The genetic basis of differences in drug response

Radboud Honours Academy – FNWI programme

May 2017, Nijmegen

Mara Nicolasen  Biology
Quinten Rutgers  Physics & Mathematics
Anne Savenije  Science
Janneke Toorians  Biology

Under guidance of Dr. J.L. Peters
Table of contents

1. Details
   1) Details of proposal ........................................ 3
   2) Details of the applicants ................................... 3
   3) Supervisor .................................................. 3

2. Summary
   1) Scientific summary ......................................... 4
   2) Summary for the broad scientific committee ............... 4
   3) Summary for the general public ............................ 4

3. Description of the proposed research
   1) Introduction ................................................ 5
   2) Thiopurines and metabolites ................................ 5
   3) Important enzymes in the thiopurine pathway .......... 7
   4) Genome-wide association studies ........................ 8
   5) Outline of the research ................................... 10
   6) Cell lines .................................................. 10
   7) siRNA ....................................................... 10
   8) Proteome analysis ......................................... 12
   9) Thiopurine metabolites analysis ......................... 13

4. Timetable of the project ....................................... 15

5. Statements by the applicant .................................. 15

6. References ....................................................... 16
1. **Details**

1a. **Details of proposal**

Title: The genetic basis of differences in drug response

Area: (bio)Chemistry  Physics and Mathematics  Health

1b. **Details of the applicants**

<table>
<thead>
<tr>
<th>Name</th>
<th>Gender</th>
</tr>
</thead>
<tbody>
<tr>
<td>Janneke Toorians</td>
<td>Male</td>
</tr>
<tr>
<td>Anne Savenije</td>
<td>Male</td>
</tr>
<tr>
<td>Quinten Rutgers</td>
<td>Male</td>
</tr>
<tr>
<td>Mara Nicolasen</td>
<td>Male</td>
</tr>
</tbody>
</table>

1c. **Supervisor**

<table>
<thead>
<tr>
<th>Name</th>
<th>Tel</th>
<th>Email</th>
<th>Institute and Department</th>
</tr>
</thead>
<tbody>
<tr>
<td>Janny Peters</td>
<td>+49 1577 1401583</td>
<td><a href="mailto:jl.peters@science.ru.nl">jl.peters@science.ru.nl</a></td>
<td>Molecular Plant Physiology</td>
</tr>
</tbody>
</table>
2. Summary

2a. Scientific summary
This research proposal focuses on the effect that a person’s genes have on the response to the thiopurine drugs. Through a genome wide association study, seventeen SNPs outside the thiopurine pathway have been found to be correlated with an adverse response to a treatment with thiopurines. Being different from the genes identified with the thiopurine pathway, the genes associated with the aforementioned SNPs and their relation to the response to thiopurines are an interesting research topic. The candidate genes will be investigated in an siRNA knockdown *in vitro* study. The response to thiopurines by cells with and without the siRNA knockdown will be compared to each other. This comparison is based on the candidate’s protein and the thiopurine metabolite levels, found with mass spectrometry and high-performance liquid chromatography respectively. The thiopurine metabolites that are evaluated are the three most important ones in the thiopurine pathway, being 6-mercaptopurine (6-MP), 6-thioguanine (6-TG) and 6-methylmercaptopurine (6-MMP).

2b. Summary for the broad scientific committee
In the medical world people are treated for the disease the doctor thinks they have. The patients are seen as a group of people who are having the same symptoms. This approach comes with disadvantages. Every individual has their own genome. This means that every person has another response to a drug and so there are lots of differences in side effects. It would not seem logical to treat people as a group, when the side effects depend on the genome of the patient. One drug that has different responses is the thiopurine drug. This drug is mainly used for patients with Crohn’s disease, where the thiopurines azathioprine (AZA), 6-mercaptopurine (6-MP) are used. It is also a treatment in acute lymphoblastic leukemia (ALL), where the 6-thioguanine (6-TG) is used. The drug has a pathway in which it is metabolised. The metabolites are active and built into the DNA. The whole pathway contributes to the metabolising of the drug. In a GWAS study was found that there are also non-thiopurine pathway SNP’s that have a correlation to the response of the patient to thiopurines. The aim of this proposal is to get to know why the non-thiopurine SNP’s found have a correlation to the response to thiopurines. In our proposal the metabolites of thiopurine are measured, so that the relation of the non-thiopurine pathway gene could be determined. By doing this research, the thiopurine pathway could be expanded. The patient’s response would become more clear. Every step that is made in the thiopurine is a step closer to getting personalised medicine for patients suffering from ALL and Crohn’s disease.

2c. Summary for the general public
Ons onderzoeksvoorstel betreft gepersonaliseerde medicijnen. Op dit moment is het zo dat er een standaard medicijn is voor bepaalde symptomen die iedereen krijgt. Het menselijk genoom is echter niet voor iedereen gelijk en dit zorgt ervoor dat niet iedereen op dezelfde manier reageert op medicijnen. Wij willen gaan onderzoeken wat veranderingen in het genoom doen met de reactie op medicijnen. Als modelmedicijn zijn thiopurines gekozen. Dit zijn medicijnen die worden gebruikt bij acute lymfoblastische leukemie en inflammatoire darmziekten. We hopen inzicht te krijgen in de thiopurine pathway en de verschillende reacties op het medicijn, om zo verschillende patiënten te kunnen helpen.
3. Description of the proposed research

3.1 Introduction

In the medical world, people are treated for the disease the doctor presumes they have. This means that almost every patient that is diagnosed with a certain disease, receives a standard medicine that is representative for that particular patient group. However, what works for one patient, does not necessarily improve another person’s health. Patients can show different responses to a drug. What can account for this difference in response is the genetic base that is different for every person. When a drug enters the body, it is metabolized by different enzymes, mainly in the liver. Some patients have many or efficient working enzymes, some have less or not-optimal working enzymes. This will result in a different metabolism of the drug. In the field of pharmacogenetics, the correlation between the genes and the specific response of a person to a drug is studied. This can eventually be applied to personalized medicine: predicting how someone will respond to a drug based on his or her genotype and using this to determine the appropriate dosage of the drug for that specific person. The consequence of such an approach would be that genomes of patients need to be analyzed before prescribing a drug treatment. In this way patients are not considered as a group with the same disease that can receive the standard drug prescription, but as individuals that benefit from a personalized approach.

One class of drugs that gives different responses in patients, is the Thiopurine drugs. These drugs are mainly used for patients with Crohn’s disease and acute lymphoblastic leukemia (ALL) occurring to 5-6 per 1000 persons in the EU [Burisch],[Population Eurostat] and less than 1 per 1000 persons in the EU [ALL Orphan-Europe], respectively. It took more than ten year to define a pathway describing the way the drug is metabolized. One way to investigate the genetic factors that possibly influence the response of a person to a drug is to perform a genome wide association study (GWAS). Such a research usually not only leads to information about the identification of SNPs that are coupled to genes inside the thiopurine pathway, but also to genes that are outside this pathway. The aim of the research proposed here is to determine whether and how Single Nucleotide Polymorphisms (SNPs) outside the thiopurine pathway, found in a previously performed GWAS, affect a person’s response to thiopurines. With the results of this study, the now considered non-thiopurine pathway genes could perhaps be added to the pathway and explain why patients respond differently to thiopurine drugs. The results could presumably contribute to the realization of personalized medicine.

3.2 Thiopurines and metabolites

Thiopurines are purine antimetabolites that are used in the treatment of autoimmune disorders like Crohn’s disease and rheumatoid arthritis. They are also used in the treatment of ALL. There are different thiopurines that are used. Azathioprine (AZA) and 6-mercaptopurine (6-MP) are widely used in Crohn’s disease (CD) [Wilhelm A J, 2015]. The thiopurine 6-thioguanine (6-TG) is a thiopurine that is used for treating acute lymphoblastic leukemia and chronic myeloid leukemia (PubChem, 2005). AZA and 6-MP are effective drugs for both the induction of remission and the retention of CD. In addition, both agents prevent exacerbation (worsening of a disease or an increase in its symptoms) of CD after surgery and in combination with other immunosuppressants these medicines have been found to be effective in another disease named perianal fistulae [Derijks].

Thiopurines are almost similar in structure to DNA-purines adenine and guanine, the only difference is an S-group (sulfur atom) that is attached to the sixth C-atom of the purine
ring. The thiopurines compete with normal DNA-purines during DNA synthesis or replication. If instead of a purine, a thiopurine is built in, the DNA synthesis stops and apoptosis of the cell is initiated [Derijks, 2005].

Thiopurines are prodrugs that need to be metabolized to get an active immunosuppressive function (Figure 1). AZA is a prodrug and can be metabolized to 6-mercaptopurine (6-MP). This in its turn can be converted into the metabolites 6-thioguanine (6-TG) and 6-methylmercaptopurine (6-MMP). It is not completely known what has the key working in the thiopurine pathway, but it is thought that either 6-MP or thioguanine nucleotides (TGNs) has the pharmacological function. That is because thioguanine is the anhydrous form of thioguanine, a synthetic guanosine analogue antimetabolite, with antineoplastic activity. This means that it acts to prevent or halt the development of a tumor. Thioguanine is phosphorylated by hypoxanthine-guanine phosphoribosyltransferase to 6-thioguananlyc acid and upon conversion to thioguanosine diphosphate and thioguanosine triphosphate, this agent is incorporated into DNA and RNA, resulting in inhibition of DNA and RNA synthesis and cell death. This agent also inhibits glutamine-5-phosphoribosylpyrophosphate amidotransferase, thereby inhibiting purine ribonucleotide synthesis [Derijks, 2005].

6-Mercaptopurine (6-MP) is a "purine antagonist," which means that it inhibits the synthesis and metabolism of purine nucleotides. DNA consists of purines (adenine and guanine) that are essential for the structure of DNA and without these purines, the function and synthesis of DNA is blocked. In cancer cells, functional DNA is necessary for growth and reproduction. 6-MP, being an inhibitor of the purine synthesis and metabolism, slows/stops the growth of cancer cells in the human body. In ALL it results in a decrease of ALL cancer cells. 6-MP can be methylated by thiopurine S-methyltransferase (TPMT) to 6-MMP, another metabolite [Zaza G, 2010].

Liver toxicity appears to be related to the presence of high concentrations of methylated metabolites, particularly 6-MMP. The level of the metabolites can cause different problems in the human system. 6-TG levels between 230 pM and 400 pM can indicate an inflammatory bowel disease, 6-TG levels over 400 pM correlate with bone marrow suppression and levels of 6-MMP above 5700 pM may correlate with liver toxicity [MacDermott R P, 2015]. Because TPMT activity in the liver is similar to the activity in the erythrocytes, this suggests that a high TPMT activity is a determining factor for possible thiopurine-induced liver toxicity. The duration of the exposure to high concentrations of MMP-6 could be an important parameter. Because of the difference in the TPMT activity, there is a different response in the patients that get the medicine thiopurine. If TPMT-activity is too high, the patient can get resistant to the thiopurine. This also works vice versa. When the patient's TPMT-activity is too low, the patient has a higher chance to get side effects. The most important side effects are bone marrow depression: leukopenia, thrombocytopenia, anemia pancreatitis and hepatotoxicity. It is thought that these side effects are correlated with the concentrations of the metabolites 6-TGN and 6-MMP in the blood [PubChem, 2005].
3.3 Important enzymes in the thiopurine pathway

Important enzymes in the thiopurine pathway are thiopurine S-methyltransferase (TPMT), xanthine oxidase (XO), hypoxanthine-guanine phosphoribosyltransferase (HPRT) and inosine triphosphate pyrophosphatase (ITPase) [M Behmanesh, 2009]. Thiopurine methyltransferase catalyzes S-methylation of thiopurines such as azathioprine and 6-mercaptopurine. This methylation is an important metabolic pathway for thiopurine drugs. Thiopurine methyltransferase (TPMT) metabolizes thiopurine drugs and influences their cytotoxicity. TPMT acts like DNA methyltransferases (DNMTs) and it transfers methyl groups from S-adenosylmethionine (SAM) and generates S-adenosylhomocysteine (SAH). Since SAM levels are dependent on de novo purine synthesis (DNPS) and the metabolic products of 6-TG and 6-MP differ in their ability to inhibit DNPS, it is postulated that 6-TG compared to 6-MP would have differential effects on changes in SAM and SAH levels and global DNA methylation, depending on TPMT status. Variations in the genes of these enzymes will lead to different metabolites and that will result in a different response to the drug. Pharmacogenomics focuses on this variation in the genomes of human beings. The aim of pharmacogenomics is to treat individuals the way they would have the best response on. In TPMT for example, different alleles will have a higher chance in developing leukopenia. When the ITP-ase is decreased in its activity, there are also more side effects. Other medicines in the thiopurine metabolism can interact with thiopurines. That is why it is necessary that the dose is not too high [M Behmanesh, 2009].

TMPT genotype

The enzyme activity of TPMT is a major factor for the determination of what the dosage of AZA should be prescribed and 6-MP and therefore also for the determination of 6-TG and 6-MMP. 89 percent of the population has the wild type genotype, which is associated with a ‘high’ enzyme activity. 11% has a heterozygous variant of TPMT, a low-activity enzyme and 0.3 percent of the population is homozygous for mutations in the TPMT gene, and has a negligible activity. This negligible activity causes the 6-MP to be metabolized to high levels of 6-TG, which correlates with bone marrow suppression [Nguyen C M, 2011].
Some individuals are more vulnerable to side effects than others due to genetic polymorphisms of TPMT and this makes dose-adjustments of azathioprine and 6-mercaptopurine riskier [Nguyen C M, 2011].

3.4 Genome-wide association studies (GWAS)

Introduction
In order to achieve the goal of personalized medicine, the genetic factors of variability in drug response have to be found. The first step in this direction is to conduct a genome-wide association study (GWAS). The aim of such a study is to identify single nucleotide polymorphisms (SNPs) that might influence the trait or disease of interest. In a GWAS, the frequency of SNPs is associated with the trait or disease by finding SNPs that occur more in the part of the population that has the trait (figure 2).

![Figure 2: Outline of GWA studies. Source: Genetics and Genomics Wiki.](image)

The disadvantage of GWAS is that it takes fundamentally a whole-genome and thus a non-gene-specific approach. A GWAS on its own is not enough to find the causal relationship between genotype (SNPs) and phenotype (trait or disease). In the research proposed here, the results of previously performed GWAS on thiopurines will be used to identify candidate genes that can then be used for a more detailed gene-specific analysis [Matimba].

Although GWAS is used here as a tool, the limitations of the techniques should also be beared in mind. This section then, is meant to clarify the methodology of GWA studies and list the most important disadvantages and limitations that are relevant for this particular study.

SNPs
Single nucleotide polymorphisms (SNPs) are the basis blocks of genetic variation that are using in GWA studies. SNPs are single-basepair differences in DNA that usually have an uncommon (1-40%) and a common allele (60-99%) [Bush]. Usually, different SNPs that are, for example, physically close on the chromosome, are also linked hereditarily. The result is that the frequencies of the two SNPs are not independent, but are correlated. This is called linkage disequilibrium (LD). LD allows you to use only a subset of all available SNPs, since using multiple SNPs that are correlated would only result in redundant data [Eberle].

A disadvantage of only using a subset is that at first hand it is not certain whether a certain SNP is actually responsible for the trait or if it is only linked with one that is. Once a SNP is found that is correlated with the trait of interest, further studies are needed to determine the functional role of this SNP. This is what the aim of our study is.
**eQTLs**

A major drawback of GWAS is that it can only identify associated SNPs, but cannot directly predict which genes play a role. GWA studies can identify a quantitative trait locus (QTL) that is associated with the trait in question, but this is not necessarily also the locus of directly involved genes, e.g. genes that are in the pathway of the particular medicine of study.

One possibility is that the SNP is associated with a locus that influences the expression of a certain transcript or protein, but that does not directly change its basepair sequence. Such a locus is called an expression quantitative trait locus (eQTL). There exist two types of eQTL. The first, called cis-eQTLs, change the expression of their gene of target locally. Trans-eQTLs change expression of a gene that is separated from it [Rockman].

Previous studies have shown that most complex-trait-associated SNPs are more likely eQTLs (expression QTLs) [Nicolae]. It is therefore necessary to take into account the possibility that any SNP identified by a GWA study is in fact an eQTL. Particularly, the non-thiopurine-pathway genes could be eQTL coupled to genes that are in the pathway.

**Data**

In a genome-wide study performed by Matimba et. al. multiple non-pathway SNPs, i.e. SNPs that are not close to genes that code for enzymes in the thiopurine pathway, were identified that were correlated with TPMT activity and concentration of 6-MP, 6-TG and other thiopurine metabolites. These SNPs are recorded in table 1. This table also records the genes that are nearby these SNPs, what chromosome they are on, and the location of the SNP with respect to these genes. The statistical significance of the SNPs is also recorded [Matimba].

<table>
<thead>
<tr>
<th>Phenotype</th>
<th>SNP ID</th>
<th>Chr</th>
<th>Genes/nearby gene(s)</th>
<th>Location</th>
<th>MAF (Direction) Increase or Decrease in R²</th>
<th>p-value</th>
</tr>
</thead>
<tbody>
<tr>
<td>TPMT activity</td>
<td>rs1040637</td>
<td>6</td>
<td>MLA60P</td>
<td>3' - FR</td>
<td>0.38 (+) -0.013</td>
<td>3.88 x 10^-3</td>
</tr>
<tr>
<td></td>
<td>rs2601149</td>
<td>6</td>
<td>NHE2, AIG1</td>
<td>5' - FR</td>
<td>0.44 (-) -0.028</td>
<td>4.90 x 10^-3</td>
</tr>
<tr>
<td>6-TG, 6-TGN</td>
<td>rs2215790</td>
<td>17</td>
<td>NME1</td>
<td>5' - FR</td>
<td>0.10 (-) -0.048</td>
<td>4.58 x 10^-3</td>
</tr>
<tr>
<td></td>
<td>rs700487</td>
<td>17</td>
<td>NME1, NME2</td>
<td>TFBS</td>
<td>0.10 (-) -0.048</td>
<td>4.58 x 10^-3</td>
</tr>
<tr>
<td></td>
<td>rs1558254</td>
<td>17</td>
<td>NME1, NME2</td>
<td>IntAct</td>
<td>0.10 (-) -0.048</td>
<td>4.58 x 10^-3</td>
</tr>
<tr>
<td>6-MP, 6-TGN</td>
<td>rs777419</td>
<td>8</td>
<td>PAG1</td>
<td>3' - FR</td>
<td>0.49 (-) -0.052</td>
<td>6.39 x 10^-3</td>
</tr>
<tr>
<td></td>
<td>rs10011976</td>
<td>10</td>
<td>KCNU1</td>
<td>IntAct</td>
<td>0.11 (+) +0.038</td>
<td>4.06 x 10^-3</td>
</tr>
<tr>
<td></td>
<td>rs17480656</td>
<td>10</td>
<td>KCNU1</td>
<td>IntAct</td>
<td>0.11 (+) +0.042</td>
<td>2.31 x 10^-3</td>
</tr>
<tr>
<td></td>
<td>rs12765334</td>
<td>10</td>
<td>KCNU1</td>
<td>IntAct</td>
<td>0.11 (+) +0.042</td>
<td>2.31 x 10^-3</td>
</tr>
<tr>
<td></td>
<td>rs17339781</td>
<td>10</td>
<td>KCNU1</td>
<td>IntAct</td>
<td>0.10 (+) +0.042</td>
<td>2.31 x 10^-3</td>
</tr>
<tr>
<td></td>
<td>rs1101997</td>
<td>10</td>
<td>KCNU1</td>
<td>IntAct</td>
<td>0.10 (+) +0.042</td>
<td>2.31 x 10^-3</td>
</tr>
<tr>
<td>6-MP, MeMPN</td>
<td>rs11707715</td>
<td>3</td>
<td>PRADMB</td>
<td>IntAct</td>
<td>0.25 (-) -0.028</td>
<td>9.86 x 10^-3</td>
</tr>
<tr>
<td></td>
<td>rs6540198</td>
<td>3</td>
<td>PRADMB</td>
<td>IntAct</td>
<td>0.25 (-) -0.031</td>
<td>3.03 x 10^-3</td>
</tr>
<tr>
<td></td>
<td>rs10473594</td>
<td>5</td>
<td>CCHL2</td>
<td>IntAct</td>
<td>0.22 (-) -0.036</td>
<td>3.62 x 10^-3</td>
</tr>
<tr>
<td></td>
<td>rs12886319</td>
<td>14</td>
<td>GNG2</td>
<td>IntAct</td>
<td>0.15 (-) -0.009</td>
<td>4.21 x 10^-3</td>
</tr>
<tr>
<td>6-MP, MeMPN-6-TGN</td>
<td>rs28683168</td>
<td>14</td>
<td>GNG2</td>
<td>IntAct</td>
<td>0.15 (-) -0.004</td>
<td>3.62 x 10^-3</td>
</tr>
<tr>
<td></td>
<td>rs5748488</td>
<td>20</td>
<td>SLC2A1</td>
<td>Folding</td>
<td>0.31 (+) +0.036</td>
<td>1.89 x 10^-3</td>
</tr>
</tbody>
</table>

Table 1. Top ‘non-thiopurine pathway’ SNPs associated with clinical phenotypes in UK ALL97 patients: TPMT activity and thiopurine metabolites. Source: [Matimba].

The genes that are associated with these correlated SNPs will be used as candidate genes for a more detailed gene-specific study to determine the functional relation between these SNPs/genes and drug response.
3.5 Outline of the research

The research question reads: how do the non-thiopurine pathway SNPs found in previously performed GWAS influence the response to and metabolism of thiopurine drugs?

To reach this goal, a few obstacles need to be overcome. First, to simulate the human drug response, an appropriate model should be chosen. For this research, several cell lines will be used, as will be explained in the next section. The correlation between SNPs/genes and the response to thiopurines will be studied under in vitro conditions using small interference RNA (siRNA). The effects of the siRNA on the cells will be investigated with a proteome and metabolite analysis. This is to determine the activity of the thiopurine pathway and where in the pathway the alterations take place per associated gene.

Methods

When SNPs or candidate genes are selected from the GWAS, validation studies have to be performed. This will be done with in vitro cell culture studies.

For our research we propose that mammalian cell lines with the selected SNPs are compared with the same cell lines that have siRNA knockdown. The effect of the SNPs on the response to the thiopurine drugs will be investigated using four groups:

1. Cells with siRNA and thiopurine
2. Cells with siRNA and without thiopurine
3. Cells without siRNA and with thiopurine
4. Cells without siRNA and without thiopurine

3.6 Cell lines

For the in vitro experiment, numerous immortalized cell lines are available. The selected cell lines for this experiment will be treated with the same protocol, so that the results of the different cell lines can be compared. The following options were chosen:

- Human embryonic kidney cells [Wielinga];
- LCL (lymphoblastoid cell line), as they are extremely susceptible for thiopurines [Couchana], HeLa (human cervical carcinoma), U87MG and U251 (human glioma) as well as OVCAR10 (ovarian cancer) cell lines [Matimba]. These four cell lines are used in a previous study executed by dr. Weinshilboum and his research group, which we have been using for our GWAS-data [Matimba].

The cell lines are also screened for their TMPT genotype and are chosen so that they have the same allele. This way the different cell lines can receive the same dosage of thiopurines and their response will not be affected by varying TMPT activity due to different alleles.

A disadvantage here is that choosing a specific cell line will give information of that specific cell type, but the response in human cells might be somewhat different from those in cell cultures due to the nature of an immortalized cell line. HeLa cells are a good example to illustrate what is meant with the latter: the many mutations that the cells have undergone as time has passed compared to their original state might have altered its phenotype [Neimark].

3.7 Small interference RNA (siRNA)

The expression of the candidate SNPs in the different cell lines will be suppressed by small interfering RNA (siRNA). Consisting of double-stranded RNA molecules of 20-25 base pairs, siRNAs bind to mRNA and thereby block translation [Agrawal]. The mechanism of siRNA, cleaved from double stranded RNA (dsRNA), is visualized in Figure 3. Before siRNA can bind to a single mRNA strand with a complementary sequence, its two strands are separated by the RNA-induced Silencing Complex (RISC) (see Figure 3). When the siRNA is bound to the mRNA, the double-stranded complex can be degraded by enzymes. The
effect of this degradation is that the mRNA can not be translated into amino acids and that is why the gene from which the mRNA is copied is said to be ‘knocked down’.

Figure 3: Schematic depiction of the mechanism of siRNA in a cell. Double stranded siRNA is formed after cleavage of double-stranded RNA and uncoupling of two siRNA strands is guided by the RNA-Induced Silencing Complex. Source: [Mocellin and Provenzano, 2004].

Like in vivo, also in vitro it is important that the used siRNA molecules do not induce a response by the immune system, so that the silencing effect is not counteracted. Double stranded RNA with a length of no longer than 30 base pairs were found to be acceptable in the use of silencing, because they do not trigger a non-specific immune response [Grunweller] [Elbashir]. An aspect that might improve the stability of the siRNA molecules and increase the effectivity is incorporation of Locked Nucleic Acids (LNAs) [Elmen], so these will be incorporated in the siRNA used in this experiment.

In addition to siRNA there is another suitable method which makes use of antisense oligonucleotides, but it is a point of discussion whether one of the two is better in silencing genes. An advantage of antisense oligonucleotides is that they have lower costs [Grunweller], but some research states that siRNA is more stable [Bertrand]. Therefore, this research will start with the use of siRNA, but if despite the LNAs this will turn out to be too ineffective, antisense molecules will be used instead. Antisense oligonucleotides also consist of a base sequence complementary to their target mRNA, but come single stranded [Watts].

The silencing of the (selected) genes can have various effects. The response can be a difference in cell morphology, cell growth or protein levels in the cell. In our case we are primarily interested in how the thiopurine pathway is affected by the candidate genes outside this pathway. This can be studied by measuring the metabolite level of 6-MP, 6-TG and 6-MMP, and the level of the proteins encoded by the candidate genes.
In order to make sure that the selected genes do not influence the basal metabolism of the cell, the transfected cells can also be tested on growth arrest. As described in [Elbashir], the condition of the nucleus is an indication of whether the cell is able to divide or not and thereby indicate whether the basal metabolism is affected. It is possible that levels of non-targeted proteins will change [Scacheri] and that is why the selection of the siRNA and a control group are so important.

In order to account for the effectiveness of the siRNA knockdown in the cells, there has to be a negative control in addition to the four groups. The group in which cells are incubated with only the target siRNA has to be compared with a control group that is transfected with non-targeted siRNA. This non-targeted siRNA should for example consist of a sequence from a species other than human and that is not known to be similar to a human gene [Scacheri]. This comparison should then indicate how effectively the targeted siRNA binds.

Measuring the mRNA, used here to determine the silencing efficiency of the siRNA, can be done by giving the mRNA molecules a fluorescent label and performing quantitative (real time) Reverse Transcription Polymerase Chain Reaction (qRT-PCR) [Scacheri]. After the mRNA with fluorescent label is extracted from the siRNA transfected cells, it is reversed transcribed to form complementary DNA (cDNA) and this is amplified by PCR, which together make up the qRT-PCR. Fluorescence levels are measured during PCR, but it is also possible to additionally analyze the PCR products by for example gel electrophoresis [Degen]. Comparing the signals found with qRT-PCR will indicate to what extent the mRNA levels differ under different circumstances. The efficiency of the siRNA silencing should be sufficient so that it is possible to draw conclusions from the candidate gene knockdown studies.

How the four groups with or without siRNA and thiopurine will be analyzed, is the subject of the following sections.

3.8 Proteome analysis

Introduction
Next to transcript and metabolite analysis, proteome analysis can also be used to determine the effect of a certain gene on the drug response. In all three cases RNAi will be used to control the expression of the gene. Proteome analysis is in between the two other methods, since the flow of (genetic) information is in the order transcript - protein - metabolite (see fig. 4). The metabolite analysis is closest to the actual drug response, so this will be the focus of the study. However, since the metabolites are higher up, there are more factors that can play a role in the drug response, other than the gene that has been knocked down. These factors are hard to control, and therefore it is also useful to take a lower level approach and look at the proteome.

![Flow of genetic information](image)

Figure 4. Flow of genetic information.

Method
The most widely used approach for proteome analysis is mass spectrometry (MS). This amounts to breaking up the proteins in the cell into polypeptides, which are parts of the
proteins, and analyzing these polypeptides with a mass spectrometer. The abundance of different polypeptides can then be used to determine the abundance (and other properties [Larance]) of the proteins present in the cells that are being studied.

The method is outlined in figure 2 [Aebersold]. It consists of five phases. The first phase is selection, where the proteins to be analysed are isolated from the lysate using two-dimensional electrophoresis (2DE). These will be the proteins in the known thiopurine pathway, since it is likely that these will have the biggest effect on the drug metabolism. Then, in the second phase, the proteins are degraded into peptides, which results in higher accuracy of the MS measurements. In the third phase the different polypeptides are separated using liquid chromatography, and an electrospray ionization (ESI) source nebulizes the peptides into charged droplets that can be analyzed by the spectrometer. The combination of liquid chromatography and ESI (LC-MS) is commonly used for the analysis of large protein samples. The fourth stage is the first mass spectrometry step, where a first selection of peptides is made. In the last step tandem MS (MS/MS) is used to isolate different peptides, and determines their spectra, which are called collision-induced (CID) spectra. The data consists of the MS and MS/MS spectra and these can be compared with a protein database to determine the abundance of the different proteins.

![Figure 5: Mass spectrometric methods used for proteomics [Aebersold].](image)

The mass analyzer is the most important part of the method used, because it separates the peptides. There are multiple types of analyzer that are commonly used for proteome analysis. For this study, a linear trap quadrupole (LTQ) analyzer will be used. This is because it is relatively inexpensive, but still has a high mass accuracy.

Modern advances in mass-spectrometric methods also allow measurements of multiple properties of the proteome, together with suitable data analysis [Larance]. These so called multidimensional methods are most likely too sophisticated for this study, but if more detailed data may be required, they could be considered.

### 3.9 Thiopurine metabolites analysis

Measuring the metabolites in the thiopurine pathway is the next step of this research. The aim is to measure 6-MP, 6-MMP and 6-TG, the most important metabolites of thiopurines. This way the activity of the pathway of the thiopurine can be determined. By doing this to the
different cell lines and different treatments, it can be proved that there is a correlation between the non-pathway gene and the thiopurine pathway. A widely used method is High-Performance Liquid Chromatography. The metabolite levels will be converted to a logarithmic scale to normalize the distribution of values [Lennard]. HPLC-technique can be used for separating 6-MP, 6-TG and 6-MMP from a cell matrix. These components will be identified and quantified. The figure below shows a construction of the HPLC method. The system relies on a pump that pumps the liquid through the column, which is filled with a solid adsorbent material. Every component (6-MP, 6-TG and 6-MMP) will react differently with this material. This causes the metabolites to each have a different flow rate and the latter means that the metabolites come out of the column at different times [Zakrewski]. With the help of this method, a Michaelis-Menten and a Lineweaver-Burk diagram will be constructed. With the concentrations of the thiopurine metabolites, determined by the HPLC method, comparisons will be made to the regular metabolite levels. This will indicate the effect that the gene knockdown has on the (regular) thiopurine pathway.

![Figure 6: The High-performance liquid chromatography method used for metabolites analysis](Waters).
4. Timetable of the project

<table>
<thead>
<tr>
<th>Activity</th>
<th>Year 1</th>
<th>Year 2</th>
<th>Year 3</th>
<th>Year 4</th>
</tr>
</thead>
<tbody>
<tr>
<td>Writing proposal</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>A. Cell line research</td>
<td></td>
<td>X</td>
<td></td>
<td></td>
</tr>
<tr>
<td>B. Knock-down and validation</td>
<td></td>
<td></td>
<td>X</td>
<td></td>
</tr>
<tr>
<td>C. Metabolite analysis</td>
<td></td>
<td></td>
<td>X</td>
<td></td>
</tr>
<tr>
<td>D. Proteome analysis</td>
<td></td>
<td></td>
<td></td>
<td>X</td>
</tr>
<tr>
<td>Writing papers/thesis</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

5. Statements by the applicant

YES  I endorse and follow the Code Biosecurity (if applicable).

YES  By submitting this document I declare that I satisfy the nationally and internationally accepted standards for scientific conduct as stated in the Netherlands Code of Conduct for Scientific Practice 2012 (Association of Universities in the Netherlands (VSNU)).

YES  I have completed this form truthfully.
6. References:


15. Chouchana, L; Fernández-Ramos, AA; Dumont, F; Marchetti, C; Ceballos-Picot, I; Beaune, P; Gurwitz, D; Loriot, M. (2015). Molecular insight into thiopurine resistance: transcriptomic signature in lymphoblastoid cell lines. Genome med.

16. Wielinga, PR; Reid, G; Challa, EE; Heijden, van der I; Deemter, van L; Haas, de M; Mol, C; Kuil, AJ; Groeneveld, E; Schuetz JD; Brouwer, C; De-Abreu, RA; Wijnholds, J; Beijnen, JH; Borst, P. (2002). Thiopurine Metabolism and Identification of the Thiopurine Metabolites Transported by MRP4 and MRP5 Overexpressed in Human Embryonic Kidney Cells. Molecular Pharmacology, 1321-1331.


22. Grunweller, A; Wyszko, E; Bieber, B; Jahnel, R; Erdmann, VA; Kurreck, J. Comparison of different antisense strategies in mammalian cells using locked nucleic acids, 2′-O-methyl

16


