimicrob. Anticancer Chemother.*, vol. 39, pp. 18–40, 2018.

[77] A. Di Paolo *et al.*, "Pharmacokinetic and pharmacogenetic predictive markers of irinotecan activity and toxicity," *Curr. Drug Metab.*, vol. 12, no. 10, pp. 932–943, Dec. 2011.

[78] G. Toffoli *et al.*, "The role of UGT1A1*28 polymorphism in the pharmacodynamics and pharmacokinetics of irinotecan in patients with metastatic colorectal cancer," *J. Clin. Oncol. Off. J. Am. Soc. Clin. Oncol.*, vol. 24, no. 19, pp. 3061–3068, Jul. 2006.

[79] J. M. Hoskins, R. M. Goldberg, P. Qu, J. G. Ibrahim, and H. L. McLeod, "UGT1A1*28 genotype and irinotecan-induced neutropenia: dose matters," *J. Natl. Cancer Inst.*, vol. 99, no. 17, pp. 1290–1295, Sep. 2007.

[80] G. Toffoli *et al.*, "Genotype-driven phase I study of irinotecan administered in combination with fluorouracil/leucovorin in patients with metastatic colorectal cancer," *J. Clin. Oncol. Off. J. Am. Soc. Clin. Oncol.*, vol. 28, no. 5, pp. 866–871, Feb. 2010.

[81] R. H. J. Mathijssen *et al.*, "Clinical Pharmacokinetics and Metabolism of Irinotecan (CPT-11)," *Clin. Cancer Res.*, vol. 7, no. 8, pp. 2182–2194, Aug. 2001.

[82] D. B. Longley, D. P. Harkin, and P. G. Johnston, "5-fluorouracil: mechanisms of action and clinical strategies," *Nat. Rev. Cancer*, vol. 3, no. 5, pp. 330–338, May 2003.

[83] T. R. Buroker *et al.*, "Randomized comparison of two schedules of fluorouracil and leucovorin in the treatment of advanced colorectal cancer," *J. Clin. Oncol. Off. J. Am. Soc. Clin. Oncol.*, vol. 12, no. 1, pp. 14–20, Jan. 1994.

[84] J. S. Macdonald, "Toxicity of 5-fluorouracil," *Oncol. Williston Park N*, vol. 13, no. 7 Suppl 3, pp. 33–34, Jul. 1999.

[85] G. Toffoli *et al.*, "Clinical validity of a DPYD-based pharmacogenetic test to predict severe toxicity to fluoropyrimidines," *Int. J. Cancer*, vol. 137, no. 12, pp. 2971–2980, Dec. 2015.

[86] N. N. Grenon and J. Chan, "Managing toxicities associated with colorectal cancer chemotherapy and targeted therapy: a new guide for nurses," *Clin. J. Oncol. Nurs.*, vol. 13, no. 3, pp. 285–296, Jun. 2009.

[87] T. Yoshino *et al.*, "Pan-Asian adapted ESMO consensus guidelines for the management of patients with metastatic colorectal cancer: a JSMO–ESMO initiative endorsed by CSCO, KACO, MOS, SSO and TOS," *Ann. Oncol.*, vol. 29, no. 1, pp. 44–70, Jan. 2018.

[88] B. Tran *et al.*, "Impact of BRAF mutation and microsatellite instability on the pattern of metastatic spread and prognosis in metastatic colorectal cancer," *Cancer*, vol. 117, no. 20, pp. 4623–4632, Oct. 2011.

[89] C. I. Müller *et al.*, "Predictive and prognostic value of microsatellite instability in patients with advanced colorectal cancer treated with a fluoropyrimidine and oxaliplatin containing first-line chemotherapy. A report of the AIO Colorectal Study Group," *Int. J. Colorectal Dis.*, vol. 23, no. 11, pp. 1033–1039, Nov. 2008.

[90] G. Toffoli *et al.*, "The genotype for DPYD risk variants in colorectal cancer patients and the related toxicity management costs in clinical practice," *Clin. Pharmacol. Ther.*, Oct. 2018.

[91] Y.-Y. Qian *et al.*, "The ERCC1 C118T polymorphism predicts clinical outcomes of colorectal cancer patients receiving oxaliplatin-based chemotherapy: a meta-analysis based on 22 studies," *Asian Pac. J. Cancer Prev. APJCP*, vol. 15, no. 19, pp. 8383–8390, 2014.

[92] G. J. Peters *et al.*, "Induction of thymidylate synthase as a 5-fluorouracil resistance mechanism," *Biochim. Biophys. Acta*, vol. 1587, no. 2–3, pp. 194–205, Jul. 2002.

[93] B. Zhao *et al.*, "Mechanisms of resistance to anti-EGFR therapy in colorectal cancer," *Oncotarget*, vol. 8, no. 3, pp. 3980–4000, Jan. 2017.

[94] Y. Y. Juo *et al.*, "Prognostic value of CpG island methylator phenotype among colorectal cancer patients: a systematic review and meta-analysis," *Ann. Oncol. Off. J. Eur. Soc. Med. Oncol.*, vol. 25, no. 12, pp. 2314–2327, Dec. 2014.

[95] J.-B. Li *et al.*, "Importance of the plasma soluble HLA-G levels for prognostic stratification with traditional prognosticators in colorectal cancer," *Oncotarget*, vol. 8, no. 30, pp. 48854–48862, Jul. 2017.

[96] C. Kirana *et al.*, "Soluble HLA-G is a differential prognostic marker in sequential colorectal cancer disease stages," *Int. J. Cancer*, vol. 140, no. 11, pp. 2577–2586, 01 2017.

[97] E. C. Castelli *et al.*, "In silico analysis of microRNAS targeting the HLA-G 3' untranslated region alleles and haplotypes," *Hum. Immunol.*, vol. 70, no. 12, pp. 1020–1025, Dec. 2009.

[98] E. C. Castelli *et al.*, "Insights into HLA-G Genetics Provided by Worldwide Haplotype Diversity," *Front. Immunol.*, vol. 5, p. 476, 2014.

[99] D. Tronik-Le Roux *et al.*, "Novel landscape of HLA-G isoforms expressed in clear cell renal cell carcinoma patients," *Mol. Oncol.*, vol. 11, no. 11, pp. 1561–1578, 2017.

[100] G. Curigliano, C. Criscitiello, L. Gelao, and A. Goldhirsch, "Molecular pathways: human leukocyte antigen G (HLA-G)," *Clin. Cancer Res. Off. J. Am. Assoc. Cancer Res.*, vol. 19, no. 20, pp. 5564–5571, Oct. 2013.

[101] C. S. Clements *et al.*, "Crystal structure of HLA-G: a nonclassical MHC class I molecule expressed at the fetal-maternal interface," *Proc. Natl. Acad. Sci. U. S. A.*, vol. 102, no. 9, pp. 3360–3365, Mar. 2005.

[102] W.-H. Yan, "HLA-G expression in cancers: potential role in diagnosis, prognosis and therapy," *Endocr. Metab. Immune Disord. Drug Targets*, vol. 11, no. 1, pp. 76–89, Mar. 2011.

[103] "Fundamental immunology - NLM Catalog - NCBI." [Online]. Available: https://www.ncbi.nlm.nih.gov/nlmcatalog/101591462. [Accessed: 20-Aug-2018].

[104] K.-Y. HoWangYin *et al.*, "Multimeric structures of HLA-G isoforms function through differential binding to LILRB receptors," *Cell. Mol. Life Sci. CMLS*, vol. 69, no. 23, pp. 4041–4049, Dec. 2012.

[105] G. Amodio and S. Gregori, "Distinctive Immunological Functions of HLA-G," *Histocompatibility*, May 2012.

[106] N. Rouas-Freiss, P. Moreau, J. LeMaoult, and E. D. Carosella, “The Dual Role of HLA-G in Cancer,” *Journal of Immunology Research*, 2014. [Online]. Available: https://www.hindawi.com/journals/jir/2014/359748/. [Accessed: 17-Oct-2018].

[107] S. Rajagopalan and E. O. Long, “KIR2DL4 (CD158d): An activation receptor for HLA-G,” *Front. Immunol.*, vol. 3, p. 258, 2012.

[108] F. Zhang *et al.*, “Inhibitory leukocyte immunoglobulin-like receptors in cancer development,” *Sci. China Life Sci.*, vol. 58, no. 12, pp. 1216–1225, Dec. 2015.

[109] G. Amodio *et al.*, “HLA-G expression levels influence the tolerogenic activity of human DC-10,” *Haematologica*, vol. 100, no. 4, pp. 548–557, Apr. 2015.

[110] A. Lin and W.-H. Yan, “Human Leukocyte Antigen-G (HLA-G) Expression in Cancers: Roles in Immune Evasion, Metastasis and Target for Therapy,” *Mol. Med. Camb. Mass*, vol. 21, no. 1, pp. 782–791, Nov. 2015.

[111] E. D. Carosella, “The tolerogenic molecule HLA-G,” *Immunol. Lett.*, vol. 138, no. 1, pp. 22–24, Jul. 2011.

[112] L. M. R. Ferreira, T. B. Meissner, T. Tilburgs, and J. L. Strominger, “HLA-G: At the Interface of Maternal–Fetal Tolerance,” *Trends Immunol.*, vol. 38, no. 4, pp. 272–286, Apr. 2017.

[113] J. S. Hunt, D. K. Langat, R. H. McIntire, and P. J. Morales, “The role of HLA-G in human pregnancy,” *Reprod. Biol. Endocrinol.*, vol. 4, no. Suppl 1, p. S10, Oct. 2006.

[114] A. González, V. Rebmann, J. LeMaoult, P. A. Horn, E. D. Carosella, and E. Alegre, “The immunosuppressive molecule HLA-G and its clinical implications,” *Crit. Rev. Clin. Lab. Sci.*, vol. 49, no. 3, pp. 63–84, Jun. 2012.

[115] V. Rebmann, F. da Silva Nardi, B. Wagner, and P. A. Horn, “HLA-G as a tolerogenic molecule in transplantation and pregnancy,” *J. Immunol. Res.*, vol. 2014, p. 297073, 2014.

[116] R. Rizzo, D. Bortolotti, O. R. Baricordi, and E. Fainardi, “New insights into HLA-G and inflammatory diseases,” *Inflamm. Allergy Drug Targets*, vol. 11, no. 6, pp. 448–463, Dec. 2012.

[117] P. Moreau, G. Mouillot, P. Rousseau, C. Marcou, J. Dausset, and E. D. Carosella, “HLA-G gene repression is reversed by demethylation,” *Proc. Natl. Acad. Sci. U. S. A.*, vol. 100, no. 3, pp. 1191–1196, Feb. 2003.

[118] K. Poláková, E. Bandžuchová, D. Kuba, and G. Russ, “Demethylating agent 5-aza-2’-deoxycytidine activates HLA-G expression in human leukemia cell lines,” *Leuk. Res.*, vol. 33, no. 4, pp. 518–524, Apr. 2009.

[119] E. C. Castelli, L. C. Veiga-Castelli, L. Yaghi, P. Moreau, and E. A. Donadi, “Transcriptional and Posttranscriptional Regulations of the HLA-G Gene,” *Journal of Immunology Research*, 2014. [Online]. Available: https://www.hindawi.com/journals/jir/2014/734068/. [Accessed: 28-Jan-2019].

[120] I. Zidi *et al.*, “Increase in HLA-G1 proteolytic shedding by tumor cells: a regulatory pathway controlled by NF-kappaB inducers,” *Cell. Mol. Life Sci. CMLS*, vol. 63, no. 22, pp. 2669–2681, Nov. 2006.

[121] P. Rousseau, M. Le Discorde, G. Mouillot, C. Marcou, E. D. Carosella, and P. Moreau, “The 14 bp deletion-insertion polymorphism in the 3’ UT region of the HLA-G gene influences HLA-G mRNA stability,” *Hum. Immunol.*, vol. 64, no. 11, pp. 1005–1010, Nov. 2003.

[122] I. Manaster *et al.*, “MiRNA-mediated control of HLA-G expression and function,” *PloS One*, vol. 7, no. 3, p. e33395, 2012.

[123] A. A. Celik, G. S. Simper, W. Hiemisch, R. Blasczyk, and C. Bade-Döding, “HLA-G peptide preferences change in transformed cells: impact on the binding motif,” *Immunogenetics*, vol. 70, no. 8, pp. 485–494, Aug. 2018.

[124] S. Yie, L. Li, R. Xiao, and C. L. Librach, “A single base-pair mutation in the 3’-untranslated region of HLA-G mRNA is associated with pre-eclampsia,” *Mol. Hum. Reprod.*, vol. 14, no. 11, pp. 649–653, Nov. 2008.

[125] V. Rebmann, K. van der Ven, M. Pässler, K. Pfeiffer, D. Krebs, and H. Grosse-Wilde, “Association of soluble HLA-G plasma levels with HLA-G alleles,” *Tissue Antigens*, vol. 57, no. 1, pp. 15–21, Jan. 2001.

[126] M. Le Discorde, C. Le Danff, P. Moreau, N. Rouas-Freiss, and E. D. Carosella, “HLA-G*0105N Null Allele Encodes Functional HLA-G Isoforms,” *Biol. Reprod.*, vol. 73, no. 2, pp. 280–288, Aug. 2005.

[127] J. Lajoie, A. Jeanneau, M.-C. Faucher, P. Moreau, and M. Roger, “Characterisation of five novel HLA-G alleles with coding DNA base changes,” *Tissue Antigens*, vol. 72, no. 5, pp. 502–504, Nov. 2008.
[128] C. T. Mendes-Junior, E. C. Castelli, D. Meyer, A. L. Simões, and E. A. Donadi, “Genetic diversity of the *HLA-G* coding region in Amerindian populations from the Brazilian Amazon: a possible role of natural selection,” *Genes Immun.*, vol. 14, no. 8, pp. 518–526, Dec. 2013.
[129] W. Y. Wang, W. Tian, X. X. Liu, and L. X. Li, “HLA-G coding region and 3’untranslated region (3’UTR) in two Chinese Han populations,” *Immunol. Lett.*, vol. 176, pp. 65–71, Aug. 2016.
[130] T. V. F. Hviid, R. Rizzo, O. B. Christiansen, L. Melchiorri, A. Lindhard, and O. R. Baricordi, “HLA-G and IL-10 in serum in relation to HLA-G genotype and polymorphisms,” *Immunogenetics*, vol. 56, no. 3, pp. 135–141, Jun. 2004.
[131] A. M. Bamberger, S. Jenatschke, H. M. Schulte, T. Löning, and M. C. Bamberger, “Leukemia inhibitory factor (LIF) stimulates the human HLA-G promoter in JEG3 choriocarcinoma cells,” *J. Clin. Endocrinol. Metab.*, vol. 85, no. 10, pp. 3932–3936, Oct. 2000.
[132] P. Paul *et al.*, “HLA-G expression in melanoma: a way for tumor cells to escape from immunosurveillance,” *Proc. Natl. Acad. Sci. U. S. A.*, vol. 95, no. 8, pp. 4510–4515, Apr. 1998.
[133] G. Singer *et al.*, “HLA-G is a potential tumor marker in malignant ascites,” *Clin. Cancer Res. Off. J. Am. Assoc. Cancer Res.*, vol. 9, no. 12, pp. 4460–4464, Oct. 2003.
[134] G. Murdaca *et al.*, “HLA-G expression in gastric carcinoma: clinicopathological correlations and prognostic impact,” *Virchows Arch. Int. J. Pathol.*, vol. 473, no. 4, pp. 425–433, Oct. 2018.
[135] G. Caocci *et al.*, “HLA-G molecules and clinical outcome in Chronic Myeloid Leukemia,” *Leuk. Res.*, vol. 61, pp. 1–5, 2017.
[136] F. Gros *et al.*, “Soluble HLA-G Molecules Are Increased during Acute Leukemia, Especially in Subtypes Affecting Monocytic and Lymphoid Lineages,” *Neoplasia N. Y. N*, vol. 8, no. 3, pp. 223–230, Mar. 2006.
[137] R.-L. Zhang *et al.*, “Predictive value of different proportion of lesion HLA-G expression in colorectal cancer,” *Oncotarget*, vol. 8, no. 64, pp. 107441–107451, Nov. 2017.
[138] C. Menier, S. Prevot, E. D. Carosella, and N. Rouas-Freiss, “Human leukocyte antigen-G is expressed in advanced-stage ovarian carcinoma of high-grade histology,” *Hum. Immunol.*, vol. 70, no. 12, pp. 1006–1009, Dec. 2009.
[139] M. J. Rutten *et al.*, “HLA-G Expression Is an Independent Predictor for Improved Survival in High Grade Ovarian Carcinomas,” *Journal of Immunology Research*, 2014. [Online]. Available: https://www.hindawi.com/journals/jir/2014/274584/. [Accessed: 20-Oct-2018].
[140] C. B. M. Bijen *et al.*, “The prognostic role of classical and nonclassical MHC class I expression in endometrial cancer,” *Int. J. Cancer*, vol. 126, no. 6, pp. 1417–1427, Mar. 2010.
[141] B. S. de Almeida, Y. C. N. Muniz, A. H. Prompt, E. C. Castelli, C. T. Mendes-Junior, and E. A. Donadi, “Genetic association between HLA-G 14-bp polymorphism and diseases: A systematic review and meta-analysis,” *Hum. Immunol.*, vol. 79, no. 10, pp. 724–735, Oct. 2018.
[142] R. L. Elliott, X. P. Jiang, J. T. Phillips, B. G. Barnett, and J. F. Head, “Human leukocyte antigen G expression in breast cancer: role in immunosuppression,” *Cancer Biother. Radiopharm.*, vol. 26, no. 2, pp. 153–157, Apr. 2011.
[143] X. He *et al.*, “HLA-G expression in human breast cancer: implications for diagnosis and prognosis, and effect on allocytotoxic lymphocyte response after hormone treatment in vitro,” *Ann. Surg. Oncol.*, vol. 17, no. 5, pp. 1459–1469, May 2010.
[144] W.-H. Yan, D. Liu, H.-Y. Lu, Y.-Y. Li, X. Zhang, and A. Lin, “Significance of tumour cell HLA-G5/-G6 isoform expression in discrimination for adenocarcinoma from squamous cell carcinoma in lung cancer patients,” *J. Cell. Mol. Med.*, vol. 19, no. 4, pp. 778–785, Apr. 2015.
[145] Z.-Y. Guo *et al.*, “Predictive value of HLA-G and HLA-E in the prognosis of colorectal cancer patients,” *Cell. Immunol.*, vol. 293, no. 1, pp. 10–16, Jan. 2015.
[146] L.-H. Gan *et al.*, “Tumor-specific upregulation of human leukocyte antigen-G expression in bladder transitional cell carcinoma,” *Hum. Immunol.*, vol. 71, no. 9, pp. 899–904, Sep. 2010.
[147] B.-L. Li *et al.*, “Characterization of HLA-G expression in renal cell carcinoma,” *Tissue Antigens*, vol. 74, no. 3, pp. 213–221, Sep. 2009.

[148] P. Khodabandeh Shahraki, Y. Zare, N. Azarpira, M. Hosseinzadeh, and S. Farjadian, “Prognostic Value of HLA-G in Malignant Liver and Pancreas Lesions,” *Iran. J. Immunol. IJI*, vol. 15, no. 1, pp. 28–37, Mar. 2018.

[149] Y.-F. Xu *et al.*, “High Expression of Human Leukocyte Antigen-G is Associated with a Poor Prognosis in Patients with PDAC,” *Curr. Mol. Med.*, vol. 15, no. 4, pp. 360–367, 2015.

[150] S.-M. Yie, H. Yang, S.-R. Ye, K. Li, D.-D. Dong, and X.-M. Lin, “Expression of HLA-G is associated with prognosis in esophageal squamous cell carcinoma,” *Am. J. Clin. Pathol.*, vol. 128, no. 6, pp. 1002–1009, Dec. 2007.

[151] A. Lin *et al.*, “Human leukocyte antigen-G expression is associated with a poor prognosis in patients with esophageal squamous cell carcinoma,” *Int. J. Cancer*, vol. 129, no. 6, pp. 1382–1390, Sep. 2011.

[152] J. Zheng *et al.*, “Human leukocyte antigen G is associated with esophageal squamous cell carcinoma progression and poor prognosis,” *Immunol. Lett.*, vol. 161, no. 1, pp. 13–19, Sep. 2014.

[153] I. J. Wastowski *et al.*, “Human leukocyte antigen-G is frequently expressed in glioblastoma and may be induced in vitro by combined 5-aza-2’-deoxycytidine and interferon-γ treatments: results from a multicentric study,” *Am. J. Pathol.*, vol. 182, no. 2, pp. 540–552, Feb. 2013.

[154] M.-B. Cai *et al.*, “Expression of human leukocyte antigen G is associated with prognosis in nasopharyngeal carcinoma,” *Int. J. Biol. Sci.*, vol. 8, no. 6, pp. 891–900, 2012.

[155] A. S. Gonçalves *et al.*, “The clinicopathologic significance of the expression of HLA-G in oral squamous cell carcinoma,” *Oral Surg. Oral Med. Oral Pathol. Oral Radiol.*, vol. 117, no. 3, pp. 361–368, Mar. 2014.

[156] A. S. Gonçalves *et al.*, “Overexpression of immunomodulatory mediators in oral precancerous lesions,” *Hum. Immunol.*, vol. 78, no. 11–12, pp. 752–757, Nov. 2017.

[157] L. M. Nunes *et al.*, “Association between the HLA-G molecule and lymph node metastasis in papillary thyroid cancer,” *Hum. Immunol.*, vol. 74, no. 4, pp. 447–451, Apr. 2013.

[158] N. L. de Figueiredo-Feitosa, G. Martelli Palomino, D. C. Cilião Alves, C. T. Mendes Junior, E. A. Donadi, and L. M. Z. Maciel, “HLA-G 3’ untranslated region polymorphic sites associated with increased HLA-G production are more frequent in patients exhibiting differentiated thyroid tumours,” *Clin. Endocrinol. (Oxf.)*, vol. 86, no. 4, pp. 597–605, Apr. 2017.

[159] W.-H. Yan *et al.*, “Unfavourable clinical implications for HLA-G expression in acute myeloid leukaemia,” *J. Cell. Mol. Med.*, vol. 12, no. 3, pp. 889–898, Jun. 2008.

[160] H. Nückel, V. Rebmann, J. Dürig, U. Dührsen, and H. Grosse-Wilde, “HLA-G expression is associated with an unfavorable outcome and immunodeficiency in chronic lymphocytic leukemia,” *Blood*, vol. 105, no. 4, pp. 1694–1698, Feb. 2005.

[161] G. Locafaro, G. Amodio, D. Tomasoni, C. Tresoldi, F. Ciceri, and S. Gregori, “HLA-G expression on blasts and tolerogenic cells in patients affected by acute myeloid leukemia,” *J. Immunol. Res.*, vol. 2014, p. 636292, 2014.

[162] M. A. Attia, N. A. Nosair, A. Gawally, G. Elnagar, and E. M. Elshafey, “HLA-G expression as a prognostic indicator in B-cell chronic lymphocytic leukemia,” *Acta Haematol.*, vol. 132, no. 1, pp. 53–58, 2014.

[163] E. Alegre, R. Rizzo, D. Bortolotti, S. Fernandez-Landázuri, E. Fainardi, and A. González, “Some Basic Aspects of HLA-G Biology,” *J. Immunol. Res.*, vol. 2014, 2014.

[164] V. Rebmann *et al.*, “Detection of soluble HLA-G molecules in plasma and amniotic fluid,” *Tissue Antigens*, vol. 53, no. 1, pp. 14–22, Jan. 1999.

[165] A.-S. Moreau *et al.*, “Clinical relevance of soluble HLA class I molecules in Waldenstrom Macroglobulinemia,” *Eur. J. Haematol.*, vol. 80, no. 6, pp. 503–509, Jun. 2008.

[166] S. Ugurel, V. Rebmann, S. Ferrone, W. Tilgen, H. Grosse-Wilde, and U. Reinhold, “Soluble human leukocyte antigen--G serum level is elevated in melanoma patients and is further increased by interferon-alpha immunotherapy,” *Cancer*, vol. 92, no. 2, pp. 369–376, Jul. 2001.

[167] A. Ben Amor, K. Beauchemin, M.-C. Faucher, A. Hamzaoui, K. Hamzaoui, and M. Roger, “Human Leukocyte Antigen G Polymorphism and Expression Are Associated with an Increased Risk of Non-Small-Cell Lung Cancer and Advanced Disease Stage,” *PloS One*, vol. 11, no. 8, p. e0161210, 2016.

[168] Y. Sebti *et al.*, “Soluble HLA-G molecules are increased in lymphoproliferative disorders,” *Hum. Immunol.*, vol. 64, no. 11, pp. 1093–1101, Nov. 2003.

[169] P. Wlasiuk *et al.*, “Expression of soluble HLA-G in multiple myeloma patients and patients with renal failure,” *Leuk. Res.*, vol. 36, no. 7, pp. 881–883, Jul. 2012.
[170] S. Farjadian, M. Tabebordbar, M. Mokhtari, A. Safaei, M. Malekzadeh, and A. Ghaderi, “HLA-G Expression in Tumor Tissues and Soluble HLA-G Plasma Levels in Patients with Gastrointestinal Cancer,” *Asian Pac. J. Cancer Prev. APJCP*, vol. 19, no. 10, pp. 2731–2735, Oct. 2018.
[171] H. Ben Yahia *et al.*, “Increased plasmatic soluble HLA-G levels in endometrial cancer,” *Mol. Immunol.*, vol. 99, pp. 82–86, Jul. 2018.
[172] P. Schütt *et al.*, “Prognostic relevance of soluble human leukocyte antigen-G and total human leukocyte antigen class I molecules in lung cancer patients,” *Hum. Immunol.*, vol. 71, no. 5, pp. 489–495, May 2010.
[173] A. Lin *et al.*, “Aberrant human leucocyte antigen-G expression and its clinical relevance in hepatocellular carcinoma,” *J. Cell. Mol. Med.*, vol. 14, no. 8, pp. 2162–2171, Aug. 2010.
[174] F. Morandi, S. Pozzi, B. Carlini, L. Amoroso, V. Pistoia, and M. V. Corrias, “Soluble HLA-G and HLA-E Levels in Bone Marrow Plasma Samples Are Related to Disease Stage in Neuroblastoma Patients,” *J. Immunol. Res.*, vol. 2016, p. 7465741, 2016.
[175] E. Lesport *et al.*, “Human melanoma cell secreting human leukocyte antigen-G5 inhibit natural killer cell cytotoxicity by impairing lytic granules polarization toward target cell,” *Hum. Immunol.*, vol. 70, no. 12, pp. 1000–1005, Dec. 2009.
[176] B. Riteau *et al.*, “Exosomes bearing HLA-G are released by melanoma cells,” *Hum. Immunol.*, vol. 64, no. 11, pp. 1064–1072, Nov. 2003.
[177] V. Rebmann, L. König, F. da S. Nardi, B. Wagner, L. F. S. Manvailer, and P. A. Horn, “The Potential of HLA-G-Bearing Extracellular Vesicles as a Future Element in HLA-G Immune Biology,” *Front. Immunol.*, vol. 7, May 2016.
[178] R. Rizzo *et al.*, “HLA-G is a component of the chronic lymphocytic leukemia escape repertoire to generate immune suppression: impact of the HLA-G 14 base pair (rs66554220) polymorphism,” *Haematologica*, vol. 99, no. 5, pp. 888–896, May 2014.
[179] M. Bielska *et al.*, “Human leukocyte antigen-G polymorphisms influence the clinical outcome in diffuse large B-cell lymphoma,” *Genes. Chromosomes Cancer*, vol. 54, no. 3, pp. 185–193, Mar. 2015.
[180] M. Garziera *et al.*, “Association of the HLA-G 3’UTR polymorphisms with colorectal cancer in Italy: a first insight,” *Int. J. Immunogenet.*, vol. 43, no. 1, pp. 32–39, Feb. 2016.
[181] E. A. Donadi, E. C. Castelli, A. Arnaiz-Villena, M. Roger, D. Rey, and P. Moreau, “Implications of the polymorphism of HLA-G on its function, regulation, evolution and disease association,” *Cell. Mol. Life Sci. CMLS*, vol. 68, no. 3, pp. 369–395, Feb. 2011.
[182] N. Ghandri *et al.*, “Association of HLA-G polymorphisms with nasopharyngeal carcinoma risk and clinical outcome,” *Hum. Immunol.*, vol. 72, no. 2, pp. 150–158, Feb. 2011.
[183] E. C. Castelli, C. T. Mendes-Junior, J. L. Viana de Camargo, and E. A. Donadi, “HLA-G polymorphism and transitional cell carcinoma of the bladder in a Brazilian population,” *Tissue Antigens*, vol. 72, no. 2, pp. 149–157, Aug. 2008.
[184] E. C. M. Zeestraten *et al.*, “Combined analysis of HLA class I, HLA-E and HLA-G predicts prognosis in colon cancer patients,” *Br. J. Cancer*, vol. 110, no. 2, pp. 459–468, Jan. 2014.
[185] M. Swets *et al.*, “Promoter methylation and mRNA expression of HLA-G in relation to HLA-G protein expression in colorectal cancer,” *Hum. Immunol.*, vol. 77, no. 9, pp. 764–772, Sep. 2016.
[186] R. B. Özgül Özdemir, A. T. Özdemir, F. Oltulu, K. Kurt, G. Yiğittürk, and C. Kırmaz, “A comparison of cancer stem cell markers and nonclassical major histocompatibility complex antigens in colorectal tumor and noncancerous tissues,” *Ann. Diagn. Pathol.*, vol. 25, pp. 60–63, Dec. 2016.
[187] M. Swets *et al.*, “HLA-G protein expression in colorectal cancer evaluated by immunohistochemistry and western blot analysis: Its expression characteristics remain enigmatic,” *Clin. Immunol. Orlando Fla*, vol. 194, pp. 80–86, Jul. 2018.
[188] M. Cao, S.-M. Yie, J. Liu, S. R. Ye, D. Xia, and E. Gao, “Plasma soluble HLA-G is a potential biomarker for diagnosis of colorectal, gastric, esophageal and lung cancer,” *Tissue Antigens*, vol. 78, no. 2, pp. 120–128, Aug. 2011.

[189] R. Rizzo *et al.*, "Role of HLA-G 14bp deletion/insertion and +3142C>G polymorphisms in the production of sHLA-G molecules in relapsing-remitting multiple sclerosis," *Hum. Immunol.*, vol. 73, no. 11, pp. 1140–1146, Nov. 2012.

[190] C.-B. Zhu, C.-X. Wang, X. Zhang, J. Zhang, and W. Li, "Serum sHLA-G levels: A useful indicator in distinguishing colorectal cancer from benign colorectal diseases," *Int. J. Cancer*, vol. 128, no. 3, pp. 617–622, Feb. 2011.

[191] R. Samadi *et al.*, "Clinical Value of Human Leucocyte Antigen G (HLA-G) Expression in the Prognosis of Colorectal Cancer," *Int. J. Cancer Manag.*, vol. 10, no. 4, 2017.

[192] B. F. Barrier, B. S. Kendall, K. L. Sharpe-Timms, and E. R. Kost, "Characterization of human leukocyte antigen-G (HLA-G) expression in endometrial adenocarcinoma," *Gynecol. Oncol.*, vol. 103, no. 1, pp. 25–30, Oct. 2006.

[193] A. Lin, H.-H. Xu, D.-P. Xu, X. Zhang, Q. Wang, and W.-H. Yan, "Multiple steps of HLA-G in ovarian carcinoma metastasis: alter NK cytotoxicity and induce matrix metalloproteinase-15 (MMP-15) expression," *Hum. Immunol.*, vol. 74, no. 4, pp. 439–446, Apr. 2013.

[194] L. Loumagne, J. Baudhuin, B. Favier, F. Montespan, E. D. Carosella, and N. Rouas-Freiss, "In vivo evidence that secretion of HLA-G by immunogenic tumor cells allows their evasion from immunosurveillance," *Int. J. Cancer*, vol. 135, no. 9, pp. 2107–2117, Nov. 2014.

[195] X. Zhang *et al.*, "Methotrexate-loaded PLGA nanobubbles for ultrasound imaging and Synergistic Targeted therapy of residual tumor during HIFU ablation," *Biomaterials*, vol. 35, no. 19, pp. 5148–5161, Jun. 2014.

[196] Y. Komohara *et al.*, "HLA-G as a target molecule in specific immunotherapy against renal cell carcinoma," *Oncol. Rep.*, vol. 18, no. 6, pp. 1463–1468, Dec. 2007.

[197] J. S. Boura *et al.*, "Evaluation of gene delivery strategies to efficiently overexpress functional HLA-G on human bone marrow stromal cells," *Mol. Ther. - Methods Clin. Dev.*, vol. 1, p. 14041, Jan. 2014.

[198] J. LeMaoult, M. Daouya, J. Wu, M. Loustau, A. Horuzsko, and E. D. Carosella, "Synthetic HLA-G proteins for therapeutic use in transplantation," *FASEB J.*, vol. 27, no. 9, pp. 3643–3651, Sep. 2013.

[199] J. M. Michot *et al.*, "Immune-related adverse events with immune checkpoint blockade: a comprehensive review," *Eur. J. Cancer Oxf. Engl. 1990*, vol. 54, pp. 139–148, Feb. 2016.

[200] F. Kroschinsky *et al.*, "New drugs, new toxicities: severe side effects of modern targeted and immunotherapy of cancer and their management," *Crit. Care Lond. Engl.*, vol. 21, no. 1, p. 89, 14 2017.

[201] J. Uetrecht and D. J. Naisbitt, "Idiosyncratic Adverse Drug Reactions: Current Concepts," *Pharmacol. Rev.*, vol. 65, no. 2, pp. 779–808, Apr. 2013.

[202] M. Pirmohamed, D. A. Ostrov, and B. K. Park, "New genetic findings lead the way to a better understanding of fundamental mechanisms of drug hypersensitivity," *J. Allergy Clin. Immunol.*, vol. 136, no. 2, pp. 236–244, Aug. 2015.

[203] T. Usui and D. J. Naisbitt, "Human leukocyte antigen and idiosyncratic adverse drug reactions," *Drug Metab. Pharmacokinet.*, vol. 32, no. 1, pp. 21–30, Feb. 2017.

[204] W. J. Pichler, D. J. Naisbitt, and B. K. Park, "Immune pathomechanism of drug hypersensitivity reactions," *J. Allergy Clin. Immunol.*, vol. 127, no. 3 Suppl, pp. S74-81, Mar. 2011.

[205] S. J. Posadas and W. J. Pichler, "Delayed drug hypersensitivity reactions – new concepts," *Clin. Exp. Allergy*, vol. 37, no. 7, pp. 989–999, 2007.

[206] C.-B. Chen *et al.*, "An Updated Review of the Molecular Mechanisms in Drug Hypersensitivity," *J. Immunol. Res.*, vol. 2018, p. 6431694, 2018.

[207] H. Luo *et al.*, "Molecular docking to identify associations between drugs and class I human leukocyte antigens for predicting idiosyncratic drug reactions," *Comb. Chem. High Throughput Screen.*, vol. 18, no. 3, pp. 296–304, 2015.

[208] A. Tailor, L. Faulkner, D. J. Naisbitt, and B. K. Park, "The chemical, genetic and immunological basis of idiosyncratic drug-induced liver injury," *Hum. Exp. Toxicol.*, vol. 34, no. 12, pp. 1310–1317, Dec. 2015.

[209] A. K. Daly *et al.*, "HLA-B*5701 genotype is a major determinant of drug-induced liver injury due to flucloxacillin," *Nat. Genet.*, vol. 41, no. 7, pp. 816–819, Jul. 2009.

[210] P. Zhou, S. Zhang, Y. Wang, C. Yang, and J. Huang, "Structural modeling of HLA-B*1502/peptide/carbamazepine/T-cell receptor complex architecture: implication for the molecular mechanism of carbamazepine-induced Stevens-Johnson syndrome/toxic epidermal necrolysis," *J. Biomol. Struct. Dyn.*, vol. 34, no. 8, pp. 1806–1817, Aug. 2016.

[211] M. McCormack *et al.*, "HLA-A*3101 and carbamazepine-induced hypersensitivity reactions in Europeans," *N. Engl. J. Med.*, vol. 364, no. 12, pp. 1134–1143, Mar. 2011.

[212] H. Miyadera, T. Ozeki, T. Mushiroda, and N. Hirayama, "In silico Analysis of Interactions between HLA-A* 31: 01 and carbamazepine-related Compounds," *Chem-Bio Inform. J.*, vol. 16, pp. 5–8, 2016.

[213] S.-I. Hung *et al.*, "HLA-B*5801 allele as a genetic marker for severe cutaneous adverse reactions caused by allopurinol," *Proc. Natl. Acad. Sci. U. S. A.*, vol. 102, no. 11, pp. 4134–4139, Mar. 2005.

[214] W.-L. Fan *et al.*, "HLA Association with Drug-Induced Adverse Reactions," *Journal of Immunology Research*, 2017. [Online]. Available: https://www.hindawi.com/journals/jir/2017/3186328/. [Accessed: 19-Feb-2019].

[215] S. Amari, R. Kataoka, T. Ikegami, and N. Hirayama, "HLA-Modeler: Automated Homology Modeling of Human Leukocyte Antigens," *International Journal of Medicinal Chemistry*, 2013. [Online]. Available: https://www.hindawi.com/journals/ijmc/2013/690513/. [Accessed: 22-Oct-2018].

[216] J. Goto, R. Kataoka, H. Muta, and N. Hirayama, "ASEDock-docking based on alpha spheres and excluded volumes," *J. Chem. Inf. Model.*, vol. 48, no. 3, pp. 583–590, Mar. 2008.

[217] A. B. Miller, B. Hoogstraten, M. Staquet, and A. Winkler, "Reporting results of cancer treatment," *Cancer*, vol. 47, no. 1, pp. 207–214, Jan. 1981.

[218] Anonymous, *Guidelines for Reporting of Adverse Drug Reactions. Bethesda, MD, National Cancer Institute, Division of Cancer Treatment, 2004*. Bethesda, MD, National Cancer Institute, Division of Cancer Treatment, 2004.

[219] E. De Mattia *et al.*, "Pharmacogenetics of ABC and SLC transporters in metastatic colorectal cancer patients receiving first-line FOLFIRI treatment," *Pharmacogenet. Genomics*, vol. 23, no. 10, pp. 549–557, Oct. 2013.

[220] D. W. Cockcroft and M. H. Gault, "Prediction of creatinine clearance from serum creatinine," *Nephron*, vol. 16, no. 1, pp. 31–41, 1976.

[221] C. Tournigand *et al.*, "FOLFIRI Followed by FOLFOX6 or the Reverse Sequence in Advanced Colorectal Cancer: A Randomized GERCOR Study," *J. Clin. Oncol.*, vol. 22, no. 2, pp. 229–237, Jan. 2004.

[222] E. C. Castelli *et al.*, "The genetic structure of 3'untranslated region of the HLA-G gene: polymorphisms and haplotypes," *Genes Immun.*, vol. 11, no. 2, pp. 134–141, Mar. 2010.

[223] M. Stephens, N. J. Smith, and P. Donnelly, "A new statistical method for haplotype reconstruction from population data," *Am. J. Hum. Genet.*, vol. 68, no. 4, pp. 978–989, Apr. 2001.

[224] M. Stephens and P. Donnelly, "A comparison of bayesian methods for haplotype reconstruction from population genotype data," *Am. J. Hum. Genet.*, vol. 73, no. 5, pp. 1162–1169, Nov. 2003.

[225] I. Poras *et al.*, "Haplotypes of the HLA-G 3' Untranslated Region Respond to Endogenous Factors of HLA-G+ and HLA-G- Cell Lines Differentially," *PLOS ONE*, vol. 12, no. 1, p. e0169032, gen 2017.

[226] M. Garziera *et al.*, "HLA-G 3'UTR Polymorphisms Impact the Prognosis of Stage II-III CRC Patients in Fluoropyrimidine-Based Treatment," *PloS One*, vol. 10, no. 12, p. e0144000, 2015.

[227] A. Sabbagh *et al.*, "Worldwide genetic variation at the 3' untranslated region of the HLA-G gene: balancing selection influencing genetic diversity," *Genes Immun.*, vol. 15, no. 2, pp. 95–106, Mar. 2014.

[228] I. Zlobec and A. Lugli, "Prognostic and predictive factors in colorectal cancer," *J. Clin. Pathol.*, vol. 61, no. 5, pp. 561–569, May 2008.

[229] X. Yan, S. Zhang, Y. Deng, P. Wang, Q. Hou, and H. Xu, "Prognostic Factors for Checkpoint Inhibitor Based Immunotherapy: An Update With New Evidences," *Front. Pharmacol.*, vol. 9, 2018.

[230] G. Amodio *et al.*, "Association of genetic variants in the 3'UTR of HLA-G with Recurrent Pregnancy Loss," *Hum. Immunol.*, vol. 77, no. 10, pp. 886–891, Oct. 2016.

[231] D. Marques *et al.*, "Association of insertion-deletions polymorphisms with colorectal cancer risk and clinical features," *World J. Gastroenterol.*, vol. 23, no. 37, pp. 6854–6867, Oct. 2017.

[232] X.-Y. Chen, W.-H. Yan, A. Lin, H.-H. Xu, J.-G. Zhang, and X.-X. Wang, "The 14 bp deletion polymorphisms in HLA-G gene play an important role in the expression of soluble HLA-G in plasma," *Tissue Antigens*, vol. 72, no. 4, pp. 335–341, Oct. 2008.

[233] T. V. F. Hviid, S. Hylenius, C. Rørbye, and L. G. Nielsen, "HLA-G allelic variants are associated with differences in the HLA-G mRNA isoform profile and HLA-G mRNA levels," *Immunogenetics*, vol. 55, no. 2, pp. 63–79, May 2003.

[234] S. G. Svendsen *et al.*, "The expression and functional activity of membrane-bound human leukocyte antigen-G1 are influenced by the 3'-untranslated region," *Hum. Immunol.*, vol. 74, no. 7, pp. 818–827, Jul. 2013.

[235] T. D. Veit and J. a. B. Chies, "Tolerance versus immune response -- microRNAs as important elements in the regulation of the HLA-G gene expression," *Transpl. Immunol.*, vol. 20, no. 4, pp. 229–231, Mar. 2009.

[236] M. M. Dias, R. A. McKinnon, and M. J. Sorich, "Impact of the UGT1A1*28 allele on response to irinotecan: a systematic review and meta-analysis," *Pharmacogenomics*, vol. 13, no. 8, pp. 889–899, Jun. 2012.

[237] M. Garziera *et al.*, "HLA-G 3'UTR Polymorphisms Predict Drug-Induced G3-4 Toxicity Related to Folinic Acid/5-Fluorouracil/Oxaliplatin (FOLFOX4) Chemotherapy in Non-Metastatic Colorectal Cancer," *Int. J. Mol. Sci.*, vol. 18, no. 7, Jun. 2017.

[238] D. Rodrigues, A. Longatto-Filho, and S. F. Martins, "Predictive Biomarkers in Colorectal Cancer: From the Single Therapeutic Target to a Plethora of Options," *BioMed Res. Int.*, vol. 2016, 2016.

[239] F. Coppedè, A. Lopomo, R. Spisni, and L. Migliore, "Genetic and epigenetic biomarkers for diagnosis, prognosis and treatment of colorectal cancer," *World J. Gastroenterol. WJG*, vol. 20, no. 4, pp. 943–956, Jan. 2014.

[240] V. Das, J. Kalita, and M. Pal, "Predictive and prognostic biomarkers in colorectal cancer: A systematic review of recent advances and challenges," *Biomed. Pharmacother.*, vol. 87, pp. 8–19, Mar. 2017.

[241] K. E. Caudle *et al.*, "Clinical Pharmacogenetics Implementation Consortium guidelines for dihydropyrimidine dehydrogenase genotype and fluoropyrimidine dosing," *Clin. Pharmacol. Ther.*, vol. 94, no. 6, pp. 640–645, Dec. 2013.

[242] S. Hylenius, A.-M. N. Andersen, M. Melbye, and T. V. F. Hviid, "Association between HLA-G genotype and risk of pre-eclampsia: a case-control study using family triads," *Mol. Hum. Reprod.*, vol. 10, no. 4, pp. 237–246, Apr. 2004.

[243] M. Garziera *et al.*, "Identification and characterization of CDH1 germline variants in sporadic gastric cancer patients and in individuals at risk of gastric cancer," *PloS One*, vol. 8, no. 10, p. e77035, 2013.

[244] M. Garziera, L. Scarabel, and G. Toffoli, "Hypoxic Modulation of HLA-G Expression through the Metabolic Sensor HIF-1 in Human Cancer Cells," *J. Immunol. Res.*, vol. 2017, p. 4587520, 2017.

[245] M. B. Suárez *et al.*, "A new HLA-G allele (HLA-G*0105N) and its distribution in the Spanish population.," *Immunogenetics*, vol. 45, no. 6, pp. 464–465, 1997.

[246] J. Di Cristofaro *et al.*, "HLA-G*01:04~UTR3 Recipient Correlates With Lower Survival and Higher Frequency of Chronic Rejection After Lung Transplantation," *Am. J. Transplant. Off. J. Am. Soc. Transplant. Am. Soc. Transpl. Surg.*, vol. 15, no. 9, pp. 2413–2420, Sep. 2015.

[247] F. Carlini *et al.*, "HLA-G UTR Haplotype Conservation in the Malian Population: Association with Soluble HLA-G," *PLOS ONE*, vol. 8, no. 12, p. e82517, dic 2013.

[248] S.-R. Ye, H. Yang, K. Li, D.-D. Dong, X.-M. Lin, and S.-M. Yie, "Human leukocyte antigen G expression: as a significant prognostic indicator for patients with colorectal cancer," *Mod. Pathol. Off. J. U. S. Can. Acad. Pathol. Inc*, vol. 20, no. 3, pp. 375–383, Mar. 2007.

[249] M. Swets *et al.*, "HLA-G and classical HLA class I expression in primary colorectal cancer and associated liver metastases," *Hum. Immunol.*, vol. 77, no. 9, pp. 773–779, Sep. 2016.

[250] I. A. Guedes, C. S. de Magalhães, and L. E. Dardenne, "Receptor–ligand molecular docking," *Biophys. Rev.*, vol. 6, no. 1, pp. 75–87, Mar. 2014.

[251] N. S. Pagadala, K. Syed, and J. Tuszynski, "Software for molecular docking: a review," *Biophys. Rev.*, vol. 9, no. 2, pp. 91–102, Apr. 2017.

[252] A. Grosdidier, V. Zoete, and O. Michielin, "SwissDock, a protein-small molecule docking web service based on EADock DSS," *Nucleic Acids Res.*, vol. 39, no. Web Server issue, pp. W270–W277, Jul. 2011.

[253] A. Grosdidier, V. Zoete, and O. Michielin, “Fast docking using the CHARMM force field with EADock DSS,” *J. Comput. Chem.*, vol. 32, no. 10, pp. 2149–2159, Jul. 2011.
[254] G. N. Murshudov *et al.*, “REFMAC5 for the refinement of macromolecular crystal structures,” *Acta Crystallogr. D Biol. Crystallogr.*, vol. 67, no. Pt 4, pp. 355–367, Apr. 2011.
[255] D. F. Swinehart, “The Beer-Lambert Law,” *J. Chem. Educ.*, vol. 39, no. 7, p. 333, Jul. 1962.
[256] M. Shiroishi *et al.*, “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. U. S. A.*, vol. 103, no. 44, pp. 16412–16417, Oct. 2006.
[257] J. Liu *et al.*, “Characterizing the binding interaction of astilbin with bovine serum albumin: a spectroscopic study in combination with molecular docking technology,” *RSC Adv.*, vol. 8, no. 13, pp. 7280–7286, 2018.
[258] H. F. Jeremias *et al.*, “Study of the interactions of bovine serum albumin with a molybdenum(II) carbonyl complex by spectroscopic and molecular simulation methods,” *PloS One*, vol. 13, no. 9, p. e0204624, 2018.
[259] S. Tartaggia, M. D. Alvau, A. Meneghello, B. Casetta, F. Polo, and G. Toffoli, “Practical fluorimetric assay for the detection of anticancer drug SN-38 in human plasma,” *J. Pharm. Biomed. Anal.*, vol. 159, pp. 73–81, Sep. 2018.
[260] M. Bobeničová, M. Valko, V. Brezová, and D. Dvoranová, “UVA generated free radicals in irinotecan (CPT-11) in the presence of copper ions,” *J. Photochem. Photobiol. Chem.*, vol. 290, pp. 125–138, Sep. 2014.
[261] N. Bijari, S. Ghobadi, and K. Derakhshandeh, “Irinotecan binds to the internal cavity of beta-lactoglobulin: A multi-spectroscopic and computational investigation,” *J. Pharm. Biomed. Anal.*, vol. 139, pp. 109–115, May 2017.
[262] P. L. Gentili, F. Ortica, and G. Favaro, “Static and Dynamic Interaction of a Naturally Occurring Photochromic Molecule with Bovine Serum Albumin Studied by UV−Visible Absorption and Fluorescence Spectroscopy,” *J. Phys. Chem. B*, vol. 112, no. 51, pp. 16793–16801, Dec. 2008.
[263] G.-Z. Li, Y.-M. Liu, X.-Y. Guo, and J.-J. Wang, “Study on the Interaction between Bovine Serum Albumins and Topotecan Hydrochloride or irinotecan hydrochloride,” *Acta Chimica Sinica -Chinese Edition-*, vol. 64, no. 7, pp. 679–685, 2006.
[264] K. Fujita, Y. Kubota, H. Ishida, and Y. Sasaki, “Irinotecan, a key chemotherapeutic drug for metastatic colorectal cancer,” *World J. Gastroenterol.*, vol. 21, no. 43, pp. 12234–12248, Nov. 2015.
[265] T. Nishimura *et al.*, “Irinotecan monotherapy as third-line treatment for advanced gastric cancer refractory to fluoropyrimidines, platinum, and taxanes,” *Gastric Cancer*, vol. 20, no. 4, pp. 655–662, Jul. 2017.
[266] K. Noda *et al.*, “Irinotecan plus Cisplatin Compared with Etoposide plus Cisplatin for Extensive Small-Cell Lung Cancer,” *N. Engl. J. Med.*, vol. 346, no. 2, pp. 85–91, Jan. 2002.
[267] D. C. Glassman *et al.*, “Nanoliposomal irinotecan with fluorouracil for the treatment of advanced pancreatic cancer, a single institution experience,” *BMC Cancer*, vol. 18, Jun. 2018.
[268] C. F. Verschraegen, “Irinotecan for the treatment of cervical cancer,” *Oncol. Williston Park N*, vol. 16, no. 5 Suppl 5, pp. 32–34, May 2002.
[269] F. Musa *et al.*, “Phase II study of irinotecan in combination with bevacizumab in recurrent ovarian cancer,” *Gynecol. Oncol.*, vol. 144, no. 2, pp. 279–284, Feb. 2017.
[270] R. H. J. Mathijssen, W. J. Loos, J. Verweij, and A. Sparreboom, “Pharmacology of topoisomerase I inhibitors irinotecan (CPT-11) and topotecan,” *Curr. Cancer Drug Targets*, vol. 2, no. 2, pp. 103–123, Jun. 2002.
[271] Y. Pommier, “Drugging topoisomerases: lessons and challenges,” *ACS Chem. Biol.*, vol. 8, no. 1, pp. 82–95, Jan. 2013.
[272] L. Anthony, “Irinotecan toxicity,” *Curr. Opin. Support. Palliat. Care*, vol. 1, no. 1, p. 35, Apr. 2007.
[273] “Irinotecan Toxicity and UGT1A: Overview, Clinical Implications, Testing for the Genetic Mutation,” Jul. 2017.
[274] T. TOPALĂ, A. BODOKI, L. OPREAN, and R. OPREAN, “Bovine Serum Albumin Interactions with Metal Complexes,” *Clujul Med.*, vol. 87, no. 4, pp. 215–219, 2014.

[275] A. Varlan and M. Hillebrand, “Bovine and human serum albumin interactions with 3-carboxyphenoxathiin studied by fluorescence and circular dichroism spectroscopy,” *Mol. Basel Switz.*, vol. 15, no. 6, pp. 3905–3919, Jun. 2010.

[276] V. D. Suryawanshi, L. S. Walekar, A. H. Gore, P. V. Anbhule, and G. B. Kolekar, “Spectroscopic analysis on the binding interaction of biologically active pyrimidine derivative with bovine serum albumin,” *J. Pharm. Anal.*, vol. 6, no. 1, pp. 56–63, Feb. 2016.

[277] M. I. Rodríguez Cáceres, I. Durán-Merás, N. E. O. Soto, P. L. L. de Alba, and L. L. Martínez, “Spectrofluorimetric determination of irinotecan in the presence of oxidant agents and metal ions,” *Talanta*, vol. 74, no. 5, pp. 1484–1491, Feb. 2008.

[278] L. A. Serrano, Y. Yang, E. Salvati, F. Stellacci, S. Krol, and S. Guldin, “pH-Mediated molecular differentiation for fluorimetric quantification of chemotherapeutic drugs in human plasma,” *Chem. Commun. Camb. Engl.*, vol. 54, no. 12, pp. 1485–1488, Feb. 2018.