Studies of metabolite–protein interactions: A review

Ryan Matsuda  
University of Nebraska-Lincoln

Cong Bi  
University of Nebraska-Lincoln

Jeanethe Anguizola  
University of Nebraska–Lincoln

Matthew Sobansky  
University of Nebraska–Lincoln

Elliott Rodriguez  
University of Nebraska–Lincoln

See next page for additional authors

Follow this and additional works at: http://digitalcommons.unl.edu/chemistryhage

Part of the Biochemistry Commons

Matsuda, Ryan; Bi, Cong; Anguizola, Jeanethe; Sobansky, Matthew; Rodriguez, Elliott; Vargas Badilla, John; Zheng, Xiwei; Hage, Benjamin; and Hage, David S., "Studies of metabolite–protein interactions: A review" (2013). David Hage Publications. 66.  
http://digitalcommons.unl.edu/chemistryhage/66

This Article is brought to you for free and open access by the Published Research - Department of Chemistry at DigitalCommons@University of Nebraska - Lincoln. It has been accepted for inclusion in David Hage Publications by an authorized administrator of DigitalCommons@University of Nebraska - Lincoln.
Studies of metabolite–protein interactions: A review

Ryan Matsuda, Cong Bi, Jeanethe Anguizola, Matthew Sobansky, Elliott Rodriguez, John Vargas Badilla, Xiwei Zheng, Benjamin Hage, David S. Hage

Department of Chemistry, University of Nebraska–Lincoln, Lincoln, NE 68588-0304, USA
Corresponding author — D. S. Hage, tel 402 472-2744, fax 402 472-9402, email dhage1@unl.edu

Abstract
The study of metabolomics can provide valuable information about biochemical pathways and processes at the molecular level. There have been many reports that have examined the structure, identity and concentrations of metabolites in biological systems. However, the binding of metabolites with proteins is also of growing interest. This review examines past reports that have looked at the binding of various types of metabolites with proteins. An overview of the techniques that have been used to characterize and study metabolite–protein binding is first provided. This is followed by examples of studies that have investigated the binding of hormones, fatty acids, drugs or other xenobiotics, and their metabolites with transport proteins and receptors. These examples include reports that have considered the structure of the resulting solute–protein complexes, the nature of the binding sites, the strength of these interactions, the variations in these interactions with solute structure, and the kinetics of these reactions. The possible effects of metabolic diseases on these processes, including the impact of alterations in the structure and function of proteins, are also considered.

Keywords: metabolomics, drug–protein interactions, hormone–protein interactions, fatty acid–protein interactions, xenobiotic–protein interactions, protein modification

1. Introduction
Metabolomics is a field that involves the study of low mass compounds (i.e., metabolites) that are produced through metabolic processes [1, 2]. Metabolites are part of a collection of chemicals known as the "metabolome", which can include small molecules that are found in cells, tissues, organs, or biological fluids. The area of metabolomics is of interest because the identity and concentration of metabolites can provide information about cellular activity and can be directly related to processes such as protein and gene expression [1, 2, 3]. This means that metabolomics can provide information on the phenotypes of individuals at the molecular level [3]. In addition, the characterization and examination of metabolites could lead to new discoveries in biomedical research and personalized medicine [1, 3].

Research in metabolomics began in the late 1990s and early 2000s, with the emphasis at that time being on the effects of different metabolites on the gene expression of bacteria and yeast [1]. The first examples of metabolomic studies utilized two-dimensional thin-layer chromatographic separations to characterize metabolites in samples. This provided researchers with evidence that variation in the concentrations of metabolites can affect cellular activity [1, 4, 5, 6]. Further progress in the area of analytical methods such as structural characterization and separation methods has resulted in the development of new instruments and techniques that can be used to provide high resolution information and data from complex samples such as tissues and cells [1, 2].

Research in metabolomics can involve either targeted or untargeted approaches [7]. In a targeted approach, researchers use techniques such as nuclear magnetic resonance (NMR) spectroscopy and mass spectrometry (MS) for the identification, quantification, and structural characterization of specific metabolites. This information can be used to examine specific classes of metabolites and to provide information on the biochemical pathways that are involved in metabolism [2]. In an untargeted approach, scientists use global profiling to analyze the group of chemicals in a metabolome as a whole. This second approach is less specific and sensitive than the targeted approach but allows for the highest possible coverage of the metabolites that may be involved in biochemical pathways [7].

A significant amount of recent research has been devoted to metabolic profiling, or the identification and measurement of the different metabolites that are present and produced in the metabolome [8]. However, it is also important to consider the interactions that occur between metabolites and biological agents, such as the binding of cofactors to enzymes, hormones to receptors, and drugs or their metabolites to proteins [8]. Information on these interactions can be combined with structural data to provide a better understanding of the regulatory networks and connections in biological pathways. Such information, in turn, could provide a better understanding of how healthy and disease states
differ at the molecular level and could provide vital data that can be used for pharmaceutical development [7, 9].

This review will look at previous studies that have examined biological interactions as related to metabolites and proteins as binding agents. This will include an overview of the various methods and techniques that have been used in this work to study metabolite–protein interactions. A summary will also be provided of the different types of metabolite–protein binding interactions that have been investigated with these approaches. In addition, the possible effects that metabolic diseases may have on these interactions will be considered.

2. Techniques for examining metabolite–protein interactions

The characterization of metabolite–protein interactions can provide a better understanding in clinical diagnostics of the cellular activity and the biochemical pathways that are present in various medical conditions [1–3, 9]. There are many methods that can be used to examine the binding of metabolites with proteins. These methods may involve the direct examination of binding that occurs between proteins and low mass drugs, hormones and their metabolites, or may involve an examination of the free concentrations of these molecules [9–12]. The approaches that are used for this purpose can be divided into three categories: in vitro, in vivo and in silico techniques [9, 11–46].

2.1. In vitro methods for studying metabolite–protein interactions

In vitro methods are the most popular techniques used to characterize metabolite–protein interactions. This approach involves the use of standard, well-controlled conditions and reagents that are used in the laboratory to mimic conditions seen in biological systems. To examine metabolite–protein interactions, in vitro methods may use a binding assay (e.g., one based on ultrafiltration or equilibrium dialysis) to examine an interaction or to identify the chemicals that are involved in this process [9]. This approach can provide information such as the strength of the interaction, as well as the thermodynamics and kinetics of binding and possible conformational changes that occur as a result of the interaction [13–15]. Alternatively, an in vitro study may make use of a method that directly examines the structure of a protein and a bound metabolite, such as occurs in X-ray crystallography or NMR spectroscopy [1, 16–20]. Other methods may examine the protein–metabolite complex, as demonstrated with mass spectrometry [24–29].

There are many in vitro approaches that can be used to examine the binding of proteins with small molecules and their metabolites. For instance, radiometry and fluorimetry can be used with a binding assay by employing labeled metabolites that contain either a radioisotopic label or fluorophore, respectively [10, 21–25]. These labeled metabolites are then incubated with proteins and the signal that is produced from the label is measured, such as through a displacement assay or a proteome microarray [10, 23]. Radioisotopic labeling has been applied to enzymes to determine their activity in metabolomic reactions [9]. One example involved the screening of potential inhibitors for an enzyme, in which the substrate was radioactively labeled and the resulting metabolite profiles were analyzed and measured [21]. Fluorescence labeling can provide similar results to radiolabeling; however, this method can also be used to identify and determine the location of a binding site for a metabolite on a protein, such as by observing the displacement of specific probes that are bound to known locations on a protein [10].

Surface plasmon resonance (SPR) and calorimetry are two other methods that can provide information on the strength of protein–metabolite binding and the thermodynamics or kinetics of this interaction [13–15]. Studies based on SPR utilize an immobilized protein on a sensor chip, in which changes in the resonance energy (e.g., from binding of the protein with a target) are detected [9]. The change in this signal is related to the mass of the bound metabolites and can be used to determine the equilibrium constants for this process or, if examined over time, the association and dissociation kinetics that occur between the metabolite and protein during binding [9]. The reaction between a metabolite and protein can result in heat being absorbed or given off [9, 13]. Calorimetry can then be used to measure the overall enthalpy of the binding reaction between a metabolite and a protein [13].

NMR spectroscopy and X-ray crystallography are two tools that have been used to characterize the structures of metabolite–protein complexes [9, 16–20]. NMR spectroscopy has often been used in recent years for characterizing and identifying metabolites in biological samples, but this method can also be used to examine conformational changes that occur during the binding of metabolites with proteins [18–20]. X-Ray crystallography can also give structural information on such interactions by providing detailed information on the binding sites and active sites for hormones, drugs and their metabolites or related compounds on proteins and enzymes [16, 17], as is illustrated in Figure 1 [30].

Mass spectrometry can not only be used as a tool for analyzing the structure and identity of metabolites, but it can be used to analyze metabolite–protein interactions in which information about enzymatic processes or binding by small molecules is generated [9]. Experiments utilizing various types of mass spectrometry, such as quadrupole mass spectrometry or matrix-assisted laser desorption/ionization time-of-flight mass spectrometry (MALDI-TOF MS), have allowed for monitoring of the resulting products and analysis of the reaction kinetics of enzyme-substrate reactions [7, 24, 25]. Further analysis through high resolution mass spectrometers (e.g., an orbitrap or Fourier transform ion cyclotron resonance mass spectrometry) has resulted in accurate analyses of enzymatic activities in which the intermediate steps in enzymatic reactions can be identified [26–29].

Various separation techniques can also be used to examine metabolite–protein interactions. Examples of traditional methods that have often been utilized for this purpose are equilibrium
Capillary electrophoresis (CE) is another separation method that can be used to examine metabolite–protein binding [9, 37, 40, 41]. One way this method can be used is to separate the free and bound metabolites through the differences in their size-to-charge ratios. This approach can be utilized alone to determine the affinity of metabolite–protein binding or combined with other methods such as mass spectrometry to examine this interaction [41]. One form of CE is affinity capillary electrophoresis (ACE), in which a biological molecule such as a protein is used as a running buffer additive, thus making it possible to obtain data on the interactions of solute components with this additive [40]. Like HPAC, ACE is a relatively fast method and can be used with small amounts of sample for the screening or analysis of metabolite–protein interactions [38, 40].

2.2. In vivo methods for studying metabolite–protein interactions

Although in vitro methods can provide detailed information about metabolite–protein interactions, in vivo analysis can provide a better representation about the metabolite–protein interactions within a biological sample [8, 9]. This is particularly true in a situation where a protein may undergo post-translational modifications that result in changes in the protein’s interactions with solutes such as drugs and their metabolites [9]. In vivo methods are often similar to techniques used for in vitro studies but must be able to work with complex samples. In many cases, clinical samples from patients can be obtained and analyzed through approaches such as labeling, NMR or MS structural characterization, and affinity separation methods. By utilizing in vivo studies, researchers are better able to understand the effect of disease states on metabolite–protein interactions, as well as related biochemical pathways and regulatory processes [9, 39].

2.3. In silico methods for studying metabolite–protein interactions

Another area of examining metabolite–protein interactions is through in silico tools [9]. These methods utilize computational schemes to determine the docking configurations of a metabolite's binding sites on proteins or enzymes, as obtained through the use of molecular modeling or quantum mechanics [42, 43]. This approach can provide information about the structure of a metabolite–protein complex at a given binding site through an analysis of the most thermodynamically favorable configurations. These computational methods can result in docking predictions that have a 1.5–2 Å accuracy with success rates of 70–80% [43]. If the location of a binding site is not known in advance, a homology method can be used to predict binding sites on a protein through the use of the protein's amino acid sequence and chemical structures of the metabolites [44]. This method can allow for accurate prediction of ligand–binding proteins and enable the development of a database for these peptide sequences selected for binding to different metabolites [9, 45]. These in silico methods can be combined with in vitro analysis to optimize the structural characterization of metabolite–protein interactions, as demonstrated in NMR experiments [46].

3. Interactions of proteins with hormones and related metabolites

Hormones are chemicals that are secreted by endocrine glands. Hormones play a significant role in many regulation pathways, including metabolism, growth and development [47, 48]. Examples of low mass hormones include various types of steroids (e.g., estrogens and testosterone) or thyroid hormones (e.g., thyroxine) [49–52]. As these chemicals enter the circulation, they are
carried to their target tissue or organ to produce an effect. Many low mass hormones are transported in the bloodstream through their binding to serum proteins [51, 52]. These transport proteins may bind to a broad range of hormones and other targets, as occurs for human serum albumin (HSA), or they may be specific for a given hormone or group of hormones, as is the case for thyroxine-binding globulin (TBG) [49]. Once it has been delivered to its target tissue or organ, the hormone can then bind with a receptor to produce an effect. This section will consider interaction studies that have been reported for several types of hormones or their metabolites with serum proteins and hormone receptors.

### 3.1. Thyroid hormones

Thyroid hormones are a group of iodothyronine compounds that are responsible for metabolism, growth, development, and the regulation of iodine within the body [47]. Many of these hormones are bound in the bloodstream to both HSA through low-to-moderate affinity interactions and to transthyretin or TBG through higher affinity processes [48, 49]. An important compound in this group is the hormone l-thyroxine (L-3,5,3',5'-tetraiodothyronine, or T₄), which can be metabolized to form L-3,5,3'-triiodothyronine (T₃) [53, 54]. Both T₄ and T₃ are actively involved in regulatory processes and are more than 99% bound to transport proteins in blood [49, 53, 54].

Several studies have explored the structural differences between thyroid hormones and related compounds as they bind to serum proteins or cell surface receptors [53–55]. One report utilized HPAC to characterize the binding of T₄, T₃, and related compounds with HSA; the results were used to examine both the affinity constants and thermodynamic properties of these compounds in their interactions with this protein [54]. Some typical results that were obtained in competition studies and through the use of site-specific probes are provided in Table 1. The results indicated that these thyroid hormones were interacting with HSA at both of the major drug-binding sites on this protein (i.e., Sudlow sites I and II) [53, 54]. A comparison of the data obtained for the thyroid hormones and their metabolites indicated that the number and position of iodines, the phenol group, and the thyronine backbone were all important during the binding of these compounds to HSA [54]. Structural studies have also been carried out through the use of modeling and crystallographic data to examine the binding of thyroxine and related compounds to a cell surface receptor for thyroid hormones on αβ3 integrin [55].

### 3.2. Steroid hormones

The protein binding of steroid hormones and their metabolites has also been characterized through a variety of techniques. As an example, the crystal structure of the serum transport protein sex-hormone binding globulin (SHBG) was determined for a complex of this protein with 5α-dihydrotestosterone [56]. SHBG is an important binding agent in blood for many sex hormones and related compounds, including estradiol, testosterone, androste-5-ene-3β,17β-diol, and 5α-dihydrotestosterone [47, 49, 56]. The information that was obtained from the crystal structure for the 5α-dihydrotestosterone/SHBG complex was compared with the results of previous binding studies for steroid hormones with SHBG [57, 58], and this allowed a model of the binding site for these compounds to be developed. This model gave good agreement with prior data from site-directed mutagenesis [59–61] and photolabeling experiments [62, 63] that have been conducted with SHBG [58].

Another structural study looked at the interactions between the human androgen receptor (AR) ligand-binding domain and several androgen-related steroid hormones and metabolites [30]. The compounds that were examined included testosterone, dihydrotestosterone, and tetrahydrogestrinone. An example of some of the results was provided earlier in Figure 1. Both the binding affinity and structural characteristics for the complexes of these agents with AR were explored. Tetrahydrogestrinone was found to have the highest affinity for the AR ligand-binding domain. This strong binding was thought to be due to the presence of multiple hydrophobic interactions for this compound than for the other steroids that were studied. Dihydrotestosterone had a higher affinity than testosterone, an effect that was proposed to be due to the stronger electrostatic interactions between the structure of dihydrotestosterone and the AR binding domain [30].

### 4. Interactions of proteins with fatty acids and related metabolites

Fatty acids can also have significant binding to proteins. These compounds are carboxylic acids that contain hydrocarbon chains with lengths of 4 to 36 carbons. In some fatty acids, the hydrocarbon chain is unbranched and fully saturated, such as myristic acid (C14:0). In others, the chain contains one or more double bonds, as is the case for linoleic acid (C18:2) [48]. Long chain fatty acids (i.e., fatty acids with chains containing 16–20 carbons) are

---

**Table 1.** Site-specific association equilibrium constants ($K_a$) for thyroxine and related compounds with HSA at 37 °C.  

<table>
<thead>
<tr>
<th>Compound</th>
<th>$K_a$ (M⁻¹ × 10⁵)</th>
<th>Sudlow site I</th>
<th>Sudlow site II</th>
</tr>
</thead>
<tbody>
<tr>
<td>Thyroxine (T₄)</td>
<td>L-T₄</td>
<td>1.4 (±0.1)</td>
<td>5.7 (±0.8)</td>
</tr>
<tr>
<td></td>
<td>D-T₄</td>
<td>5.5 (±0.9)</td>
<td>29. (±2)</td>
</tr>
<tr>
<td>Triodothyronine (T₃)</td>
<td>L-T₃</td>
<td>0.170 (±0.001)</td>
<td>0.25 (±0.02)</td>
</tr>
<tr>
<td></td>
<td>D-T₃</td>
<td>1.45 (±0.06)</td>
<td>1.2 (±0.2)</td>
</tr>
<tr>
<td>Reverse triiodothyronine (rT₃)</td>
<td>L-rT₃</td>
<td>1.99 (±0.07)</td>
<td>3.3 (±0.6)</td>
</tr>
<tr>
<td>Diiodothyronine (T₂)</td>
<td>L-T₂</td>
<td>0.53 (±0.03)</td>
<td>1.16 (±0.06)</td>
</tr>
<tr>
<td>Thyronine (T₁)</td>
<td>L-T₁</td>
<td>0.18 (±0.02)</td>
<td>–</td>
</tr>
</tbody>
</table>

*Each value in parentheses represents a range of ±1 S.D. All association equilibrium constants were measured at 37 °C and pH 7.4. The data shown in this table was obtained from Refs. [53, 54]. Sudlow site I and Sudlow site II are also known as the warfarin site and indole site, respectively.*
particularly critical for a diverse set of cellular and metabolic functions. For instance, long chain fatty acids act as fuel that can be stored as triacylglycerols (or triglycerides) and that can be used to generate ATP through β-oxidation in mitochondria and peroxisomes. In addition, fatty acids are the precursors of phospholipids and glycolipids, which are needed for the construction of membranes [64].

Long chain fatty acids such as oleic (C18:1), palmitic (C16:0), linoleic (C18:2), stearic (C18:0), arachidonic (C20:4) and palmitoleic (C16:1) acid are crucial intermediates in lipid metabolism [65]. They tend to have low solubility in water and are typically bound in plasma to proteins, with less than 0.1% being present as non-bound, or “free”, fatty acids. Most of the long chain fatty acids in the blood are transported by HSA [65–68]. HSA carries between 0.1 and 2 mol of fatty acids under normal physiological conditions. However, this value can rise to as high as 6 mol fatty acid per mol of HSA in the peripheral vasculature during fasting or exercise or disease states such as diabetes, liver and cardiovascular disease [67].

Many recent studies have attempted to locate fatty acid binding sites on HSA by using X-ray crystallography or NMR spectroscopy [67, 69–73]. Some typical results of such studies are provided in Figure 3 [74, 75]. In addition, site-directed mutagenesis has been utilized with these methods to see how specific changes in the peptide sequence of HSA will affect its binding properties and structure [67, 69]. Such studies have revealed that five to seven binding sites on HSA may be occupied by medium and long chain fatty acids [71]. These binding sites are asymmetrically distributed across the three domains of HSA, with three of these sites overlapping Sudlow sites I and II [70]. All of these sites have similar structural interactions with fatty acids, providing a deep hydrophobic pocket for the methylene tail and containing two or three polar surface residues nearby which provide a binding location for the carboxylic head group of the fatty acid.

A variety of techniques have been employed to estimate the binding constants for fatty acids at their sites on HSA. The strongest of these interactions have association equilibrium constants that range from 10^5 and 10^6 M^{-1} [66, 74–77]. It has been observed for fatty acids with multiple binding sites on HSA that the value of the individual association constants for each mole of added fatty acid increases as the length of the fatty acid chain was raised [71]. It was later found that the association equilibrium constant for the first bound fatty acid increases with chain length but that this increase does not necessarily occur in a linear fashion; in stead, the affinity is generally dependent on the hydrophobic portion of the fatty acid and how it interacts with HSA [69]. It has further been demonstrated that some fatty acids can have direct competition with drugs on HSA or can lead to allosteric effects during these binding processes [72, 74, 75, 78].

5. Interactions of proteins with drugs and related metabolites

Numerous studies have examined the interactions of drugs and their metabolites with proteins. Like low mass hormones, many drugs and their metabolites are transported throughout the body through the use of serum transport proteins. Approximately 43% of the 1500 most commonly used pharmaceuticals have at least 90% binding to such binding agents [35, 79]. These interactions usually involve proteins that can bind to a broad range of targets, such as HSA and α1-acid glycoprotein (AGP), and can play a significant role in determining the activity, distribution, rate of excretion or metabolism, and toxicity of many pharmaceutical agents in the body [80]. In recent years there has also been interest in how the presence of drug metabolites can affect the distribution, apparent activity, and protein interactions of the parent drug [81].

5.1. General effects of metabolites on drug–protein interactions

Many studies have investigated the difference between drugs and their metabolites in their overall binding in serum or to specific serum proteins. For instance, equilibrium dialysis was used to examine the binding of propionate and its major metabolite to human serum and isolated serum proteins such as AGP [82] and the binding of acetohexamide and its metabolite (−)-hydroxyhexamidine to HSA [83]. Another study utilized a similar approach to investigate the binding by tolterodine and its 5-hydroxymethyl or N-dealkylated metabolites to human serum, HSA and AGP [84]. Equilibrium dialysis was further used to measure the binding of tizoxanide, an active metabolite of the drug nitazoxanide, with albumin and AGP [85].

A few studies have been conducted to provide a more detailed comparison of the binding regions and binding constants for drugs and their metabolites on serum proteins. As an example, HPAC and competition studies have been used to compare the binding regions on HSA for the drug phenytoin and its two major metabolites: 5-(3-hydroxyphenyl)-5-phenylhydantoin and 5-(4-hydroxyphenyl)-5-phenylhydantoin (i.e., m-HPPH and p-HPPH, respectively) [81, 86]. In an examination of both the major and minor drug binding regions of HSA, phenytoin was found to have direct binding at Sudlow site II and the digoxin site, with association equilibrium constants at these regions in the range of 0.65–1.04 × 10^5 M^{-1} at 37 °C and pH 7.4 (see Table 2). The same drug had allosteric effects plus possible direct binding at Sudlow site I and the tamoxifen site [86]. However, m-HPPH and p-HPPH only had significant interactions with Sudlow site II, with binding constants of 0.32–0.57 × 10^5 M^{-1} for this region [81]. Thus, the parent drug and its metabolites not only had different affinities for HSA but also had differences in the number of their interaction sites with this protein [81, 86].

5.2. Effects of chirality on drug metabolite–protein binding

Another factor to consider for drug- and drug metabolite–protein binding is the effect of chirality on these interactions. Chiral drugs have been estimated to represent 40–50% of all drugs that are currently on the market [87, 88]. The separate chiral forms for some drugs can exhibit a wide variation in their toxicology, pharmacokinetics and metabolism. In the extreme case, one enantio-
mer may produce the desired function in treatment while another may be inactive or even produce undesired or toxic effects. This is because within the body numerous compounds (i.e., enzymes, plasma proteins, and other biomolecules) work as chiral selectors, causing them to sometimes bind or metabolize each chiral form of a drug differently [89–95].

These differences have made it possible in the past to use protein-based HPLC columns, such as those containing serum proteins, for separating the various chiral forms of many drugs [93–95]. The same approach has been utilized to separate and measure chiral drugs and their metabolites in biological samples. For instance, an AGP column was recently used with fluorescence detection to measure the enantiomers of tramadol and its two major metabolites, O-desmethyltramadol and N-desmethyltramadol, in plasma samples (see Figure 4). This method was then used to examine the pharmacokinetics for each of these compounds in the body [96]. A similar approach has been used with LC-MS to examine the chiral forms of methadone and its metabolites 2-ethylidene-1,5-dimethyl-3,3-diphenylpyrrolidine and 2-ethyl-5-methyl-3,3-diphenyl-1-pyrroline in hair samples from patients undergoing methadone maintenance therapy [97].

As has been observed for their parent drugs, the different forms of a chiral metabolite can differ in how they interact with proteins. One report compared the chiral forms of oxybutynin and its metabolite, N-desethyloxybutynin, in their binding and competition on HSA and AGP. The results showed that the affinity of oxybutynin enantiomers on AGP was much higher than on HSA, and that the enantiomers of N-desethyloxybutynin and oxybutynin were all bound by the same site on AGP [33]. Another study involving the phenytoin metabolites m-HPPH and p-HPPH compared the dissociation rates of the chiral forms of these metabolites from an HPLC column containing immobilized HSA [98]. Dissociation rate constants of 8.2–9.6 s\(^{-1}\) were obtained at pH 7.4 and 37° C for the enantiomers of m-HPPH, while values of 3.2–4.1 s\(^{-1}\) were obtained for the enantiomers of p-HPPH. These results were then used along with separate estimates of the association equilibrium constants to also compare the association rate constants for these metabolites and their enantiomers [98].

### 5.3. Use of binding data to characterize protein interaction sites for drug metabolites

A number of reports have used binding data for drugs, their metabolites and related analogs to learn about the binding sites of these compounds on a protein. Binding and retention data that have been acquired by HPAC have been used to examine the binding of several types of compounds with immobilized serum proteins. This approach has been used to examine the binding of warfarin and coumarin compounds to HSA [99–101], as well as the

---

**Table 2. Association equilibrium constants (K\(_a\)) and types of binding for phenytoin and its metabolites to various regions on HSA.**

<table>
<thead>
<tr>
<th>Binding region on HSA</th>
<th>Drug or drug metabolite</th>
<th>Sudlow site I(^a)</th>
<th>Sudlow site II(^b)</th>
<th>Digitoxin site</th>
<th>Tamoxifen site</th>
</tr>
</thead>
<tbody>
<tr>
<td>Phenytoin</td>
<td>Allosteric effects + possible direct binding</td>
<td>No binding</td>
<td>Direct binding</td>
<td>Direct binding</td>
<td>No binding</td>
</tr>
<tr>
<td>m-HPPH</td>
<td>Direct binding</td>
<td>Direct binding</td>
<td>Direct binding</td>
<td>No binding</td>
<td></td>
</tr>
<tr>
<td>p-HPPH</td>
<td>Allosteric effects + possible direct binding</td>
<td>No binding</td>
<td>No binding</td>
<td>No binding</td>
<td></td>
</tr>
</tbody>
</table>

a. All of these results were obtained at 37 °C in pH 7.4, 0.067 M phosphate buffer and are based on data from References [81, 86].

b. Sudlow sites I and II are also known as the warfarin site and indole site, respectively.

---

**Figure 4.** Chiral separation and analysis of tramadol and its major metabolites using HPLC and a column containing immobilized AGP as the stationary phase. The results in (a) are for a blank human plasma sample. The results in (b) are for a plasma sample taken from a volunteer 2.5 h after receiving a 100 mg dose of racemic tramadol. Symbols: enantiomers of tramadol, +(T) and −(T); enantiomers of the metabolite O-desmethyltramadol, +(M1) and −(M1); enantiomers of the metabolite N-desmethytramadol, +(M2) and −(M2); and internal standard (fluconazol), IS. Adapted with permission from Reference [96].
binding of l-tryptophan and various indole compounds to this protein [102, 103, 104]. The same general method has been used to compare the binding of several sulfonylurea drugs with HSA and various preparations of glycated HSA [39, 105–110].

If a relatively large group of compounds is considered in a binding study, the results can be used to create a quantitative structure-retention (or reactivity) relationship (QSSR) to describe the site at which these agents are binding to a protein [111–114]. For instance, binding studies based on HPLC or CE using serum proteins can be used to mimic biological systems and to quickly study how changes in the structure of an applied drug or analog will alter these interactions [115, 116]. This format has been used to build models that describe the binding of HSA with benzodiazepines [117–119]. Such an approach has also been utilized to examine the binding of AGP with beta-adrenolytic drugs, antihistamines, amino alcohols, cyclic vinca alkaloid analogs, and quinazolone derivatives [116, 120–125].

6. Interactions of proteins with xenobiotics and related metabolites

The term “xenobiotics” refers to chemicals that are produced synthetically and that are not normally found in biological organisms [126]. Drugs represent one type of xenobiotic, but others include environmental pollutants and food additives [126–128]. When they enter the body, xenobiotics can be metabolized through various enzymatic processes. The resulting metabolites, in turn, can sometimes interact with proteins and compete for endogenous compounds for common binding agents [126, 129].

Several studies have examined the effects that xenobiotics and their metabolites may have on hormone–protein binding [127, 130, 131]. For example, the effect of polychlorinated diphenyl ethers (PBDEs) on the binding of thyroid hormones to serum proteins has been examined [131]. It has been suggested in several studies that environmental exposure to PBDEs can result in decreased thyroid hormone concentrations in serum, leading to possible neurotoxicity and behavioral effects [131–133]. This effect may be linked to the fact that, when metabolized, PBDEs become hydroxylated and produce a chemical structure similar to that of T₄ and its metabolites. It has been further found that PBDE metabolites are able to bind to T₄-binding proteins in serum, which could result in the displacement of thyroid hormones. One study examined the binding of transthyretin and TBG with fourteen hydroxylated PBDE compounds through various methods. A fluorescence displacement assay indicated that hydroxylated PBDEs could compete with T₄ for binding sites on transthyretin, while work with circular dichroism indicated that hydroxylated PBDEs could bind to the same sites as T₄ on TBG and transthyretin [131].

Another report examined the effects for a number of xenobiotics and their metabolites on the binding of 17β-estradiol to the estrogen receptor and on the binding of 5α-dihydrotestosterone to the androgen receptor, androgen-binding protein, and SHBG [127]. Compounds that were tested included hexachlorocyclohexane, DDT, methoxychlor, pentachlorophenol, and nonylphenol. It was found that some of these xenobiotics and metabolites could cause a significant decrease in the binding of 5α-dihydrotestosterone or 17β-estradiol to their binding proteins. It was further found that binding by these xenobiotic agents could be selective for the steroid receptors and binding proteins that were tested [127].

Polyphenolic compounds are flavonoids that are often used as dietary supplements [128]. Ultrafiltration and CE were used to examine the binding of these compounds to the human serum proteins HSA and AGP. Although similar in structure, these compounds did vary in their affinity toward HSA, with a high level of binding being observed for those compounds with hydrophobic properties and a carbonyl at position C(4) in their structure. It was further noted that these hydrophobic properties did not play a major role in the ability of polyphenolic compounds to bind with AGP [128].

7. Variations in protein structure and binding due to metabolic processes

Another way in which changes in metabolites may affect solute protein interactions is through changes that are created in the structure of the protein. In some cases, these changes may be a direct result of the modification of a protein by a metabolite (e.g., glycation, as discussed in the next section) [39, 80, 93, 134]. In others, this change may be a response to differences in a protein’s environment that are created as the metabolic profile is altered (e.g., as might occur through oxidation) [135, 136]. This section will discuss both types of effects using changes that have been observed in serum transport proteins and binding agents as examples.

7.1. Human serum albumin

One protein that has been found to be altered by some metabolic diseases is HSA. As has been indicated earlier, HSA is a serum protein that plays a fundamental role in the reversible binding and transport of metabolites, drugs and various endogenous ligands, such as fatty acids [65, 137]. HSA is normally found in blood at concentrations ranging from 30–50 g/L and accounts for approximately 60% of the total serum protein content [65]. Binding to HSA is known to greatly influence the pharmacokinetics and activity of many common drugs [49, 138–140]. In addition, HSA can increase the solubility of lipophilic drugs, sequester toxins, and act as an important antioxidant in plasma [49, 65].

Several past studies have noted that the chemical modification of HSA can alter its binding to drugs, hormones and other solutes. For instance, the reaction of HSA with p-nitrophenyl acetate, which is thought to mainly modify Tyr-411 at Sudlow site II, can change the binding of various solutes with this protein [141]. The modification of Trp-214 by p-nitrophenylsulphenyl chloride has been demonstrated to change the stereoselectivity and binding affinity of Sudlow site I of HSA [142]. Similar work has been presented that has examined the effects of modifying the lone free cysteine group on HSA by reacting this protein with ethacrynic acid [143, 144].

Diabetes is a metabolic disease in which the structure of HSA can be modified. This disease is actually a group of disorders that are characterized by abnormally high levels of blood glucose (i.e., hyperglycemia) that result from insulin deficiency and/or insulin resistance [145]. Many of the long term complications of diabetes, such as heart disease and nerve damage, are associated with the non-enzymatic glycation of proteins [145, 146]. Glycation starts with the nucleophilic attack of a reducing sugar (e.g., glucose) onto some of the primary amine groups on proteins to form a reversible Schiff base (see Figure 5). This intermediate can then slowly rearrange to form a more stable Amadori product [145–147]. Oxidation of the Amadori products or free sugars can also generate reactive α-oxoaldehydes that can react with both lysines and arginines on proteins to form advanced glycation end-products (AGEs) [147].

In recent years, it has been found that glycation can affect the binding of several endogenous and exogenous solutes with HSA. For example, L-tryptophan is an essential amino acid [148] and has been extensively used as a site-selective probe for Sudlow site II of glycated HSA and normal HSA [105, 106–108, 149]. Recent binding studies using glycated HSA with levels of modification
similar to those found in diabetes found an increase of 4.7–5.8-fold in the affinity of L-tryptophan for this protein at 37 °C [106, 149]. Sulfonylurea drugs are a group of anti-diabetic drugs that are used in the management of type 2 diabetes; these drugs are also highly bound to serum proteins such as HSA. Binding studies based on HPAC have found that glycation can affect the binding strength of these drugs to HSA, with both the degree of glycation and the specific type of drug influencing the size of the change [39, 105–110].

As indicated in the last section, fatty acids are the major endogenous ligands of HSA and are also known to have many binding sites on this protein [75]. Reports that have examined the combined effect of glycation and the presence of various fatty acids on the binding of sulfonylurea drugs to HSA have found that glycation increases the overall affinity of these drugs to HSA, while the addition of increasing amounts of fatty acids causes a decrease in affinity [74, 76]. It has further been noted that glycation could produce changes of at least 3–5-fold in the affinities of some fatty acids at their sites of competition with sulfonylurea drugs when comparing the binding of these solutes to normal HSA [76].

Methylglyoxal is a highly reactive metabolite of glucose that has been implicated in several chronic diseases associated with diabetes [150, 151]. The elevated concentrations of methylglyoxal in diabetes patients can also lead to protein modification and the formation of AGEs through the reaction of methylglyoxal with arginine or lysine residues. A recent report using quantitative MS and multiple reaction monitoring found that a major site for modification by methylglyoxal on HSA occurs at Arg-257, which is located in Sudlow site I. Molecular modeling conducted in the same study indicated that a decrease in binding by warfarin may occur due to these modifications when comparing glycated HSA and normal HSA [151].

### 7.2. Alpha-<br>acid glycoprotein

A second type of serum transport protein that can be affected by metabolic diseases is AGP. AGP is an acute-phase protein that is responsible for binding and delivering numerous basic and neutral drugs in the bloodstream [121]. The concentration of AGP in blood can vary over a wide range and is affected by systemic tissue injury, inflammation and infection. In addition, the glycosylation of AGP can be altered in some disease states, such as rheumatoid arthritis, systemic lupus erythematosus and autoimmune thyroid disease [152]. These changes are important because they can also alter the binding of drugs to AGP. As an example, the affinity of disopyramide for AGP has been found to be affected by the biantennary glycan structures for this protein [153, 154]. It has also been reported that genetic variants of AGP can have a significant effect on binding by chiral drugs such as disopyramide and warfarin [153, 155].

A number of reports have looked at how changes in AGP binding can affect parent drugs compared to their metabolites. One study evaluated the effect of AGP on lidocaine and its active metabolites monoethylglycinexylidide and glycinexylidide during continuous epidural anesthesia in infants and young children. The results indicated the AGP concentration in plasma could be used as an index to monitor and prevent the toxicity caused by the accumulation of monoethylglycinexylidide during the continuous administration of lidocaine [156]. Another report looked at the concentrations of vecuronium and its metabolite 3-OH desacetylvecuronium in children who were receiving phenytoin or carbamazepine for chronic anticonvulsant therapy [157]. These last two drugs were of interest because many anticonvulsant drugs have been shown to increase the concentration of AGP in plasma, which can then increase protein binding to cationic drugs and alter their distribution. It was found that the increase in AGP concentration associated with the anticonvulsant therapy did not significantly contribute to resistance to vecuronium [158].

### 7.3. Lipoproteins

Lipoproteins are another set of binding agents in serum that can be affected by metabolic diseases. Lipoproteins are macromolecular complexes of proteins and lipids that transport hydrophobic lipids and related compounds, such as cholesterol and triglycerides, throughout the body [158–161]. Lipoproteins are also known to interact with several basic and neutral hydrophobic drugs in blood [99, 162–172]. Examples of drugs that bind to lipoproteins are propranolol and verapamil [37, 162–175].

Lipoprotein concentrations in blood can vary with different disease states. For example, the levels of low-density lipoprotein (LDL) in plasma can increase in diseases such as atherosclerosis and hyperlipidemia [176]. In addition to changes in the levels of lipoproteins in the circulation, metabolic diseases often result in modifications in lipoprotein structures. For instance, increased amounts of LDL that have been modified by AGEs are found in individuals with diabetics and non-diabetics with renal failure. Glycation of LDL may lead to the formation of foam cells and an increase in atherosclerosis. In addition, glycated LDL is more susceptible to further modifications due to oxidation [135].

The oxidation of lipoproteins occurs through free radicals, such as peroxyl radicals, which are released from cells and chemical reactions [136]. These radicals can react with lipoproteins, depleting the particle’s antioxidant defense and initiating oxidation of the lipid core. In the later stages of this process, the surface protein also becomes modified. The oxidation of lipoproteins, specifically LDL, leads to atherosclerosis [136]. In addition to the increased risk of atherosclerosis, oxidized lipoproteins may also impact the
ability of the complexes to bind and carry basic and neutral drugs throughout the body [176].

The effects of LDL oxidation on drug binding have been evaluated by using CE and using verapamil and nilvadipine as models for basic and neutral drugs, respectively [176]. It was found that the affinity of these drugs increased with the amount of LDL oxidation. In addition, the binding of verapamil was increased more than it was for nilvadipine, suggesting that basic drugs were more sensitive to oxidation effects. No stereoselective binding was detected between LDL and these model drugs at any oxidation state [176]. However, other studies based on HPAC have noted different binding for the chiral forms of some drugs to LDL [174, 175].

8. Conclusion

The field of metabolomics has seen great growth in recent years because of the wealth of information it can provide about biochemical pathways and processes. This review examined previous reports that have looked at the interactions of metabolites with proteins. The first topic discussed was an overview of techniques that have been used to characterize and study metabolite–protein binding. These methods have been used in vitro and in vivo to provide information on the structures of metabolite–protein complexes and to examine the nature of metabolite–protein interactions. Computational studies using in silico tools have been used to provide additional data on metabolite–protein complexes and interactions.

This review next described numerous studies that have investigated the binding of various types of small solutes and their metabolites with proteins. This included work that has been carried out with hormones, fatty acids, drugs or other xenobiotics, and their metabolites with transport proteins and receptors. These examples have considered the structures of the resulting solute–protein complexes, the nature of the binding sites, the strength of these interactions, the variations in these interactions with solute structure, and the kinetics of these reactions. Studies that have examined the effects of various metabolic processes on the structure and activities of proteins, and on the corresponding interactions of solutes with these proteins, were also summarized.

Although most past work in metabolomics has been concerned with the structure and analysis of metabolites, research in metabolite–protein interactions is still a relatively new area. Based on the research that has already been carried out, it is already clear that data on metabolite–protein interactions can provide useful information on biological processes that involve hormones, drugs and other low mass solutes. It is further expected that this type of research will continue to grow in the future as metabolomics becomes more widely used in biomedical research, pharmaceutical science, and personalized medicine.

Acknowledgments — This research was supported by the National Institutes of Health under grants R01 GM044931 and R01 DK069629 and was conducted in facilities that were renovated under NIH grant RR015468-01. R. Matsuda was supported, in part, under a fellowship through the Molecular Mechanisms of Disease program at the University of Nebraska.

References
