Review article

An overview of albumin and alpha-1-acid glycoprotein main characteristics: highlighting the roles of amino acids in binding kinetics and molecular interactions

Michel Bteich*

Department of Environmental and Occupational Health, School of Public Health, Université de Montréal, Montréal, Québec, Canada

ARTICLE INFO

Keywords:
Plasma protein binding
Protein-ligand complex
Albumin
Alpha-1-acid glycoprotein
Molecular interactions

ABSTRACT

Although Albumin (ALB) and alpha-1-acid glycoprotein (AGP) have distinctive structural and functional characteristics, they both play a key role in binding a large variety of endogenous and exogenous ligands. An extensive binding to these plasma proteins could have a potential impact on drugs disposition (e.g. bioavailability, distribution and clearance), on their innocuity and their efficacy. This review summarizes the common knowledge about the structural and molecular characteristics of both ALB and AGP in humans, and about the most involved amino acids in their high-affinity binding pockets. However, the variability in residues found in binding pockets, for the same species, allows each plasma protein to interact differently with the ligands. The protein-ligand interaction influences differently the disposition of drugs that bind to either of these plasma proteins.

The content of this review is useful for the design of new drug entities with high-binding characteristics, in qualitative and quantitative modelling (e.g. in vitro-in vivo extrapolations, 3D molecular docking, interspecies extrapolations), and for other interdisciplinary research.

1. Introduction

In the human body, there are many plasma proteins that allow the binding of xenobiotics (drugs and/or contaminants) such as albumin (ALB), alpha-1-acid glycoprotein (AGP), lipoproteins (HDL, LDL, VLDL, and chylomicrons), fibrinogen, C-reactive protein, transferrin, α, β, γ-globulins, etc (Bowman, 1993; Kerns and Di, 2008; Putnam, 1984). ALB and AGP are the most important for two reasons: first, their main role is carriers and storage depot of the majority of exogenous and endogenous substances into the bloodstream (Lehman-McKeeman, 2013); second, ALB and AGP may be good biomarkers for inflammation and liver disease. They are also potential biomarkers for all-cause mortality, along with citrate and VLDL particle size (Fischer et al., 2014).

Although both ALB and AGP have a key role in clinical pharmacokinetics (PK) and pharmacodynamics (PD) research, these plasma proteins have never been presented by contrasting their main characteristics from top to bottom. Therefore, the main objective of this review is to give an overview of the common knowledge about ALB and AGP in humans by contrasting their structural and functional characteristics. The composition in amino acids is also addressed for main sites of both plasma proteins, in an effort to better understand the differences in the molecular interactions occurring between each plasma protein and its ligands.

2. Common knowledge about albumin and alpha-1-acid glycoprotein

2.1. Albumin

ALB is a single chain, non-glycosylated protein, primarily synthesized in the liver (Miller, 1951; Peters and Anfinsen, 1950) and rapidly secreted in vesicles (13.3 g per day) to the extracellular space (Fries et al., 1984; Peters Jr., 1996). In humans, a mature ALB has 585 amino acids (AA) (Meloun et al., 1975; Minghetti et al., 1986) and a molecular weight of 66.5 kDa (Putnam, 1984) with great interspecies similarities despite the variability in some residues (Brown, 1976; Majorek et al., 2012). It is the most abundant protein in mammals, with almost 60% of the total amount of plasma proteins, and a physiological concentration varying between 3.0 to 5.0 g/dL (500 μM–750 μM, mean value ~600 μM or ~4.5 % w/w) (Peters Jr., 1996). It has acidic characteristics, a negative net charge (-15 for humans) at the physiological pH of 7.4, and an isoelectric

* Corresponding author.
E-mail address: michel.bteich@umontreal.ca.

https://doi.org/10.1016/j.heliyon.2019.e02879
Received 20 October 2018; Received in revised form 4 January 2019; Accepted 15 November 2019
2405-8440/© 2019 The Author. Published by Elsevier Ltd. This is an open access article under the CC BY-NC-ND license (http://creativecommons.org/licenses/by-nc-nd/4.0/).
point of 4.7 (Peters Jr., 1996; Rothschild et al., 1988).

The 17 conserved disulfide bridges of cysteine (Cys) bonds ensure the flexibility of its conformation (normal or N state) (Weber, 1975) and the stability of its structure at pH 7.4 (Markus and Karush, 1957), even in extreme conditions, noticeably in all mammalian serum ALB (Majorek et al., 2012). These bridges contribute in prolonging its biological lifetime (Peters Jr., 1996), and in conserving the three-dimensional structure among different species (Nurdiansyah, Rifa‘i, & Widodo, 2016). The neonatal Fc receptor (FeRn) may also have a role in prolonging the lifespan of ALB (Chaudhury et al., 2003), as it binds to ALB (domain III) in a pH-dependent internalization process of the complex with immunoglobulin G (IgG) (Andersen et al., 2006; Goebl et al., 2008; Sand et al., 2014). Thus, FeRn receptor plays a fundamental role in homeostatic regulation of both IgG and ALB, and in helping these proteins escape intracellular degradation by recycling them back to the systemic circulation (Chaudhury et al., 2003; Sand et al., 2014).

The disulfide bonding patterns give ALB a heart-like crystal structure in humans (3.2 Å) (He & Carter, 1992) and resolved at higher resolution (2.5 Å) (Curry et al., 1998). It has about 67% of α-helices, 10% β-turns, but no β-sheet (Peters Jr., 1996). The flexibility of ALB allows for the binding with a high affinity of a wide variety of xenobiotics at its two primary sites: Sudlow I and II (Fanali et al., 2012; Peters Jr., 1996; Sudlow, Birkett, & Wade, 1975, 1976). Sudlow's site I is known as the warfarin-azapropazone binding cleft (Fehske et al., 1982), while Sudlow's site II is the ibuprofen-diazepam (or indole/benzodiazepine) binding cleft (Ascenzi and Fasano, 2010; Fanali et al., 2012; Fasano et al., 2005).

The structural representation contains three homologous domains (Brown, 1976; Dockal et al., 1999): I (25–222 AA), II (223–414 AA), and III (415–609 AA) with nine double loops repeated in triplet fashion (Peters Jr., 1996; Rothschild et al., 1988). Double loops 1–2, 4–5, and 7–8 are grouped as subdomains IA, IIA, and IIIA, respectively. Single loops 3, 6, and 9 are represented by subdomains IB, IIB, and IIIB, respectively. In each domain consisting of about 10 α-helices, there is a pair of helicoidal subdomains named A (6 α-helices) and B (4 α-helices) that are connected by random coils (Ascenzi and Fasano, 2010). Interactions in-between domains favour an asymmetrical environment for allosteric ligand binding (Curry et al., 1998; Fasano et al., 2005; Peters Jr., 1996).

ALB is characterized by its two hydrophobic cavities in site I (subdomain IIA) and site II (subdomain IIIA), with polar residues in the inner surroundings and at the entrance of the two pockets (Ghumant et al., 2005). Bulky heterocyclic anions preferentially bind to the first main site, and the second is preferred by aliphatic and aromatic carboxylates (Ascenzi and Fasano, 2010). Although a third major binding site (subdomain IB, or lidocaine site) has been recently suggested to exist by Zsila (2013); further studies on binding with a high affinity to this site are needed before ALB can be represented as a conventional three-site model.

ALB displays up to seven binding sites with different affinities for free long-chain fatty acids (FAs) which may be associated to a local conformational change of the secondary structure (e.g. palmitic acid, oleic acid and arachidonic acid) (Fanali et al., 2012; Ghumant et al., 2005). The FAI to FA7-occupied sites accommodate the binding of endogenous and exogenous substances to the primary sites I and II (Ascenzi and Fasano, 2010; Fanali et al., 2012). This kind of accommodation would increase the affinity of the ligand, making it a cooperative interaction (Lehman-Koeneman, 2013). Conversely, some studies reported a possible competition between FAs and some ligands (Chakrabarti, 1978), implying a possible displacement from the binding site due to allosteric modulation (Ascenzi and Fasano, 2010; Fasano et al., 2005). Because the binding of small ligands might not completely fill the hydrophobic pocket in Sudlow's site I; the site is predisposed to the insertion of water molecules, and thus the polar residues by hydrogen bonding (H-bonds) (Curry, 2009; Ghumant et al., 2005). This effect influences the adaptability of the site to the binding interaction. Multiple monolayers of water molecules reversibly bond their hydrogen atoms to the residues at the surface of the protein and navigate along with it in the extracellular space (Peters Jr., 1996).

ALB has many prominent physiological functions (Peters Jr., 1996). It is a leading carrier and a depot in blood circulation for endogenous substances (e.g. bilirubin, FAs, hormones, bile acids, etc.) (Fanali et al., 2012; Kurtzhal et al., 1997; Roda et al., 1982), exogenous substances (e.g. heavy metals Hg, Cd and Pb, drugs, contaminants) (Li et al., 2007) and other essential ligands (e.g. transition metal ions Zn²⁺, Ca²⁺, nitric oxide, metallotheonine, vitamins, haem) (Ascenzi and Fasano, 2009; Bal et al., 2013; Hou et al., 2008; Quiming et al., 2005; Sadler and Viles, 1996). ALB regulates the coloidal osmotic pressure with its high concentration in plasma by keeping the blood within vessels. It has enzymatic properties that are useful in metabolism and detoxification of exogenous and endogenous ligands (Fitzpatrick and Wynalda, 1983; Kragh-Hansen et al., 2002; Yang et al., 2002). ALB has a buffering capacity of the blood but to a lesser extent than haemoglobin (Pettifer, 2003). ALB has the largest thiol pool in plasma (Turell et al., 2013). It plays an antioxidant role (Tiu, 1994) as it scavenges reactive oxygen species and reactive nitrogen species (Quinlan et al., 2005). This scavenging activity of ALB is largely dependent on the redox state of the residue Cys34 (Anraku et al., 2013; Stamler et al., 1992).

ALB is a negative acute-phase protein (Bowman, 1993), as its concentration decreases in case of inflammation under the influence of cytokines (IL6, IL1 and TNFα) (Kniec, 2001). Thanks to ALB’s many appreciable characteristics, inter alia, in vivo half-life (~19 days), stability, and binding versatility (Peters Jr., 1996); medical and pharmaceutical applications that include ALB in in-vitro and in vivo conditions (e.g. nanomaterials, biomarkers, toxicokinetics, preclinical, etc.) are becoming numerous with technological advancements in health science disciplines (e.g. genetics, oncology, nanotechnology, biochemistry, toxicology, therapeutics, etc.). For further exploration of the use of ALB, examples are given in two reviews by Peters Jr. (1996) and Otagiri and Chuang (2016).

2.2. Alpha-1-acid glycoprotein

The plasma protein AGP, also known as orosomucoid (ORM), is a glycosylated single chain protein (Fournier, Medjoubi-N, & Porquet, 2000; Kremer et al., 1988; Schmid, 1950, 1989; Weimer et al., 1950). Like ALB, AGP is mainly synthesized (10 mg/kg/day) (Lentner, 1984) and catalysed in the liver (Sarcione, 1963). However, asialo- or agalacto-AGP is internalized into hepatocytes using a specific membrane asialoglycoprotein receptor (ASGP-R) for its degradation (Taguchi et al., 2013). This process is only possible when its binding site is vacant and after AGP's conformation changes (Meijer and Nijsen, 1991; Meijer and van der Suijs, 1989). Its half-life is 2–3 days, so a high AGP concentration is easily cleared from the body even after its induced synthesis (Bree et al., 1986). Unlike ALB, AGP is scarce in plasma as it accounts for only 3% of plasma proteins. In healthy adults, the physiological concentration varies between 0.05 to 0.1 g/dL (10 μM–20 μM; mean value ~15 μM) (Kremer et al., 1988); some references even give a maximum normal value of 0.140 g/dL (31 μM). The plasma concentration of AGP increases with age and is sex-dependent (Issari and Dayton, 2001). AGP is an acidic (pKa = 2.6) and highly soluble protein in water and other polar organic solvents (Schmid, 1989). Its net charge is negative and depends on the nature of its carbohydrate entities. Its isoelectric point varies between 2.8 and 3.8, at the physiological pH of 7.4 (Fournier et al., 2000).

In humans, the mature form of AGP consists of 183 AA. It has a molecular weight of about 44 kDa (Kremer et al., 1989) or 41–43 kDa according to Hochepied et al. (2003). The AGP protein consists of about 59% peptide residues and 41% carbohydrates - of which about 11% are sialic acids (or 12% according to Hochepied et al. (2003)). This high sialic content contributes to the negative charge on its surface (Kremer et al., 1988; Schmid et al., 1977) and determines the nature of interactions that this glycoprotein can have with biological membranes.
There are five structurally heterogeneous carbohydrate entities (five N-glycans) with mono-, bi-, tri-, and tetra-antennary glycans covalently linked to five asparagine (Asn) residues, which form N-glycosylation sites (Asn-15, Asn-38, Asn-54, Asn-75, Asn-85) to the most active gene ORM1 (variant F1*S) between the two main genes of AGP (i.e. ORM1 and ORM2) (Dente et al., 1987; Fournier et al., 2000; Taguchi et al., 2013; Yoshima et al., 1981; Yuasa et al., 1997).

Physiological (e.g. inflammation vs healthy status) and environmental conditions influence the variability in glycosyl branching (Gornik et al., 2009). The heterogeneity of glycosylation does not only influence the binding of xenobiotics but also affects the kinetics of the protein itself (half-life, uptake, and catabolism) (Fernandes et al., 2015; Gross et al., 1989). The carbohydrate moiety contributes to this protein's stability, native conformational structure, and solubility (Bürgi, 1989). This moiety of asialo-AGP interacts with plasma membranes and receptors (i.e. ASGP-R), and is internalised by endosomes under lysosomal action or recycled to the extracellular space (Meijer and van der Slujs, 1989; Taguchi et al., 2013).

There are two main variants (F1*S and A) implicated in high affinity binding (Eap et al., 1988). The difference in AA residues between these variants is due to 21 AA substitutions, which affects the affinity and the stereoselectivity of the binding to AGP (Dente et al., 1987; Fitos et al., 2010; Hervé et al., 1998; Hervé et al., 1993). The variant F1*S is predominant (Monteil et al., 1990), and in normal conditions its expression is two-to-three-fold higher than variant A (Yuasa et al., 1997).

AGP contains two disulphide bonds in all its main allelic variants that sustain the stability of its secondary structure (Kremers et al., 1988; Schmid et al., 1973). These bonds, like sialic acids, contribute to AGP's potential interactions with biomembranes (Nishi et al., 2006). The native protein consists of 15% α-helices, 41% β-sheets, 12% β-turns, 8% bends, and 24% unordered structure at pH 7.4 (Kopecký et al., 2003). The α-helical content may increase upon disulfide reduction and in the presence of biomembranes (Nishi et al., 2006). AGP only has one primary saturable site that could be for some ligands of high-affinity and low-capacity characteristics formed by the tertiary structure (Müller, 1989), and located in the hydrophobic crevice consisted of three lobes (Maruyama et al., 1990; Schonfeld et al., 2008). While many other sites with different capacities and low affinities exist, only the main site is of clinical relevance (Müller, 1989).

Although the biological functions of AGP are not clear, it is worth noting a few of its physiological activities. Like ALB, AGP is a leading plasma protein in transporting various endogenous (e.g. serotonin, compounds with centrally located negative or electronegative features and glucocorticoids (Baumann et al., 1989; Israili and Dayton, 2001; van Dijk et al., 1991), or in response to various stress stimuli (e.g. infection, burn, surgery, trauma, or cancer) (Israili and Dayton, 2001; Jackson et al., 1982; Kremers et al., 1988; Piafsky, 1980). In the case of acute phase reactions, the normal ratio between the two main variants (F1*S and A) increases more than three-fold, marking a significant production of the F1*S variant (Taguchi et al., 2013). This variation could be about eight times higher in case of lymphoma, melanoma, and ovarian cancer (Taguchi et al., 2013). In cancer biomonitoring, specific glycans on the surface of AGP could be used as biomarkers in the progression of breast cancer (Abd Hamid et al., 2008).

3. Amino acids in high affinity binding sites

The affinity binding is a descriptive characteristic estimated for a bundle of amino acids, in a site of a plasma protein, that interacts and accommodates ligands to bind into binding cavities. In plasma proteins, some residues have a more important role than others for binding ligands in high-affinity binding cavities, and this discrepancy in importance depends, inter alia, on the functional groups of each ligand. Also, some changes in residues may occur and significantly modulate the inner composition of high-affinity binding pockets and their peptide outer surroundings. For example, polar AAs may be replaced with neutral or hydrophobic ones, or positive residues may be replaced with negative ones and vice versa. Consequently, the alteration of nucleotides and peptides influences the extent of binding to plasma proteins (Huang and Ung, 2013; Otagiri and Chuang, 2016). To better understand this influence, the composition in AAs of high-affinity binding sites of most frequent variants for human ALB and AGP is discussed next (see Tables 1 and 2).

3.1. Mature ALB in humans

The free Cys34, known for its antioxidant role, is well conserved in humans (position 58 in native ALB). About 70–80% of Cys34 in plasma contains a free sulphydryl group, while the other 25% forms a disulfide bond with small sulphydryl compounds, such as cysteine, homocysteine, or glutathione (Oettl and Stauber, 2007). Another well-conserved residue in ALB is Trp214. It is situated in a major interdomain cluster of hydrophobic residues (position 238 in native ALB), and precisely in a conserved sequence of loop 4 (binding pocket of Sudlow I). This hydrophobic residue is important, just like disulfide bridges, for holding together the two halves of the heart-shaped ALB (Carter and Ho, 1994). The absorption of UV light by ALB is particularly attributed to Trp214, which is quite appreciated as a fluorescent residue in binding studies for the measurement of thermodynamic parameters (e.g. fluorescence quenching) (Tayeh et al., 2009).

In subdomain IIa of the mature ALB, site I has the voluminous cavity and it is constituted from 177 AA to 291 AA. In subdomain IIia, site II has the smaller cavity and it is constituted from 367 AA to 491 AA. The residues most involved in binding ligands to the two high-affinity binding cavities of human ALB, other than Trp214 and Cys34, are presented in Table 2. Each binding site includes a hydrophobic cavity with distinct polar or ionic residues within it or its surroundings. Therefore, hydrophobic, polar, and electrostatic interactions play a key role in determining the affinity of ligands for both sites I and II. The interior Site I’s pocket cavity is predominantly hydrophobic delimited by residues Trp214, Leu219, Phe223, Leu238, His242, Leu260, Ile264, Ser287, Ile290, and Ala291; so mainly hydrophobic interactions are dominant (Yang et al., 2014). The site also contains two clusters of polar or ionic residues, an inner one towards the bottom of the pocket composed of Tyr150, His242, and Arg257; and an outer strictly-cationic cluster at the cavity entrance composed of Lys195, Lys199, Arg218, Arg222 (Ghuman et al., 2005). Such placement of residues orchestrates binding of aromatic compounds with centrally located negative or electronegative features (Curry, 2009). Thus, bulky heterocyclic anions preferentially bind to site

TNFα, and glucocorticoids (Baumann et al., 1989; Israili and Dayton, 2001; van Dijk et al., 1991), or in response to various stress stimuli (e.g. infection, burn, surgery, trauma, or cancer) (Israili and Dayton, 2001; Jackson et al., 1982; Kremers et al., 1988; Piafsky, 1980). In the case of acute phase reactions, the normal ratio between the two main variants (F1*S and A) increases more than three-fold, marking a significant production of the F1*S variant (Taguchi et al., 2013). This variation could be about eight times higher in case of lymphoma, melanoma, and ovarian cancer (Taguchi et al., 2013). In cancer biomonitoring, specific glycans on the surface of AGP could be used as biomarkers in the progression of breast cancer (Abd Hamid et al., 2008).
### Table 1
Summary of distinct characteristics of ALB and AGP in humans.

| Characteristics                                      | Human plasma proteins | AGP<sup>a</sup>                                                                 |
|------------------------------------------------------|-----------------------|---------------------------------------------------------------------------------|
| Gene (allelic variants); chromosome location         | 1 gene ALB (number of variants depends on ethnic group); Located on chromosome 4 (4q11-q13) | 2 genes with main variants and 21 amino acid differences: ORM1 or AGP-A (variants F1, F2 and/or S), collectively referred as F1*S; ORM2 or AGP-B/B<sup>+</sup> (variant A) with different drug properties; Located on human chromosome 9 (9q31–34.1) |
| Native polypeptide sequence (N-terminal signal peptide; propeptide; mature protein) | 609 AA (18 AA; 6 AA; 585 AA) | 201 AA (18 AA; 183 AA) |
| Family                                               | Albumin superfamily   | Lipocalin subfamily (immunocalin family) |
| Synthesis                                            | Synthesized by the liver (13.3 g/day) with mild extrahepatic synthesis | Synthesized by the liver (20mg/Kg/day) with extrahepatic synthesis in normal conditions and increased synthesis in case of inflammation, infection or disease |
| Half-life                                            | Around 19 days        | At least 2–3 days, around 5 days |
| Structure                                            | Glycosidic residues   | Glycosidic residues. In the case of pathology, some glycosidic residues may be present on the protein |
| Glycosidic residues                                  | No, 100% peptide residues. In the case of pathology, some glycosidic residues may be present on the protein | Yes, 59% peptide residues and 41% carbohydrates, of which about 11% are sialic acids, 14% neutral hexoses, 14% hexosamine, and 1% fructose. Five carbohydrate entities (N-glycans) covalently linked to Asn residues (A15, A38, A45, A67, A85) |
| Internalization mechanism                           | Internalisation that includes a FcRn- IgG- ALB complex for its degradation or recycling | Internalization after binding to ASGP-R at the sinusoidal for its degradation in endosomes |
| Concentration                                        | 3.0–5.0 g/dl (500 μM–750 μM); 60% of total plasma proteins | 0.05 and 0.1 g/dl (10 μM–20 μM); 3% of total plasma proteins |
| Main binding sites                                   | Two primary sites with high affinity, low capacity; secondary sites with low affinity | One primary site with high affinity and low capacity; many secondary sites with low affinity (five sites for endogenous and some drugs, one site for drugs and the other for steroids) |
| Domains                                              | Domain I (25–222 AA), domain II (223–414 AA) and domain III (415–609 AA); secondary and tertiary structures | Variant F1*S (3 lobes; main lobe I and two smaller lobes II and III) and variant A (2 lobes); secondary and tertiary structures |
| Net charge                                           | -15 at pH 7.4          | Depends on glycosylation (up to 20 types of glycans structures) |
| Isoelectric point                                     | 4.7                    | -2.8 and 3.8 for AGP with high content of sialic acid |
| Main roles                                           | - Negative acute-phase protein; its concentration decreases significantly | - Activities are dependent on the carbohydrates' composition |
|                                                       | - Antioxidant activity, protecting cells and tissues | - Positive acute-phase protein; its concentration increases significantly |
|                                                       | - Predominant binding protein for transport and depot | - Immunomodulatory activities |
| Main binding and common high affinity xenobiotics with| Acidic (anionic) characteristics | Stereoselective binding depends on glycosylation and variant type |
| Notable features for PTMs                             | - 17 disulfide bonding; free sulphydryl group of Cys34; one Tryptophan (W214) | - Two disulfide bonding (C5-C147; C72-165) |
|                                                       | - Electrostatic, polar and hydrophobic interactions with ligands or membrane phospholipids may induce conformational changes | - Three Tryptophan (W25, W122 and W160) |
|                                                       | - Cys34 (antioxidant role and ROS scavenger) | - Electrostatic, polar and hydrophobic interactions with ligands or membrane phospholipids may induce conformational changes |
| Biomarker                                            | Low ratio of ALB to globulin is a biomarker for several types of cancer (neck, breast, prostate, lung, gastric etc.) | Effect biomarker for inflammation, cancer and various pathologies |
| Main factors influencing the binding                 | Plasma protein concentration itself (regulation and feed-back) | Ratio of AGP to prealbumin biomarker for carcinoma of the prostate |
|                                                       | - Low levels in foetuses and infants | Biomarker for all-cancer mortality |
|                                                       | - Decrease during aging | Plasma protein concentration itself (regulation and feed-back) |
|                                                       | - Decrease of ALB under the effect of cytokines and in case of diseases (e.g. liver) | Low levels in foetuses and infants, and in pregnant women |
|                                                       | - Decrease in case of burns, infections, cancer, inflammation (e.g. rheumatoid arthritis), pregnancy and other stress stimuli | Increase during aging |
|                                                       | - Analbuminemia: extremely low concentration if there is no expression of mRNA due to mutations | Increase of AGP under the effect of cytokines and various diseases |
|                                                       |                         | Increase in case of burns, infections, cancer, inflammation (e.g. rheumatoid arthritis), and other stress stimuli |

<sup>a</sup> AA amino acids; AGP alpha-1-acid glycoprotein; ALB albumin; An Assparagine; ASGP-R asialoglycoprotein receptor; Cys Cysteine; IgG Immunoglobulin G; LSEC liver endothelial cells; ORM orosomucoid; PTMs post-translational modifications; ROS reactive oxygen species; RNS reactive nitrogen species.

<sup>b</sup> (Ascenzi and Fasano, 2009; Fanelli et al., 2012; Fasano et al., 2005; He and Carter, 1992; Kragh-Hansen, 2016; Otagiri and Chuang, 2016; Peters Jr., 1996).

<sup>c</sup> (Baumann et al., 1989; Fournier et al., 2000; Hocepied et al., 2003; Israili and Dayton, 2001; Kremer et al., 1988; Piafsky, 1980; Piafsky and Boga, 1977; Schmid et al., 1977; Schonfeld et al., 2008; Taguchi et al., 2013).
Phenylalanine; Pro; Leu; isoleucine; aspartic acid; subdomain IB facing the central interdomain crevice (Zsila, 2013). For interactions play a crucial role in preferentially binding aliphatic and solvent; it is composed of Arg410, Tyr411, Lys414, and Ser489. There- dures is near the pocket entrance which makes it more exposed to the I between apolar Leu238 and Ala291 in the centre of the pocket.

Table 2
Summary of most involved amino acids in high-affinity binding sites.

| Human ALB | Human AGP |
|-----------|-----------|
| Sudlow’s site I: warfarin-binding site (subdomain IIA) | One main site |
| Sudlow’s site II: indole/ benzodiazepine site (subdomain IIIA) | |
| Helix | Residue | Helix | Residue | Stranded β-barrel | Residue |
|--------|---------|--------|---------|-----------------|---------|
| H1     | Asp187  | H1     | Pro384  | BB1             | Tyr27   |
|        | Lys190  |        | Leu387  |                  |         |
|        | Lys195  |        | Ile388  |                  |         |
|        | Lys199  |        | Asn391  |                  |         |
|        |         |        | Cys392  |                  |         |
|        |         |        | Phe395  |                  |         |
| H2     | Phe211  | H2     | Arg410  | BB2             | Phe49   |
|        | Trp214  |        | Tyr411  |                  |         |
|        | Ala215  |        | Lys414  |                  |         |
|        | Arg218  |        |         |                  |         |
|        | Leu219  |        |         |                  |         |
|        | Arg222  |        |         |                  |         |
|        | Phe223  |        |         |                  |         |
| H3     | Leu234  | H3     | Leu430  | BB3             | Glu64   |
|        | Leu238  |        | Val433  |                  |         |
|        | His242  |        | Cys438  |                  |         |
|        | Arg257  | H4     | Ala449  | BB4             | Leu79   |
|        | Leu260  |        | Glu450  |                  |         |
|        | Ala261  |        | Leu453  |                  |         |
|        | Ile264  |        |         |                  |         |
| H5     | H5      | H5     | BB5     | Ile88           |         |
|        |         |        |         | Ser89           |         |
|        |         |        |         | Arg90           |         |
|        |         |        |         | Leu92           |         |
| H6     | Ile290  | H6     | Arg485  | BB6             | His97   |
|        | Ala291  |        | Ser489  |                  | His100  |
|        |         |        |         | Leu101          |         |
| None   | Arg114  |        | BB7     | Tyrl10          |         |
|        | Tyr150  |        |        | Leu112          |         |
|        | Glu292  |        |        | Ala113          |         |
|        |         |        |        | Phe114          |         |
|        |         |        |        | Ile13           |         |
|        |         |        |        | Leu18           |         |
|        |         |        |        | Ser40           |         |
|        |         |        |        | Asp115          |         |
|        |         |        |        | His172          |         |

Ala alanine; AGP alpha-1-acid glycoprotein; ALB albumin; Arg arginine; Asp aspartic acid; BB β-barrel; Glu glutamic acid; H1-6 Helix 1–6; His histidine; Ile isoleucine; Leu leucine; Lys lysine; Ser serine; Trp tryptophan; Tyr tyrosine; Phe Phenylalanine; Pro proline.

Adapted from Peters Jr. (1996) page 37. The information for human ALB is completed from Ascenzi and Fasano (2010), Curry (2009), Hein et al. (2010) and Ghuman et al. (2005).

Residues that are not found in a helix nor in a β sheet of ALB site I.

Information for AGP-F1*S in humans is found in Nishi et al. (2006) and Taguchi et al. (2013). The same important As constituting and surrounding AGP’s main binding site can be viewed in the 3D crystal structure of human AGP (pdb 3KQO) at: http://www.rcsb.org/pdb/submit.do?pdbid=3KQO&preset=validationReport.

I between apolar Leu238 and Ala291 in the centre of the pocket.

Regarding the site II, only one main cluster of polar or cationic residues is near the pocket entrance which makes it more exposed to the solvent; it is composed of Arg410, Tyr411, Lys414, and Ser489. Therefore, a combination of hydrophobic, hydrogen bonding, and electrostatic interactions play a crucial role in preferentially binding aliphatic and aromatic carboxylic acid residues with peripheral located negative or electronegative features to site II (Curry, 2009).

Additionally, there is a third unique site formed by residues from subdomain IB facing the central interdomain crevice (Zsila, 2013). For example, the binding of lidocaine (caticonic drug) to this site is mainly the result of cation-π interaction with Arg114, polar interaction with Lys190, and electrostatic attraction to Asp187 (Hein et al., 2010). According to Carter and Ho (1994), other potential salt bridges exist between Lys190-Glu425; Lys205-Glu465; Asp451-Arg218; and Asp187-Lys432-Arg521; with other hydrophobic interactions that involve the interdomain cluster (Phe206, Leu481, Val482, Trp214, Leu347, Val343, Val344, Leu331, Ala217, and Tyr452)—these should be considered when assessing ALB binding.

Plasma proteins of distinct species have antigenic variants, which induce immunological reactions. According to Majorek et al. (2012), subdomain IB has two regions with significant differences between human and bovine ALB. Residues in these two regions are surface-exposed and may be recognized as epitopes by human immunoglobulins (e.g. in allergies). Other species such as rabbits have different regions that can induce the cross-reactivity of alloantibodies to its antigenic variants.

3.2. Mature AGP in humans

Human AGP’s sequence is less conserved than that of human ALB. In the solely high-affinity and low-capacity site of AGP, there are three distinct lobes (Nishi et al., 2006). The deep main lobe I is voluminous and apolar; it provides sufficient space to accommodate ligands with hydrophobic characteristics. The entrance of the cavity has two positively charged side residues, Arg68 and Arg90, which favour interactions with ligands having polar and negative characteristics. The entrance of human AGP-variant A is narrower than that of F1*S (Taguchi et al., 2013). Thus, the variant F1*S binds a larger variety of ligands.

AGP’s lobe I is approximately 18 Å deep and is constituted of Tyr27, Phe49, Phe51, Leu62, Leu79, Ile88, Leu101, Tyr110, Leu112, and Tyr127 (Table 2). Other apolar residues Ile13, Leu18, and Val81 form its bottom. The hydrophobic lobe I in AGP is located between two negatively charged smaller lobes (lobes II and III). Because the overall surface of AGP is negatively charged, it is highly soluble in the extracellular space. Sialic acids and carbohydrate moiety contribute to the negative charge of glycosylated AGP, consequently attracting xenobiotics with positive charges at pH 7.4. The negative charge of peptide sequence is sufficiently exposed in non-glycosylated human AGP due to the presence of nine Asp and eight Glu side chains that surround the β-barrel (Nishi et al., 2006). The negative charge of lobe II is due to residue Glu64 in its lower part, while the negative charge in lobe III is due to Asp115. All these residues are responsible for the strong binding of basic-characteristic ligands into AGP pockets.

4. Differences in the plasma protein binding

4.1. Factors related to the structural characteristics of the plasma proteins

Many factors could be behind the differences in the binding kinetics of drugs to the same plasma protein or to different plasma proteins. As presented in section 2, these factors could be related, inter alia, to the health status of the individual, the species (e.g. human, rat or bovine), the medium conditions (e.g. pH, temperature, salinity and presence of FAs), the structural characteristics, the physiological abundance, the polymorphism and the variants of the plasma proteins, and to the physicochemical of the drugs.

In this review, there is an emphasis on the structural and molecular characteristics (i.e. glycosylated/non-glycosylated forms (Wu et al., 2018), the charge/polarity of residues and the molecular volume of the site). In studies on binding kinetics, the type of ligands (i.e. base, acid or neutral), the type of molecular attractions and the values of their affinity coefficients to the binding cavities depend mainly on the structural characteristics of ALB and AGP. Particularly, the ionic, polar and hydrophobic clusters constituting high-affinity binding pockets and their surroundings play a major role in plasma protein binding for both proteins. For example, the sites of ALB and AGP have hydrophobic character-istics which make these pockets propitious for binding organic compounds (Ajmal, Nusrat, et al., 2017a; Hanada, 2017; Yang et al., 2014).

Helyon 5 (2019) e02879
The in-sit interactions would depend on the functional groups of the ligands. Therefore, acidic, base or neutral molecules may have either hydrophobic or hydrophilic characteristics (e.g. drugs, organic contaminants and heavy metals). Despite their differences in functional groups, the distinct molecules may bound to the same site if the accommodating-bundle of residues is found. For example, the drugs that preferentially bind to site II may be an acidic molecule like ibuprofen (Ascenzi et al., 2009), or a basic molecule like clofazimine (Ajmal, Zaidi, et al., 2017b). As presented in section 3, the charged residues found in binding cavities for both proteins influence the extent of binding of each ligand. It is most likely possible that the charged residues on the surfaces of both proteins may also have a role in attracting ligands with opposite charges towards sites. On their surfaces, an ALB molecule is rich in clusters of charged residues (Lys and Arg) which attracts acidic probes, while AGP has glycosyl groups that mainly attract base probes and influence their binding affinities (Fernandes et al., 2015). Thus, base drugs may have high affinities for AGP (Asp and Glu residues), and most probably low affinities for ALB (Lys and Arg). Conversely, acidic drugs may have higher affinities to ALB than to AGP.

The computational methods and molecular modelling are powerful tools to estimate protein binding affinity and to simulate the potential interactions in the binding pockets based on the physicochemical properties of compounds, and of the residues involved in the plasma protein binding (Wanat et al., 2018). These tools may also give an idea on how different the binding of a ligand to human ALB from its binding to AGP and which protein-ligand complex is more stable under the same experimental conditions, particularly that the residues involved in the molecular interactions are not the same for the two proteins (Ajmal, Nusrat, et al., 2017a; D. P. Yeggoni, Manidhar, Suresh Reddy and Subramanyam, 2017; Daniel Pushparaju Yeggoni, Rachamallu and Subramanyam, 2016). Consequently, the interactions simulated between functional groups of the ligand and the residues of the plasma proteins are distinct which could be interesting to explore a priori of PK/PD studies in different species.

4.2. Post-translational modifications

The plasma proteins go through post-translational modifications (PTMs) which may affect molecular interactions between residues in binding residues and their ligands. PTMs are biochemical mechanisms that are crucial for normal cell differentiation (Wang et al., 2014), and during which residues at the side chains covalently modify their properties (Walsh et al., 2005). The PTMs may thus affect binding of ligands to plasma protein (i.e. stereoselective allosteric binding) and/or may be used as biomarkers for some diseases (e.g. glycation of ALB) (Macizk-Jurczyk et al., 2018). Therefore, we also emphasize the consideration of PTMs of plasma proteins (e.g. disulphide bridges and N-linked glycosylation) in binding and PK/PD studies. Their inclusion could further guide in vitro and in vivo experiments (Audagnotto and Dal Peraro, 2017), computational approaches and molecular docking. The most common types of PTMs are phosphorylation, acetylation, N-linked and O-linked glycosylation, methylation, and ubiquitylation (Walsh et al., 2005; Wang et al., 2014). PTMs can be classified according to the type of residue (i.e. Lys, Cys, Asn, Arg, Ser, Thr and Tyr); the category of the enzyme making one of the two covalent modifications (i.e. additions, cleavage); and the extent of reversibility of the modification.

5. Summary

Both plasma proteins ALB and AGP have key roles in the transporting and storing of drugs and contaminants in plasma. Structural and molecular characteristics of ALB and AGP may have a direct influence on binding kinetics of xenobiotics (i.e. drugs and contaminants). Change in binding kinetics would cause some alterations in drugs disposition (i.e. distribution, uptake, clearance) and efficacy. Thus, understanding the differences in the functions and the structural and molecular characteristics of these plasma proteins, and addressing the changes in their residues are necessary for interdisciplinary research in PK/PD that is gaining place in the 21st century. For qualitative/quantitative modelling purposes, researchers designing new preclinical and clinical studies should address the intra- and interspecies differences (i.e. type of molecular attractions) not only in plasma protein binding extent, but more particularly in the composition of binding cavities and their peptide surroundings.

Declarations

Author contribution statement

All authors listed have significantly contributed to the development and the writing of this article.

Funding statement

This research did not receive any specific grant from funding agencies in the public, commercial, or not-for-profit sectors.

Competing interest statement

The authors declare no conflict of interest.

Additional information

No additional information is available for this paper.

Acknowledgements

This review makes a part of the author's Ph.D. thesis. The author acknowledges the support given by his main supervisor Dr. Sami Haddad and co-supervisor Dr. Patrick Poulin.

References

Abd Hamid, U.M., Royle, L., Saldova, R., Radcliffe, C.M., Harvey, D.J., Storr, S.J., Rudd, P.M., 2008. A strategy to reveal potential glycan markers from serum glycoproteins associated with breast cancer progression. Glycoconjugate J. 18 (10), 1105–1118.

Ajmal, M.R., Nusrat, S., Alam, P., Zaidi, N., Khan, M.V., Zaman, M., Khan, R.H., 2017. Interaction of anticancer drug clofarabine with human serum albumin and human alpha-1 acid glycoprotein. Spectroscopic and molecular docking approach. J. Pharm. Biomed. Anal. 135, 106–115.

Ajmal, M.R., Zaidi, N., Alam, P., Nusrat, S., Siddiqi, M.K., Badr, G., Khan, R.H., 2017. Insight into the interaction of anti-tubercular and anticancer compound clofazimine with human serum albumin: spectroscopy and molecular modelling. J. Biomol. Struct. Dyn. 35 (1), 46–57.

Andersen, J.T., Dee Qian, J., Sandlie, I., 2006. The conserved histidine 166 residue of the human neonatal Fc receptor heavy chain is critical for the pH-dependent binding to albumin. Eur. J. Immunol. 36 (11), 3044–3051.

Anraku, M., Chuang, V.T., Maruyama, T., Otagiri, M., 2013. Redox properties of serum albumin. Biochim. Biophys. Acta 1830 (12), 5465–5472.

Ascenzi, P., di Masì, A., De Sanctis, G., Coletta, M., Fasano, M., 2009. Ibuprofen modulates allosterically NO dissociation from ferrous nitrosylated human serum heme-albumin by binding to three sites. Biochem. Biophys. Res. Commun. 381 (1), 83–86.

Ascenzi, P., Fasano, M., 2009. Serum heme-albumin: an allosteric protein. JUBMB Life 61 (12), 1118–1122.

Ascenzi, P., Fasano, M., 2010. Allostery in a monomeric protein: the case of human serum albumin. Biochem. Biophys. Chem. 148 (1-3), 16–22.

Audagnotto, M., Dal Peraro, M., 2017. Protein post-translational modifications: in silico prediction tools and molecular modeling. Comput. Struct. Biotechnol. J. 15, 307–319.

Baumann, W., Sokolowska, M., Kurowska, E., Faller, P., 2013. Binding of transition metal ions to albumin: sites, affinities and rates. Biochim. Biophys. Acta 1830 (12), 5444–5455.

Baumann, P., Eaps, G.B., Müller, W.E., Tillement, J.P., 1989. Alpha 1 – Acid Glycoprotein (Genetics, Biochemistry, Physiological Functions and Pharmacology). Alan R. Liss, New York.

Baumann, Held, W.A., 1981. Biosynthesis and hormone-regulated expression of secretory glycoproteins in rat liver and hepatoma cells. Effect of glucocorticoids and inflammation. J. Biol. Chem. 256 (19), 10145–10155.
