Accepted Manuscript

Functional characterization of the ABCG2 5' non-coding exon variants: stem cell specificity, translation efficiency and the influence of drug selection

Sára Sándor, Theodora Jordanidisz, Anita Schamberger, György Várady, Zsuzsa Erdei, Ágota Apáti, Balázs Sarkadi, Tamás I. Orbán

PII: S1874-9399(16)30097-9
DOI: doi: 10.1016/j.bbagrm.2016.05.007
Reference: BBAGRM 1036

To appear in: *BBA - Gene Regulatory Mechanisms*

Received date: 23 February 2016
Revised date: 26 April 2016
Accepted date: 13 May 2016

Please cite this article as: Sára Sándor, Theodora Jordanidisz, Anita Schamberger, György Várady, Zsuzsa Erdei, Ágota Apáti, Balázs Sarkadi, Tamás I. Orbán, Functional characterization of the ABCG2 5' non-coding exon variants: stem cell specificity, translation efficiency and the influence of drug selection, *BBA - Gene Regulatory Mechanisms* (2016), doi: 10.1016/j.bbagrm.2016.05.007

This is a PDF file of an unedited manuscript that has been accepted for publication. As a service to our customers we are providing this early version of the manuscript. The manuscript will undergo copyediting, typesetting, and review of the resulting proof before it is published in its final form. Please note that during the production process errors may be discovered which could affect the content, and all legal disclaimers that apply to the journal pertain.

ACCEPTED MANUSCRIPT

Functional characterization of the ABCG2 5' non-coding exon variants: stem cell specificity, translation efficiency and the influence of drug selection

Sára Sándor [a], Theodora Jordanidisz [a], Anita Schamberger [a], György Várady [a], Zsuzsa Erdei [a], Ágota Apáti [a], Balázs Sarkadi [a,b], and Tamás I. Orbán [a,*]

[a] *Institute of Enzymology, Research Centre for Natural Sciences, Hungarian Academy of Sciences, Budapest, Hungary*
[b] *Department of Biophysics and Radiation Biology, Semmelweis University, MTA-SE Molecular Biophysics Research Group, Budapest, Hungary*

* **Correspondence to**:
Tamás I. Orbán
Institute of Enzymology, Research Centre for Natural Sciences,
Hungarian Academy of Sciences
Budapest, 1117. Magyar Tudosok korutja 2.
Hungary
Phone: +36 1 382 6638
Fax: +36 1 382 6295
E-mail: orban.tamas@ttk.mta.hu / orbant@biomembrane.hu

Keywords: ABCG2; noncoding exons; 5' UTR; alternative promoters; MCF-7; stem cell

Short title: Functional characterization of ABCG2 mRNA variants

ACCEPTED MANUSCRIPT

ACCEPTED MANUSCRIPT

Abstract

ABCG2 is a multidrug transporter with wide substrate specificity, and is believed to protect several cell types from various xenobiotics and endobiotics. This "guardian" function is important in numerous cell types and tissue barriers but becomes disadvantageous by being responsible for the multidrug resistance phenotype in certain tumor cells. ABCG2 regulation at the protein level has already been extensively studied, however, regulation at the mRNA level, especially the functional role of the various 5' untranslated exon variants (5' UTRs) has been elusive. In the present work, we describe a comprehensive characterization of four ABCG2 mRNA variants with different exon 1 sequences, investigate drug inducibility, stem cell specificity, mRNA stability, and translation efficiency. Although certain variants (E1B and E1C) are considered as "constitutive" mRNA isoforms, we show that chemotoxic drugs significantly alter the expression pattern of distinct ABCG2 mRNA isoforms. When examining human embryonic stem cell lines, we provide evidence that variant E1A has an expression pattern coupled to undifferentiated stem cell stage, as its transcript level is regulated parallel to mRNAs of Oct4 and Nanog pluripotency marker genes. When characterizing the four exon 1 variants we found no significant differences in terms of mRNA stabilities and half-lives of the isoforms. In contrast, variant E1U showed markedly lower translation efficiency both at the total protein level or regarding the functional presence in the plasma membrane. Taken together, these results indicate that the different 5' UTR variants play an important role in cell type specific regulation and fine tuning of ABCG2 expression.

1. Introduction

ABCG2, also known as BCRP (Breast Cancer Resistance Protein), MXR (Mitoxantrone Resistance protein) and ABCP (Placental ABC transporter), is one of the 48 known members of the human ATP-binding-cassette (ABC) protein family, and its gene is located on chromosome 4q22, encoding a 655 amino acid long protein. Originally cloned from placenta [1] and cells selected for multidrug resistance [2, 3], ABCG2 was confirmed to create an "atypical" drug resistance phenotype when overexpressed in tumor cells [4]. This characteristic is attributed to the transport function of this transmembrane protein that expels a wide range of chemicals from cells, including typical chemotherapy agents such as mitoxantrone, topotecan [5] or flavopiridol [6], giving a particular relevance to the expression of this protein in cancer treatment [7]. ABCG2 is expressed in various differentiated tissues, including the ovary, the kidney, the liver, breast epithelial cells, intestinal epithelia, the blood–brain barrier, and the placenta [8, 9]. According to its wide substrate specificity and distribution among various tissues, including important physiological barriers, it is hypothesized to play physiological role in the chemoimmunity defense system that protects cells from harmful xenobiotics or endogenous metabolites [8]. This idea is supported by the finding that null mutant mice, albeit being viable, exhibit sensitivity to certain drugs and severe phototoxicity caused by dietary pheophorbide A accumulation [10]. The same group also demonstrated protoporphyria in these animals, indicating a role of ABCG2 in hematopoiesis, possibly through interaction with protoporphyrin IX, which was supported by further studies. Moreover, the functional relevance of ABCG2 was demonstrated in other mammals, being responsible for the accumulation of chemicals in milk [11], or being an important protector of retinal cells against toxic agents in domestics cats [12]. In human populations, certain SNPs in the ABCG2 gene have been linked to the occurrence of gout [13]. The gene's other functions and its physiological substrates are

ACCEPTED MANUSCRIPT

not reassuringly determined as yet, but ABCG2 most likely protects different cell types from various toxins, especially from certain harmful degradation products formed as a result of oxidative stress or hypoxia [14]. An important aspect of ABCG2 is that the gene is also expressed in different stem cell types: it was proven to be the molecular basis for the so called “side population” phenotype [15] and was also found to be present in diverse human embryonic stem cell lines by various research groups [16-19]. Although the demonstration of its functional presence could be methodologically challenging [20], we could undoubtedly provide evidence that ABCG2 is expressed in a wide variety of human embryonic stem cell lines [21]. Moreover, we have shown that a dynamic heterogeneity of expression exists at the population level of stem cells which correlates with the presence of various stress conditions in the cellular environment [22]. For these reasons, several research groups focused their efforts on describing the gene structure and transcript species of ABCG2 to gain a better understanding of its function and genetic regulation.

Regarding the mRNA structure of ABCG2, the initial cloning studies predicted 16 exons of the gene, the first one being a non-coding exon, as the canonical translational start codon is located in exon 2 [1-3]. Subsequent work, aiming to study the ABCG2 promoter structure and regulation, predicted a number of putative transcription factor binding sites (e.g.: AP1, AP2 and several Sp1 sites) and described a CpG island but no TATA-box in the investigated promoter region [23]. A later study, however, described the presence of two other potential transcriptional initiation sites, which – analogously to the mouse ortholog [24] – result in the formation of two additional leader exons and thereby two mRNA species with different 5’ untranslated regions (5’ UTRs) [25]. In our previous study, we were able to refine the proposed leader exons, and analyzed their expression patterns in various tissue types [18]. We have detected the previously described exon 1a leader sequence and provided evidence that the originally described exon 1b and 1c by Nakanishi et al. [25] are not functionally distinct variants but rather result from transcriptional leakage and alternative splicing events from the same promoter (1b). On the other hand, we confirmed the existence of an “exon 1c” containing ABCG2 transcript in dendritic cells of hematopoietic origin, resulting from the activity of a promoter located the farthest from the originally predicted 5’ regions [18]. Our results were confirmed by later studies, further proving that these sequence variants are *bona fide* exons of certain ABCG2 messages [20, 26, 27].

Although the presence of ABCG2 5’ UTR variants has been clearly described and accepted, little is known about their potentially distinct function(s). In this study, we aimed to characterize the specific role of ABCG2 exon 1 variants in different cell types, including human embryonic stem cell lines. We studied the expression patterns of these variants, specifically addressing the effects of drug treatments on the cell type specific ABCG2 mRNA profiles. We also investigated whether the different exon 1 sequences have any influence on the mRNA stability or the translation efficiency of ABCG2. Our results point to the stem cell specific presence of E1A variant, further supporting the particular role that ABCG2 plays in such unique cell types. Moreover, we demonstrated a lower translation efficiency when E1U is present in the ABCG2 mRNA. Taken together, these results suggest that the variability of the 5’ UTR sequence has an important role in fine tuning the expression of this multidrug transporter at the posttranscriptional level.

2. Materials and methods

2.1. Cell cultures, drug-selection and transfection

ACCEPTED MANUSCRIPT

HEK293, MCF-7, HUES1, HUES4 and HUES9 cell lines were used in this study. The HEK293 cell line was purchased from ATCC, whereas the MCF-7 cell line, as well as its mitoxantrone (Mx) [28] and flavopiridol (Flv) [6] resistant sublines were kind gifts from Susan Bates, NIH, Bethesda, USA. HEK293 and MCF-7 cells, as well as drug resistant derivatives of the latter one were maintained in DMEM medium supplemented with 10% fetal bovine serum and 1% penicillin/streptomycin, and cells were cultured at 37°C in 5% CO_2 atmosphere.

The human embryonic stem cell lines HUES9, HUES4 and HUES1 were kind gifts from Dr. Douglas Melton, Harvard University, USA and were cultured as described previously [18]. Briefly, the following passage numbers were used during the experiments: 27-33 passage numbers for HUES1, 23-29 passage numbers for HUES4, and 24-30 passage numbers for HUES9. Embryonic stem cell lines were grown on mitomycin-C treated mouse embryonic fibroblast (MEF) feeder cells, in special KoDMEM medium supplemented with 15% serum replacement reagent, 1% nonessential amino acids, 1 mM L-glutamine, 0.1 mM beta-mercaptoethanol, and 4 ng/mL human fibroblast growth factor. Cells were passaged every 2 days regularly unless indicated otherwise (such as 72h treatment with Mx). For selection and differentiation induction, cells were treated with 10 nM Mx. Since the proportion of ABCG2 positive stem cells shows a dynamic fluctuation [22], the IC50 value of Mx is in a range between 3-5 nM. The applied concentration of 10 nM is higher than the IC50 value, and based on its cytotoxic effect, it allows for the selection and differentiation induction of HUES9 cells. For FACS analysis (see later), trypsinized HUES cells were separated from MEFs by gating out Sca1 positive cells (a mouse marker). For RNA analysis (see also later), collagenase was used to collect HUES cells which leaves most MEF cells behind. Nevertheless, the used human specific real-time PCR assays assured that MEF contamination was excluded in the analysis. All cell lines were regularly tested for Mycoplasma infection with the MycoAlert detection kit from Lonza.

For transfection experiments, HEK293 cells were seeded onto 6- or 12-well plates 24 hours before using FuGENE® 6 or FuGENE® HD reagents, according to the manufacturer's protocol (Promega). Cells were harvested 48 hours following transfection using the appropriate sampling methods for different purposes.

2.2. RNA isolation and real-time PCR

Total cellular RNA was isolated using the TriFast reagent from PEQLAB Biotechnologie GmbH or the TRIzol reagent from Thermo Fisher Scientific. Preparation steps were performed according to the manufacturer's protocol, the quality of obtained RNA samples was checked by gel electrophoresis and concentration was determined by NanoDrop 2000 Spectrophotometer (Thermo Fisher Scientific). 1 μg of total RNA was reverse-transcribed using the High Capacity cDNA Reverse Transcription Kit with random oligomers (Thermo Fisher Scientific). cDNA samples were diluted 10x before quantitation of gene expression by real-time PCR using gene specific TaqMan® assays; measurements were performed on StepOne™ or StepOnePlus™ Real-Time PCR platforms, according to the manufacturer's instructions (Thermo Fisher Scientific). The following pre-designed TaqMan® assays (catalog numbers) were used: "total" ABCG2 detecting the exon3/exon 4 boundary (Hs01053790_m1), Nanog (Hs02387400_g1), Oct4 (Hs03005111_g1); for endogenous controls, PolR2A (Hs00172187_m1) and RPLP0 (Hs99999902_m1). To detect the four ABCG2 mRNA 5' isoforms, specific TaqMan® assays were designed where the forward and reverse primers bind to the different exon 1 sequences and the common exon 2 sequence, respectively, whereas the probe

ACCEPTED MANUSCRIPT

binds to the unique exon 1/exon 2 junction sequences of the variants. Sequences of these primers and probes can be found in Table 1, primer locations are indicated on Fig. 1A. To avoid false detection, cross-reactivity of the assays was tested on isoform specific sequence containing plasmids, and minor cross reactions with atypical amplification curves were detected between E1B and E1C variants, due to partial sequence overlaps. These TaqMan® assays could detect but not reliably quantitate variants E1B and E1C separately, therefore in quantitative measurements, the expression levels are indicated as E1B/C.

2.3. mRNA stability measurements

MCF-7/Flv cells, which express all four exon 1 variants, were seeded onto 12-well plates and were grown under conditions described above for 48 hours before initiation of the experiment. At the starting time point, fresh media containing 5 μg/ml actinomycin D (from Sigma-Aldrich) were added to the cells. Samples were collected in TRIzol reagent at predefined intervals through the following 24 hours. mRNA-isolation, reverse-transcription and quantitative PCR were performed as described above. Expression of total ABCG2 mRNA and all four 5' UTR variants were measured at every time point and rate of decay for each isoform was determined by polynomial regression using the Microcal Origin software (OriginLab Corporation).

2.4. Construction of expression plasmids

cDNA isolated from MCF-7/Flv cells were used to amplify sequences of the four ABCG2 exon 1 variants. The forward primers were designed to selectively bind to the specific 5' end of each sequence and one common reverse primer, specific for the UTR region in exon 2, was used in the cloning reactions (Table 1). The expression plasmids used were established in our laboratory previously, all containing two separate transcription units with various promoters and transgenes [29]. In this case, one unit expressed the CMV promoter driven human ABCG2 coding sequence, whereas the other expressed the CMV promoter driven eGFP (see also Fig. 4A). The NheI and BspEI restriction sites indicated in the cloning primers were used to provide appropriate insertion of the amplified exon 1 variant sequences into the plasmid backbone, right before the coding sequence of ABCG2. The separate eGFP expression cassette served as a transfection and protein expression normalization control. In the end, we obtained four plasmids, each simultaneously expressing GFP and one particular ABCG2 mRNA isoform (E1x-ABCG2/GFP plasmids).

2.5. Western blot

Cell lysates were collected 48 hours following transfection with the E1x-ABCG2/GFP expression plasmids. Samples were sonicated prior to determination of protein concentration by Lowry method and gel electrophoresis. Two parallels were run separately from each sample: one was boiled to allow detection of soluble proteins (for GFP), and the other one left untreated for detection of membrane bound proteins (for ABCG2). 30 μg of protein samples were run on 10% acrylamide gels and electroblotted onto PVDF membranes (BioRad, Hungary). After washing with TBS-Tween, membranes were blocked by 5% milk/TBS-Tween, and subsequently incubated in primary antibody solution on 4°C overnight. The mouse monoclonal antibody BXP-21 was used for detection of ABCG2, and a rabbit monoclonal antibody was used for GFP detection (ab290 from

ACCEPTED MANUSCRIPT

Abcam). Next day, membranes were washed three times with TBS-Tween and then incubated in secondary antibody solution for 1 hour at room temperature. Anti-mouse and anti-rabbit secondary antibodies conjugated with horse radish peroxidase were used. Prior to detection with ECL reagent (from Thermo Fisher Scientific), membranes were washed twice with TBS-Tween, and the membranes were exposed to Agfa x-ray films. GFP signals were used to normalize ABCG2 expression, also allowing to control for transfection efficiency. Expression levels were determined by densitometry of the scanned images using the GelAnalyser software.

2.6. Immunostaining and mitoxantrone efflux assay

For immunofluorescent labeling of cell-surface ABCG2, we used the conformation-sensitive mouse monoclonal antibody 5D3 conjugated with Alexa647 as described earlier [30]. The 5D3 antibody was diluted in BSA/PBS to a final concentration of 2 μg/ml and cells were incubated in the solution for 1 hour at 37°C in the presence of 1 μM Ko143 inhibitor to enhance antibody binding. Control staining with appropriate isotype-matched control mAbs were also included in the experiment.

To evaluate transport activity of ABCG2 positive cells, samples were collected the same way as for immunofluorescence studies, and were incubated in PBS containing 17 μg/ml mitoxantrone for 30 minutes at 37°C. Negative control samples were obtained by the addition of the Ko143 inhibitor to a final concentration of 1 μM. Following incubation, cells were collected for fluorescence intensity measurements by flow cytometry in the presence of 2 μg/ml propidium-iodide for identification of dead cells. Samples were measured on a FACSCalibur or FACSCanto instrument (BD Biosciences*)*, data collection and analysis were performed using the BD FACSDiva Software (version 6.1.3) from BD Biosciences or the WinMDI software (developed by Joe Trotter).

2.7. Statistical analysis

For all the experiments, the presented representative measurements show the mean values ± standard deviations of technical replicates. For statistical analysis, a two-sample Student's t-test was performed on the dataset of at least 3 independent biological replicates, and $p<0.05$ values were considered as statistically significant.

3. Results

3.1. Expression patterns of ABCG2 exon 1 variants in drug resistant cells

Previous studies demonstrated that exposure to cytotoxic agents can dramatically alter the expression and function of ABCG2, potentially leading to a multidrug resistance phenotype. We addressed the question how these treatments influence the expression pattern of the ABCG2 mRNA variants having different 5' noncoding sequences. We examined the MCF-7 human breast adenocarcinoma cell line and its Mx or Flv resistant sublines. MCF-7 is used in several studies and shows a medium level of ABCG2 expression [31], whereas the resistant sublines were established by drug selection [6, 28]. We used a TaqMan® assay designed for the 3' mRNA region that detects

ACCEPTED MANUSCRIPT

the entire pool of 5' UTR variants (denoted as "total" mRNA level). We could reproduce previous results in MCF-7 cells showing elevated expression of ABCG2 in drug resistant cells, as the total mRNA levels were increased by >600 fold in MCF-7/Mx and by >700 fold in MCF-7/Flv cells (Fig. 1B).

Next we determined the expression of the ABCG2 mRNA isoforms containing the four exon 1 variants by qRT-PCR using specifically designed TaqMan® assays (see Materials and methods). The nomenclature of these exon 1 variants has not been standardized yet, therefore we included a table with the names and genomic positions of the sequences represented in earlier studies (Table 2), and proposed a uniform nomenclature of the four investigated sequences as E1A, E1B, E1C and E1U (Fig. 1A). Exons E1U and E1A both have distinct transcription start sites and 3' exon ends (splice donor sites), whereas E1B and E1C share the third transcription start site but differ in their 3' exon ends and therefore in their length due to alternative splice donor sites (see Fig. 1A). Contrary to previous studies, we were able to design real-time PCR assays to distinctly detect variants E1B and E1C, although it is difficult to quantitate their individual expression levels due to strong sequence overlaps.

Variants E1A, E1B and E1C were present in all samples examined, and their expression levels increased massively as a result of drug treatment (Fig. 1C), which is in concordance with previous findings in other human cell types [18, 25]. The E1B/C variants seemed to be the most abundant species in the ABCG2 mRNA pool, and they were always present whenever ABCG2 was transcribed. Interestingly, we also detected the E1U variant in MCF-7/Flv but not in MCF-7/Mx cells (Fig. 1C), suggesting that the two drugs may influence distinct signaling pathways and thus differentially activate the promoters in the ABCG2 gene.

3.2. Normal and drug induced expression of ABCG2 mRNA isoforms in embryonic stem cell lines

To investigate the basic ABCG2 expression and the effect of chemotoxic stress on the expression of ABCG2 mRNA isoforms in human embryonic stem cells, we incubated HUES9 cells with 10 nM mitoxantrone for three days and analyzed changes in transcript levels by qRT-PCR. Cells were also analyzed to determine ABCG2 protein level in the plasma membrane by immunostaining and flow cytometry. Using these methods we could confirm the expression of both ABCG2 mRNA and protein in untreated HUES9 cells. Real-time PCR results indicated the presence of E1B/C and E1A isoforms in the mRNA pool, however, the E1U variant could not be detected (Fig. 2A).

We found that the levels of total ABCG2, as well as variant E1B and variant E1C mRNAs were decreased during the 3 days (between two passages) in the control cells, while variant E1A retained its mRNA level, similarly to those of the Oct4 and Nanog major pluripotency markers (Fig. 2B-C). On the other hand, the level of another pluripotency marker, SSEA4, started decreasing in these control cells (Fig. S1D-E), and the expression of variants E1B/C was correlated with this cell surface lipopolysaccharide marker. These data indicate that the HUES9 cells between passages show early signs of differentiation, while still retaining the key pluripotency markers.

Next, we examined the effects of 3 days of mitoxantrone drug treatment in the HUES9 cells, and found that the expression of ABCG2 and all of its mRNA isoforms were strongly decreased after mitoxantrone administration. Variant E1A showed a little delayed response, but also decreased to 20% of the base level by the end of the experiment (Fig. 2D). Changes in protein levels in the plasma membrane also indicated negative regulation of ABCG2 expression (Fig. S1A-C). As flow

ACCEPTED MANUSCRIPT

cytometry measurements ensured the examination of live cells by applying Topro3 staining exclusion (see Fig. 1S), drug-induced differentiation under these conditions is clearly indicated by a decline in Oct4 and Nanog mRNA levels (Fig. 2E), and a reduction in the SSEA4 level (Fig. S1F).

These results are in agreement with previous findings of ABCG2 being expressed in pluripotent cells [18, 22], and further support the hypothesis of induced differentiation as a consequence of exposure to chemotoxic stress. The investigation of other human embryonic stem cell lines (HUES1 and HUES4) also revealed a relatively high level of variant E1A as compared to variants E1B/C (Fig. 2A). These data altogether suggest that variant E1A might play an important role in stem cells, and its expression is uncoupled from isoforms E1B/C, but correlates with the expression of the key pluripotency markers, Oct4 and Nanog.

3.3. Effect of 5'UTR variance on the stability of ABCG2 mRNA isoforms

In the following experiments we investigated whether sequence variability in the 5' UTR may contribute to posttranscriptional regulation of ABCG2 expression through modulating mRNA stability, as in many other cases when different untranslated regions are present in mRNA species [32]. We performed transcription-inhibition via actinomycin D treatment on MCF-7/Flv cells that endogenously express all four isoforms. Although it was demonstrated previously that actinomycin D was not a substrate of ABCG2 and does not interfere with the planned experiment [33], we also performed ATPase assay measurements to confirm the absence of such interactions (data not shown). Steady-state levels of mRNA isoforms were determined by qPCR at various time points after treatment, and the relative expression values for each isoform were normalized to that in control cells. Decay curves for each variant, as well as for the "total" ABCG2 mRNA pool are shown on Fig. 3. The average half-lives of the isoforms did not significantly differ from one another and they were estimated to be around 11 hours (Fig. 3). Although variant E1U and E1A had a tendency to show the lowest levels at each time point, the results indicated no significant stability differences among the 5' UTR isoforms of ABCG2 mRNA.

3.4. Effects of mRNA 5' UTR variability on ABCG2 protein expression

In order to investigate whether exon 1 variability has any detectable effect on overall protein production, we transfected HEK293 cells (with very low endogenous ABCG2) with the E1x-ABCG2/GFP expression plasmids described above. The presence of an eGFP expression cassette on a distinct transcription unit provided the basis to control for transfection efficiency and protein expression (Fig. 4A). The average transfection efficiency was around 80%. The relative amount of cell surface ABCG2 protein in each transfected population was determined by immunostaining with a conformation sensitive antibody ("5D3"), whereas functionality was tested by mitoxantrone efflux assay by flow cytometry. As shown on Fig. 4B-D, we found no significant differences in surface protein expression or in functional transport activity between isoforms containing exon 1A, exon 1B and exon 1C. However, in the presence of exon 1U lower ABCG2 signal was detected in the plasma membrane (Fig. 4E), suggesting lower protein expression from this mRNA variant. On the other hand, no major differences were detected in the transport activities measured by mitoxantrone extrusion.

In order to examine regulation at the RNA level, ABCG2 mRNA variants were quantified by qRT-PCR in the transfected cells. In the case of the E1U sequence variant, we detected a lower

ACCEPTED MANUSCRIPT

mRNA steady state level (Fig. 4G), which was approximately 70% of that measured for variants E1A/B/C, and which was in line with the immunofluorescence results. To examine protein expression levels, we performed Western blot analysis on cell lysates obtained from transfected populations. As indicated on Fig. 4H, mRNA isoform with E1U showed a clearly weaker protein expression, showing 50% reduction as compared to variants E1A and E1B. In addition, protein expression from variant E1C showed a slightly higher steady state level but it was not statistically significant. Expression levels were normalized to GFP signal, also excluding the variability caused by transfection efficiency differences. Taking together, these data indicated lower mRNA level and lower translational activity from the ABCG2 mRNA isoform containing exon 1U.

3.5. Possible upstream alternative translation regulations: the presence of uORF or alternative translation initiation?

After detecting variant specific differences, namely the lower translation efficiency of E1U and the stem cell specific expression of E1A, we addressed the question whether any sequence motif is present in the mRNA, indicating a mechanism that could influence translation. In the case of E1U, *in silico* analyses indicated an upstream open reading frame (uORF) beginning with an AUG codon 245 base pairs upstream of the 3' end of the exon. Based on literature data [34], we hypothesized that this uORF might play a role in the down-regulation of translation from E1U containing ABCG2 mRNA. To investigate this question we created an expression plasmid containing the E1U isoform in which the upstream AUG was changed to GGG to abolish the putative translation initiation activity. Next, we compared protein expression in cells transfected with either the wild type or the mutagenized isoform.

As shown on Fig. 4E-F and 4H-I (E1U versus E1Umut), the disruption of the upstream AUG did not result in a detectable increase in total protein amount measured by Western blot, and the more sensitive immunofluorescence method indicated a small but not significant increase in surface ABCG2 signal. In conclusion, we found that the uORF present in exon 1U cannot be responsible for the inhibition of the main translation initiation codon by itself.

Next, we looked for sequential elements in variant E1A, and found a potential in frame GUG codon 90 bp upstream from the canonical AUG codon in exon 2. According to the literature [35, 36], such upstream GUG codons can initiate translation under certain circumstances, resulting in a longer N-terminus, which might contain peptide signals potentially modifying the function and/or the localization of the protein. To test if this alternative initiation is functional, we created plasmid constructs expressing the wild type E1A mRNA, or mutants where either the canonical AUG codon or both the AUG and GUG codons were changed to GGG. ABCG2 expression and function was measured 48 hours after transfection by immunostaining and mitoxantrone efflux assay. Our expectation was that if the upstream GUG codon exhibits translational initiation activity then the disruption of the canonical AUG will not result in a total loss of protein expression. Results showed that mutation in the AUG codon caused a strong decrease in ABCG2 surface expression and mitoxantrone transport activity, and no further decline was detected in cells transfected with the double mutant plasmid (data not shown). This finding led to the conclusion that the GUG codon present in exon 1A does not function as an alternative translation initiation codon in HEK293 cells.

4. Discussion

ACCEPTED MANUSCRIPT

In the present study we aimed to expand the understanding of the role of alternative promoter utilization and the function of ABCG2 5' UTR isoforms in regulating gene expression. This phenomenon was investigated by several groups in recent years, both in the case of human and mice ABCG2, and led to the identification of at least two exon 1 isoforms and their corresponding promoter regions besides the one described initially [23]. All exon 1 sequences contribute to the 5' UTR of the mRNA, and are spliced to a common exon 2 where the translation initiation codon is localized. The fact that such variance in the 5' UTR region is conserved between the mouse and the human ABCG2 suggests an evolutionary conserved role, yet little is known about the potential function(s) of these splice variants. Previous studies mainly focused on transcriptional regulation by investigating transcription factor binding sites in promoters [37-42], and little attention was paid to the mRNA isoforms produced by such alternative transcription events.

In order to investigate these questions, we designed real-time PCR (TaqMan®) probes that can detect the main four variants independently of each other by specifically binding to the unique exon 1/exon 2 junction sites, and determined the ratio of these mRNA species in ABCG2 positive cell lines. Variants E1B and E1C can be considered as constitutive variants, and they could always be detected when the ABCG2 gene is expressed. Variants E1A and E1U, however, are more subjects to specific regulation: we could confirm that E1A is expressed in stem cells and could strongly be induced in drug selected cell lines (both in MCF-7/Mx and MCF-7/Flv, see Fig. 1). These findings were in line with earlier observations on stem cells [18, 38], and in agreement with the results of Nakanishi et al. [25] describing the expression of variant E1A in the parental MCF-7, and in a higher level, in the MCF-7/AdrVp selected cell line. In previous studies, however, variant E1U could not be detected even in drug selected cell lines [25], and was found to be expressed only in cells of hematopoietic origin, namely in dendritic cells [18], the CHRF megakaryoblast cell line [18] and in a subtype of pediatric acute megakaryoblastic leukemia [27]. Here we show that MCF-7 cells treated with certain drugs (such as flavopiridol) also express the E1U variant (Fig. 1), presumably due to the modulation of specific pathways by the drug, and affecting different promoter regions, as suggested by our previous studies [38].

We also examined the effect of drug treatment (Mx exposure) in a human embryonic stem cell line, HUES9. ABCG2 expression has been shown to be responsible for the generation of side population phenotype in many stem cell types [15] and it also contributes to resistance against different environmental stresses [22]. We could show that variants E1A/E1B/E1C but not E1U are expressed in untreated HUES9 cells, and the expression of all variants strongly decreased in Mx-induced stem cell differentiation. However, variant E1A seems to have specific regulation in stem cells which is distinct from the other variants: its expression is coupled to the Oct4 and Nanog pluripotency markers. In our previous work we demonstrated that when stem cells enter an early stage of differentiation, a decrease in the levels of total ABCG2 and SSEA4 can be detected [18, 22]. In the present experiments, however, we found that ABCG2 mRNA isoforms are regulated differently, and an upheld expression of the E1A variant could be responsible for the residual ABCG2 expression in the early differentiated cells. The distinct, stem cell specific presence of E1A is further supported by the high expression of this variant relative to E1B and E1C in other human embryonic stem cell lines, HUES1 and HUES4 (Fig. 2).

In order to detect potential functional differences among the variants, we first examined whether there is a difference in their mRNA stability, providing potential regulatory role for controlling the amount of the functional transporter level. The MCF-7/Flv cell line seemed appropriate for such experiments as it expresses all four variants, which can be measured

ACCEPTED MANUSCRIPT

simultaneously. In this model system, however, we could not see significant mRNA stability differences among the variants, although there seemed to be a tendency of the E1U variant (and to some extent, of the E1A variant) to have lower steady state level than the others (Fig. 3). This indication was further supported when the variants were separately expressed in HEK293 cells, where E1U containing mRNA levels were significantly lower than those of the other variants (Fig. 4G).

When investigating ABCG2 protein abundance and plasma membrane functionality, an almost 50% reduction in ABCG2 protein level was detected when expression was driven from the E1U variant, indicating the presence of another translation regulatory mechanism. In a GFP fusion construct, a similar destabilizing role of E1U was suggested earlier by Campbell et al. [27]. Interestingly, sequence analysis pointed to the presence of an uORF in the E1U sequence, however, mutating its potential start codon resulted only in a minor increase in protein expression (Fig. 4H-I), suggesting that this motif by itself cannot be responsible for the lower protein abundance. Another possibility is that sequence motifs that increase the level of translation are present in the other mRNA variants. Therefore, we analyzed the other exon 1 sequences but did not find any indication for positively regulating signals in variants E1B and E1C. A putative alternative translation initiation sequence was found upstream in variant E1A, being a potential explanation for the emphasized role of this isoforms in stem cells. However, mutagenesis studies ruled out this possibility, with the question remaining what mechanisms could stand behind the translation efficiency differences among the ABCG2 mRNA variants.

In conclusion, our work presents a systematic comparison of the role and functionality of the four ABCG2 mRNA isoforms differing in their exon 1 sequence. We demonstrate that beyond the two constitutive variants E1B and E1C, drug selection can also induce the expression of the other two mRNA isoform. Furthermore, we provide evidence for the stem cell specific expression of variant E1A, and demonstrate for the first time that the most recently described mRNA isoform of ABCG2, denoted as E1U, exhibits lower translational activity than the other three variants, and this down-regulation could partly be due to the presence of an uORF motif in the E1U core region. The results suggests that the 5’ UTR variants play an important role in fine tuning of ABCG2 expression, however, the exact mechanism behind the lower translation efficiency of E1U, as well as the significance of the E1A variant still await further investigations.

Acknowledgement

The authors would like to thank Kornélia Némethy for excellent technical assistance, Dr. Susan Bates for providing the MCF-7 sublines, and Dr. Douglas Melton for originally providing the HUES cell lines. Tamás I. Orbán is a recipient of the János Bolyai Scholarship of the Hungarian Academy of Sciences. This work was supported by the Hungarian Scientific Research Fund (OTKA) grants number K112112 and K115375.

References

[1] R. Allikmets, L.M. Schriml, A. Hutchinson, V. Romano-Spica, M. Dean, A human placenta-specific ATP-binding cassette gene (ABCP) on chromosome 4q22 that is involved in multidrug resistance, Cancer Res, 58 (1998) 5337-5339.

[2] L.A. Doyle, W. Yang, L.V. Abruzzo, T. Krogmann, Y. Gao, A.K. Rishi, D.D. Ross, A multidrug

ACCEPTED MANUSCRIPT

resistance transporter from human MCF-7 breast cancer cells, Proc Natl Acad Sci U S A, 95 (1998) 15665-15670.

[3] K. Miyake, L. Mickley, T. Litman, Z. Zhan, R. Robey, B. Cristensen, M. Brangi, L. Greenberger, M. Dean, T. Fojo, S.E. Bates, Molecular cloning of cDNAs which are highly overexpressed in mitoxantrone-resistant cells: demonstration of homology to ABC transport genes, Cancer Res, 59 (1999) 8-13.

[4] H. Lage, M. Dietel, Effect of the breast-cancer resistance protein on atypical multidrug resistance, Lancet Oncol, 1 (2000) 169-175.

[5] T. Litman, M. Brangi, E. Hudson, P. Fetsch, A. Abati, D.D. Ross, K. Miyake, J.H. Resau, S.E. Bates, The multidrug-resistant phenotype associated with overexpression of the new ABC half-transporter, MXR (ABCG2), J Cell Sci, 113 (Pt 11) (2000) 2011-2021.

[6] R.W. Robey, W.Y. Medina-Perez, K. Nishiyama, T. Lahusen, K. Miyake, T. Litman, A.M. Senderowicz, D.D. Ross, S.E. Bates, Overexpression of the ATP-binding cassette half-transporter, ABCG2 (Mxr/BCrp/ABCP1), in flavopiridol-resistant human breast cancer cells, Clin Cancer Res, 7 (2001) 145-152.

[7] O. Polgar, R.W. Robey, S.E. Bates, ABCG2: structure, function and role in drug response, Expert Opin Drug Metab Toxicol, 4 (2008) 1-15.

[8] B. Sarkadi, L. Homolya, G. Szakacs, A. Varadi, Human multidrug resistance ABCB and ABCG transporters: participation in a chemoimmunity defense system, Physiol Rev, 86 (2006) 1179-1236.

[9] P. Krishnamurthy, J.D. Schuetz, Role of ABCG2/BCRP in biology and medicine, Annu Rev Pharmacol Toxicol, 46 (2006) 381-410.

[10] J.W. Jonker, M. Buitelaar, E. Wagenaar, M.A. Van Der Valk, G.L. Scheffer, R.J. Scheper, T. Plosch, F. Kuipers, R.P. Elferink, H. Rosing, J.H. Beijnen, A.H. Schinkel, The breast cancer resistance protein protects against a major chlorophyll-derived dietary phototoxin and protoporphyria, Proc Natl Acad Sci U S A, 99 (2002) 15649-15654.

[11] J. Tallkvist, Y. Yagdiran, L. Danielsson, A. Oskarsson, A model of secreting murine mammary epithelial HC11 cells comprising endogenous Bcrp/Abcg2 expression and function, Cell biology and toxicology, 31 (2015) 111-120.

[12] C.J. Ramirez, J.D. Minch, J.M. Gay, S.M. Lahmers, D.J. Guerra, G.J. Haldorson, T. Schneider, K.L. Mealey, Molecular genetic basis for fluoroquinolone-induced retinal degeneration in cats, Pharmacogenet Genomics, 21 (2011) 66-75.

[13] H. Matsuo, T. Takada, K. Ichida, T. Nakamura, A. Nakayama, Y. Ikebuchi, K. Ito, Y. Kusanagi, T. Chiba, S. Tadokoro, Y. Takada, Y. Oikawa, H. Inoue, K. Suzuki, R. Okada, J. Nishiyama, H. Domoto, S. Watanabe, M. Fujita, Y. Morimoto, M. Naito, K. Nishio, A. Hishida, K. Wakai, Y. Asai, K. Niwa, K. Kamakura, S. Nonoyama, Y. Sakurai, T. Hosoya, Y. Kanai, H. Suzuki, N. Hamajima, N. Shinomiya, Common defects of ABCG2, a high-capacity urate exporter, cause gout: a function-based genetic analysis in a Japanese population, Sci Transl Med, 1 (2009) 5ra11.

[14] J. Susanto, Y.H. Lin, Y.N. Chen, C.R. Shen, Y.T. Yan, S.T. Tsai, C.H. Chen, C.N. Shen, Porphyrin homeostasis maintained by ABCG2 regulates self-renewal of embryonic stem cells, PLoS ONE, 3 (2008) e4023.

[15] S. Zhou, J.D. Schuetz, K.D. Bunting, A.M. Colapietro, J. Sampath, J.J. Morris, I. Lagutina, G.C. Grosveld, M. Osawa, H. Nakauchi, B.P. Sorrentino, The ABC transporter Bcrp1/ABCG2 is expressed in a wide variety of stem cells and is a molecular determinant of the side-population phenotype, Nat Med, 7 (2001) 1028-1034.

[16] B. Bhattacharya, T. Miura, R. Brandenberger, J. Mejido, Y. Luo, A.X. Yang, B.H. Joshi, I.

ACCEPTED MANUSCRIPT

Ginis, R.S. Thies, M. Amit, I. Lyons, B.G. Condie, J. Itskovitz-Eldor, M.S. Rao, R.K. Puri, Gene expression in human embryonic stem cell lines: unique molecular signature, Blood, 103 (2004) 2956-2964.

[17] I. Ginis, Y. Luo, T. Miura, S. Thies, R. Brandenberger, S. Gerecht-Nir, M. Amit, A. Hoke, M.K. Carpenter, J. Itskovitz-Eldor, M.S. Rao, Differences between human and mouse embryonic stem cells, Dev Biol, 269 (2004) 360-380.

[18] A. Apati, T.I. Orban, N. Varga, A. Nemeth, A. Schamberger, V. Krizsik, B. Erdelyi-Belle, L. Homolya, G. Varady, R. Padanyi, E. Karaszi, E.W. Kemna, K. Nemet, B. Sarkadi, High level functional expression of the ABCG2 multidrug transporter in undifferentiated human embryonic stem cells, Biochim Biophys Acta, 1778 (2008) 2700-2709.

[19] X.W. Ding, J.H. Wu, C.P. Jiang, ABCG2: A potential marker of stem cells and novel target in stem cell and cancer therapy, Life Sci, 86 (2010) 631-637.

[20] H. Zeng, J.W. Park, M. Guo, G. Lin, L. Crandall, T. Compton, X. Wang, X.J. Li, F.P. Chen, R.H. Xu, Lack of ABCG2 Expression and Side Population Properties in Human Pluripotent Stem Cells, Stem Cells, 27 (2009) 2435-2445.

[21] B. Sarkadi, T.I. Orban, G. Szakacs, G. Varady, A. Schamberger, Z. Erdei, K. Szebenyi, L. Homolya, A. Apati, Evaluation of ABCG2 expression in human embryonic stem cells: crossing the same river twice?, Stem Cells, 28 (2010) 174-176.

[22] Z. Erdei, B. Sarkadi, A. Brozik, K. Szebenyi, G. Varady, V. Mako, A. Pentek, T.I. Orban, A. Apati, Dynamic ABCG2 expression in human embryonic stem cells provides the basis for stress response, European biophysics journal : EBJ, 42 (2013) 169-179.

[23] K.J. Bailey-Dell, B. Hassel, L.A. Doyle, D.D. Ross, Promoter characterization and genomic organization of the human breast cancer resistance protein (ATP-binding cassette transporter G2) gene, Biochim Biophys Acta, 1520 (2001) 234-241.

[24] Y. Zong, S. Zhou, S. Fatima, B.P. Sorrentino, Expression of mouse Abcg2 mRNA during hematopoiesis is regulated by alternative use of multiple leader exons and promoters, J Biol Chem, 281 (2006) 29625-29632.

[25] T. Nakanishi, K.J. Bailey-Dell, B.A. Hassel, K. Shiozawa, D.M. Sullivan, J. Turner, D.D. Ross, Novel 5' untranslated region variants of BCRP mRNA are differentially expressed in drug-selected cancer cells and in normal human tissues: implications for drug resistance, tissue-specific expression, and alternative promoter usage, Cancer Res, 66 (2006) 5007-5011.

[26] B. Poonkuzhali, J. Lamba, S. Strom, A. Sparreboom, K. Thummel, P. Watkins, E. Schuetz, Association of breast cancer resistance protein/ABCG2 phenotypes and novel promoter and intron 1 single nucleotide polymorphisms, Drug Metab Dispos, 36 (2008) 780-795.

[27] P.K. Campbell, Y. Zong, S. Yang, S. Zhou, J.E. Rubnitz, B.P. Sorrentino, Identification of a novel, tissue-specific ABCG2 promoter expressed in pediatric acute megakaryoblastic leukemia, Leuk Res, 35 (2011) 1321-1329.

[28] R.W. Robey, Y. Honjo, A. van de Laar, K. Miyake, J.T. Regis, T. Litman, S.E. Bates, A functional assay for detection of the mitoxantrone resistance protein, MXR (ABCG2), Biochim Biophys Acta, 1512 (2001) 171-182.

[29] O. Kolacsek, V. Krizsik, A. Schamberger, Z. Erdei, A. Apati, G. Varady, L. Mates, Z. Izsvak, Z. Ivics, B. Sarkadi, T.I. Orban, Reliable transgene-independent method for determining Sleeping Beauty transposon copy numbers, Mob DNA, 2 (2011) 5.

[30] C. Ozvegy-Laczka, G. Varady, G. Koblos, O. Ujhelly, J. Cervenak, J.D. Schuetz, B.P. Sorrentino, G.J. Koomen, A. Varadi, K. Nemet, B. Sarkadi, Function-dependent conformational

ACCEPTED MANUSCRIPT

changes of the ABCG2 multidrug transporter modify its interaction with a monoclonal antibody on the cell surface, J Biol Chem, 280 (2005) 4219-4227.
[31] J.F. Deeken, R.W. Robey, S. Shukla, K. Steadman, A.R. Chakraborty, B. Poonkuzhali, E.G. Schuetz, S. Holbeck, S.V. Ambudkar, S.E. Bates, Identification of compounds that correlate with ABCG2 transporter function in the National Cancer Institute Anticancer Drug Screen, Mol Pharmacol, 76 (2009) 946-956.
[32] T.A. Hughes, Regulation of gene expression by alternative untranslated regions, Trends Genet, 22 (2006) 119-122.
[33] C.R. Hill, D. Jamieson, H.D. Thomas, C.D. Brown, A.V. Boddy, G.J. Veal, Characterisation of the roles of ABCB1, ABCC1, ABCC2 and ABCG2 in the transport and pharmacokinetics of actinomycin D in vitro and in vivo, Biochem Pharmacol, 85 (2013) 29-37.
[34] K. Wethmar, The regulatory potential of upstream open reading frames in eukaryotic gene expression, Wiley interdisciplinary reviews. RNA, 5 (2014) 765-778.
[35] J.L. Wegrzyn, T.M. Drudge, F. Valafar, V. Hook, Bioinformatic analyses of mammalian 5'-UTR sequence properties of mRNAs predicts alternative translation initiation sites, BMC Bioinformatics, 9 (2008) 232.
[36] I.P. Ivanov, A.E. Firth, A.M. Michel, J.F. Atkins, P.V. Baranov, Identification of evolutionarily conserved non-AUG-initiated N-terminal extensions in human coding sequences, Nucleic Acids Res, 39 (2011) 4220-4234.
[37] R.R. Singh, K. Kunkalla, C. Qu, E. Schlette, S.S. Neelapu, F. Samaniego, F. Vega, ABCG2 is a direct transcriptional target of hedgehog signaling and involved in stroma-induced drug tolerance in diffuse large B-cell lymphoma, Oncogene, 30 (2011) 4874-4886.
[38] H. de Boussac, T.I. Orban, G. Varady, B. Tihanyi, C. Bacquet, A. Brozik, A. Varadi, B. Sarkadi, T. Aranyi, Stimulus-induced expression of the ABCG2 multidrug transporter in HepG2 hepatocarcinoma model cells involves the ERK1/2 cascade and alternative promoters, Biochem Biophys Res Commun, 426 (2012) 172-176.
[39] S. Hamada, K. Satoh, M. Hirota, A. Kanno, J. Umino, H. Ito, A. Masamune, K. Kikuta, K. Kume, T. Shimosegawa, The homeobox gene MSX2 determines chemosensitivity of pancreatic cancer cells via the regulation of transporter gene ABCG2, J Cell Physiol, 227 (2012) 729-738.
[40] K.K. To, L. Yu, S. Liu, J. Fu, C.H. Cho, Constitutive AhR activation leads to concomitant ABCG2-mediated multidrug resistance in cisplatin-resistant esophageal carcinoma cells, Molecular carcinogenesis, 51 (2012) 449-464.
[41] W.J. Yang, M.J. Song, E.Y. Park, J.J. Lee, J.H. Park, K. Park, J.H. Park, H.P. Kim, Transcription factors Sp1 and Sp3 regulate expression of human ABCG2 gene and chemoresistance phenotype, Molecules and cells, 36 (2013) 368-375.
[42] Y. Xie, T. Nakanishi, K. Natarajan, L. Safren, A.W. Hamburger, A. Hussain, D.D. Ross, Functional cyclic AMP response element in the breast cancer resistance protein (BCRP/ABCG2) promoter modulates epidermal growth factor receptor pathway- or androgen withdrawal-mediated BCRP/ABCG2 transcription in human cancer cells, Biochim Biophys Acta, 1849 (2015) 317-327.

Figure legends

Figure 1. Genomic location of exon 1 isoforms and their expression in MCF-7 cells
(A) The figure illustrates the relative genomic positions of the four exon 1 variants investigated in the present study. Variant B and C are splice variants transcribed from the first described initiation

ACCEPTED MANUSCRIPT

site of the ABCG2 gene. Variant A is located less than 1 kb upstream of E1B/C and is likely to share some regulation elements in its promoter region. Variant U is located 72 kilobases upstream of E1B/C likely having a unique and independently regulated promoter region. Small arrows indicate the positions of the specific forward primers and the common reverse primer, whereas the location of the TaqMan® probes are not shown (binding to the unique exon 1/exon 2 junctions, see Materials and methods). (B-C) qRT-PCR with TaqMan® method were used to determine expression of ABCG2 total mRNA and its isoforms in parental and in drug resistant MCF-7 cells. The probes were designed to recognize the unique exon 1/exon 2 junction of the mRNA isoforms while overall ABCG2 expression was measured by a pre-designed assay. mRNA levels were normalized to endogenous control genes (RPLP0 and PolR2A). One representative measurement is illustrated and error bars represent standard deviation of technical replicates. (B) Expression of ABCG2 was greatly enhanced in mitoxantrone and flavopiridol resistant derivatives of MCF-7 cells with a fold change of 600 – 700x. (C) Expression of variants E1A/B/C followed a similar fold change as total ABCG2, however, variant E1U was only present in MCF-7/Flv cells and was undetectable in parental MCF-7 and MCF-7/Mx cells.

Figure 2. Expression of ABCG2 and its mRNA isoforms in human embryonic cells
Expression patterns of ABCG2 exon 1 mRNA isoforms in different HUES cell lines under normal conditions or exposed to chemotoxic stress were determined by qRT-PCR. (A) qRT-PCR results showed that variant E1A was present in several HUES cell lines in different expression levels, and its ratio relative to variants E1B/C was higher than that found in MCF-7 (see Fig. 1). (B-C) In HUES9 control cells grown in parallel with the Mx treated populations for 3 days, the expression of pluripotency markers Oct4 and Nanog did not change during the period of the experiment. Contrary, the mRNA levels of total ABCG2 and variants E1B/C declined during the three days, while variant E1A retained its expression level. (D-E) In HUES9 cells treated with 10 nM Mx for 3 days, we detected a decrease in the levels of all ABCG2 mRNA species, as well as pluripotency marker genes Oct4 and Nanog, indicating that mitoxantrone induced differentiation of the cells.
Representative measurements are shown, error bars indicate standard deviation of technical replicates. Measurements were carried out on samples from HUES cell lines of indicated passage numbers, assuring proper exclusion of MEF cell contamination (see Materials and methods).

Figure 3. Evaluating possible effects of exon 1 variance on ABCG2 mRNA stability
Decay rates of mRNA variants were measured by qRT-PCR on samples from actinomycin D treated MCF-7/Flv cells. Expression levels were normalized to the non-treated controls at each time point. Curves from one representative experiment are shown for each variant. Error bars are not drawn in order to maintain clarity of the diagram, however, no significant differences were found in the half-lives of mRNA isoforms.

Figure 4. Expression of ABCG2 protein and mRNA isoforms in transfected HEK293 cells
(A) The four exon 1 variants were cloned into expression vectors to produce the four ABCG2 mRNA isoforms investigated in the study. The plasmids also contained an eGFP coding sequence to serve as control for transfection efficiency. Two CMV promoters were used to drive the expression of ABCG2 and eGFP, respectively, as illustrated on the schematic figure. Exon 1 variants are represented in scale to each other. (B-F) Immunostaining (with 5D3 antibody) and Mx efflux assay in transfected HEK293 cells measured by flow cytometry. Left panels show ratio of ABCG2

ACCEPTED MANUSCRIPT

positive cells in the GFP positive population, indicated by darker gray color. Right panels show overlays of Ko143 treated and untreated samples, and indicate the transport activity of ABCG2 in each sample. Protein levels produced from E1U or E1Umut were significantly lower than from the other variants (see also Fig. S2). (G) Levels of mRNA isoforms relative to GFP were determined by qRT-PCR. The figure shows one representative measurement, error bars indicate standard deviation of technical replicates. (H-I) Expression of ABCG2 protein in transfected cells were determined by Western blot. Amount of loaded proteins were normalized to GFP expression to exclude differences caused by unequal transfection efficiencies. Relative protein amount of the exon 1 isoforms (I) were determined by semi-quantitative densitometry. Error bars represent standard deviations of technical replicates.

*: $p<0.05$ (n=3, two-sample Student's t-test)

ACCEPTED MANUSCRIPT

ACCEPTED MANUSCRIPT

Function	Forward 5' – 3'	Reverse 5' – 3'	TaqMan® Probe 5' – 3'
E1A cloning	*AGCTAGC*GGGCAACCTGTGCGTCAGCG *NheI*	*AGCTAGC*CTGGAGAGTTTTTATCTTTCTCG *NheI*	---
E1A Real-Time PCR	GTGACGGCGACCAAACC	ACATTACTGGAAGACATCTGGAGAGT	TAGGTCAGACGAGAAAGATA
E1B cloning	*AGCTAGC*CCACTGCGTTCAGCTCTGGC *NheI*	*ATCCGGA*CTGGAGAGTTTTTATCTTT *BspEI*	---
E1B Real-Time PCR	GAGAGACGCGGCAAGGA	TGACACTGGGATAAAAACTTCGACAT	AACTGGAAAGATAAAAACTC
E1C cloning	*AGCTAGC*CCACTGCGTTCAGCTCTGGC *NheI*	*ATCCGGA*CTGGAGAGTTTTTATCTTT *BspEI*	---
E1C Real-Time PCR	CCTTTGGTTAAGACCGAGCTCTATT	TGACACTGGGATAAAAACTTCGACAT	AAGCTGAAAAGATAAAAACTC
E1U cloning	*AGCTAGC*GAAGAGGATCCCACGCTGAC *NheI*	*ATCCGGA*CTGGAGAGTTTTTATCTTT *BspEI*	---
E1U Real-Time PCR	GAGATTTGGGCTGCTTTGCTT	ACATTACTGGAAGACATCTGGAGAGT	CACATCATAACTGAGAAAGAT
GFP Real-Time PCR	GAGCGCACCATCTTCTTCAAG	TGTCGCCCTCGAACTTCAC	ACGACGGCAACTACA

ACCEPTED MANUSCRIPT

Table 1. Collection of primers specifically designed for this study
Cloning primers were used to obtain exon 1 sequences from genomic DNA. The amplicons were integrated into expression plasmids using the restriction endonuclease sites represented on the primer sequences (NheI or BspI). The primers used for qRT-PCR were designed to bind the unique exon 1 / exon 2 junction region of each ABCG2 mRNA isoform.

ACCEPTED MANUSCRIPT

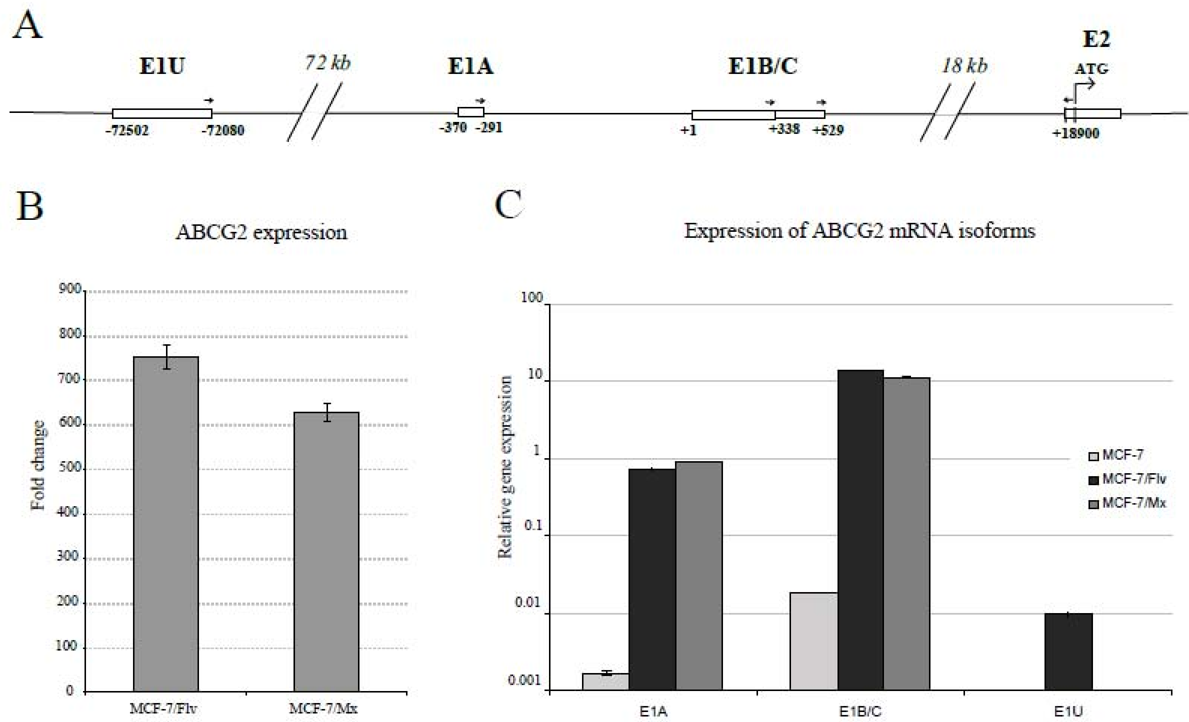
Variant	3’ end position in genome relative to TSS defined by Bailey-Dell et al., 2001 [23]	Exon length in our study (base pairs)	Nakanishi, 2006 [25]	Apati, 2008 [18]	Campbell, 2011 [27]
E1A	- 291	99	E1a	e1a	e1A
E1B	+ 338	338	E1c	-	e1B
E1C	+ 529	529	E1c	e1b	e1C
E1U	- 72080	423	-	e1c	e1U

Table 2. Nomenclature of the known human ABCG2 exon 1 isoforms.
The table contains the four exon 1 isoforms investigated in our study with their corresponding genomic location, sequence length and alternative names found in the literature.

ACCEPTED MANUSCRIPT

ACCEPTED MANUSCRIPT

Fig. 1



ACCEPTED

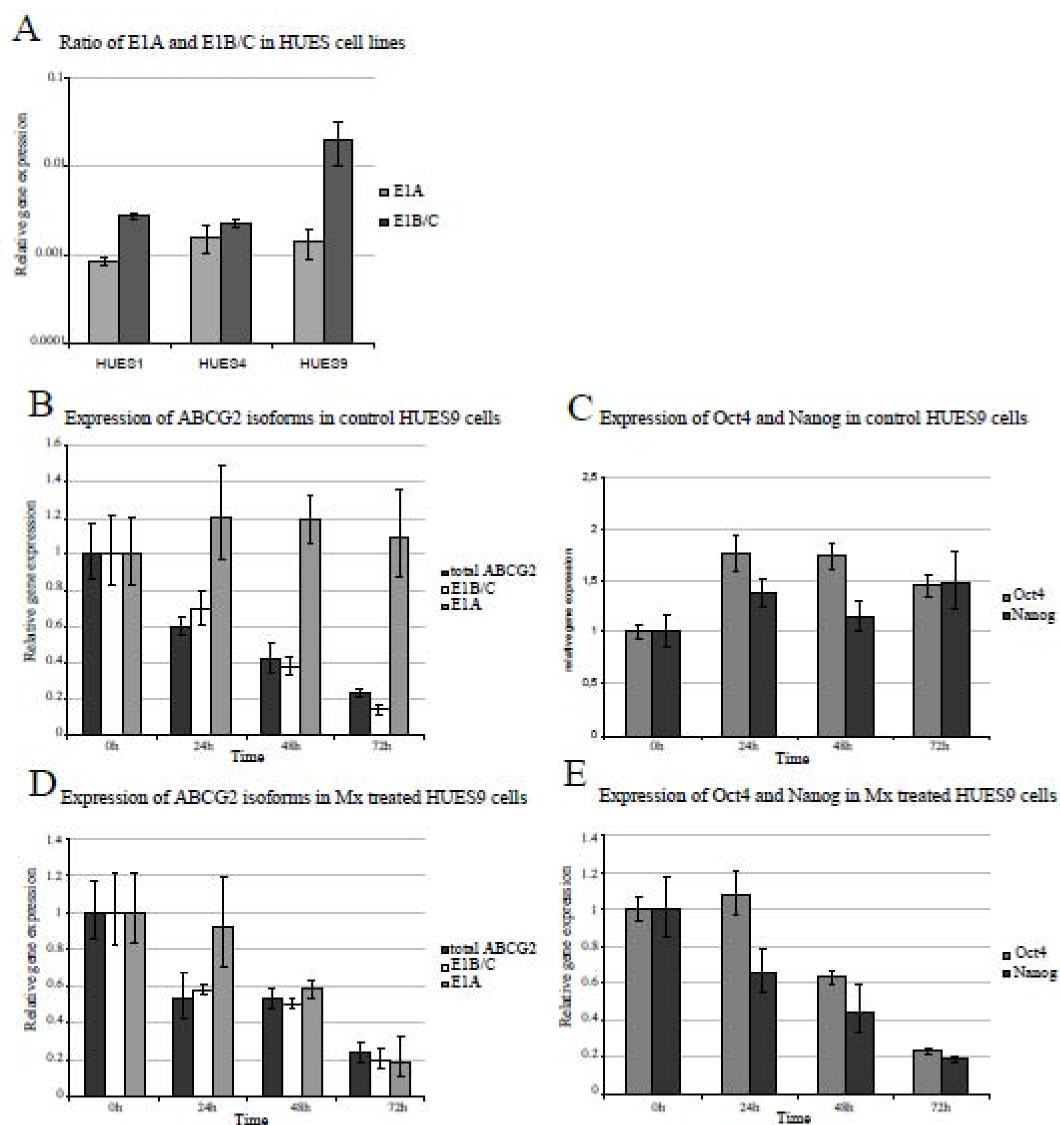


Fig. 2

ACCEPTED MANUSCRIPT

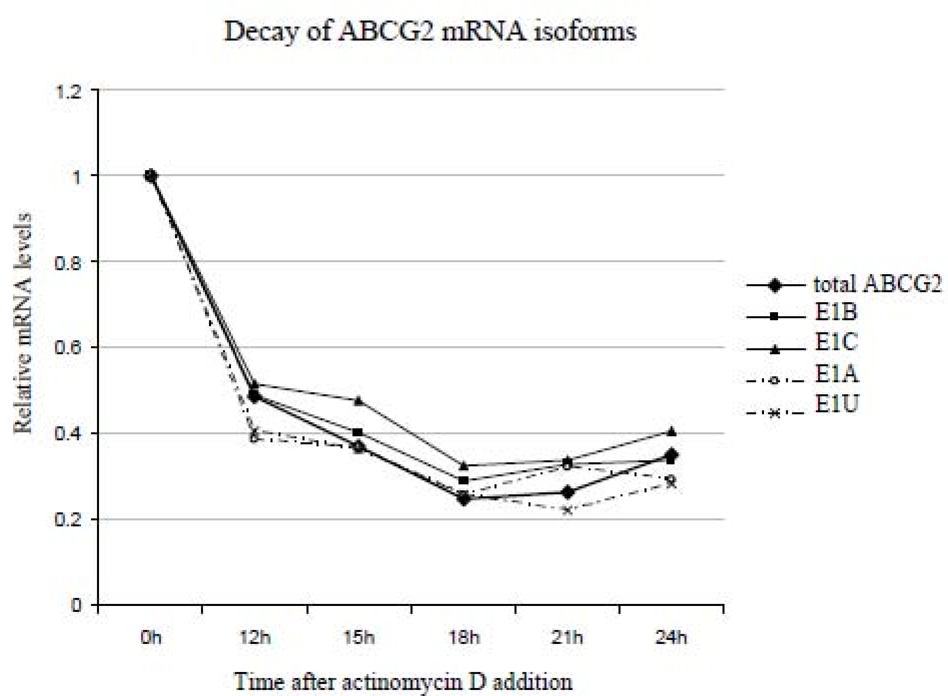


Fig. 3

ACCEPTED

ACCEPTED MANUSCRIPT

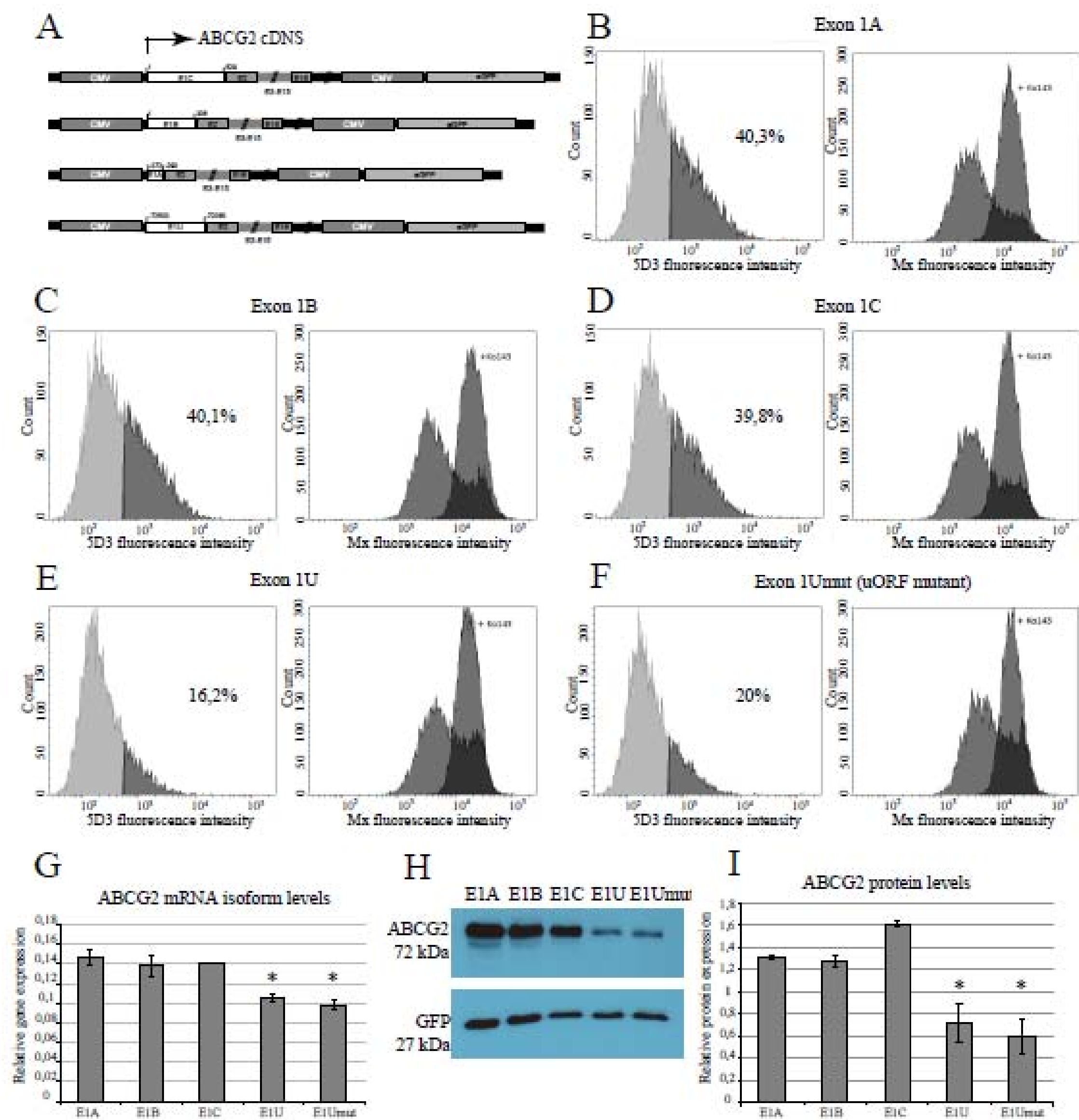
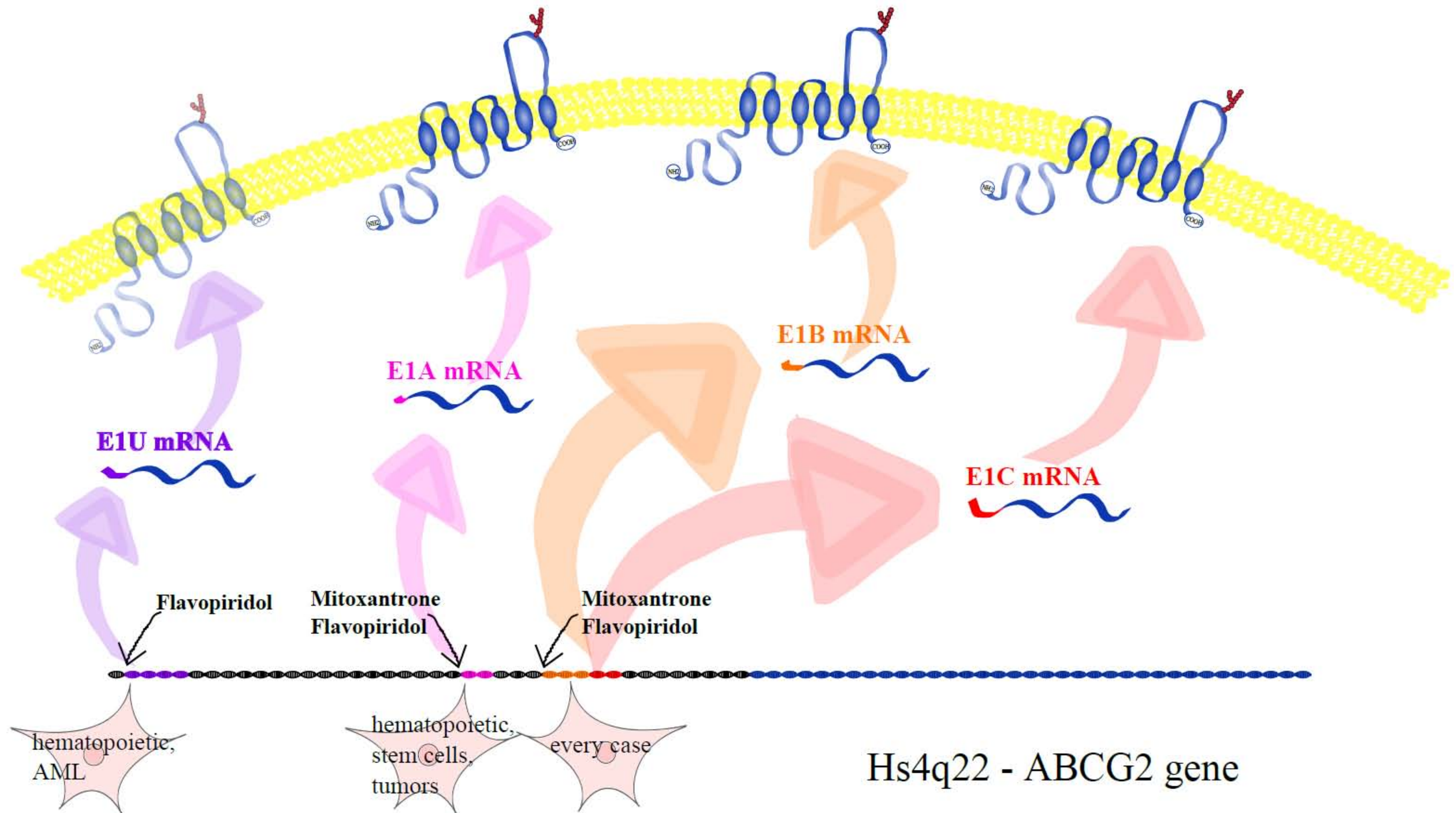


Fig. 4

ACCEPTED MANUSCRIPT

Graphical abstract



ACCEPTED MANUSCRIPT

Highlights

- The human ABCG2 multidrug transporter gene has four exon 1 mRNA isoforms which differ in their 5' UTR sequences.
- E1B and E1C variants are constitutively expressed if ABCG2 is turned on, whereas E1A and E1U show more tissue specific expression patterns.
- Chemotherapeutic drug treatment can alter the expression profile of ABCG2 mRNA isoforms.
- E1A variant is expressed in stem cells and its expression is coupled to the key pluripotency markers Oct4 and Nanog.
- There is no detectable difference among the half-lives of the 4 mRNA variants; however, the translation efficiency of the E1U isoform is significantly lower than those of the other variants.

ACCEPTED MANUSCRIPT