Chapter 1

General introduction
1.1 ADVERSE DRUG REACTIONS AND DRUG–PROTEIN ADDUCTS

The number of new drugs reaching the market has been in decline for decades despite ever increasing investments into pharmaceutical R&D. [1] Drug toxicity and more stringent regulatory requirements are seen as a major cause of the high attrition rates in drug discovery and development. [2] Even approved drugs may eventually be withdrawn from the market due to adverse events that only become apparent when the drug is used by a larger population. [3-4]

Traditionally, adverse drug reactions (ADRs) were classified into type A and B reactions, which have been described in several ways, such as predictable and non-predictable, respectively. [5] Type A reactions are related to the pharmacological effect of the drug and are usually dose-related and predictable. These reactions occur quite often, especially with individuals that are at the extremes of the dose-response curves, but they are seldom life-threatening. [5] Conversely, type B reactions are bizarre and abnormal effects that are not related to the pharmacological effect of the drug. These often severe effects occur only in a few individuals (idiosyncratic) and are, therefore, difficult to predict. [5] Due to the diverse range of ADRs, these two classes are insufficient to describe the diverse range of ADRs. As a result, other types of adverse drug reactions were distinguished over the years, leading to a total of six categories labeled A-F, see Table 1.1. [6]

Type C reactions are a combination of dose- and time-related (chronic) effects that are uncommon and related to the cumulative dose. [6] Type D reactions, originally proposed by Royer in 1997 [7], are very uncommon delayed effects that become apparent some time after use of the drug, such as adenocarcinoma in daughters of women who have taken stilbestrol during pregnancy. Type E reactions can also be considered as a delayed response to drug treatment, but are seen as a separate category because the reaction occurs directly after drug withdrawal. [6] The most recent addition to the classification of ADRs is the type F reaction that is described as therapeutic failure which may be caused by a number of reasons, such as inadequate dosage that is too low to obtain an effect or drug-drug interactions. [6, 8] Even with this expanded classification, not every ADR can be classified into one of the proposed categories. However, the classification is under constant revision and may be adapted further as more becomes known about the mechanisms underlying adverse effects.

Several theories describing the mechanisms behind ADRs have been proposed in literature and have been compiled in reviews. [9-10] ADRs have long since been associated with drug metabolism. [11,12] During this biological process, the drugs are metabolized mainly into more polar metabolites that can be easily excreted via the kidneys and urine (Figure 1.1). However, many drugs also undergo bioactivation, which leads to reactive metabolites that may bind to other molecules present in the liver. The reactive metabolites can be detoxified via covalent binding to glutathione (GSH). This conjugation reaction results in increased polarity and the drug-GSH conjugates are subsequently excreted via the urine (and feces, via the bile). When large amounts of reactive metabolites are produced, it may lead to depletion of GSH. Once the GSH stores are depleted, the surplus of reactive metabolites may covalently bind to proteins. Reversible binding
### Table 1.1 Classification of adverse drug reactions, adapted from Edwards et al. [6]

<table>
<thead>
<tr>
<th>Type of reaction</th>
<th>Mnemonic</th>
<th>Features</th>
<th>Examples</th>
</tr>
</thead>
<tbody>
<tr>
<td>A: Dose-related</td>
<td>Augmented</td>
<td>- Common&lt;br&gt;- Related to the pharmacological &lt;br&gt;action of the drug&lt;br&gt;- Predictable&lt;br&gt;- Low mortality</td>
<td>Toxicity: Hepatotoxicity due to acetaminophen overdose&lt;br&gt;Side effects: Anticholinergic effects of tricyclic antidepressants</td>
</tr>
<tr>
<td>B: Non-dose-related</td>
<td>Bizarre</td>
<td>- Uncommon&lt;br&gt;- Not related to the pharmacological &lt;br&gt;action of the drug&lt;br&gt;- Unpredictable&lt;br&gt;- High mortality</td>
<td>Immunological reactions: Penicillin hypersensitivity&lt;br&gt;Idiosyncratic reactions: Pseudoallergy (ampicillin rash)</td>
</tr>
<tr>
<td>C: Dose- and time-related</td>
<td>Chronic</td>
<td>- Uncommon&lt;br&gt;- Related to the cumulative dose</td>
<td>Hypothalamic-pituitary-adrenal axis suppression by corticosteroids</td>
</tr>
<tr>
<td>D: Time-related</td>
<td>Delayed</td>
<td>- Uncommon&lt;br&gt;- Usually dose-related&lt;br&gt;- Occurs or becomes apparent some time after use of the drug</td>
<td>Carcinogenesis&lt;br&gt;Teratogenesis (vaginal adenocarcinoma with diethylstilbestrol)</td>
</tr>
<tr>
<td>E: Withdrawal</td>
<td>End of use</td>
<td>- Uncommon&lt;br&gt;- Occurs soon after withdrawal of drug</td>
<td>Opiate withdrawal syndrome&lt;br&gt;Myocardial ischaemia (β-blocker withdrawal)</td>
</tr>
<tr>
<td>F: Unexpected failure of therapy</td>
<td>Failure</td>
<td>- Common&lt;br&gt;- Dose-related&lt;br&gt;- Often caused by drug-drug reactions</td>
<td>Inadequate dosage of an oral contraceptive, particularly when used with specific enzyme inducers</td>
</tr>
</tbody>
</table>
of drugs to proteins is considered a critical factor in drug efficacy [13], whereas covalent binding of reactive drug metabolites to proteins and the hereby formed drug–protein adducts have been associated with drug toxicity. [3-4] This hypothesis is known as the hapten theory, where it is assumed that low-molecular weight compounds (< 1000 Da), such as drugs and their metabolites are not to able to elicit ADRs because, presumably, they are too small to be recognized by the immune system. [14]

On the other hand, it has been observed that chemically inert drugs, not susceptible to bioactivation and the formation of reactive metabolites, still may elicit an immune response. This observation has led to the formation of the ‘pharmacological interaction with immune receptors’ or p-i concept that was introduced more recently by Pichler. [15] Pichler proposed that the drugs themselves may bind directly and reversibly to immune receptors and, subsequently, to T-cell receptors, thereby stimulating a T-cell response. The p-i model has proven to be useful in explaining the observed drug hypersensitivity reactions after treatment with chemically inert drugs, such as lidocaine, but also some drugs that are known to be bioactivated, for example sulfamethoxazole and carbamazepine. [16]
Other theories, such as the danger hypothesis and non-immune-mediated mechanisms have also been proposed, but the hapten theory remains the major working hypothesis. [10] Consequently, drug metabolism and bioactivation as well as covalent binding of reactive drug metabolites to proteins have received a great deal of scientific attention, which has been summarized in an extensive review by Zhou. [13] Perhaps the best known example and most likely the basis of the hapten theory is the covalent protein binding of N-acetyl-p-benzoquinoneimine (NAPQI), the reactive metabolite of acetaminophen (APAP). [11] In the 1970s, APAP was one of the first drugs of which its metabolism and covalent protein binding was studied. [12] A decade later, free protein thiols were determined to be the major binding sites of NAPQI in in vitro experiments using bovine serum albumin (BSA). [17] Similar to human serum albumin (HSA), BSA contains a free cysteine residue on position 34 that is not involved in a disulphide bridge and, thus, is available for binding NAPQI. [17] Due to its high abundance in vivo, serum albumin was considered as a promising target of NAPQI. Identification of 27 potential target proteins of NAPQI was achieved several years later in mouse studies [18-21] but the list of identified proteins did not include serum albumin. Recently, the NAPQI–albumin adduct was identified in human serum samples from patients that were exposed to high levels of APAP [22].

Apart from cysteine thiols, also other amino acids in proteins are susceptible to nucleophilic attack by reactive metabolites, including lysine amines, histidine imidazoles and protein N-terminal amines. [23] Many adducts of reactive metabolites of drugs and other xenobiotics, and their protein targets have been identified over the years and are listed in the reactive metabolite target protein database, a web-accessible resource. [24] In addition to elucidation of the mechanisms behind ADRs, covalent binding of reactive drug metabolites to proteins also is a critical factor for the evaluation of lead compounds in pharmaceutical R&D. [4, 25] The assessment of covalent protein binding and identification of drug–protein adducts is, therefore, of utmost importance in the pharmaceutical industry.

1.2 CHALLENGES AND METHODOLOGIES FOR THE DETECTION AND IDENTIFICATION OF DRUG–PROTEIN ADDUCTS

The analysis of drug–protein adducts is faced with considerable challenges. The major complication for the detection and identification of drug–protein adducts is their extremely low abundance. Drugs often do not exclusively bioactivate to a single reactive metabolite, but may produce several reactive metabolites as well as be metabolized into one or more stable metabolites. [26] Additionally, the reactive metabolites that were formed are, for the most part, detoxified by GSH, which leaves only a small portion, if anything, for covalent binding to proteins. Even in the case of a drug overdose where large amounts of reactive metabolites are produced and GSH is depleted [27], only a minute amount of the target protein is modified in vivo. This is illustrated by the detection of low pmol/mL serum levels of NAPQI–HSA in a patient exposed to high levels of APAP and suffering from severe liver toxicity [22].
Since only a very small percentage of the target protein is modified, an excess of 99% of this protein will still be present in its non-modified form, which represents another difficulty in the detection of drug–protein adducts. The separation of the modified from non-modified protein is extremely challenging because the small reactive metabolite that has covalently bound to the protein often has a negligible effect on the molecular characteristics of the macromolecule, i.e., mass, charge and size. Conventional separation techniques, such as liquid chromatography (LC), do not provide the required selectivity for this purpose. Capillary electrophoresis (CE) can efficiently separate protein isoforms caused by small modifications. Recent improvements of the CE–mass spectrometry (MS) interface allows for accurate detection and identification of the separated isoforms and protein modifications. CE–MS has also been applied to the analysis of synthetically prepared drug–protein conjugates, present at relatively high levels of 5% of the total protein content. Although CE–MS possesses the potential for the detection and identification of clinically relevant levels of drug–protein adducts of 0.1% or less, this has never been actually demonstrated.

Another method of separation is based on thiol affinity resins to achieve the selective enrichment of HSA-cysteine34 adducts. Using this method, non-modified HSA was efficiently removed from a freshly isolated HSA sample through binding to the resin via the available cysteine34. In the remaining non-bound fraction, constituting 2.9% of the total HSA, the cysteine34 residues were modified by reactive systemic electrophiles, i.e., by cysteinylation and glycosylation. Although this methodology is promising for selective extraction of drug-HSA adducts, it remains to be seen if it is applicable to 15-fold or even lower adduct levels, as is the case with drug–protein adducts. Separation of the modified from non-modified protein in itself is not straight-forward, but biological tissues and fluids also contain many other proteins that are present in a very wide dynamic range and may mask the presence of the drug–protein adduct. Consequently, analytical methodologies for the detection of drug–protein adducts must possess a high degree of selectivity and sensitivity.

Analytical methods for the detection and identification of drug–protein adducts can roughly be divided into global and targeted approaches, see Figure 1.2. Global approaches are usually based on the use of radiolabeled drugs in combination with two-dimensional gel electrophoresis (2D-GE) and bottom-up proteomics, and have been successfully applied to the identification of protein targets of reactive drug metabolites. In this way, 42 cytosolic and 24 microsomal protein targets of thiobenzamide (TB) were identified in rat liver. The identified proteins serve a broad range of biological functions, but, in the context of cytotoxicity, several groups could be distinguished, i.e., enzymes of intermediary metabolism and heat shock and stress response proteins. Covalent binding of reactive drug metabolites to these types of proteins may play a role in enzyme inhibition and misfolding of proteins, which may have grave consequences for cell viability or may lead to apoptosis. Similarly, 15 rat liver proteins were identified as likely targets of tienilic acid, 12 of which were previously unidentified targets. Again, the majority of the identified proteins are enzymes that participate in metabolic and catabolic cellular processes and their modification or inhibition may lead to cellular stress that may be used, in combination with the immune system, to explain the mechanism involved in tienilic acid ADRs.
Currently, more than 540 proteins of various sources are listed in the reactive metabolite target protein database, 358 of which are non-redundant. [24] A considerable overlap in protein targets of different drugs exists, for example protein disulfide-isomerase A3 and glyceraldehyde-3-phosphate dehydrogenase are known targets of the reactive metabolites of APAP, TB, mycophenolic acid (MA) and several other xenobiotics. These three drugs combined have 147 protein targets of which two are shared by all three drugs, a further five targets are shared only by APAP and TB, five other targets are shared only by TB and MA, and three are shared only by APAP and MA, see Figure 1.3. Commonalities in protein targets may imply that the protein adducts of different drug metabolites may affect similar pathways, leading to the same ADRs. More commonalities between target proteins and drugs may be discovered in the future when the methods become more sensitive and repeatable, perhaps by standardization. Hepatic protein targets of reactive drug metabolites have been studied most extensively, since covalent binding is primarily evaluated

**Figure 1.2 Main approaches for the analysis of drug–protein adducts.**
in liver and mostly in animal models. The list of identified proteins, therefore, contains mainly hepatic targets, but may easily be expanded with other protein targets by application of global approaches to other tissues and biofluids. Expansion of the list of protein targets may contribute to the elucidation of the role of drug–protein adducts in biological pathways, such as inflammation, and to a better understanding of the mechanism behind ADRs.

Unfortunately, radiolabeled drugs can only be used in *in vitro* experiments or *in vivo* animal models and this type of approach cannot be applied to patient samples or samples from clinical trials in humans. Additionally, global approaches often do not shed any light about the identity of the drug metabolite or the site of adduct formation due to lack of detection of the proteolytic peptide containing the adduct site. [26] Such detailed characterization of drug–protein adducts and unequivocal confirmation that adduct formation has taken place requires a higher degree of selectivity and sensitivity that may be provided by a targeted approach. The in-depth information obtained by this type of approach is crucial for understanding how reactive drug metabolites modify proteins, how the modification of specific amino acid residues may alter the protein function and affects downstream signaling pathways. [35]

A targeted adduct-proteomics method is focused on a single protein target that is selectively purified from a complex matrix, e.g., tissue or serum. Selection of potentially interesting protein targets is performed based on information obtained through global approaches using selective labeling, such as radiolabels. Selective purification of the target protein, e.g., by affinity chromatography, is often a prerequisite for qualitative or quantitative proteomic approaches and has proven to be useful for studying low-abundant proteins, as well as protein interactions, e.g. protein-protein complexes, and modifications, such as post-translational modifications (PTMs) and glutathione conjugation. [36-37] With all extraction methods, the selected protein is purified to a certain extent and although there will always be some low-abundant proteins in the

![Figure 1.3 Overview and commonalities between the protein targets of the reactive metabolites of acetaminophen (APAP), mycophenolic acid (MA) and thiobenzamide (TB).](image-url)
background, the sample complexity and interferences from other proteins will be significantly reduced. The purified protein fraction, i.e., containing both the adducted and non-modified species, is subsequently treated further using conditions specifically chosen for this target in order to maximize the information content obtained from the analysis. A priori knowledge of the number of cysteine residues and the protein sequence enables educated selection and optimization of the treatment conditions, such as reduction, alkylation and digestion of the protein. Trypsin is considered as the golden standard in enzymatic protein digestion due to its unparalleled specificity and efficiency [38], but a lack or over-abundance of tryptic cleavage sites may call for the use of another enzyme or possibly a multiple enzyme digestion [39-40]. Assessment of the proteolytic enzyme to be used in combination with knowledge of the protein sequence and prediction of possible adduct formation sites may lead to improved detection and identification of drug–protein adducts.

In contrast to the great number of identified target proteins of reactive drug metabolites, only a few drug–protein adducts have actually been characterized in vivo. A targeted approach employing albumin affinity chromatography was developed for the identification of NAPQI–albumin adducts in human serum [22]. Different enzymatic digestion approaches were developed, using either pronase for NAPQI–HSA digestion or trypsin for NAPQI–MSA digestion. In another publication, HSA adducts of several sulfur mustards were identified in whole blood after precipitation of the albumin fraction in a multi-step procedure followed by pronase digestion. [41] In all three above mentioned cases, the cysteine34 residue of serum albumin was identified as the site of adduct formation and the respective reactive metabolite or drug could be identified based on the observed shift in the peptide mass. Finally, adducts of flucloxacillin and its 5-hydroxymethyl metabolite primarily to lysine190 and lysine212 residues of serum albumin were successfully identified in plasma samples of patients treated with this synthetic penicillin using albumin affinity chromatography and peptide mass mapping. [42] In addition, more adduct modification sites on different proteins have been characterized in detail using other xenobiotics, such as bromobenzene (BB) [26], acrylonitrile [43] and acrylamide [44]. For example, bromobenzoquinone adducts to cysteine-111 of GST-A1 and -A2 were identified in liver tissue samples obtained from a BB metabolism study performed in rats. [26] For this purpose, the researchers used 14C-labeled BB and developed an analytical strategy employing GST affinity chromatography and further separation of the GST subclasses by reversed-phase HPLC, tryptic digestion and performed the identification by a combination of 1D-GE, intact protein MS and peptide mass mapping. These potentially relevant target proteins of other xenobiotics may aid in the identification of targets of reactive drug metabolites.

Whereas identification of protein targets in global approaches is more efficiently achieved by protein database searching, in targeted approaches manual searches may be performed based on a priori predicted adducts. The structure and mass of many reactive metabolites were identified in numerous drug bioactivation studies, predominantly performed in vitro using (mutant) cytochrome P450 enzymes. [45-46] Subsequent characterization of the covalent binding of reactive metabolites to clinically relevant proteins, such as GST [47], greatly aids the detection and identification of drug–protein adducts in vivo. Information regarding potential adduct formation sites on proteins may also be obtained from selective labeling of protein mixtures, for instance using radiolabels, as discussed earlier. The enhanced selectivity of this type of strategy and the information obtained from its application may contribute to the confirmation...
of drug–protein adduct formation in vivo. Therefore, the two types of approaches for drug–protein adduct identification described here are not mutually exclusive, in fact, their power lies in their combined or successive use and the increased level of information gained from their combined results. The identified potential protein targets of (labeled) reactive (drug) metabolites in global approaches are the starting point of targeted approaches where actual drug–protein adduct confirmation may be achieved. The combination of these two methods is very powerful and may be the only route to understanding the mechanisms behind ADRs and how they are linked to drug–protein adducts.

As stated earlier, some ADRs are dose-dependent, usually meaning that an overdose of a drug will cause the adverse reaction, while others are assumed not to be related to the dose. The latter may imply that an individual may respond adversely to a drug due to genetic predisposition and/or that even a small dose may lead to severe ADRs. Therefore, the quantitative relationship between ADRs and drug–protein adduct formation requires scientific attention. Quantification of protein biomarkers in vivo as readout of biological response in clinical proteomics has been the subject of research for many years. [48] Indeed, up- or down-regulation of proteins presented in a patient may be an indication for a disease state [49], e.g. HER2/NEU for staging of breast cancer [50]. The measurement of certain biomarkers, such as serum albumin in plasma as a marker for (mal)nutrition and other disease states, is straightforward and routinely performed in the clinical setting. [51]

However, the readout of other protein biomarkers may require more complex proteomics-based quantification approaches involving MS techniques, such as selected reaction monitoring (SRM), and/or isotopic labeling strategies, such as isotope coded affinity tags (ICAT) and isobaric tags for relative and absolute quantification (iTRAQ), that are not yet standardized or unsuited for clinical use. [49, 52] Additionally, the inter- (during the day) and intra-individual biological variation, the wide dynamic range of proteins in biological samples, and the (absence of) linearity of the relationship between the disease and biomarker are hurdles that still have to be overcome in the analysis and quantification of biomarkers. [52-53]

In recent years, changes in the post-translational modification of proteins have been implied as a possible biomarker for disease states, such as cancer [54-55] and Alzheimer’s disease [56-57] Determination of the level of PTMs of proteins may be useful for early detection of a disease and following its progress, but is extremely challenging due to absence of standard reference peptides, inefficiency of isotopic labeling, and large variation. [58] Many of the challenges associated with the quantification of proteins and protein PTMs also apply to drug–protein adducts. To our best knowledge, only one method for drug–protein adduct quantification in vivo has previously been reported in literature, specifically for the NAPQI–HSA adduct [22]. Damsten et al. utilize a synthetic reference peptide approach for absolute quantification, where a NAPQI–HSA pronase digest sample spiked with the synthetic NAPQI–CPF peptide is compared to a non-spiked sample. [22] A NAPQI–CPF serum concentration 35 pmol/mL was detected in a sample from a patient that had taken an APAP overdose of 40 g, while for two other patients having taken 10 or 12 g of APAP this level was extrapolated to be 3 - 4 pmol/mL serum. From these results, it appears that the NAPQI–CPF serum concentration is related to the dose, but this relationship may not be linear. Absolute quantification has advantages over other approaches, however, the one-point calibration described in this paper is generally not very accurate, but could easily be improved by extending it to a standard addition.
method. Application of the above described methodology resulted in the detection of extremely low concentrations of the NAPQI–HSA adduct in serum, but could not establish a clear relationship to the dose. Taking into account that the subjects in these studies received relatively high drug (over)doses means that the detection and quantification methods for drug–protein adducts in vivo require further improvement of the detection limit and sensitivity. Especially in cases where severe ADRs occur at therapeutic doses of a drug, the drug–protein adducts may be present at even lower levels that may be undetectable with the current methodologies.

To summarize, more research is still needed to unravel the precise role of drug–protein adducts in the mechanisms behind ADRs. Their analysis, especially in the low-abundant protein fractions of biological tissues and fluids, is challenging, but essential for identification of the affected biological pathways involved in adverse responses. The development of analytical methodologies specifically for drug–protein adduct detection, identification and quantification is indispensible for gaining a better understanding of ADRs.

1.3 SCOPE OF THE THESIS

The work in this thesis was performed within the framework of project D3-201 “Towards novel translational safety biomarkers for adverse drug toxicity” of the Dutch Top Institute Pharma. The project consortium, consisting of five academic and seven industrial partners, studied the relationship between drug bioactivation and immune system-dependent adverse drug reactions. The research groups performed translational research in their respective expertise areas aimed at correlation of pre-clinical in vitro and in vivo models to ADR patient data. This research includes: 1) Mechanistic studies of drug-induced liver injury through investigation of intracellular stress signaling using high content imaging assays, 2) Characterization of reactive drug metabolites through low-molecular weight trapping agents and adduct formation to selected proteins using in vitro bioactivation systems, 3) Precision-cut liver slices as an ex vivo model for translation of human data to in vitro human and mouse in vitro and in vivo models, 4) Validation of mechanism-based mouse in vivo models of drug-induced hepatotoxicity and systemic immunosensitization, and 5) Identification of predictive biomarkers of ADRs in plasma and urine of ADR patients through protein profiling.

This thesis is focused on the development and optimization of proteomics-based methods for the detection and identification of drug–protein adducts in vivo to be applied to samples from the other research groups. Due to the challenges involved in the detection of low-abundant drug–protein adducts in complex biological samples, sample preparation plays a crucial role. Protein digestion is a fundamental part of bottom-up proteomics, which is the method of choice for the analysis of proteins and their modifications. An overview of digestion strategies and recent novel developments in this area is given in the review in Chapter 2.
The analysis of low-abundant proteins and their modifications in general and, more relevant to the current context, drug–protein adducts can benefit considerably from efficient sample preparation protocols. Chapter 3 describes the optimization of sample pretreatment and digestion conditions specifically for enhanced detection of cysteine34 adducts of HSA. In this study, an experimental design approach was applied to the optimization of tryptic and thermolytic digestion of a model HSA adduct, which led to optimum conditions that differed from those found in literature, but performed significantly better.

Sample preparation for bottom-up proteomics consists of multiple treatment steps, prior to digestion, including protein purification, denaturation, reduction, alkylation and multiple sample clean-up steps. Especially for large sample sets, these workflows are often very laborious and time consuming, in addition to the high consumption of chemicals and reagents. Chapter 4 concerns the development of a high-throughput sample preparation method for bottom-up proteomics employing 96-well filter plates that reduces sample transfer to a minimum and allows for the simultaneous preparation of 96 samples.

Chapter 5 describes the application of a targeted approach for identification of albumin adducts, incorporating the optimized sample preparation and digestion conditions, to mouse serum samples resulting from an APAP exposure experiment to study the kinetics of ADRs caused by this drug. An adduct quantification methodology was developed for effective evaluation of the relationship between the ADRs and the formation of NAPQI–MSA adducts. Chapter 6 presents a summary of this thesis as well as a discussion of the conclusions and directions for further research in the context of analytical challenges.

REFERENCES


[31] Role of covalent binding


[34] Labenski, M.T., Fischer, A.A., Lo, H.H., Monks, T.J., Lau, S.S., Protein electrophile-binding motifs: Lysine-rich proteins are preferential targets of quinones. Drug Metab. Dispos., 2009, 37, (6), 1211-1218.


