Biological characteristics of two lysines on human serum albumin in the high affinity binding of

m4Z,15Z-bilirubin-Ixα revealed by phage display

Ai Minomo<sup>a</sup>, Yu Ishima<sup>a</sup>, Ulrich Kragh-Hansen<sup>b</sup>, Victor Tuan Giam Chuang<sup>c</sup>, Makiyo Uchida<sup>d</sup>,

Kazuaki Taguchi<sup>a</sup>, Hiroshi Watanabe<sup>a</sup>, Toru Maruyama<sup>a</sup>, Hiroshi Morioka<sup>d</sup>, and Masaki Otagiri<sup>a,e\*</sup>

<sup>a</sup>Department of Biopharmaceutics, Graduate School of Pharmaceutical Sciences, Kumamoto University

<sup>b</sup>Department of Medical Biochemistry, University of Aarhus

<sup>c</sup>School of Pharmacy, Curtin Health Innovation Research Institute, Curtin University

<sup>d</sup>Department of Analytical and Biophysical Chemistry, Graduate School of Pharmaceutical Sciences,

Kumamoto University

<sup>e</sup>Drug Delivery System Research Institute, Faculty of Pharmaceutical Sciences, Sojo University

Running title: Bilirubin binding site on human serum albumin

List of abbreviations

HSA, human serum albumin; 4Z,15Z-BR, 4Z, 15Z-Bilirubin-IXα; ELISA, enzyme-linked immunosorbent

assay; CD, circular dichroism; Pfu, plaque forming unit; TU, transducing unit; HRP, horseradish peroxidase.

\*Corresponding author: Masaki Otagiri, Faculty of Pharmaceutical Sciences, Sojo University, 1-22-4 Ikeda,

Kumamoto-shi, Kumamoto 860-0082, Japan;

Tel: +81 96 371 4150; Fax: +81 96 362 7690;

Email: otagirim@ph.sojo-u.ac.jp

This is an Accepted Article that has been peer-reviewed and approved for publication in the FEBS Journal, but has yet to undergo copy-editing and proof correction. Please cite this article as an "Accepted Article"; doi: 10.1111/j.1742-4658.2011.08316.x

Received Date: 12-Apr-2011

**Revised Date** 

: 30-Jun-2011

Accepted Date: 24-Aug-2011

Article type

: Original Article

Corresponding Author E-mail id: y-ishima@kumamoto-u.ac.jp

Abstract

4Z,15Z-bilirubin-IXα (4Z,15Z-BR), an endogenous compound that is sparingly soluble in water, binds

human serum albumin (HSA) with a high-affinity in a flexible manner. A phage library displaying

recombinant HSA-domain II was constructed, after three rounds of panning against immobilized 4Z,15Z-BR,

and eight clones with a high affinity for the pigment were found to contain conserved basic residues such as

lysine or arginine in positions 195 and 199. A wild type and two mutants, K195A and K199A, of whole HSA

as well as standalone domain II, were expressed with *Pichia pastoris* for ligand binding studies. The binding

of 4Z,15Z-BR to the K195A and K199A mutants was decreased in both whole HSA and the domain II

proteins. The P-helicity conformer (P-form) of 4Z,15Z-BR was found to preferentially bind to the wild types

and the K195A proteins whilst the M-form binds to the K199A mutants. Photoconversion experiments

showed that the P-form of 4Z,15Z-BR was transformed into highly water soluble isomers at a much faster

rate than the M-form. In addition, the M-form of 4Z,15Z-BR was found to bind to domain I with a lower

affinity than domain II. The present findings suggest that whilst both Lys195 and Lys199 in subdomain IIA

are important for the high affinity binding of 4Z,15Z-BR, Lys199 plays a more prominent role in the

elimination of 4Z,15Z-BR.

**KEY WORDS:** human serum albumin; phage display; bilirubin; high affinity binding; domain II

#### Footnote

Database: model data are available in the PMDB database under the accession number 2VUE and 2BXA.

# Introduction

Bilirubin, which is produced during the catabolism of hemoglobin, can exist in several isomeric forms [1]. 4Z,15Z-bilirubin-IX $\alpha$  (4Z,15Z-BR) is the main isomer in the body and can cause kernicterus in newborns. The neurotoxicity of 4Z,15Z-BR is most probably related to its extremely low solubility in water at neutral pH [2]. The neurotoxicity of the pigment is diminished, and its solubility in aqueous media is increased on binding to human serum albumin (HSA), to which it binds with a high-affinity constant in the order of  $10^7$ - $10^8$  M<sup>-1</sup>. Hydrogen bonds, aromatic  $\pi$ - $\pi$  and hydrophobic interactions are important for binding [3, 4]. On the other hand, the conversion of 4Z,15Z-BR to 4Z,15E-BR by photosentization results in the formation of a geometric isomer that has a lower binding affinity to HSA but a higher water solubility, which permits it to be excreted more easily than the precursor 4Z,15Z-BR [5].

HSA is the most abundant protein in plasma and is responsible for most of the oncotic pressure. It is a simple protein without prosthetic groups and does not contain any glycan chains and, hence, is not a glycoprotein. It is a heart-shaped protein comprised of three homologous domains, each of which is further divided into two subdomains A and B [6]. HSA binds a variety of different endogenous substances such as fatty acids, bilirubin, various uremic toxins and thyroxines. Two discrete drug binding sites designated as Sudlow Site I and Site II have also been found on subdomains IIA and IIIA, respectively [6].

Affinity binding experiments of 4Z,15Z-BR with albumin fragments point to a fragment comprised of amino acid residues from about 180 to 250, which roughly correspond to subdomain IIA [7]. Petersen *et al.* [8] studied the binding of 4Z,15Z-BR to eight different HSA mutants in which amino acids assumed to be important for binding had been mutated. Surprisingly, none of the single-amino acid substitutions had any significant effect on the binding of 4Z,15Z-BR. This result was taken to indicate the existence of a dynamic, unusually flexible high-affinity binding site for 4Z,15Z-BR in subdomain IIA. In addition, the bisignate CD

spectra of 4Z,15Z-BR becomes inverted when 4Z,15Z-BR binds to various mammalian albumins with more than an 80% homology with HSA [9].

In thermodynamic studies of interactions of 4Z,15Z-BR and HSA,  $\Delta H^{\circ}$  and  $\Delta S^{\circ}$  were found to be -13.5 kcal / mol and -0.0085 kcal / mol / K, respectively [3]. The large negative free energy for the interaction, which is mainly the result of a large enthalpy change, suggests the importance of electrostatic rather than hydrophobic interactions. Earlier studies proposed that the bound BR interacts with histidine, arginine, tyrosine [10] and lysine residues [11-14], especially buried lysine residues [15], but not with Lys240 [16]. Hence, salt linkages between the carboxyl groups of BR and ε-NH<sub>2</sub> groups of lysine residues of HSA appear to be necessary for the binding of BR [10, 17]. Therefore, it is highly possible that the multiple pattern binding mode is due to the formation of a flexible 4Z,15Z-BR-HSA complex, and that no single crucial binding mode exists, but multiple interactions are ongoing between 4Z,15Z-BR and HSA [11-17]. In order to conduct a binding site topology analysis, it becomes necessary to simultaneously introduce multiple mutations into domain II. However, one of the inherent limitations of point mutation techniques is that it requires the selection of a reasonably small number of residues as targets. To overcome this limitation, phage display technology, which can be used to efficiently perform multiple mutations, represents an alternative and potentially useful approach. Phage display is a powerful technology for identifying polypeptides with novel properties, and altering the properties of existing ones [18-22]. Phage display has been used in epitope mapping and studies of protein-protein interactions. This technique involves the insertion of a foreign DNA fragment in the structural gene of a bacteriophage, which leads to the expression of a peptide or protein at the surface of the viral particle. The rapid isolation of specific ligands by phage display is advantageous in many applications, including the selection of inhibitors for active and allosteric sites of enzymes, receptor agonists and antagonists, and specific ligands isolated from phage libraries can also be used in therapeutic target validation, drug design and vaccine development. The use of this technology to identify amino acid residues that are involved in ligand binding is not uncommon. The purpose of this study was to identify the functional amino acid residues in HSA that play a role in high-affinity 4Z,15Z-BR binding by constructing a phage display library using the M13 phage. We selected domain II that contains the binding site of 4Z,15Z-BR as the display protein. Since electrostatic interactions have been shown to be an important factor in the binding of 4Z,15Z-BR to HSA, we screened the role of two positively charged amino acid residues, namely lysine and arginine, which could potentially participate in electrostatic interactions with the carboxylate groups of 4Z,15Z-BR.

#### Results

# Construction of a Phage Library displaying HSA-domain II

We initially displayed the wild type HSA-domain II (residues 187-385, 22 kDa) on the M13KO7 phage, as described in the Experimental Procedures section. The expression of the HSA-domain II on the phage was confirmed by an enzyme-linked immunosorbent assay (ELISA). Both the HSA-domain II displaying phage and the M13KO7 helper phage were equally reactive against the anti-M13KO7 antibody, while only the HSA-domain II displaying phage showed specific activity for the anti-HSA polyclonal antibody (Fig. 1). These results confirm that the HSA-domain II is displayed on the M13 phage. In the second phase of this study, we constructed a phage library displaying mutated versions of recombinant HSA-domain II. Because 4Z,15Z-BR contains two carboxyl groups, previous researchers have proposed that basic residues, lysine and/or arginine, are key residues that interact with the carboxyl groups of 4Z,15Z-BR via salt bridge formation [11-16]. To test this hypothesis and to identify the positions of such key residues, the eight lysines and four arginines located in subdomain IIA were randomly substituted with basic lysine or arginine, neutral glycine or acidic glutamic acid residues. To produce these substitutions, the lysine and arginine codons were replaced with RRA (where R represents A / G) in the cDNA library. Because RRA encodes only four codons (AAA: Lys, AGA: Arg, GAA: Glu, GGA: Gly), the bias of this library can be assumed to be insignificant. As shown in Figure S3, the sequences of unselected 24 phages, displayed little or no bias. The library contained approximately  $1.7 \times 10^7$  phage clones, a size that covers all potential clones that can be expressed.

# Biopanning and Evaluation of 4Z,15Z-BR Binding Capacity of Selected Clones

In order to identify the recombinant HSA-domain II displaying phages with high 4*Z*,15*Z*-BR binding affinities, the phage library was subjected to three rounds of panning against 4*Z*,15*Z*-BR immobilized on an EAH Sepharose 4E gel. The recovery rate of each round was calculated and are listed in Table 1. After the third panning, 111 phage clones were isolated, and their binding affinity to 4*Z*,15*Z*-BR was evaluated by an

HRP assay. Among these 111 clones, 15 were found to bind 4Z,15Z-BR with the same or a higher affinity as wild type domain II (data not shown). A DNA sequencing analysis of these clones showed that 7 clones had frame shifts caused by nucleotide deletions or mutations and did not retain a complete domain II sequence. These deletions and mutations are the result of the fact that the phagemid vector contains a single-stranded DNA, which is relatively unstable, and/or that the HSA sequence has a stretch of five A bases which is a spontaneous mutational "hot spot" [23]. The remaining 8 clones retained the sequence of domain II and were numbered serially (Fig. 2A). As seen from the figure 2A, one of the clones (No. 7) has the same amino acid sequence as wild type domain II. It is noteworthy that a DNA sequencing analysis of these 8 clones showed that 7 of them contain a basic amino acid residue at position 195, and that all of the 8 clones contain a basic residue at position 199 (Fig. 2A).

The above findings were confirmed by the following experiments. First, the proteins of the 8 selected clones were expressed using the *Pichia pastoris* protein expression system and purified. Their 4Z,15Z-BR binding capabilities were then evaluated by the HRP assay. As can be seen in figure 2B, all of the mutated domains bound 4Z,15Z-BR equally well and with an affinity comparable to that of wild type domain II (No. 7). These results suggest that Lys195 and Lys199 in domain II of HSA are important amino acid residues in the binding of 4Z,15Z-BR via electrostatic interactions.

# Effects of Lys195 and Lys199 mutations to alanine on the binding of 4Z,15Z-BR to domain II and HSA

Wild type domain II and domains with the K195A or K199A mutations were produced, and their 4Z,15Z-BR binding activities determined by a HRP assay. Both mutants, but especially domain II-K199A, had a significantly lower affinity for 4Z,15Z-BR than wild type domain II (Fig. 3A). Consistent with the domain II studies, both point mutations, K195A and K199A, resulted in significantly weaker 4Z,15Z-BR binding to full body HSA mutants (Fig. 3B).

The interaction of 4Z,15Z-BR with full body HSA and mutants was also investigated by fluorescence spectroscopy. The binding constant (Ka) for wild type HSA was determined to be  $5.03 \pm 0.53 \times 10^7 \text{ M}^{-1}$  (Table 2). This value is reasonably consistent with the value for the high-affinity binding constant of  $1.56 \times 10^7 \text{ M}^{-1}$  reported by Petersen *et al.* [8], who used a peroxidase protection assay, and to that  $(1.10 \times 10^8 \text{ M}^{-1})$ 

determined by Minchiotti *et al.* [16], who used fluorescence quenching titration. The Ka-values for binding to mutants K195A and K199A were calculated to be  $1.10 \pm 0.32 \times 10^7$  M<sup>-1</sup> and  $1.01 \pm 0.15 \times 10^7$  M<sup>-1</sup>, respectively (Table 2). Thus, the two point mutations both resulted in a 5-fold reduction in binding affinity for 4Z,15Z-BR.

# Binding conformation of 4Z,15Z-BR to wild type and mutant proteins of domain II and HSA examined with CD spectroscopy

Figure 4B shows that the binding of 4Z,15Z-BR to wild type domain II results in an induced 4Z,15Z-BR -CD spectra typical for the P enantiomer of 4Z,15Z-BR (Fig. 4B). Introduction of the K195A mutation had little effect on the spectra. In contrast, in the case of the K199A substitution, the 4Z,15Z-BR -CD spectra was altered and resembled the form typical for the M enantiomer. As illustrated in figure 4C, essentially the same observations were made for full body HSA, i.e., the substitution of Lys195 with an alanine has a small effect on the CD spectra, whereas the same type of substitution for Lys199 has a dramatic, inverting effect on the 4Z,15Z-BR -CD spectra. The relatively small amplitudes observed for the wild type and K195A domain II preparations can be attributed to the weaker binding of the P form of 4Z,15Z-BR to these domain preparations and / or that the bend angle between the two dipyrrole chromophores of 4Z,15Z-BR had been modified.

A similar mutation-induced trend was found for the domain II preparations as for the full body HSAs both with respect to changes in binding (Fig. 3A and B) and with respect to changes in the CD spectra for bound 4Z,15Z-BR (Fig. 4B and C). These findings indicate that the high-affinity 4Z,15Z-BR site is contained within domain II. However, because the changes are quantitatively different, the detailed structure of the 4Z,15Z-BR -domain II complex is most probably different from those formed with full body HSA.

# 4Z,15Z-BR binding to recombinant domain I

The binding of 4Z,15Z-BR to wild type domain I (residues 1-186) is significantly lower than binding to wild type domain II. Furthermore, 4Z,15Z-BR, when bound to domain I, adopts an unusual M-conformation which is quite different from the typical P-conformation when binding to domain II and full body HSA (Fig.

# 4Z,15Z-BR binding to albumins of different species

HSA shows the highest binding affinity for 4Z,15Z-BR, follow by albumins from bovine and sheep. Dog albumin shows the weakest binding affinity for 4Z,15Z-BR (Fig. 6A). As shown in figure 6B and S4, the preferred binding conformation for dog and rat albumins is the P-form, whilst the M-form is the preferred binding conformation for bovine and sheep albumins.

#### Photoconversion of 4Z,15Z-BR bound to HSA, albumins of different species or HSA mutants

4Z,15Z-BR bound to wild type HSA and the K195A mutant showed a more rapid photoconversion rate than that for the K199A mutant (Fig. 7A). On the other hand, HSA and dog albumin showed a faster photoconversion rate than bovine and sheep albumins (Fig. 7B).

# Discussion

# Identification of the key 4Z,15Z-BR binding amino acid residues in domain II of HSA with phage display

The use of a phage display permitted the presentation of (poly) peptides as fusion bodies to capsid proteins on the surface of bacterial viruses or phage particles. The novel display formats, in vitro selection via innovative library designs and screening strategies, enabled the rapid identification and optimization of proteins based on their structural or functional properties. Recombinant techniques have been developed to increase the diversity of the library. Phage display technology has been used to develop polypeptides with novel properties and small antibodies such as scFv. There are some examples of "non-antibody" proteins that are macromolecular proteins displayed on a phage such as protective antigen protein domain 4 (24kDa) [19], insulin-like growth factor I (14.5 kDa) [20], human growth hormone (22 kDa) [21] and glutathione S-transferase (25 kDa) [22].

HSA is a single stranded polypeptide that folds to form a heart shape protein containing 3 homologous domains. A HSA fragment comprised mainly of domain II has been shown to house the primary binding site for 4Z,15Z-BR [24]. The aim of this study was to identify amino acid residues in HSA-domain II that are of importance for the high-affinity binding of the 4Z,15Z-BR. In a previous study, Dockal *et al.* successfully expressed and characterized each individual HSA domain [25]. Therefore, taking advantage of the fact that each of the three HSA domains can be independently produced using recombinant techniques without significant disruption of the secondary and tertiary structures, it was possible to display a large number of domain II mutants on the phage and screen for candidate residues for 4Z,15Z-BR binding.

Libraries with various biases can be created for specific purposes when the sequence of the peptide that binds to the target is known, and its affinity can be selectively increased by screening libraries created with limited mutagenesis of the peptide. In this study, the lysine and arginine codons located in subdomain IIA were replaced with RRA (where R represents A / G) in the cDNA library. Because RRA encodes only four codons (AAA: Lys, AGA: Arg, GAA: Glu, GGA: Gly), the bias of this library is considered to be insignificant. Figure S3 shows the sequences of the unselected library phages that display little or no bias. This phage library was composed of approximately  $1.7 \times 10^7$  clones. A reasonably sized, diverse library ensures that every possible clone is represented in the library. In fact, the input phage number at round 1 was  $3.0 \times 10^{13}$  (TU) which is sufficient to cover the diversity requirement of a phage library (Table 1). Using the phage library, the screening results indicated that Lys195 and Lys199 in the native HSA-subdomain IIA appeared to play an important role in the binding of BR, which contains two free carboxyl groups. Based on the phage display results, the role of Lys195 and Lys199 in binding 4Z,15Z-BR was further examined by site-directed mutagenesis.

# Binding affinity of 4Z,15Z-BR to HSA

The findings reported so far indicate that mutations of Lys195 and Lys199 in subdomain IIA of HSA to alanine residues caused an increase in the concentration of unbound 4Z,15Z-BR, indicating that both lysine residues are involved in the binding of 4Z,15Z-BR (Table 2). However, a similar investigation of the wild type standalone domain II protein showed the binding affinity to be lower, only  $1.93 \pm 0.26 \times 10^7$  M<sup>-1</sup> (Table

2). A possible explanation for the weaker binding to the standalone domain II protein is that the BR binding site is more exposed to the solvent in the absence of domain I and III. A similar observation has been reported by Dockal *et al.* on the binding of site I ligand, warfarin, where the primary warfarin binding site is centered around subdomain IIA with indispensable structural contributions from subdomain IIB and domain I [25]. We also reported previously that the recombinant albumin domain II protein binds warfarin and DNSA only negligibly [26]. This indicates that the integrity of the binding pocket in domain II is affected by interdomain interactions between domains I and III.

Goncharova and Urbanová [27] proposed that ligands with a carboxylate group may interact with the positively charged Lys199, Arg222 and Arg257 residues in subdomain IIA. On the other hand, in an x-ray crystallographic study, 3-carboxy-4-methyl-5-propyl-2-furanpropanoic acid (CMPF), a dicarboxylic acid-containing renal toxin, has been shown to have five hydrogen-bond or salt-bridge interactions with Tyr150, Lys199, Arg222, His242 and Arg257 [28]. In a previous study, we found that CMPF competitively inhibited the binding of 4Z,15Z-BR to HSA [11]. Inspection of the CMPF-HSA x-ray crystallographic structure revealed that CMPF interacts electrostatically with Lys199 but obvious interactions with Lys195 are lacking (Fig. 8B). These findings provide further support for the essential role of Lys199 in the binding of 4Z,15Z-BR.

The importance of Lys199 over Lys195 was clearly observed from the 3<sup>rd</sup> panning results where clone 3 retained an arginine at 199 but a glycine was located at position 195. In addition, the concentration of free 4Z,15Z-BR for K199A was higher than that for K195A and wild type HSA in this study. In an x-ray crystallographic study of drug-HSA complexes, Ghuman reported that the polar groups of Lys199 and Lys195 were clustered with other polar amino acid residues at the entrance of the binding pocket of subdomain IIA, whilst only the aliphatic portion of Lys199, which contributes to the delineation of the subchamber, protrudes from the front of the pocket [28]. This shows that Lys199 plays more than one role in the binding of 4Z,15Z-BR.

Petersen *et al.* [8] studied the binding of 4Z,15Z-BR to HSA preparations with several single-amino acid substitutions, including K195M and K199M, and found that none of the mutants had a significantly altered affinity for 4Z,15Z-BR. This is in contrast to the present study, in which we found that the binding of

4Z,15Z-BR to K195A and K199A was diminished (Table 2). A comparison of the binding results for methionine and alanine mutants illustrate the importance of having the space at position 195 and 199 occupied by an aliphatic side chain in order to retain 4Z,15Z-BR at the binding pocket. This can most likely be attributed to hydrophobic interactions. In a study using molecular dynamics modeling, Diaz *et al.* reported that Lys195 and Lys199 could play a combined and comparable chemical role in maintaining a positive charge, whereby a potential Lys195  $\rightarrow$  Lys199 proton-transfer process via water bridges has been proposed [15]. The model illustrates the complexity of the key residue, Lys199, ionization state and the nearby Lys195 has on the structure and dynamics of the IIA binding site.

# **Enantioselective binding of 4Z,15Z-BR to HSA**

The fully protonated unconjugated 4*Z*,15*Z*-BR diacid and the dianion adopt a folded biplanar ridge-tile structure analogous to a half-opened book which is stabilized through intramolecular hydrogen bonding of each of the two carboxylate groups to the oxygen and nitrogen of the opposing dipyrrole (Fig. 4A). This explains the considerable lipophilicity of the pigment compared to its biogenetic precursor biliverdin and is a dominant factor in its metabolism, antioxidant behavior, and toxicity [1]. The right-handed configuration is denoted as P (plus) and the left handed one as M (minus), referred to as the P-helicity folded conformer (P-form) and the M-helicity folded conformer (M-form) [29]. In solution at 37°C, the two mirror-image forms interconvert rapidly, forming a 50:50 mixture of P and M conformers. Nevertheless, when a chiral complexation agent such as serum albumin is added, the equilibrium shifts toward either the M or P enantiomer. In such cases one observes typically intense bisignate CD cotton effects. The conformational structure of the bound pigment, hence the shape and nature of the binding site on HSA can be inferred from the CD data [13].

As evidenced from the 4Z,15Z-BR-induced CD spectra, mutations of Lys195 and Lys199 to alanine effectively changed the binding microenvironment of the 4Z,15Z-BR binding site to different extents depending on the location of the mutation and, thus, the preferred binding conformation of 4Z,15Z-BR. The preferred binding conformation of BR by wild type HSA is the P-form, as shown in figure 4C. The mutation on K195A had only a minor decreasing effect on the intensities of the spectra, and the P-form of 4Z,15Z-BR

remains the preferred binding conformation at the binding site (Fig. 4C). However, when Lys199 is substituted with an alanine, the spectra for 4Z,15Z-BR changed dramatically, indicating that the M-form was the preferred binding conformation in the absence of a lysine aliphatic side chain at location 199. A similar binding chirality preference could also be observed for the wild type, K195A and K199A recombinant domain II proteins (Fig. 4B).

Interestingly, albumins of different species also showed such preferred binding conformation of 4*Z*,15*Z*-BR depending of the species. Pistolozzi *et al.* reported that induced CD spectra of 4*Z*,15*Z*-BR bound to rat albumin was observed as P-form [9], which is in agreement with our finding. As shown in figure 6B and S4, the P-form is the preferred binding conformation for human, dog and rat albumins whilst the M-form is the preferred binding conformation for bovine and sheep albumins (Fig. 6B and S4). A trend was found when the amino acid sequences of the albumins were compared where the preferred P-form of albumins of humans and dogs contain lysines at positions 195 and 199, the rat albumin contains an arginine at 195 [9], while bovine and sheep albumins, with the M-form preferred, contain arginines at position 195 and 199 (Table 3). Hence, Lys199 is a key amino residue responsible for P-form preferred binding conformation of 4*Z*,15*Z*-BR by albumin. Since the mutation of lysine to methionine at position 199 had no effect on changing the P-form binding preference of HSA in the point mutation study reported by Petersen [8], this provides a demonstration of the role of an aliphatic side chain at position 199 in determining the helicity of 4*Z*,15*Z*-BR in the binding preference of albumin.

Albumin has been found to facilitate a conformational change in 4Z,15Z-BR into a more water soluble form during phototherapy [5]. The physiological significance of the preference for a 4Z,15Z-BR binding conformation by albumin was evident in a further investigation of the effect of photoirradiation on 4Z,15Z-BR and albumin mixtures. In general, the P-form preferred binding albumins facilitated a faster transformation of the low solubility 4Z,15Z-BR into more water soluble forms of 4Z,15Z-BR, as shown in figure 7B for the human and dog. A similar observation was obtained when both wild type and K195A mutant HSA that preferentially bind the P-form exhibited a faster transformation of 4Z,15Z-BR. In contrast, the mutation of Lys199 to alanine effectively changed the binding preference to the M-form, and thus a slower photoconversion rate of the low solubility 4Z,15Z-BR (Fig. 7A). This observation also points to the

importance of an aliphatic side chain at position 199 for the efficient photoconversion of 4Z,15Z-BR to occur.

Zunszain *et al.* [30] recently published a crystallographic analysis of HSA complexed with 4Z,15E-BR (Fig. 8A). The analysis revealed that this isomeric form of the pigment binds with a high affinity to a pocket in subdomain IB, and that the carboxylate groups enter into salt-bridges with arginine residues at position 117 and 186 but not to lysine residues. Our results showed that 4Z,15Z-BR binds to domain I with a lower affinity than to domain II (Fig. 5A) and in an atypical M-helicity conformation instead of the usual P-form (Fig. 5B). Thus it is possible that subdomain IB is a binding site for 4Z,15E-BR, but may not be the primary high affinity binding site for 4Z,15Z-BR. However, further investigations will be needed to investigate the possible photoconversion of 4Z,15Z-BR to 4Z,15E-BR that leads to a shift in the binding site from subdomain IIA to subdomain IB.

# Potential clinical usefulness of domain II in hyperbilirubinemia

The present work laid the groundwork for development of new therapeutic agents for treating jaundiced newborns. The major treatment for 4*Z*,15*Z*-BR-induced encephalopathy in newborns and especially in premature infants is phototherapy, which converts 4*Z*,15*Z*-BR, which primarily binds to its high affinity HSA binding site into more soluble nontoxic structural isomers that are more easily eliminated from the circulatory system. The rate of formation of isomers and the isomeric composition of the reaction products have been shown to be highly dependent on the conformation and electronic environment of 4*Z*,15*Z*-BR bound to HSA [8, 31]. The results of this study show that the P-form of 4*Z*,15*Z*-BR is more readily converted into photoisomers with higher water solubility than the M-form (Fig. 7). In a previous study, we reported that there was about a 50-fold difference in total clearance between HSA and individual recombinant domains including domain II [26]. Similar observations have been reported by Sheffield *et al.* using recombinant domains of rabbit serum albumin [32]. This indicates that HSA-domain II with preserved Lys199 could be useful in treating jaundiced newborns without being dependent on a premature liver, because it preferentially binds the P-form of 4*Z*,15*Z*-BR with a comparable high affinity, and more importantly it would be cleared from the body via the kidney at a rate 50 times faster than whole albumin.

# **Concluding remarks**

Using a phage display system, we have successfully displayed domain II of HSA on M13 phages and identified two key lysine residues, K195 and K199, that are involved in the binding of 4Z,15Z-BR. Further investigations indicated that K199 plays additional roles, compared to K195, in that it is a determinant of the helicity of the P-form of 4Z,15Z-BR, which affects the binding preference of albumin, thus leading to a more rapid photoconversion of 4Z,15Z-BR into a more water soluble molecule for subsequent renal elimination.

# **Materials and Methods**

# Construction of a Phage Library Displaying Recombinant HSA-domain II

HSA-domain II cDNA containing sites for restriction enzymes *Sfi* I and *Not* I, located at the N and C terminus, respectively, was constructed. The plasmid pCANTAB 5E also contained *Sfi* I and *Not* I sites, and both HSA-domain II cDNA and pCANTAB 5E were digested with *Sfi* I / *Not* I and ligated using T4 DNA ligase. Plasmid pCANTAB 5E encoding HSA-domain II, in which the C terminus of HSA-domain II was fused to the N terminus of the M13 phage gene 3 signal sequence, was confirmed by DNA sequence analysis.

Oligos 1-6 were designed to have the sequence "RRA" (or TYY in the complementary strand), where R and Y represent G / A and C / T, respectively, at the HSA-subdomain IIA codons for Lys190, Lys195, Arg197, Lys199, Lys205, Arg209, Lys212, Arg218, Arg222, Lys225, Lys233 and Lys240. The sequence RRA can randomly encode for Lys, Arg, Gly and Glu, without bias.

- Oligo 1: 5'-CCGGCCATGTCCGATGAAGGGRRAGCTTCGTCGTCTGCC-3'
- Oligo 2: 5'-TTGGAGACTGGCACATYYGAGTYYCTGTYYGGCAGACGAAGC-3'
- Oligo 3: 5'-GTGCCAGTCTCCAARRATTTGGAGAARRAGCTTTCRRAGCATGGGCAGTAGCT-3'
- Oligo 4: 5'-ACTTCTGCAAACTCAGCTYYGGGAAATYYCTGGCTCAGTYYAGCTACTGCCCAT

Oligo 5: 5'-GCTGAGTTTGCAGAAGTTTCCRRATTAGTGACAGATCTT-3'

Oligo 6: 5'-GCAGCATTCCGTGTGGACTYYGGTAAGATGTGTCACTAA-3'

HSA\_d II\_Sfi I: 5'-CCTTTCTATGCGGCCCAGCCGGCCATGTCCGATGAA-3'

HSA\_d II\_120r: 5'-GGAAACTTCTGCAAACTCAGC-3'

HSA\_d II\_Nco I: 5'- CAGATCTCCATGGCAGCATTCCGTGTGGAC-3'

Mutagenic double-strand DNAs were synthesized from these oligonucleotides using Klenow fragment (TaKaRa, Kyoto, Japan). Prepared double-strand DNAs with oligo 1-2, 3-4 and 5-6 are denoted as fragment 1, 2, 3, respectively (Fig. S1).

To identify key residues of importance for the binding of BR to HSA, we prepared a HSA display phage library. Because it is most likely that the present isomer of BR binds to domain II, we proceeded to display that domain. As seen in figure S1, the synthesized double-strand DNA fragments were elongated and amplified by PCR using PfuUltra High-Fidelity DNA polymerase (Stratagene, La Jolla, CA, USA). Finally, mutagenized HSA-subdomain IIA cDNA was ligated into pCANTAB 5E.

# **Biopanning**

200  $\mu$ l of BR coupling gel was blocked with 1 ml of 0.5% gelatin overnight at room temperature. 1 ml of the phage library was added to the gel, which was then allowed to bind for 2 h at 37°C. After binding, the gel was washed 3 or 10 times with PBS containing 0.005% Tween 20 and eluted with 1 mL of 0.1 M Glycine-HCl (pH 2.2) containing 1 mg / mL BSA. The eluate was collected and neutralized with 60  $\mu$ l of 2 M Tris-HCl. The eluted phages were amplified and used for next-round selection.

#### Acknowledgements

This work was supported in part by Grants-in-Aid from the Japan Society for the Promotion of Science (JSPS). Thanks are due to the members of the Gene Technology Center at Kumamoto University for their important contributions to the experiments.

# Reference

- 1. Brodersen R (1980) Binding of bilirubin to albumin. CRC Crit Rev Clin Lab Sci 11, 305-399.
- Brodersen R (1979) Bilirubin. Solubility and interaction with albumin and phospholipid. *J Biol Chem* 254, 2364-2369.
- Jacobsen J (1977) Studies of the affinity of human serum albumin for binding of bilirubin at different temperatures and ionic strength. *Int J Pept Protein Res* 3, 235-239.
- Berde CB, Hudson BS, Simoni RD & Sklar LA (1979) Human serum albumin. Spectroscopic studies of binding and proximity relationships for fatty acids and bilirubin. *J Biol Chem* 254, 391-400.
- 5. Maisels MJ & McDonagh AF (2008) Phototherapy for neonatal jaundice. N Engl J Med 358, 920-928.
- 6. Peters T Jr. (1981) The Albumin Molecule: Its Structure and Chemical Properties In *All about albumin*:

  \*Biochemistry, Genetics, and Medical Applications (Peters T Jr. eds), pp. 9-75. Academic Press, San Diego, CA.
- Kragh-Hansen U (1981) Molecular aspects of ligand binding to serum albumin. *Pharmacol Rev* 33, 17-53.
- 8. Petersen CE, Ha CE, Harohalli K, Feix JB & Bhagavan NV (2000) A dynamic model for bilirubin binding to human serum albumin. *J Biol Chem* **275**, 20985-20995.
- Pistolozzi M & Bertucci C (2008) Species-dependent stereoselective drug binding to albumin: A circular dichroism study. *Chirality* 20, 552-558.
- 10. Jacobsen C (1972) Chemical modification of the high-affinity bilirubin-binding site of human-serum albumin. *Eur J Biochem* **27**, 513-519.
- 11. Tsutsumi Y, Maruyama T, Takadate A, Goto M, Matsunaga H & Otagiri M (1999) Interaction between two dicarboxylate endogenous substances, bilirubin and an uremic toxin,

- 3-carboxy-4-methyl-5-propyl-2-furanpropanoic acid, on human serum albumin. *Pharm Res* **16**, 916-923.
- 12. Tayyab S, Ahmad B, Kumar Y & Khan MM (2002) Salt-induced refolding in different domains of partially folded bovine serum albumin. *Int J Biol Macromol* **30**, 17-22.
- 13. Khan MM & Tayyab S (2001) Understanding the role of internal lysine residues of serum albumins in conformational stability and bilirubin binding. *Biochim Biophys Acta.* **1545**, 263-277.
- Díaz N, Suárez D, Sordo TL & Merz KMJ (2001) Molecular dynamics study of the IIA binding site in human serum albumin: influence of the protonation state of Lys195 and Lys199. *J Med Chem* 44, 250-260.
- 15. Mir MM, Fazili KM & Qasim MA (1992) Chemical modification of buried lysine residues of bovine serum albumin and its influence on protein conformation and bilirubin binding. *Biochim Biophys Acta* 1119, 261-267
- 16. Minchiotti L, Galliano M, Zapponi MC & Tenni R (1993) The structural characterization and bilirubin-binding properties of albumin Herborn, a [Lys240-->Glu] albumin mutant. *Eur J Biochem* **214**, 437-444.
- Lightner DA, Wijekoon WMD & Zhang MH (1988) Understanding bilirubin conformation and binding.
   Circular dichroism of human serum albumin complexes with bilirubin and its esters. *J Biol Chem* 263, 16669-16676.
- 18. Smith GP (1985) Filamentous fusion phage: novel expression vectors that display cloned antigens on the virion surface. *Science* **228**, 1315-1317.
- 19. Chen KH, Liu S, Bankston LA, Liddington RC & Leppla SH (2007) Selection of anthrax toxin protective antigen variants that discriminate between the cellular receptors TEM8 and CMG2 and achieve targeting of tumor cells. *J Biol Chem* 282, 9834-9845.

- Dubaquié Y & Lowman HB (1999) Total alanine-scanning mutagenesis of insulin-like growth factor I (IGF-I) identifies differential binding epitopes for IGFBP-1 and IGFBP-3. *Biochemistry (Mosc)* 38, 6386-6396.
- Lowman HB & Wells JA (1993) Affinity maturation of human growth hormone by monovalent phage display. J Mol Biol 234, 564-578.
- 22. Widersten M & Mannervik B (1995) Glutathione transferases with novel active sites isolated by phage display from a library of random mutants. *J Mol Biol* **250**, 115-122.
- 23. Streisinger G, Okada Y, Emrich J, Newton J, Tsugita A, Terzaghi E & Inouye M (1966) Frameshift mutations and the genetic code. This paper is dedicated to Professor Theodosius Dobzhansky on the occasion of his 66th birthday. *Cold Spring Harbor Symp Quant Biol* 31, 77-84.
- 24. Geisow MJ & Beaven G.H (1977) Physical and binding properties of large fragments of human serum albumin, *Biochem J* **163**, 477-84.
- Dockal M, Carter DC & Rüker F (1999) The three recombinant domains of human serum albumin.
   Structural characterization and ligand binding properties. *J Biol Chem* 274, 29303-29310.
- 26. Matsushita S, Isima Y, Chuang VTG, Watanabe H, Tanase S, Maruyama T & Otagiri M (2004)
  Functional analysis of recombinant human serum albumin domains for pharmaceutical applications.
  Pharm Res 21, 1924-1932.
- 27. Goncharova I & Urbanová M (2008) Stereoselective bile pigment binding to polypeptides and albumins: a circular dichroism study. *Anal Bioanal Chem* **392**, 1355-1365.
- 28. Ghuman J, Zunszain PA, Petitpas I, Bhattacharya AA, Otagiri M & Curry S (2005) Structural basis of the drug-binding specificity of human serum albumin. *J Mol Biol* **353**, 38-52.
- 29. Boiadjiev SE, Person RV, Puzicha G, Knobler C, Maverick E, Trueblood KN & Lightner DA (1992)

  Absolute configuration of bilirubin conformational enantiomers. *J Am Chem Soc* **114**, 10123-10133.

- 30. Zunszain PA, Ghuman J, McDonagh AF & Curry S (2008) Crystallographic analysis of human serum albumin complexed with 4Z,15E-bilirubin-IXalpha. *J Mol Biol* **381**, 394-406.
- 31. Khan MM, Muzammil S & Tayyab S (2000) Role of salt bridge(s) in the binding and photoconversion of bilirubin bound to high affinity site on human serum albumin. *Biochim Biophys Acta* **1479**, 103-113.
- 32. Sheffield WP, Marques JA, Bhakta V & Smith IJ (2000) Modulation of clearance of recombinant serum albumin by either glycosylation or truncation. *Thromb Res* **99**, 613-621.

# **Supporting information**

Doc. S1 Supplementary materials and methods.

Fig. S1 Construction of a phage library displaying recombinant HSA-domain II.

Fig. S2 Substitution sites of a phage library displaying recombinant HSA-domain II.

Fig. S3 A probability of codons appearance in 24 clones in unselected phage library.

Fig. S4 Induced CD spectra of 4Z,15Z-BR bound to albumin of different species.

TABLE 1 Enrichment of the phage library from round 1 to 3 expressed as recovery rate.

| Rounds | Input phage (TU)     | Output phage (TU)   | Recovery rate<br>(Per 10 <sup>5</sup> phages) |  |
|--------|----------------------|---------------------|---|--|
| 1      | $3.0 \times 10^{13}$ | $1.2 \times 10^9$   | 4.0   |  |
| 2      | $2.2 \times 10^{12}$ | $1.4 \times 10^{7}$ | 0.63  |  |
| 3      | $5.0 \times 10^{11}$ | $1.2 \times 10^{7}$ | 2.4   |  |

TABLE 2 BR affinity binding constants of recombinant domain II and HSA (WT, K195A, K199A).

|       | <b>HSA</b> $Ka \ [\times 10^7 \ (M^{-1})]$ | <b>Domain II</b> $Ka \left[ \times 10^7 \left( M^{-1} \right) \right]$ |
|-------|--|--|
| WT    | $5.03 \pm 0.53$                            | $1.93 \pm 0.26$  |
| K195A | $1.10 \pm 0.32$                            | 0.327  |
| K199A | 1.01 ± 0.15                                | 0.209  |

BR binding to different recombinant domain II and HSA preparations studied with the fluorescence enhancement titration technique. The protein solutions were excitated at 460 nm, and the emission was measured at 512 nm in the case of wild type HSA and K199A mutant, and at 507 nm in the case of the K195A mutant. The protein concentration was kept constant at 3  $\mu$ M, whereas the concentration of BR was varied from 0 to 3.0  $\mu$ M; the medium was 67 mM sodium phosphate buffer, pH 7.4. Data are expressed as means  $\pm$  SEM (n = 3).

TABLE 3 Amino acid residues at positions 195 and 199 of albumins of different species and the corresponding BR binding helicity form.

| Mammalian | Homology (%) | BR helicity      | 195 | 199 | 222 |
|-----------|--------------|------------------|-----|-----|-----|
| Human     | 100          | Р                | K   | K   | R   |
| Dog       | 80           | P                | K   | K   | R   |
| Rat       | 73           | P (Fig. S1, [9]) | R   | K   | R   |
| Bovine    | 75           | M                | R   | R   | K   |
| Sheep     | 75           | M                | R   | R   | K   |

#### Figure Legends

**Fig. 1.** Detection of domain II display phage. (A) Various concentrations of phage solutions were plated on 96-well microtiter plates at  $100 \,\mu\text{L}$  / well. After an overnight incubation, the wells were washed and blocked with 0.5% gelatin, and treated with a HRP-conjugated anti-M13KO7 antibody. (B) Various concentrations of phage solutions were plated on 96-well microtiter plates at  $100 \,\mu\text{L}$  / well. After an overnight incubation, the coated wells were washed and blocked with gelatin, followed by reaction with anti-HSA polyclonal antibody. Color development was with 5-amino-2-hydroxybenzoic acid, and  $0.02\% \,\text{H}_2\text{O}_2$  was used as a chromogen.

Fig. 2. Selection and confirmation of domain II clones that bind 4Z,15Z-BR. (A) Partial amino acid sequences of domain II clones selected after 3 rounds of panning on 4Z,15Z-BR. (B) 4Z,15Z-BR binding to domain II clones as determined by a HRP assay. The molar ratio between 4Z,15Z-BR and protein was 2:1, and the concentration of unbound ligand is shown. Protein solutions were plated on 96-well microtiter plates at  $100 \,\mu\text{L}$  / well. After a 2 hr incubation, the coated wells were washed, and  $60 \,\mu\text{M}$  4Z,15Z-BR was added to each well. Then,  $10 \,\mu\text{L}$  of 1.75 mM H<sub>2</sub>O<sub>2</sub> and  $10 \,\mu\text{l}$  of 1 ng / mL HRP were added, and the absorbance at 450 nm was recorded. The free 4Z,15Z-BR concentrations were estimated as described in Experimental Procedures. Data are expressed as means  $\pm$  SEM (n = 3). When the results obtained for the mutated domain II clones were compared to those obtained for the wild type domain II clone (No. 7), the *p*-values were in all cases > 0.05.

**Fig. 3.** 4*Z*,15*Z*-BR binding properties of recombinant domain II and HSA (WT, K195A, K199A). The concentration of unbound 4*Z*,15*Z*-BR in the presence of recombinant domain II proteins (A) or recombinant HSA proteins (B) determined by the HRP assay. 60  $\mu$ M 4*Z*,15*Z*-BR was incubated with 30  $\mu$ M of domain II proteins or to 15  $\mu$ M HSA proteins, each dissolved in 67 mM phosphate buffer (pH7.4). The free 4*Z*,15*Z*-BR concentrations were estimated as described in Experimental Procedures. Data are expressed as means  $\pm$  SEM (n = 3). \*\*, p < 0.01, compared with corresponding wild type polypeptide (WT).

**Fig. 4.** Induced CD spectra of 4*Z*,15*Z*-BR bound to recombinant domain II and HSA (WT, K195A, K199A). (A) Structures of the M and P enantiomers of 4*Z*,15*Z*-BR. The "ridge-tile" conformations are indicated. (B) CD spectra of 4*Z*,15*Z*-BR bound to different versions of domain II. (C) CD spectra of 4*Z*,15*Z*-BR bound to different HSAs. Common to (B) and (C), 800 μL of a 40 μM domain II solution or a 20 μM HSA solution in 67 mM sodium phosphate buffer (pH7.4), was placed in a 1.0-cm pathlength cell and scanned from 350 to 550 nm. 4*Z*,15*Z*-BR was subsequently added to a final concentration of 60 μM or 30 μM, respectively. After mixing and then standing for 10 min, the samples were re-scanned from 350 to 550 nm.

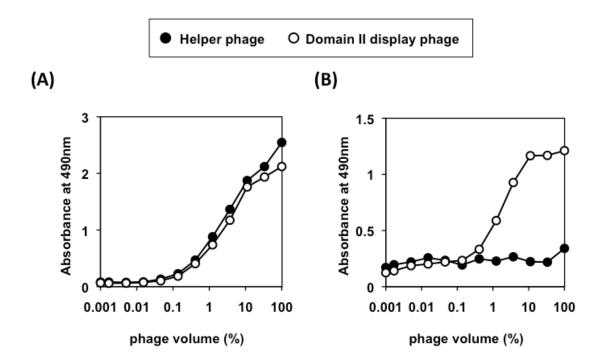
**Fig. 5.** 4*Z*,15*Z*-BR binding properties of recombinant domain I compared to domain II and HSA. (A) The concentration of unbound 4*Z*,15*Z*-BR in the presence of domain I or II was determined by the HRP assay. 60 μM 4*Z*,15*Z*-BR was incubated with 30 μM of recombinant HSA, domain I and domain II in 67 mM phosphate buffer (pH7.4). The free 4*Z*,15*Z*-BR concentrations were estimated as described in Experimental Procedures. Data are expressed as means  $\pm$  SEM (n = 3). (B) Induced CD spectra of 4*Z*,15*Z*-BR bound to wild type domain I or II or HSA. Firstly, the CD spectra were obtained by placing 800 μL of a 40 μM domain I or II solution or a 20 μM HSA solution in 67 mM sodium phosphate buffer, pH7.4, in a 1.0-cm pathlength cell and scanned from 350 to 550 nm. 4*Z*,15*Z*-BR was then added to a final concentration of 60 μM or 30 μM, respectively. After mixing and standing for 10 min, the samples were scanned from 350 to 550 nm.

**Fig. 6.** 4*Z*,15*Z*-BR binding properties of albumins of different species. (A) The concentration of unbound 4*Z*,15*Z*-BR in the presence of albumins of different species was determined by the HRP assay. 15 μM 4*Z*,15*Z*-BR was incubated with 30 μM of albumin in 67 mM phosphate buffer (pH7.4). The free 4*Z*,15*Z*-BR concentrations were estimated as described in Experimental Procedures. Data are expressed as means  $\pm$  SEM (n = 3). (B) Induced CD spectra of 4*Z*,15*Z*-BR bound to albumin of different species. The CD spectra were initially obtained by placing 800 μL of a 20 μM albumin solution in 67 mM sodium phosphate buffer, pH 7.4, in a 1.0-cm pathlength cell and the solution scanned from 350 to 550 nm. 4*Z*,15*Z*-BR was then added to a

final concentration of 20  $\mu$ M. After mixing and standing for 10 min, the samples were scanned from 350 to 550 nm.

**Fig. 7.** Rate of photoconversion of 4*Z*,15*Z*-BR into other more water soluble forms of BR. (A) Recombinant HSA and its K195A, K199A mutants and (B) albumin from different species.

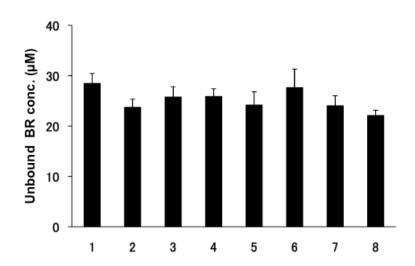
**Fig. 8.** Binding regions of 4*Z*,15*E*-BR and CMPF on HSA. Magnification of the binding region of 4*Z*,15*E*-BR (A) and CMPF (B), with the 4 amino acid residues lining the entrance of the site I binding pocket in subdomain IIA, namely K195, K199, R218 and R222, shown as ball and stick models, CPK color style. Domain I (grey) and II (green) were displayed in back bone style while domain III (grey) in strand style. Both diagrams were generated with RASMOL, using the x-ray crystallographic structure data files from protein data bank, PDB ID: **2VUE** for 4*Z*, 15*E*-BR-HSA (A) and PDB ID: **2BXA** for CMPF-HSA (B). The insert in (A) shows whole HSA structure.

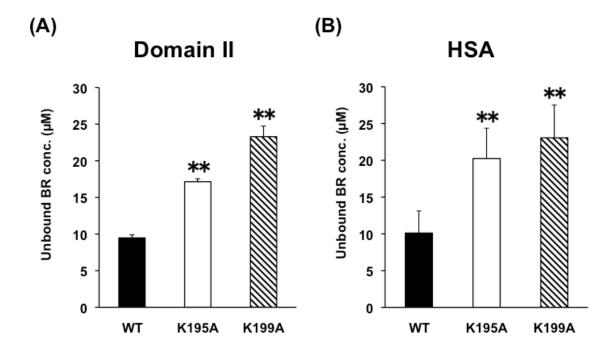


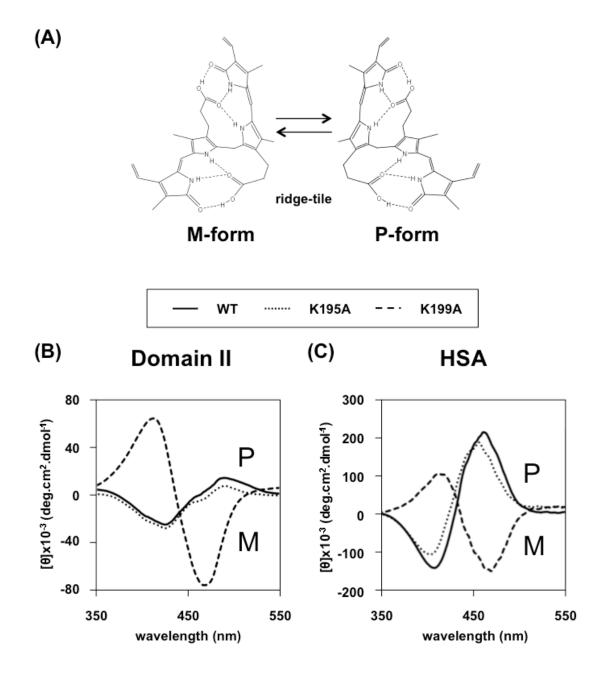
# FIGURE 2

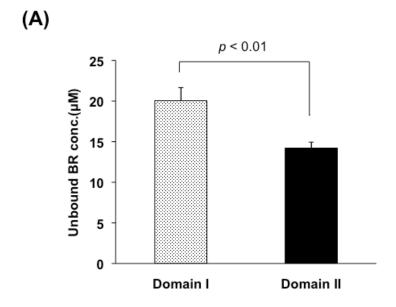
(A) 195197199 205 209 212 218 222 225 233 240 WΤ DEGKASSAKQRLKCASLQKFGERAFKAWAVARLSQRFPKAEFAEVSKLVTDLTK 1 DEGGASSAKQRLRCASLQGFGEGAFEAWAVAGLSQEFPKAEFAEVSGLVTDLTE 2 DEGEASSARQELRCASLQRFGEEAFKAWAVAGLSQRFPRAEFAEVSGLVTDLTG 3 DEGRASSAGORLRCASLQEFGEEAFGAWAVAELSQGFPEAEFAEVSRLVTDLTK 4 DEGKASSARQGLRCASLQEFGEGAFRAWAVAELSQGFPRAEFAEVSELVTDLTE 5 DEGEASSARQGLKCASLQRFGEGAFRAWAVAGLSQRFPGAEFAEVSELVTDLTR 6 DEGKASSARQRLRCASLQGFGEGAFRAWAVAGLSQGFPRAEFAEVSELVTDLTE 7 DEGKASSAKQRLKCASLQKFGERAFKAWAVARLSQRFPKAEFAEVSKLVTDLTK DEGKASSARQKLRCASLQGFGEGAFGAWAVAELSQGFPRAEFAEVSGLVTDLTK

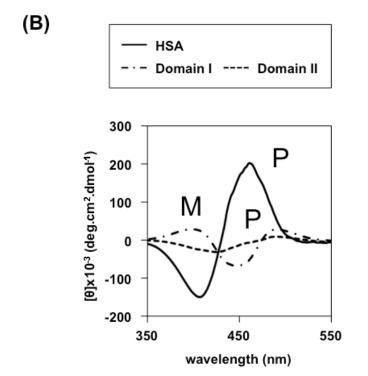




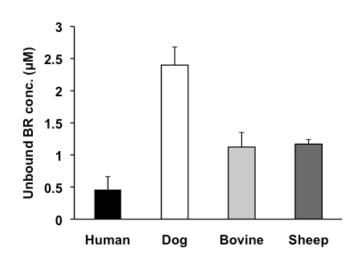




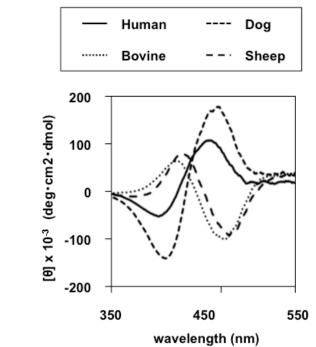




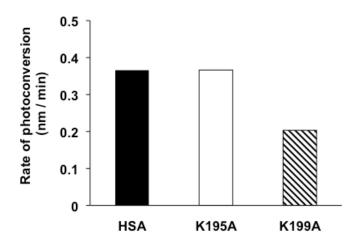




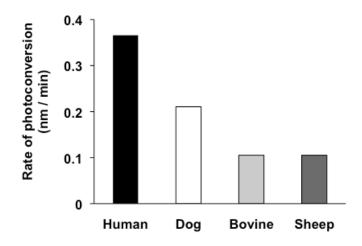




(A)



(B)



# FIGURE 8

# (A) 4*Z*,15*E*-BR

