Acta Crystallographica Section D
Biological Crystallography
ISSN 0907-4449

Structures of bovine, equine and leporine serum albumin

Anna Bujacz

Institute of Technical Biochemistry, Lodz University of Technology, Stefanowskiego 4/10, 90-924 Lodz, Poland

Received 14 May 2012
Accepted 14 June 2012

PDB References: bovine serum albumin, 4f5t; equine serum albumin, 4f5s; 4f5u; leporine serum albumin, 4f5v.

Serum albumin first appeared in early vertebrates and is present in the plasma of all mammals. Its canonical structure supported by a conserved set of disulfide bridges is maintained in all mammalian serum albumins and any changes in sequence are highly correlated with evolution of the species. Previous structural investigations of mammalian serum albumins have only concentrated on human serum albumin (HSA), most likely as a consequence of crystallization and diffraction difficulties. Here, the crystal structures of serum albumins isolated from bovine, equine and leporine blood plasma are reported. The structure of bovine serum albumin (BSA) was determined at 2.47 Å resolution, two crystal structures of equine serum albumin (ESA) were determined at resolutions of 2.32 and 2.04 Å, and that of leporine serum albumin (LSA) was determined at 2.27 Å resolution. These structures were compared in detail with the structure of HSA. The ligand-binding pockets in BSA, ESA and LSA revealed different amino-acid compositions and conformations in comparison to HSA in some cases; however, much more significant differences were observed on the surface of the molecules. BSA, which is one of the most extensively utilized proteins in laboratory practice and is used as an HSA substitute in many experiments, exhibits only 75.8% identity compared with HSA. The higher resolution crystal structure of ESA highlights the binding properties of this protein because it includes several bound compounds from the crystallization solution that provide additional structural information about potential ligand-binding pockets.

1. Introduction

The properties and functions of serum albumins have been described in many publications and books (Peters, 1996), but until now structural investigations have concentrated on human serum albumin (HSA; Curry, 2009). Serum albumin is a 67 kDa transport protein that is present in the body fluids of vertebrates and belongs to the minor group of proteins produced by the liver that do not undergo glycosylation. It is one of the most abundant proteins in mammalian plasma, ranging between 35 and 55 g l^{-1} (Schreiber & Urban, 1978). Human liver cells produce about 12 g of albumin per day, with the half-life of the protein reaching 19 d. During that time, a single albumin molecule passes through the circulatory system almost 15 000 times carrying various ligands, since serum albumin is the main carrier in the blood for metabolites, hormones and drugs. Albumin collects drugs from the place of application and hormones from secretion sites and delivers them to other parts of the body (Rothman & Orci, 1992). To be able to bind a variety of ligands with different chemical properties, albumins have developed a number of binding

sites, such as seven sites for long and medium fatty acids (FAs; Ashbrook *et al.*, 1972, 1975), four additional sites for short FAs (Bhattacharya *et al.*, 2000), two main sites for drugs (Sudlow *et al.*, 1975, 1976), one site for bilirubin (Zunszain *et al.*, 2008), which also binds haem in primates (Zunszain *et al.*, 2003), five sites with affinity for thyroxine (Petitpas *et al.*, 2003) and a few sites that are selective for single-ligand binding (Ghuman *et al.*, 2005). It is also capable of binding biologically important metal cations (Masuoka *et al.*, 1993). HSA is also capable of binding the protein fragment known as the GA module originating from the anaerobic bacterium *Finegoldia magna* (Lejon *et al.*, 2004, 2008). Changes in protein sequence dictate the shape and charge distribution of particular pockets and modulate the affinity of albumin for its ligands. Some of these pockets are selective for one group of ligands, but most of them are capable of binding ligands from various groups (Varshney *et al.*, 2010). The sequence identity between mammalian albumins varies from 65% for evolutionarily distant species to a single variation for the closest relatives (Fanali *et al.*, 2012).

Albumin was one of the first proteins to be described. Hippocrates of Cos noted in the 5th–4th century BC that patients with kidney disease had foamy urine, which is now known to be caused by the presence of albumin. In the 16th century, Paracelsus precipitated albumin from urine using vinegar. In 1840 Denis performed the first dialysis of serum albumin and in 1894 Gurber first crystallized equine serum albumin (Peters, 1996). The crystal structure of equine serum albumin (ESA) has been reported but has not been deposited in the Protein Data Bank (PDB; Ho *et al.*, 1993). However, to date structural investigations of albumin have mostly been based on human serum albumin (Carter & Ho, 1992), with more than 50 crystal structures of HSA complexes deposited in the PDB, mainly from the Curry laboratory (Curry, 2009).

1.1. Evolution of albumins

Serum albumin belongs to the four-member albumin superfamily, which also includes vitamin D-binding protein, α-foetoprotein and α-albumin (Cooke & David, 1985; Verboven *et al.*, 2002). Albumin architecture is predominately helical and consists of three domains with very similar conformations that create an overall heart-like shape. The domain borders are found in the middle of the longest helices (residues 193 and 382). Each domain possesses two subdomains: subdomain A, which contains six helices, and subdomain B with four helices.

The ancestral gene underwent triplication during the evolutionary process about 525 million years ago (Harper & Dugaiczyk, 1983), when vertebrates first appeared. The only documented living fossil is lamprey, which contains an albumin consisting of seven domains (Gray & Doolittle, 1992). Owing to its presence in all vertebrates and the availability of amino-acid sequences from a multitude of organisms, serum albumin serves as a time indicator of the evolution of a species (Metcalf *et al.*, 2003). For example, scientists investigating the phylogenetic tree of primate albumins have shown that orangutans separated early from the primates, followed by gorillas, later chimpanzees and finally humans (Sarich & Wilson, 1967).

Serum albumin is one of the most evolutionarily variable proteins, with 70–80% of its residues differing between the domains in a particular species; this could be the cause of the development of its highly advanced binding properties. During their evolution, albumins have developed binding ability for new groups of ligands such as hormones, metabolites or toxins. In contrast, retinol-binding protein has changed by 40% on average and histones by less than 10% (Doolittle, 1992). Although albumin is a rapidly changing protein, it has two features that are highly conserved not only among its domains but also among different species. The tertiary structure, which is highly helical and does not include any β-sheet fragments, is generally conserved, as well as the characteristic pattern of disulfide bridges.

1.2. Applications of albumins

In 1940, Cohn prepared human and bovine albumin solutions for intravenous injection to minimize osmotic shock after serious bleeding of wounded soldiers (Cohn, 1948). However, bovine albumin given intravenously caused sickness and in some cases death from kidney failure. Purified HSA is now widely used in medical studies and procedures and all expired transfusion blood is utilized for its production.

The concentration of serum albumin in the blood is a diagnostic parameter that can indicate many diseases (Peters, 1996). Owing to their easy accessibility, high stability, ability to bind various ligands and similarity to HSA, mammalian serum albumins, especially bovine serum albumin (BSA), are utilized in kinetic and affinity drug tests as replacements for HSA, as well as in many biochemical and pharmacological applications (Sohl & Splittgerber, 1991). BSA is used as a standard in many procedures, including vaccine preparation, as a drug by itself and as a test antigen to check the validity of assays (Leong & Fox, 1988), as a molecular-weight standard (Peters, 1996; Birnbaum & Nilsson, 1992; Sinibaldi *et al.*, 2008) and in immunological tests such as ELISA and others (Herzog *et al.*, 1983). ESA can replace BSA in radioimmunological and immunoenzyme analyses and its use in virological practice has been raised (Chan & Porter, 1967). Differing ligand-binding properties of serum albumins have been noted (Kandagal *et al.*, 2006), but no structural information has been reported to justify them; thus, the characterization of different albumins becomes an important practical issue. The crystal structures of the mammalian serum albumins described here allow an explanation of the dissimilarities between HSA, BSA, ESA and leporine serum albumin (LSA) at the molecular level.

1.3. Recognition by the immune system

Pure serum albumins display relatively low immunogenicity, which is useful in antivenom production (Lovrecek & Tomić, 2011), but in some cases a strong immunological response is observed after the injection of a serum albumin from another species. Such an effect is used to produce monoclonal antibodies against serum albumins from a particular species, which

is widely used in laboratory practice (Mogues *et al.*, 2005). To distinguish 'own/foreign' serum albumin, the immune system must use the surface fragments which differentiate these relatively similar proteins. A structural study may allow the identification of the protein fragments that are responsible for immunogenicity and the postulation of a mechanism for antigen–antibody recognition.

2. Materials and methods

As albumins from various mammalian species are easily obtained from commercial sources, the main challenges in performing crystallographic investigations were to find crystallization conditions that gave single crystals suitable for diffraction experiments and to identify the best cryprotection methods for preserving crystals from radiation damage.

2.1. Purification and crystallization of serum albumins

All of the albumins used in this study were isolated from mammalian blood and obtained from commercial sources. Bovine and leporine serum albumins were purchased from Sigma–Aldrich (St Louis, Missouri, USA), whereas equine albumin was purchased from Equitech (Kerrville, Texas, USA). The albumins were sold as 98–96% pure; however, they required additional purification, *i.e.* the elimination of fatty acids bound to the protein (Rosseneu-Motreff *et al.*, 1970) and of dimers formed during protein isolation caused by the unpaired cysteine that is present at position 34 in all mammalian serum albumins. Removal of FA was accomplished by adsorption onto charcoal (Fluka, St Louis, Missouri, USA; Chen, 1967). The disulfide-bridged dimer formed during protein isolation was removed by gel-filtration chromatography on an ÄKTA FPLC system using an XK 16/100 column (Amersham Biosciences, Uppsala, Sweden) filled with Superdex 200 prep grade. The elution buffer consisted of 100 m*M* NaCl, 10 m*M* Tris pH 7.3. The main fraction containing monomeric protein (Fig. 1) was concentrated on Vivaspin filters (Sartorius, Göttingen, Germany) with a 10 kDa cutoff using a centrifuge at 6500*g* (Eppendorf, Hamburg, Germany).

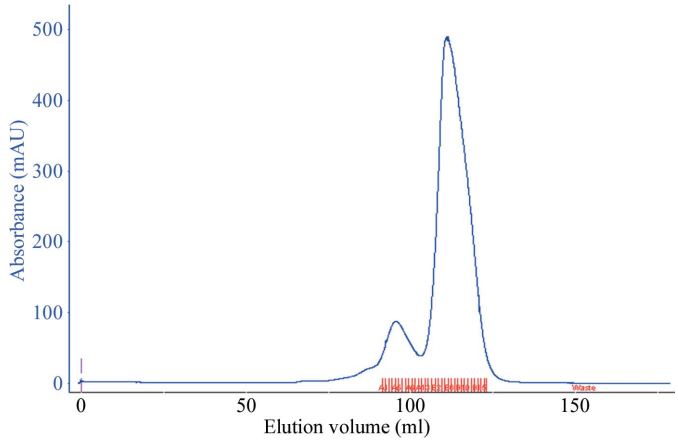


Figure 1
Gel-filtration chromatography of BSA.

Crystallization was set up by the hanging-drop vapour-diffusion method using 24-well plates and crystal screens from Hampton Research, Aliso Viejo, California, USA. Several crystal screens were tested and promising microcrystals of BSA were obtained using PEG 3350 with calcium acetate, potassium nitrate and ammonium chloride (from the PEG/Ion screen). Crystals of ESA were obtained using ammonium sulfate (from the Index screen) and crystals of LSA were obtained using PEG 3350 with ammonium acetate (from the PEG/Ion screen). The differences in the crystallization conditions are indicative of the differences between the albumins from these species. During optimization of the crystallization conditions, the concentration of PEG and its molecular weight, the salt concentration, the buffer type, the pH and the ratio of protein solution to well solution in the crystallization drop were adjusted. Additionally, it was observed that the concentration of each albumin seemed to be a crucial factor in crystal formation.

The BSA crystals grew from different PEG conditions with various salts. They always had undesirable morphology: thick elongated plates with a groove running along the largest surface of the crystal. A wide range of protein concentrations (40–120 mg ml^{-1}) were tested to find the best crystal shape. The best crystallization conditions for BSA were 20–24% PEG MME 5K, 0.15–0.3 *M* NH_4Cl and 0.1 *M* MES pH 6.5.

The physiological pH for serum albumin in plasma is in the neutral range. Initally, crystals of ESA were obtained under acidic conditions (0.1 *M* acetate buffer pH 4.5–5.0, 1.8–2.1 *M* ammonium sulfate) with a protein concentration of 60–120 mg ml^{-1}. However, during crystallization optimization, neutral conditions for crystal growth were found using 60–80% Tacsimate pH 7.0–7.5 at a protein concentration of 60–80 mg ml^{-1}; this crystal form is referred to in the text as ESA-t.

Crystals of LSA that initially grew from PEG 3350 and ammonium acetate had poor morphology in the form of flower petals with rounded edges. Several additives from Additive Screen (Hampton Research) improved their morphology, but also changed the unit-cell parameters. Three additives, glutathione, glycine and polypropylene glycol 400, were used for optimization of the crystallization conditions and the best crystals of LSA were grown from 14–20% PEG 3350, 6–10% PPG 400, 0.1–0.3 *M* ammonium acetate, 0.1 *M* HEPES pH 7.0–7.5 at a protein concentration of 40–80 mg ml^{-1}.

It should be stressed that the crystals of the albumins were grown from quite concentrated protein solutions with no adverse effects such as precipitation, although the actual range of concentrations useful for crystallization was rather wide.

2.2. X-ray diffraction data collection

Preliminary diffraction data were collected on various European synchrotron beamlines, including Max-lab in Lund, BESSY in Berlin and DESY in Hamburg, whereas the data for the final four structures presented here were measured on

Table 1
Data-collection and structure-refinement statistics.

Values in parentheses are for the last resolution shell.

	BSA	ESA	ESA-t	LSA
Data collection				
Radiation source	BL14.2, BESSY, Berlin	X13, DESY, Hamburg	X12, DESY, Hamburg	BL14.2, BESSY, Berlin
Year	2009	2010	2011	2011
Wavelength (Å)	0.9184	0.8123	0.9320	0.9400
Measurement temperature (K)	100	100	100	100
Space group	$C2$	$P6_1$	$P6_1$	$P2_12_12_1$
Unit-cell parameters (Å, °)	a = 217.8, b = 45.0, c = 143.1, $\alpha = \gamma$ = 90.0, β = 114.1	a = 96.6, b = 96.6, c = 143.8, $\alpha = \beta$ = 90.0, γ = 120.0	a = 89.2, b = 89.2, c = 134.8, $\alpha = \beta$ = 90.0, γ = 120.0	a = 74.2, b = 79.6, c = 104.2, $\alpha = \beta = \gamma$ = 90.0
Resolution range (Å)	40.0–2.47 (2.57–2.47)	40.0–2.31 (2.39–2.31)	50–2.04 (2.11–2.04)	50–2.27 (2.35–2.27)
Reflections collected	187936	204748	269009	173837
Unique reflections	43883	31470	38668	27095
Completeness (%)	95.2 (84.9)	98.9 (94.14)	99.9 (99.2)	93.1 (91.0)
Multiplicity	4.3 (3.0)	6.5 (2.8)	7.0 (5.4)	6.4 (4.0)
$\langle I\rangle/\langle\sigma(I)\rangle$	11.4 (1.8)	17.4 (1.9)	16.7 (2.1)	15.3 (1.7)
$R_{\rm int}$†	0.091 (0.549)	0.087 (0.521)	0.110 (0.639)	0.109 (0.499)
Refinement				
No. of reflections in working/test set	42367/1369	31495/1033	37399/1211	25667/1379
$R/R_{\rm free}$‡ (%)	19.69/25.88	19.71/25.85	19.61/24.29	19.29/25.76
No. of atoms (protein/solvent/ligands)	9338/327/10	4618/270/91	4659/353/63	4653/193/40
R.m.s.d.				
Bond lengths (Å)	0.021	0.022	0.021	0.020
Bond angles (°)	2.006	2.093	1.943	1.929
$\langle B\rangle$ (Å^2)	44.86	27.32	25.44	35.40
Residues in Ramachandran plot (%)				
Most favoured regions	91.9	92.1	93.2	91.0
Allowed regions	8.1	7.9	6.8	9.0

† $R_{\rm merge} = \sum_{hkl}\sum_i |I_i(hkl) - \langle I(hkl)\rangle|/\sum_{hkl}\sum_i I_i(hkl)$, where $I_i(hkl)$ is the intensity of observation i of reflection hkl. ‡ $R = \sum_{hkl} ||F_{\rm obs}| - |F_{\rm calc}||/\sum_{hkl} |F_{\rm obs}|$ for all reflections, where $F_{\rm obs}$ and $F_{\rm calc}$ are observed and calculated structure factors, respectively. $R_{\rm free}$ is calculated analogously for the test reflections, which were randomly selected and excluded from the refinement.

BL14.2 at BESSY (Mueller *et al.*, 2012) and X12 and X13 at DESY. The measurements were carried out on MAR CCD 165 and 225 detectors under a stream of cold nitrogen (100 K). The collected diffraction images were indexed and integrated with *DENZO* and the intensity data were scaled with *SCALEPACK* using the *HKL*-2000 program package (Otwinowski & Minor, 1997). The statistics of the diffraction data are presented in Table 1.

Diffraction experiments for albumin crystals were conducted utilizing different methods of cryoprotection. The BSA crystals did not require cryoprotectant, but had to be mounted on a very thin film from the crystallization buffer. Excess liquid was removed by gently touching the mounting loop to the plate surface. The ESA crystals were grown from ammonium sulfate and required the use of 30% glycerol as a cryoprotectant. Crystals obtained in the presence of 70% Tacsimate were mounted without additional cryoprotection, confirming the fact that carboxylate salts present in the crystallization drop can successfully act as cryoprotective agents during diffraction experiments (Bujacz *et al.*, 2010). These new conditions were not only highly cryoprotective, but also improved the initial resolution from 2.32 to 2.04 Å. For the LSA crystals, a mixture of well solution and 50% PEG 400 in a 1:1 ratio was used as the cryoprotectant.

Owing to the high concentration of the protein solutions, which caused a thick 'skin' on crystallization drops that contained PEG, a special technique had to be used to 'fish' the crystals from the drop. Using a Hampton Research tool, a triangular window was cut in the 'skin' of the drop and flipped open. Next, using a mounting loop through the window, the target crystal was separated from the remaining crystals and from the skin and pulled out of the drop. This method allowed the safe harvesting of a number of crystals from a single drop covered with a dried shell.

2.3. Structure determination and refinement

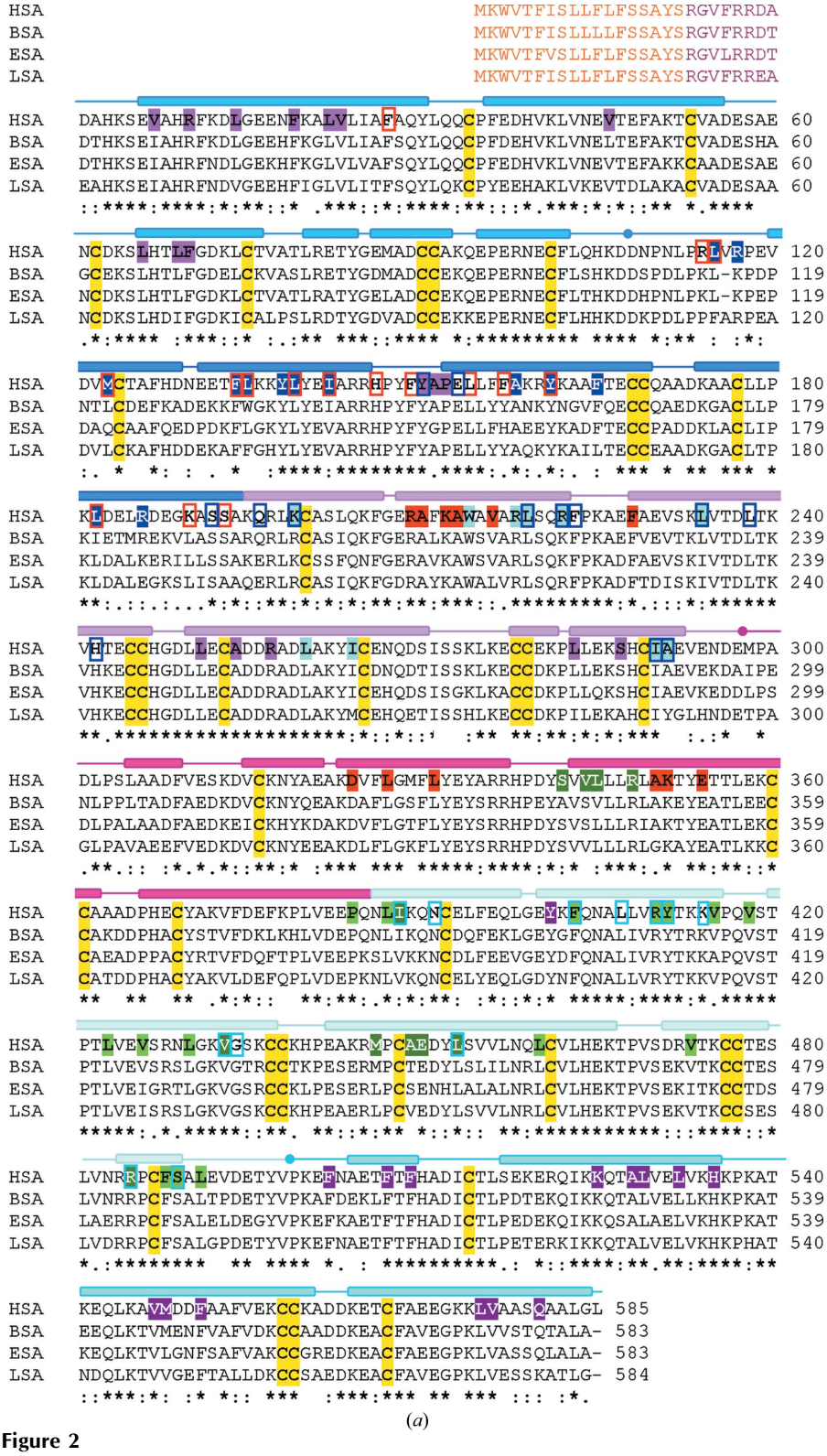
The structure of BSA was solved at 2.47 Å resolution by molecular replacement (Rossmann, 1990) with the program *Phaser* (McCoy *et al.*, 2007) using the HSA data set with PDB code 1ao6 (Sugio *et al.*, 1999) as a model. The amino-acid sequence was modified with the program *CHAINSAW* (Stein, 2008; Schwarzenbacher *et al.*, 2004); the side chains were removed from the model, leaving only the part common to the sequence of the target protein. The crystal of BSA belonged to a centred monoclinic system in space group $C2$; the Matthews coefficient (Matthews, 1968) indicated the presence of two monomers in the asymmetric unit.

The ESA crystal structures were solved at resolutions of 2.32 and 2.04 Å using crystals obtained under different crystallization conditions, although both crystal forms belonged to the hexagonal space group $P6_1$. The refined crystal structure of BSA was used as a model in molecular replacement to determine the ESA structure for the crystal grown from ammonium sulfate. The second set of diffraction data (ESA-t) was re-indexed to the same set of axes as the apo ESA form.

Next, the model of the ESA crystal structure was refined against the ESA-t diffraction data by a rigid-body procedure (Murshudov *et al.*, 2011). Despite the relatively large differences in their unit-cell parameters, the refinement reached convergence and produced a model with similar positions of the protein molecules in the unit cell.

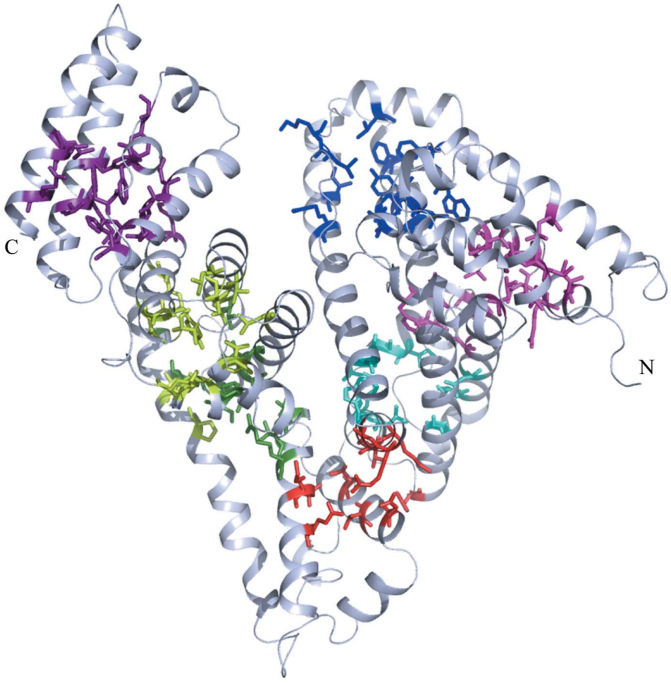
The fourth crystal structure presented here, that of LSA, was solved by molecular replacement in *Phaser* using the HSA and BSA structures as search models with automatic implementation of the sequence of LSA. The structure was determined to 2.27 Å resolution in an orthorhombic system, space group $P2_12_12_1$, with a single monomer in the asymmetric unit (Table 1). This structure presented considerable difficulties in fitting the coordinates to the electron-density map because some loops and subdomain IIIB had a different conformation compared with the starting model. This part of the structure was removed after the first cycle of refinement and built from scratch into the $F_o - F_c$ OMIT maps.

In all cases, structure refinement was carried out in *REFMAC*5 (Murshudov *et al.*, 2011) using maximum-likelihood targets including TLS parameters (Winn *et al.*, 2001) defined for each subdomain and separately for both polypeptide chains in the BSA structure. Model rebuilding was carried out in *Coot* (Emsley & Cowtan, 2004); water molecules were introduced manually. The progress of refinement was monitored and the models were validated using R and R_{free} tests, the r.m.s.d.s of bonds and angles, the FOM and the map quality (Brünger, 1992). Complete amino-acid sequences could be modelled in each structure. The main-chain atoms and a majority of the side chains were covered by a $2F_o - F_c$ map contoured at 1σ, but in rare cases it was necessary to decrease the σ level to 0.7 during model rebuilding of the flexible side chains. It was attempted to complete the model side chains even if the atomic displacement parameters were becoming high, since the side chains, especially on the surface, determine properties of the molecules such as antigen recognition,



(*a*)

Figure 2
(*a*) Sequence alignment of HSA, BSA, ESA and LSA. The first row shows the prepropeptide fragments. The numbering corresponds to the mature albumins. The residues creating the main FA binding pockets in HSA (PDB entry 2e7g) are marked in the following colours: FA1, blue; FA2, pink; FA3, green; FA4, light green; FA5, violet; FA6, red; FA7, cyan. The residues creating the haemin-binding site in HSA (PDB entry 1n5u) are marked with red frames. The amino acids in the Sudlow sites are marked by dark blue (DSI) and cyan (DSII) frames (based on PDB entries 2bxc and 2bxg). The bars over the sequence represent helices according to the BSA secondary structure, with the domains coloured as follows: subdomain IA/IB, blue/dark blue; subdomain IIA/IIB, pink/raspberry; subdomain IIIA/IIIB, pale cyan/cyan.



(*b*)

Figure 2 (continued)
(*b*) BSA monomer with FA pockets coloured analogously to (*a*).

ligand binding and surface potential. All crystallographic calculations were performed using the *CCP*4 suite of programs (Winn *et al.*, 2011). Molecular illustrations were prepared using *PyMOL* (DeLano, 2002). The quality of the final structures was assessed using *PROCHECK* (Laskowski *et al.*, 1993). The final refinement statistics are shown in Table 1.

The refined atomic coordinates and structure factors for the crystal structures presented here have been deposited in the PDB with accession codes 4f5t, 4f5s, 4f5u and 4f5v for BSA, ESA, ESA-t and LSA, respectively.

3. Results and discussion

3.1. Comparison of the sequences and crystal structures of HSA, BSA, ESA and LSA

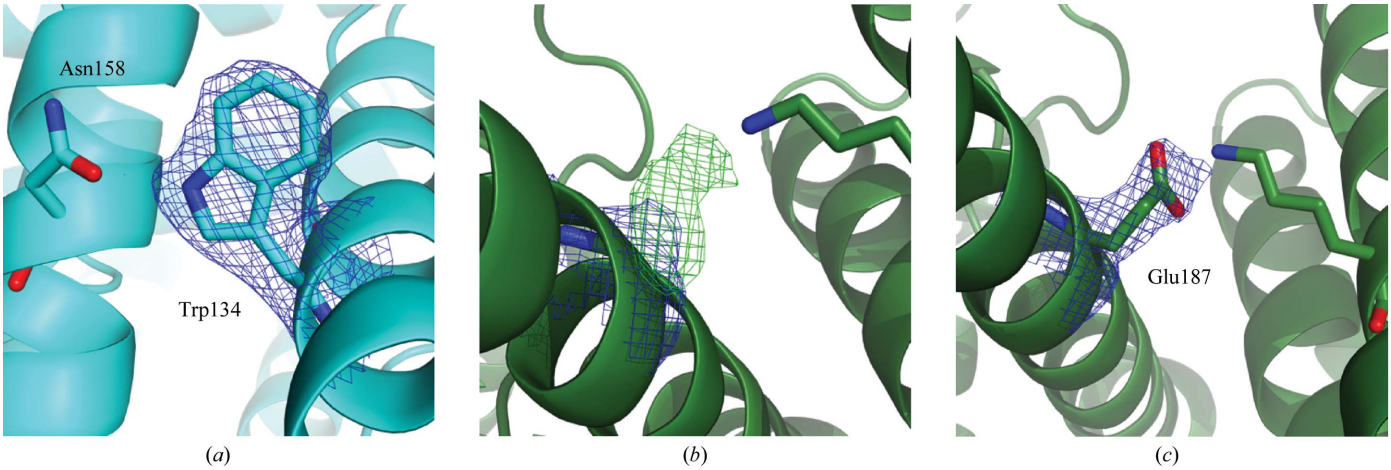
A comparison of the sequences of the investigated proteins shows differences in the amino-acid composition of the prepropeptide fragment which is removed before the release of the albumins from the liver into the circulatory system (Patterson & Geller, 1977). The cleavage process has two stages. In the first step the prepropeptide is removed in the cell producing the albumin and in the second the shorter fragment, the propeptide, is hydrolyzed just before the mature serum albumin is secreted into the blood (Rothman & Orci, 1992). The changes in the sequence of serum albumin are highly correlated with the evolution of the species and even closely related organisms can be distinguished by their SA sequences (Metcalf *et al.*, 2003).

Mature HSA contains 585 amino acids, LSA contains 584 amino acids and BSA and ESA contain 583 amino acids each; the identity between the enzymes is in the range 71–76%. A sequence alignment of the serum albumins investigated here produced using the *ClustalW*2 server (Thompson *et al.*,

Table 2
Structural comparison of HSA, BSA, ESA and LSA aligned on C^{α} atoms in *Coot*.

The upper part above the diagonal shows the r.m.s.d. (Å) for alignment of whole chains, with the sequence identity/similarity in parentheses, and the lower part below the diagonal shows the r.m.s.d. (Å) of individual aligned domains (domains I/II/III). For HSA and BSA monomer *A* was chosen arbitrarily.

	HSA (*A*)	BSA (*A*)	ESA	ESA-t	LSA
HSA (*A*)	—	1.53 (75.8/88.0)	1.44 (76.3/88.2)	1.76 (76.3/88.2)	1.43 (74.3/88.0)
BSA (*A*)	0.93/1.16/1.07	—	1.84 (73.9/88.0)	2.18 (73.9/88.0)	1.44 (71.5/86.3)
ESA	0.89/0.73/1.14	0.89/1.35/1.09	—	1.81 (100)	1.39 (71.0/85.2)
ESA-t	1.32/1.17/1.42	1.28/1.53/1.51	1.10/1.05/1.31	—	1.75 (71.0/85.2)
LSA	0.93/0.61/1.43	0.89/1.22/1.70	0.98/0.68/1.85	1.37/1.05/1.93	—



(*a*) (*b*) (*c*)

Figure 3
(*a*) Electron-density map for Trp134 in the BSA structure. (*b*) $F_o - F_c$ map showing the density for Glu187 in the LSA structure; (*c*) the same residue after refinement. The $2F_o - F_c$ maps are displayed in blue at 1.0σ and the $F_o - F_c$ maps are displayed in green at 3.0σ.

Table 3
Domain comparison (r.m.s.d. and sequence identity) for each investigated serum albumin.

	R.m.s.d. (Å)/sequence identity (%)				
	HSA (*A*)	BSA (*A*)	ESA	ESA-t	LSA
I–II	2.87 (25.7)	2.49 (30.0)	2.53 (28.3)	1.89 (30.5)	2.27 (26.1)
I–III	2.44 (15.5)	2.85 (16.4)	2.72 (18.7)	2.56 (15.1)	2.46 (14.1)
II–III	2.99 (19.6)	2.87 (17.0)	2.71 (20.1)	2.71 (20.4)	2.99 (19.4)

1994) is presented in Fig. 2 and Table 2. UniProt reports two natural variants of BSA that differ by a A190T mutation (Brown, 1975). The electron-density map for the reported crystal structure unequivocally shows a threonine at this position. BSA possesses an additional Trp134 (Fig. 3) that is not found in HSA, ESA and LSA. This residue is located in the FA1 binding pocket and is important from the point of view of spectrofluorometric studies of binding ligands. The residues in the FA binding sites of HSA (Bhattacharya *et al.*, 2000) are marked in different colours in this figure. The three-domain structure and the distribution of 17 disulfide bridges is highly conserved among all mammalian serum albumins (Brown, 1974). The only cysteine that is not involved in the creation of a disulfide bridge is found at position 34. This is the key amino acid that is involved in forming dimers by making a covalent connection *via* an intermolecular disulfide bridge, which is a very undesirable side effect during albumin isolation.

The two sequences of LSA deposited in SwissProt, P49065 and G1U9S2, differ by three residues and the initial model of LSA was built based on the first sequence. Positive electron density appeared as a 'side chain' of Gly187 after a few cycles of refinement and disagreement with the assumed sequence was observed at some other residues. The second variant of the deposited sequence differs from the first by E186K, G187E and S189A substitutions. The second sequence is referred to as that from the particular rabbit strain Thorbecke.

Structural alignment of the crystal structures of albumins using the *SSM* function in *Coot* shows exactly the same locations of disulfide bridges despite the differences in the main peptide chains and the subdomain rotations (Fig. 4). The r.m.s. deviations of aligned pairs of C^{α} atoms between the crystal structure of the apo form of HSA (PDB entry 1ao6; Sugio *et al.*, 1999) and the structures of the mammalian serum albumins presented here are in the range 1.43–1.76 Å. The lowest r.m.s.d. between these albumins is that between ESA and LSA (1.39 Å), whereas the largest is that between BSA and ESA-t (2.18 Å). The structural similarity is not correlated with sequence identity. Albumin molecules are very flexible and thus even the structures of ESA crystallized under different conditions exhibited an r.m.s.d. of as high as 1.81 Å (Table 2). Although the sequence similarity is quite high, flexibility of the domains is the cause of the relatively high r.m.s.d.s of the aligned structures.

3.2. Structure and evolution of serum albumin domains

Serum albumins are built from three structurally similar domains. They are rare examples of proteins in which the domain divisions can be assigned to the middle of secondary-structure elements: long α-helices (Fig. 5*a*). Such atypical domain linkages are the result of gene triplication in the early stages of evolution. Each domain is composed of ten helices and can be further divided into two subdomains, A and B, containing six and four helices, respectively. The two subdomains are connected by a long loop. The conformational flexibility between the domains is based on helix bending, whereas loop flexibility is responsible for the change of orientations of the subdomains. The space between the domains and the subdomains is utilized to create binding pockets for various ligands.

It can be assumed that in early vertebrates, just after gene multiplication, all three domains had identical sequences and a very similar structure. During evolution the domains differentiated, changing both their sequence and structure. Table 2 shows the r.m.s.d. between individual domains and the sequence identity for all of the investigated albumins using the sequence of HSA as a reference. A graphical representation of the superposed domains of BSA is presented in Fig. 5(*b*). The lowest sequence identity and (usually) highest r.m.s.d. for the superposition of domain III on the other two domains shows that the evolution of this fragment was the fastest. The sequence identity in the range 14.1–30.7% shows that the albumins have exchanged 70–85% of their residues in 500 million years of evolution. Although a majority of the residues had mutated during the evolutionary process, the positions of the cysteines and the locations of the six disulfide bridges in the two domains remained unchanged. The exception was domain I, which had lost one of its cysteines and one of the disulfide bridges. The remaining cysteine in position 34 is the only unpaired cysteine in the whole

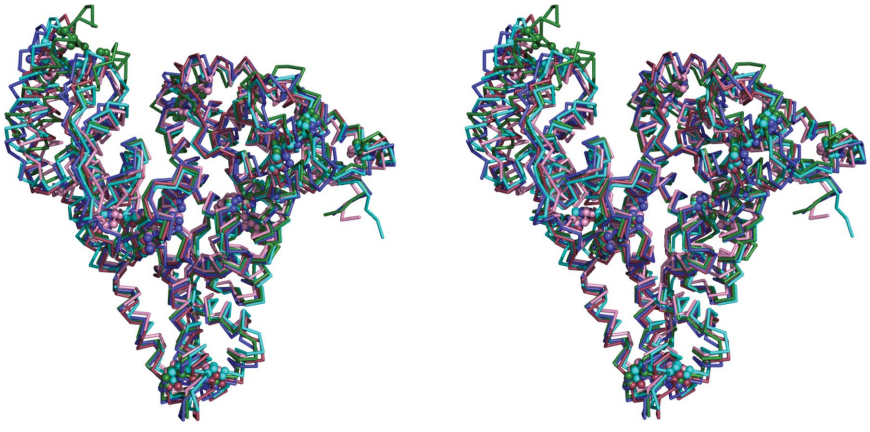


Figure 4
A stereoview of the aligned crystal structures of serum albumins. HSA, burgundy (PDB entry 1ao6; Sugio *et al.*, 1999); BSA, cyan; ESA, blue; ESA-t, violet; LSA, green. The disulfide bridges are shown as spheres (the N-terminus is on the right).

molecule in all of the investigated albumins (Fig. 2). The r.m.s.d. between the superposed domains is in the range 1.85–3.01 Å and shows that although the domains have changed significantly they have maintained their overall architecture and distribution of secondary-structure elements.

Superposition of the three domains of BSA, as well as of those of ESA and LSA, shows significant differences in their standard deviations. Sequence similarity among the domains is no greater than 30% and the r.m.s.d. of C^{α} atoms ranges from 1.9 to 3.0 Å (Table 3). The highest sequence identity and usually the lowest r.m.s.d. (with the exception of HSA) are observed on superposition of domains I and II (Fig. 5). The lowest sequence identity is between domains I and III, but the highest r.m.s.d. indicates that the largest structural differences are between domains II and III. This shows that lower sequence identity is not directly correlated with lower structural similarity. Domains I and II are relatively similar and domain III is most different sequentially from domain I and most different structurally from domain II. This indicates that the sequence evolution of individual domains proceeds in the same direction, but the fastest evolutionary change occurred in domain III (Table 3). Additionally, domain III in all albumins shows the largest conformational flexibility, which is indicated by the high r.m.s.d. after individual domain superposition for all investigated albumins (Table 2).

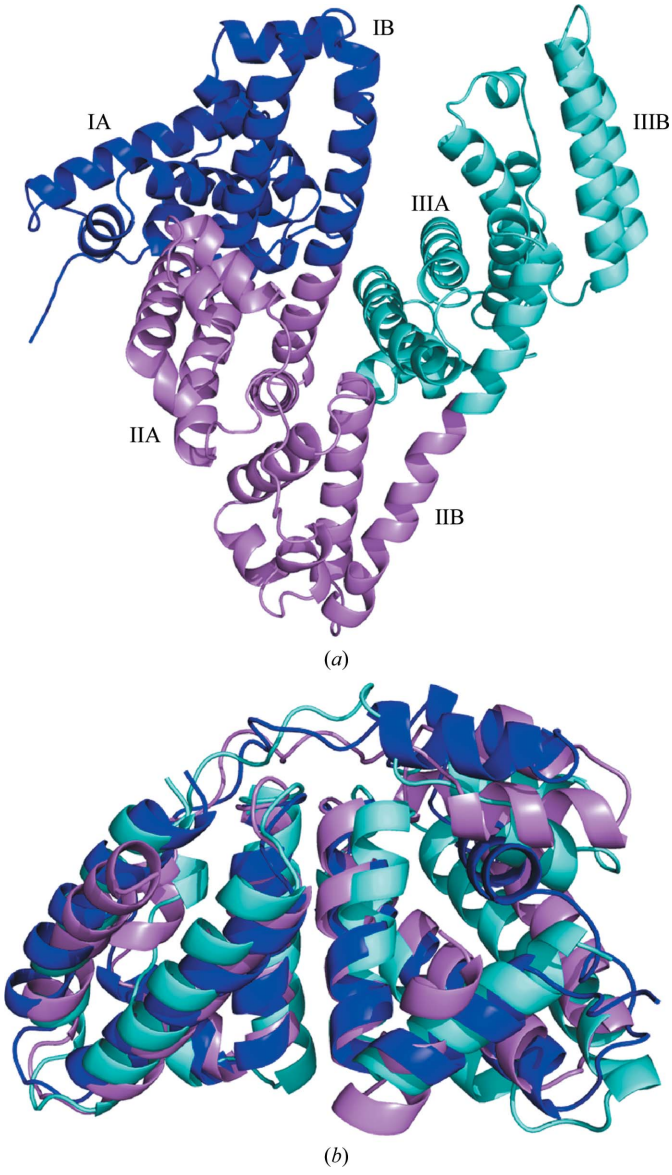


Figure 5
(*a*) Domain division in the BSA structure and (*b*) aligned domains I (blue,) II (pink) and III (cyan).

3.3. Flexible regions and domain movement

The tight network of disulfide bridges makes the internal structure of the subdomains rigid. The A subdomains are composed of six helices stapled by four disulfide bridges in all of the crystal structures discussed here. Subdomain IA is an exception since it contains only three disulfide bridges. The number of disulfide bridges in the B subdomains is lower, with four helices rigidified by two disulfide bridges. The linkage between subdomains A and B in each domain is created by a long flexible loop. An analysis of temperature factors indicates that the most flexible regions are found in these loops. This is also reflected in the differences in the conformations of the superposed structures. The movement of the subdomains modulates the volumes of the binding pockets that are located between them. Mutual orientation of the domains is regulated by a different mechanism. Domains I–II and II–III are connected by the longest helices and the division between the domains is in their centre. A single bend in the vicinity of the domain division in these two helices causes increased thermal motion of residues 183–187 and 381–384 and may be responsible for the domain movement (Fig. 5).

3.4. Comparison of the binding pockets

Structural studies of HSA show that this protein possesses seven binding pockets with high affinity for long-chain and medium-chain fatty acids (Bhattacharya *et al.*, 2000). Similar biochemical data confirm that analogous binding pockets are present in other albumins (Urlich, 1994). The residues forming these binding pockets in HSA complexed with myristic acid (PDB entry 1e7g; Bhattacharya *et al.*, 2000) are marked in different colours in Fig. 2. A highly conserved motif consisting of residues belonging to the same pockets in all the investigated albumins can be observed. Of the 89 residues that create the FA pockets, 21 are replaced by similar residues, five are replaced by different residues and the remaining 63 are the same in HSA, BSA, ESA and LSA. This high level of similarity allows it to be postulated that the positions of the FA binding pockets may be the same in other mammalian albumins and that the affinity for FA is similar among the investigated serum albumins. The fatty-acid binding pockets FA1, FA3, FA4 and FA7 are located inside subdomains IB, IIIA, IIIA and IIA, respectively. Although the subdomain helical fragments are fixed by the disulfide bridges, the loops on the opposite side of the helix bundle are still flexible and provide a mechanism for controlling the volume of these binding pockets.

In HSA, the bilirubin-binding pocket is also responsible for accommodation of haemin (PDB entries 1n5u, 1o9x and 2vue; Wardell *et al.*, 2002; Zunszain *et al.*, 2003, 2008) and partially overlaps with long and medium FA binding site 1. Comparing the number of residues interacting with bilirubin or haemin, it is clear that this pocket has the ability to extend to accommodate large ligands. Among mammals, only primate serum albumins are capable of binding haemin owing to the presence of Lys190 at the entrance to this pocket. This residue replaces the hydrophobic leucine that is found in this position in other albumins. This example shows that a single amino-acid substitution can considerably change the binding properties of a particular pocket (Adams & Berman, 1980; Wardell *et al.*, 2002).

The amino acids in the Sudlow drug-binding sites DSI and DSII concurrently create pockets for FA7 and FA3/FA4, respectively (Ghuman *et al.*, 2005). The hydrophobic residues located at the bottom of these pockets are more highly conserved then the polar residues situated at their entrances. The changes in protein sequence adjust the shape and charge distribution of the pockets and modulate the affinity of particular albumins for selective ligands. The residues that interact with phenylbutazone in DSI (PDB entry 2bxc) and with ibuprofen in DSII (PDB entry 2bxg) are marked in Fig. 2.

3.5. Crystallization components as markers of the binding sites

A number of small molecules were found in the crystals of bovine, equine and leporine serum albumin and were identified as the ions of acetic, malonic, malic and succinic acids, fragments of polyethylene glycol and sulfate ions. The composition of the crystallization solution, the shape of the electron-density maps and the character of the surrounding amino-acid side chains were taken into consideration in modelling these ligands into the structure. Their binding reveals the high affinity of albumins for binding various ligands. BSA and LSA were crystallized in the presence of different PEGs, and fragments of these polymers are visible in these structures. They are bound to the FA5 pocket in BSA monomer *A* and to the FA6 pocket in the LSA structure. One of the ESA structures was crystallized from a solution containing ammonium sulfate. A number of sulfate ions are bound in positions corresponding to the carboxylate groups of fatty acids in the FA3 and FA7 pockets of the HSA–myristic acid complex (PDB entry 1e7g). The acetate buffer anions are bound to the FA1, FA4 and FA7 pockets, and glycerol from the cryoprotection solution is bound to the FA3 pocket. A number of acetate and sulfate anions and glycerol molecules are bound on the surface of this protein. The ESA-t structure was determined from a crystal grown in solution containing Tacsimate. Dicarboxylic acids such as succinic, malic and malonic acids are bound in the FA2, FA3, FA4, FA6 and FA7 pockets (Fig. 6). A single malonate ion is bound in all of these binding pockets and the position of its carboxyl group creates similar interactions with protein side chains as the fatty-acid carboxyl group in the HSA–myristic acid complex (PDB entry 1e7g), with the exception of FA7, corresponding to the Sudlow DSI site, which is occupied by malonate, succinate and malate ions. Additionally, the two malonate ions in the FA4 and FA6 pockets have one of their carboxylate groups in a position similar to the location of the carboxylate group of ibuprofen in the HSA–ibuprofen complex (PDB entry 2bxg).

3.6. Surface potential

The ligand-binding pockets are conserved in all serum albumins discussed here, but the surface residues show relatively low similarity, which is visible after analyzing the surface potentials of the investigated proteins (Fig. 7). The surface potential was calculated by the Adaptive Poisson–Boltzmann Solver method (Baker *et al.*, 2000) and was visualized using *PyMOL* (DeLano, 2002). Differences between residues on the protein surface can cause an immunological response after the injection of an albumin from a foreign species into the host organism. Antibodies against a particular albumin are widely used in laboratory procedures, allowing signal multiplication and albumin identification.

A comparison of the C^{α} traces of serum albumins indicates that the backbones of the proteins are similar (Fig. 4 and Table 2). A large fraction of the charged long side chains on the surface of the molecules create a coating which modulates the shape and charge distribution. These surface residues are the least conserved between species and may be responsible for specific molecule recognition, *e.g.* by the immune system. The surface potentials of the serum albumins discussed here are presented in Fig. 7. The sides of the molecules presented in the left panel show a similar pattern of charge distribution. Much larger differences are visible in the right panel; a different potential on this side of the molecules is especially visible on the surfaces of domains I and III, as well as at the edges of the molecules. Additionally, the surface roughness of

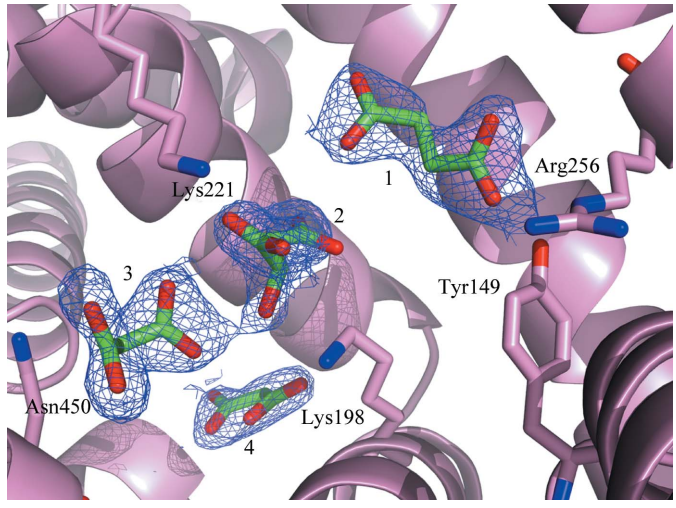


Figure 6
Tacsimate components in the FA7 binding pocket of ESA: succinate ion (1) interacts with Arg256, Tyr149 and Lys221, malate ion (2) interacts with Lys221 and Lys198, and two malonate anions (3 and 4) interact with Arg217, Lys194 and Ser201. The $2F_o - F_c$ maps are displayed at 1.0σ.

this side differs between the investigated structures more than for the other side, as shown in the left panel. The epitopes for anti-BSA murine monoclonal antibody are localized on various patches of the surface, especially on the outer edge of domain I (Ueno *et al.*, 1994), which supports the hypothesis that fragments with unique charge distribution and surface shape are utilized by the immune system for specific recognition.

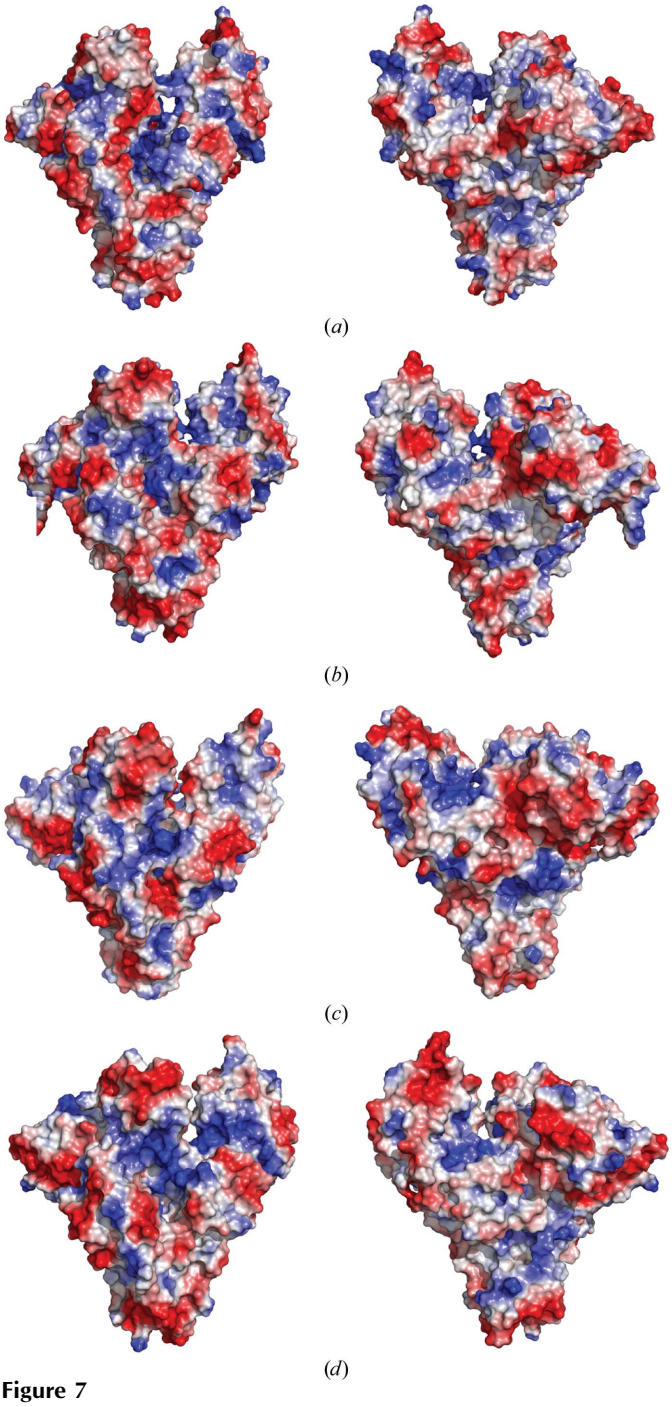


Figure 7
Charge distribution on the surface of serum albumins: (*a*) HSA, (*b*) BSA, (*c*) ESA and (*d*) LSA. The left panel shows molecules in the orientation with domain I on the left and further domains arranged anticlockwise, whereas in the right panel domain I is on the right and the domains are arranged clockwise.

3.7. Influence of the surface residues on crystal packing

Two crystal forms of HSA have been utilized in most of the previously published studies. Unliganded protein has been crystallized in a primitive triclinic cell with two molecules present in the asymmetric unit, whereas complexes with fatty acids (usually myristic) have been crystallized in a *C*-centred monoclinic system with a single monomer in the asymmetric unit. BSA crystallizes in a monoclinic system in space group *C*2 with two monomers in an asymmetric unit (Fig. 8). These two monomers create a dimer in the crystal lattice related by a noncrystallographic twofold axis which is approximately parallel to the (1 0 0.5) direction of the unit cell. The dimer interface is located on the surface of domains I and III and is relatively large: 17 110 Å^2 of the buried surface. The interactions between subdomains IB and IIIB of both monomers

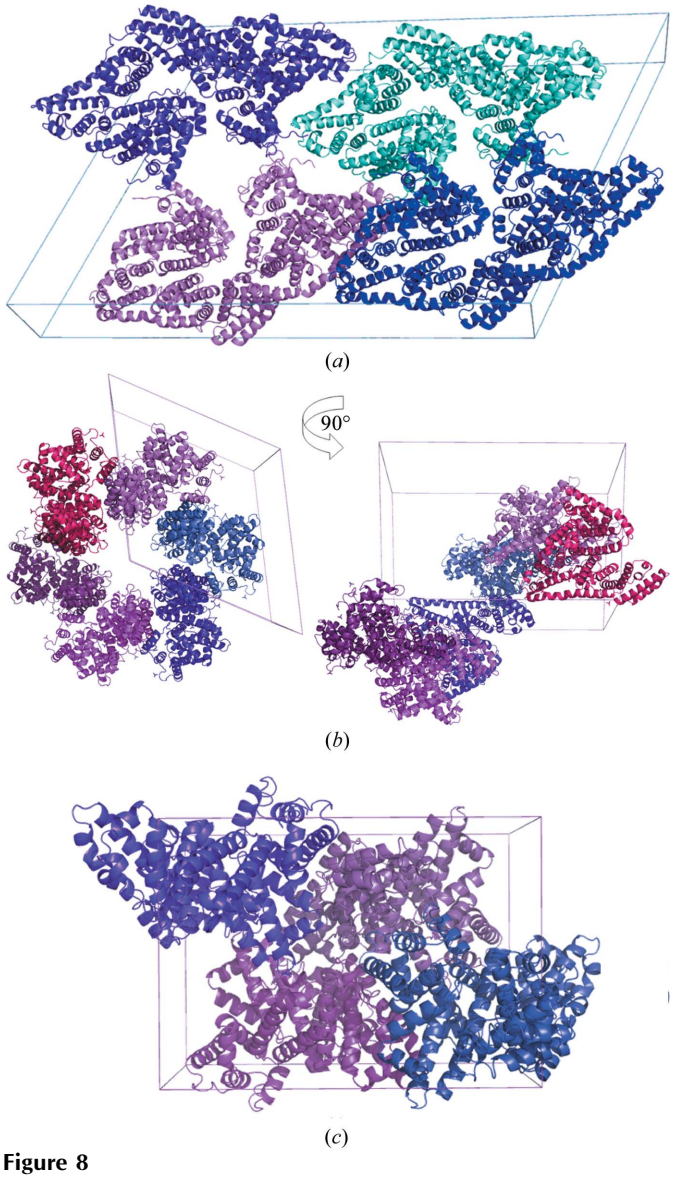


Figure 8
Crystal packing (*a*) of BSA showing four dimers in the unit cell, (*b*) of ESA with six monomers in the unit cell arranged around a 6_1 screw axis and (*c*) of four monomers related by a twofold screw axis in the LSA unit cell.

are polar. Crystals of BSA obtained under different crystallization conditions were always isomorphous, which means that the interactions that are responsible for dimer formation are stronger than in HSA and promote this crystal packing. The dimer present in the crystal structure of BSA differs from the covalent sulfide-bridged dimer that was removed during purification (Fig. 1). Cys34 is evidently unbound and makes a hydrogen bond to Tyr84 in both monomers.

Both ESA crystal forms belonged to the hexagonal system with space group $P6_1$ and the helical assembly of the molecules around the sixfold axis is visible in Fig. 5(*b*). A similar arrangement of molecules exists in both crystals and is based on interactions between subdomain IA and subdomain IIIB of the next molecule in the spiral assembly. The main interactions responsible for the creation of the sixfold screw axis are maintained in both forms, but a few additional intermolecular contacts are visible in the more densely packed form (ESA-t).

LSA crystallizes in an orthorhombic system in space group $P2_12_12_1$ with one monomer in the asymmetric unit. An additive screen was used during the optimization of the crystallization conditions, resulting in two other orthorhombic crystal forms of LSA being obtained which had the same space group but different unit-cell parameters. An analysis of the crystal packing showed that the molecules of LSA in these crystal forms had slightly different orientations with respect to the crystallographic twofold axes.

The differences in amino-acid sequences on the surface of the investigated structures caused different intermolecular contacts between symmetry-related molecules, resulting in the observed variety of space groups for such closely related proteins.

4. Conclusions

The crystal structures of BSA, ESA and LSA presented in this paper confirmed the overall similarity of these proteins to HSA and revealed a number of differences in the binding pockets, as well as a considerable variation in surface structure and charge distribution. All serum albumins have a mostly helical architecture, with identical division into domains and subdomains. However, the three domains originating from a common ancestor of all vertebrates have evolved significantly. The well refined structures allowed structural comparison of these domains, showing discrepancies between their subdomains and allowing the largest structural and sequence changes to be pinpointed. Structural alignment of the whole albumin molecules shows their conformational lability. This flexibility is connected to the function of albumins, which requires their adjustability for binding a wide variety of ligands.

This work has identified crystallization conditions for a number of serum albumins and describes methods for cryoprotection of the crystals, creating a good starting point for further research on the complexes of albumin with important groups of ligands. Direct structural comparisons of the complexes of various albumins with the same ligands will highlight differences in the binding properties of particular pockets and species.

Note added in proof: While this paper was under review, an independent publication describing structural studies of bovine, horse and rabbit albumins became available (Majorek *et al.*, 2012).

This work was supported by grant No. N N405 3639 39 from the Polish Ministry of Science and Higher Education. I would like to thank K. Zielinski and B. Sekula for help in crystallization and my husband, G. Bujacz, for synchrotron measurements and scientific discussions. I am grateful to A. Wlodawer, NCI for the gift of several albumin samples as well for his help in editing the manuscript.

References

Adams, P. A. & Berman, M. C. (1980). *Biochem. J.* **191**, 95–102.
Ashbrook, J. D., Spector, A. A. & Fletcher, J. E. (1972). *J. Biol. Chem.* **247**, 7038–7042.
Ashbrook, J. D., Spector, A. A., Santos, E. C. & Fletcher, J. E. (1975). *J. Biol. Chem.* **250**, 2333–2338.
Baker, N., Holst, M. & Wang, F. (2000). *J. Comput. Chem.* **21**, 1343–1352.
Bhattacharya, A. A., Grüne, T. & Curry, S. (2000). *J. Mol. Biol.* **303**, 721–732.
Birnbaum, S. & Nilsson, S. (1992). *Anal. Chem.* **64**, 2872–2874.
Brown, J. R. (1974). *Fed. Proc.* **33**, 1389.
Brown, J. R. (1975). *Fed. Proc.* **34**, 591.
Brünger, A. T. (1992). *Nature (London)*, **355**, 472–475.
Bujacz, G., Wrzesniewska, B. & Bujacz, A. (2010). *Acta Cryst.* D**66**, 789–796.
Carter, D. & Ho, J. X. (1992). *Nature (London)*, **358**, 209–215.
Chan, P. C. Y. & Porter, R. R. (1967). *Immunology*, **13**, 633–640.
Chen, R. F. (1967). *J. Biol. Chem.* **242**, 173–181.
Cohn, E. J. (1948). *Advances in Military Medicine*, edited by E. C. Andrus, 364–443. Boston: Little, Brown.
Cooke, N. E. & David, E. V. (1985). *J. Clin. Invest.* **76**, 2420–2424.
Curry, S. (2009). *Drug Metab. Pharmacokinet.* **24**, 342–357.
DeLano, W. L. (2002). *PyMOL*. http://www.pymol.org.
Doolittle, R. F. (1992). *Protein Sci.* **1**, 191–200.
Emsley, P. & Cowtan, K. (2004). *Acta Cryst.* D**60**, 2126–2132.
Fanali, G., Ascenzi, P., Bernardi, G. & Fasano, M. (2012). *J. Biomol. Struct. Dyn.* **29**, 691–701.
Ghuman, J., Zunszain, P. A., Petitpas, I., Bhattacharya, A. A., Otagiri, M. & Curry, S. (2005). *J. Mol. Biol.* **353**, 38–52.
Gray, J. E. & Doolittle, R. F. (1992). *Protein Sci.* **1**, 289–302.
Harper, M. E. & Dugaiczyk, A. (1983). *Am. J. Hum. Genet.* **35**, 565–572.
Herzog, P. J., Shaw, A., Lindsay-Smith, J. R. & Garner, R. C. (1983). *J. Immunol. Methods*, **62**, 49–58.
Ho, J. X., Holowachuk, E. W., Norton, E. J., Twigg, P. D. & Carter, D. C. (1993). *Eur. J. Biochem.* **215**, 205–212.
Kandagal, P. B., Ashoka, S., Seetharamappa, J., Shaikh, S. M., Jadegoud, Y. & Ijare, O. B. (2006). *J. Pharm. Biomed. Anal.* **41**, 393–399.
Laskowski, R. A., MacArthur, M. W., Moss, D. S. & Thornton, J. M. (1993). *J. Appl. Cryst.* **26**, 283–291.
Lejon, S., Cramer, J. F. & Nordberg, P. (2008). *Acta Cryst.* F**64**, 64–69.
Lejon, S., Frick, I. M., Björck, L., Wikström, M. & Svensson, S. (2004). *J. Biol. Chem.* **279**, 42924–42928.
Leong, M. M. & Fox, G. R. (1988). *Anal. Biochem.* **172**, 145–150.
Lovrecek, D. & Tomić, S. (2011). *Coll. Antropol.* **35**, 249–258.

Majorek, K. A., Porebski, P. J., Dayal, A., Zimmerman, M. D., Jablonska, K., Stewart, A. J., Chruszcz, M. & Minor, W. (2012). *Mol. Immunol.* **52**, 174–182.
Masuoka, J., Hegenauer, J., Van Dyke, B. R. & Saltman, P. (1993). *J. Biol. Chem.* **268**, 21533–21537.
Matthews, B. W. (1968). *J. Mol. Biol.* **33**, 491–497.
McCoy, A. J., Grosse-Kunstleve, R. W., Adams, P. D., Winn, M. D., Storoni, L. C. & Read, R. J. (2007). *J. Appl. Cryst.* **40**, 658–674.
Metcalf, V., Brennan, S. & George, P. (2003). *Appl. Bioinformatics*, **2**, S97–S107.
Mogues, T., Li, J., Coburn, J. & Kuter, D. J. (2005). *J. Immunol. Methods*, **300**, 1–11.
Mueller, U., Darowski, N., Fuchs, M. R., Förster, R., Hellmig, M., Paithankar, K. S., Pühringer, S., Steffien, M., Zocher, G. & Weiss, M. S. (2012). *J. Synchrotron Rad.* **19**, 442–449.
Murshudov, G. N., Skubák, P., Lebedev, A. A., Pannu, N. S., Steiner, R. A., Nicholls, R. A., Winn, M. D., Long, F. & Vagin, A. A. (2011). *Acta Cryst.* D**67**, 355–367.
Otwinowski, Z. & Minor, W. (1997). *Methods Enzymol.* **276**, 307–326.
Patterson, J. E. & Geller, D. M. (1977). *Biochem. Biophys. Res. Commun.* **74**, 1220–1226.
Peters, T. (1996). *All About Albumin: Biochemistry, Genetics, and Medical Applications.* San Diego: Academic Press.
Petitpas, I., Petersen, C. E., Ha, C. E., Bhattacharya, A. A., Zunszain, P. A., Ghuman, J., Bhagavan, N. V. & Curry, S. (2003). *Proc. Natl Acad. Sci. USA*, **100**, 6440–6445.
Rosseneu-Motreff, M., Blaton, V., Declercq, B., Vandamme, D. & Peeters, H. (1970). *J. Biochem.* **68**, 369–377.
Rossmann, M. G. (1990). *Acta Cryst.* A**46**, 73–82.
Rothman, J. E. & Orci, L. (1992). *Nature (London)*, **355**, 409–415.
Sarich, V. M. & Wilson, A. C. (1967). *Proc. Natl Acad. Sci. USA*, **58**, 142–148.
Schreiber, G. & Urban, J. (1978). *Rev. Physiol. Biochem. Pharmacol.* **82**, 27–95.
Schwarzenbacher, R., Godzik, A., Grzechnik, S. K. & Jaroszewski, L. (2004). *Acta Cryst.* D**60**, 1229–1236.
Sinibaldi, R., Ortore, M. G., Spinozzi, F., de Souza Funari, S., Teixeira, J. & Mariani, P. (2008). *Eur. Biophys. J.* **37**, 673–681.
Sohl, J. L. & Splittgerber, G. (1991). *J. Chem. Educ.* **68**, 262–264.
Stein, N. (2008). *J. Appl. Cryst.* **41**, 641–643.
Sudlow, G., Birkett, D. J. & Wade, D. N. (1975). *Mol. Pharmacol.* **11**, 824–832.
Sudlow, G., Birkett, D. J. & Wade, D. N. (1976). *Mol. Pharmacol.* **12**, 1052–1061.
Sugio, S., Kashima, A., Mochizuki, S., Noda, M. & Kobayashi, K. (1999). *Protein Eng.* **12**, 439–446.
Thompson, J. D., Higgins, D. G. & Gibson, T. J. (1994). *Nucleic Acids Res.* **22**, 4673–4680.
Ueno, H., Masuko, T., Wang, J. & Hashimoto, Y. (1994). *J. Biochem.* **115**, 1119–1127.
Urlich, K. (1994). *Comparative Animal Biochemistry*. New York: Springer-Verlag.
Varshney, A., Sen, P., Ahmad, E., Rehan, M., Subbarao, N. & Khan, R. H. (2010). *Chirality*, **22**, 77–87.
Verboven, C., Rabijns, A., De Maeyer, M., Van Baelen, H., Bouillon, R. & De Ranter, C. (2002). *Nature Struct. Biol.* **9**, 131–136.
Wardell, M., Wang, Z., Ho, J. X., Robert, J., Ruker, F., Ruble, J. & Carter, D. C. (2002). *Biochem. Biophys. Res. Commun.* **291**, 813–819.
Winn, M. D. *et al.* (2011). *Acta Cryst.* D**67**, 235–763.
Winn, M. D., Isupov, M. N. & Murshudov, G. N. (2001). *Acta Cryst.* D**57**, 122–133.
Zunszain, P. A., Ghuman, J., Komatsu, T., Tsuchida, E. & Curry, S. (2003). *BMC Struct. Biol.* **3**, 6.
Zunszain, P. A., Ghuman, J., McDonagh, A. F. & Curry, S. (2008). *J. Mol. Biol.* **381**, 394–406.