

The Three Recombinant Domains of Human Serum Albumin

STRUCTURAL CHARACTERIZATION AND LIGAND BINDING PROPERTIES*

(Received for publication, March 15, 1999, and in revised form, May 12, 1999)

Michael Dockal[‡], Daniel C. Carter[§], and Florian Rüker^{‡¶}

From the [‡]Institute of Applied Microbiology, University of Agricultural Sciences, Muthgasse 18,
A-1190 Vienna, Austria and [§]New Century Pharmaceuticals Inc., Huntsville, Alabama 35824

In an attempt to systematically dissect the ligand binding properties of human serum albumin (HSA), the gene segments encoding each of its three domains were defined based on their conserved homologous structural motifs and separately cloned into a secretion vector for *Pichia pastoris*. We were able to establish a generally applicable purification protocol based on Cibacron Blue affinity chromatography, suggesting that each of the three domains carries a binding site specific for this ligand. Proteins were characterized by SDS-polyacrylamide gel electrophoresis, isoelectric focusing, gel filtration, N-terminal sequencing, and matrix-assisted laser desorption/ionization time-of-flight mass spectrometry, as well as near- and far-UV CD. In addition to the affinity chromatography ligand Cibacron Blue, binding properties toward hemin, warfarin, and diazepam, each of which represents a standard ligand for HSA, respectively, were investigated by the measurement of induced circular dichroism. Clear experimental evidence is provided here for the location of the primary hemin binding site to be on domain I of HSA, and for the primary diazepam binding site to be on domain III. Further, secondary binding sites were found for hemin to be located on domains II and III, and for diazepam on domain I. The warfarin binding site was located primarily on domain II, while on domain I, a secondary binding site and/or parts of the primary binding site were found.

Human serum albumin (HSA)¹ is a globular protein of 585 amino acids, which accounts for about 60% of the total protein in blood serum. In the serum of human adults, the concentration of albumin is ~40 mg/ml. It has an exceptional binding capacity for a wide range of endogenous and exogenous ligands, and its abundance makes it an important determinant of the pharmacokinetic behavior of many drugs (1, 2).

HSA is composed of three structurally similar globular domains, each of which contains two subdomains, denominated subdomain IA, IB, IIA, IIB, IIIA, and IIIB (Fig. 1). It is generally accepted today that albumin evolved from an ancestral protein of about 190 amino acids by triplication of this primordial domain. This view is most impressively documented by the

striking internal homology of the atomic structure of the protein (3, 4), by the unique arrangement of disulfide loops that repeat as a series of triplets (1), and by the symmetrical location of the introns and exons of the albumin gene on chromosome 4 (5).

Although still partly controversial, consensus exists today that there are two high affinity binding sites (6) for small heterocyclic or aromatic compounds (located on subdomains IIA and IIIA), two to three dominant long-chain fatty acid binding sites (located on subdomains IB and IIIB) (1), and two distinct metal-binding sites, one involving Cys-34 and the other the N terminus, making a total of six dominant areas of ligand association to albumin (2). Recently, a crystal structure was reported that showed five molecules of the medium-chain-length fatty acid myristate bound in different pockets of the HSA molecule (7).

Based on the observation of the quasi-independence of the three HSA domains, attempts have been made early on to use fragments of HSA produced by proteolytic or chemical cleavage to define the binding sites for several ligands (Refs. 8–10; see Ref. 11 for a review). However, these studies have been dictated by the location of the cleavage sites on the primary sequence and did not allow for a deliberate selection of the exact size of the domains and their boundaries. Bos and colleagues (12–14) used a peptic fragment of HSA encompassing domain I and II (residues 1–387) and a tryptic fragment encompassing domain II and III (residues 198–585) to study the binding of warfarin and diazepam. Hrkál *et al.* (15) have investigated the binding of hemin to the three large cyanogen bromide fragments of HSA. Compagnini *et al.* (16) have used cyanogen bromide fragments of HSA in an attempt to define the Cibacron Blue binding sites by a chromatographic procedure and reported that two molecules of Cibacron Blue may be located in the hydrophobic pocket corresponding to subdomain IIA, with a third molecule in the hydrophobic site corresponding to subdomain IIIA. In a recent study, Kjeldsen *et al.* (17) have expressed domains I and III in *Saccharomyces cerevisiae*, but were unable to produce domain II in this expression system. Ligand binding studies with myristate and an insulin analogue acylated with myristic acid showed that the latter bound to both domains while myristate alone interacted only with domain III.

Details of the molecular interactions of ligand binding to albumin are known for only a small number of cases (1). Insight gained from such studies shows promise in improving the design of many pharmaceuticals and diagnostic agents.

In the present study, all three domains of HSA were produced by molecular cloning in a completely predetermined manner, based solely on the natural boundaries of the domains. This definite advance over the previous studies using either albumin fragments produced by proteolytic or chemical cleavage or incomplete sets of recombinant domains is expected to allow an in-depth study of the folding of the domains and of

* This work was supported by Fonds zur Förderung der wissenschaftlichen Forschung Grant P11280-MED, Jubiläumsfonds der Österreichischen Nationalbank Grant P6875, and EU Concerted Action BIO4-CT96-0389 "Phage Club." The costs of publication of this article were defrayed in part by the payment of page charges. This article must therefore be hereby marked "advertisement" in accordance with 18 U.S.C. Section 1734 solely to indicate this fact.

¶ To whom correspondence should be addressed. Tel.: 43-1-360066240; Fax: 43-1-360066652; E-mail: ruker@mail.boku.ac.at.

¹ The abbreviations used are: HSA, human serum albumin; ICD, induced circular dichroism; PBS, phosphate-buffered saline; PAGE, polyacrylamide gel electrophoresis; MALDI-TOF, matrix-assisted laser desorption/ionization time-of-flight mass spectrometry; DOM, domain.

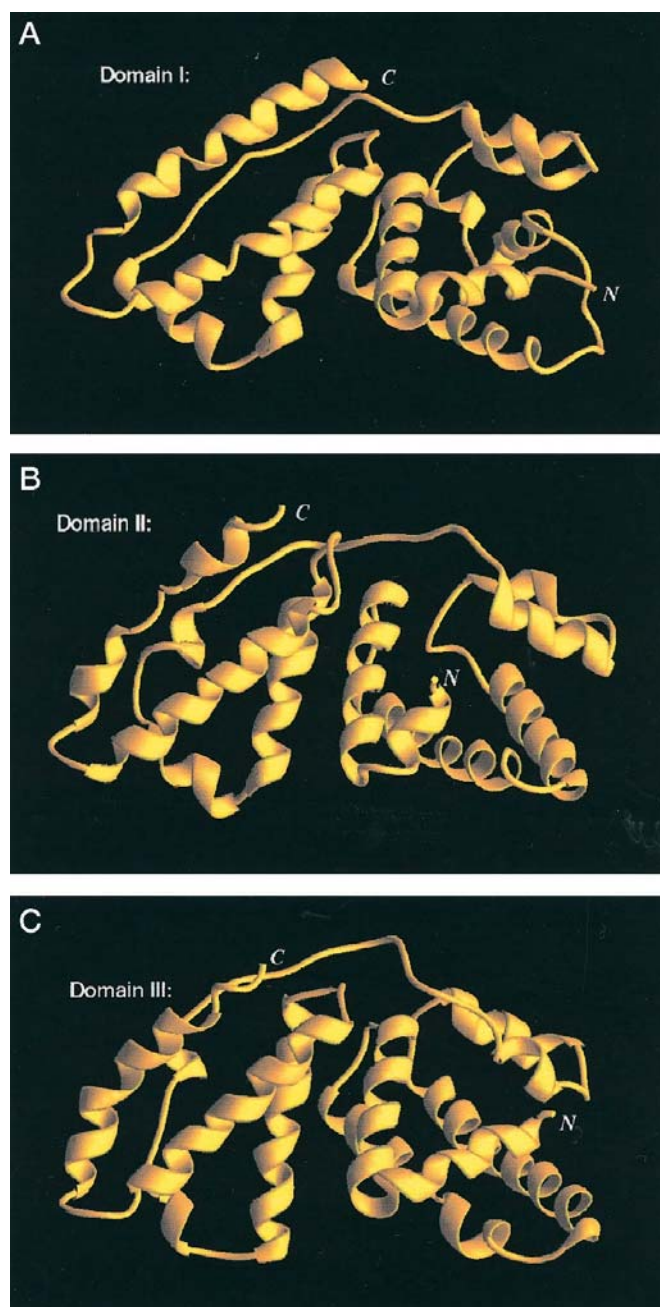


FIG. 1. Ribbon models of the three HSA domains: domain I (A), domain II (B), and domain III (C). InsightII/Discover (Biosym) was used as molecular modeling software. This figure was produced using the program Ribbons (42).

their viability as “stand-alone” proteins. Furthermore, ligand-binding studies with these domains will help in defining the exact location of the various binding sites. This paper reports on the cloning and expression of the three domains, their purification, and biochemical as well as structural characterization, providing a qualitative estimate of the chemistry of some important binding events and their comparability to natural albumin.

EXPERIMENTAL PROCEDURES

Materials—HSA (essentially free of fatty acids and globulin), warfarin, and hemin were from Sigma (catalog nos. A3782, A2250, and H2250, respectively), and diazepam was from Sigmapharm, Vienna, Austria. Drugs were used as received, and stock solutions were prepared by dissolving warfarin in sodium phosphate buffer, hemin in 10 mM NaOH (final concentration), and diazepam in absolute ethanol. Blue Sepharose 6 Fast Flow and Superdex 75 prep grade were from

Amersham Pharmacia Biotech. Lipidex-1000 was purchased from Canner-Packard. All other reagents were of analytical grade.

Cloning—The gene segments coding for the HSA domains were amplified by polymerase chain reaction with *Pfu* polymerase (Stratagene) using a cDNA clone of HSA as template. *EcoRI* restriction sites for the subsequent cloning into the secretion vector pPICZαA (Invitrogen) were attached at the same time. Recombinant plasmids were transfected into *Pichia pastoris* GS115 by electroporation following standard procedures (Invitrogen). All clones were confirmed by polymerase chain reaction and by DNA sequencing. Domains encompassed the following amino acids: HSA-DOM I, 1–197; HSA-DOM II, 189–385; HSA-DOM III, 381–585. All clones contained an extra two N-terminal amino acids (Glu-Phe) due to the restriction site used for cloning.

Expression—Standard protocols and media formulations were used (Invitrogen). Briefly, clones were grown up on YPD medium and transferred to BMMY medium for induction with methanol in baffled shake flasks. All incubations were performed at 30 °C on an orbital shaker at 150 rpm.

Purification—All manipulations were performed at 4 °C unless stated otherwise. Culture supernatants were harvested by centrifugation, followed by filtration and dialysis against PBS. 1800 ml of protein containing dialysate were passed through a Blue Sepharose column (150 ml gel volume in a XK50 column) preequilibrated with PBS. After washing with 10 volumes of PBS, elution was performed by 1 M KSCN in PBS. After dialysis against PBS and concentration by ultrafiltration (Centriprep 10, Amicon), samples were adjusted to a final NaCl concentration of 0.25 M and were applied to a Superdex 75 prep grade XK26/60 column (Amersham Pharmacia Biotech) using PBS containing 0.25 M NaCl as running buffer. For delipidation, a chromatography over Lipidex 1000 was performed at 37 °C according to Glatz and Veerkamp (18). Then, the protein samples were extensively dialyzed against water. Finally, the preparations were concentrated by ultrafiltration as above and lyophilized. All purification steps were monitored with protein assay and SDS-PAGE. The Coomassie-stained SDS-PAGE gels were scanned, and the purity of the preparations was evaluated using the program Non-Linear-Dynamic LTD (Phoretics).

Protein Solutions—Protein concentrations were measured by their absorbance at 278 nm on a Hewlett Packard HP8453 spectrophotometer. For HSA, an absorption coefficient ($A_{278\text{ nm}, 1\text{ cm}}^{0.1\%}$) of 0.58 was used (12). For the HSA domains, the molar extinction coefficients were determined using the method of Gill and Hippel (19). A Bio-Rad protein assay with HSA standard solutions (Sigma) was used to confirm the photometric results.

Analytical Gel Filtration—Analysis was performed on an LKB instrument equipped with a variable wavelength monitor set at 214 nm for detection. Data were analyzed using a Nelson Analytical 900 Series interface and Turbochrom3 software. Samples (100 µg/ml) were run in 0.1 M sodium phosphate buffer, pH 7.0, containing 0.15 M sodium chloride at a flow rate of 1 ml/min on a Superdex 75 HR10/30 column (Amersham Pharmacia Biotech). For calibration, a low molecular weight calibration kit (Amersham Pharmacia Biotech) was used.

Matrix-assisted Laser Desorption/Ionization Time-of-flight Mass Spectrometry (MALDI-TOF)—Relative masses of the proteins were detected by MALDI-TOF with a DynaMo instrument (Thermo BioAnalysis) using 3,5-dimethoxy-4-hydroxycinnamic acid as the UV-absorbing matrix. Calibration of the instrument was performed with carbonic anhydrase.

Far UV-CD Measurements—Measurements were made using a Jasco J-600 spectropolarimeter using 1-mm cells at 25 °C in a thermostatic cell holder. Scans were made from 250 to 190 nm. The slit width was programmed for a half-bandwidth of 1 nm, and the dynode voltage never exceeded 0.6 kV. All spectra were recorded at least twice. The data were expressed as mean residue ellipticity ($[\theta]_{\text{MRW}}$, degrees $\times$ cm² $\times$ dmol⁻¹), using the mean residue weights of 113.7 for the intact molecule, 114.9 for HSA-DOM I, 113.2 for HSA-DOM II, and 113.0 for HSA-DOM III. The secondary structure elements of the proteins were determined from the CD spectra (1-nm interval) using the procedure of Provencher and Glöckner (20, 21).

Near UV-CD and Induced CD Measurements—Measurements were made using a Jasco J-600 spectropolarimeter using 10-mm cells at 25 °C in a thermostatic cell holder. The slit width was programmed for a half-bandwidth of 1 nm and the dynode voltage never exceeded 0.4 kV. Over 250 nm, the results are expressed in molar ellipticity ($[\theta]$, degrees $\times$ cm² $\times$ dmol⁻¹). In the ICD experiments, all solutions were scanned from a wavelength at which no induced ellipticity was observed (hemin, 450–360 nm; warfarin, 360–260 nm; diazepam, 360–250 nm). Stock solutions of the ligands were prepared as follows and added stepwise to the protein solutions: 2 mM hemin in 10 mM NaOH; 4 mM

warfarin in 0.1 M sodium phosphate buffer, pH 7.4; 4 mM diazepam in absolute ethanol (ethanol concentration in the ICD measurements never exceeded 1% and the effects of the organic solvent on the CD measurements were undetectable). Protein solutions were 10 μ M for the hemin measurements and 20 μ M for the experiments with warfarin and diazepam, in 0.1 M sodium phosphate buffer, pH 7.4. All spectra were recorded at least twice.

The induced ellipticity was defined as the ellipticity of the drug/protein mixture minus the ellipticity of the protein alone at the same wavelength and was expressed in degrees. Intrinsic ellipticity was shown to be undetectable for all ligands at the relevant wavelengths.

RESULTS

Cloning, Expression, Purification, and Biochemical Characterization

Expression of the domains was achieved as secreted proteins (due to the α -factor leader sequence present in the vector pPICZ α A) in the supernatant of the recombinant *Pichia* clones with yields between 65 and 300 mg/liter. Based on the assumption that HSA harbors several Cibacron Blue binding sites, with each of the domains containing at least one of them, however weak, we successfully employed Blue Sepharose affinity chromatography as the central step in our purification procedure. Binding of HSA-DOM II and HSA-DOM III to the Blue Sepharose column was strongest, since hardly any protein could be detected in the flow-through of the column. The efficiency was somewhat lower for HSA-DOM I, but the overall yield of purified protein was in the same order of magnitude in all three cases. Evaluation of scans of the Coomassie-stained SDS-PAGE gel indicated a purity of the domain preparations greater than 98% (Fig. 2).

The chromatographic elution profile (data not shown) of the analytical gel filtration indicated that the HSA domains are present in our preparations as monomeric proteins. MALDI-TOF analysis gave the expected molecular masses of the proteins, and N-terminal sequencing resulted in the expected sequences (leading Glu-Phe residues are due to the *Eco*RI restriction site used for cloning), indicating both correct cleavage of the α -factor leader sequence as well as the absence of unexpected posttranscriptional modification events. The meas-

ured isoelectric points are in close agreement with the values predicted from the sequence. A summary of the physicochemical properties of the HSA domains is given in Table I.

Far and Near UV-CD Measurements

The CD spectra in the far ultraviolet region relate to the polypeptide backbone structures. A comparison of the spectra of the HSA domains with that of intact HSA on a mean residue weight basis is shown in Fig. 3. Minima of the spectra are given in Table I, and the results of the quantitative analysis using the CONTIN program (21) are given in Table II. The results obtained with defatted HSA are in good agreement with published data (22–24). The spectra of HSA-DOM I and HSA-DOM III are similar in shape with HSA, but have reduced intensity, indicating reduced α -helix content, which is supported by the quantitative analysis. This is more pronounced in HSA-DOM I than in HSA-DOM III. Evaluation of the results obtained with HSA-DOM II indicates that its structure resembles that of an $\alpha + \beta$ protein rather than that of an all α protein. This could be due to increased instability of the newly exposed long helical regions at the N and C termini of this domain, as can be seen in Fig. 1.

To obtain insight into the molecular environment of the aromatic side chains of HSA as well as the domains, near UV-CD measurements were performed. As can be seen in Fig. 4, HSA and its domains all have negative ellipticity in this wavelength region. For both HSA and each of the domains, we found negative extrema at 268 nm and 262 nm, confirming data that have previously been reported for HSA (25). The sum of the molar ellipticities of the three domains shows the same minima and shape as HSA, especially at wavelengths above 265 nm. Below 265 nm, deviations are observed, which could be due to the additional phenylalanine residue that has been added to the N terminus of each of the domains as a result of the cloning procedure.

Taken together, the results of the CD measurements indicate that the domains have secondary and the tertiary structures with significant contribution of α -helical conformation. However, it is clear that β -sheet and random structure elements have increased in percentage as compared with native HSA. This observation is most pronounced for HSA-DOM II. The observed differences to the HSA signals, in particular as seen in the far UV-CD measurements, are probably the result of the disruption of non-covalent bonds stabilizing the albumin tertiary structure in the native molecule rather than of incorrect folding or irreversible denaturation of the domains, as further supported by the ligand binding data shown below.

Induced CD Measurements

Hemin—The method of ICD is based on the observation that an optical activity arises from dissymmetry in the ligand induced by its binding to the protein, since the free ligand has either no asymmetric center (hemin, diazepam) or is a racemic mixture (*R,S*-warfarin) and therefore gives no signal in solution.

Fig. 5A shows the ICD spectra of hemin complexed with the

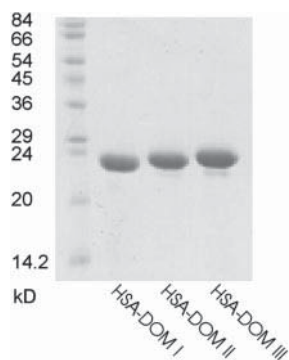


FIG. 2. **SDS-PAGE.** Polyacrylamide (15%) gel electrophoresis was performed in the presence of 0.1% SDS (according to Laemmli (Ref. 43)). Left lane, wide molecular weight standards (Sigma); lane 1, HSA-DOM I; lane 2, HSA-DOM II; lane 3, HSA-DOM III (3 μ g each per lane).

TABLE I

Summary of the physicochemical properties of the HSA and its recombinant domains

Theoretical or predicted values are given in parenthesis.

	HSA (defatted)	HSA-DOM I	HSA-DOM II	HSA-DOM III
Molecular mass (MALDI-TOF)	(66,479)	22,966 (22,860)	22,413 (22,519)	23,320 (23,383)
Isoelectric point, pI	5.8 (Ref. 2)	5.2 (5.3)	5.4 (5.1)	6.6 (6.5)
Residue range	1–585	1–197	189–385	381–585
Far UV-CD minima (nm)	222, 209	222, 209	209	223, 209
Near UV-CD minima (nm)	268, 262	268, 262	268, 262	268, 262
$A_{278nm,1cm}^{0.1\%}$	0.58 (Ref. 10)	0.50	0.79	0.30

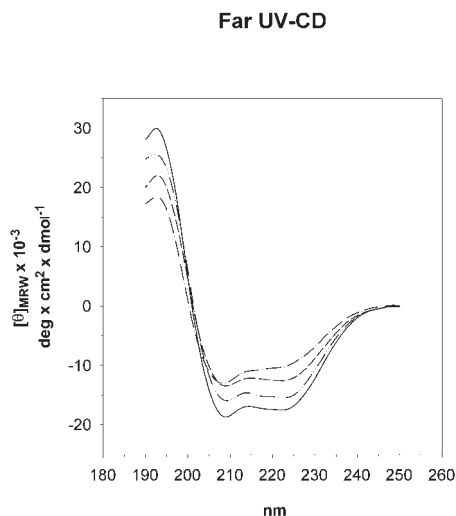


FIG. 3. Far UV-CD signals of HSA (—), HSA-DOM I (---), HSA-DOM II (·····), and HSA-DOM III (- · - · -). Proteins were dissolved in 10 mM sodium phosphate buffer, pH 7.4, at a concentration of 2.1 μ M (HSA) or 4.5 μ M (domains).

TABLE II
Fraction of secondary structure elements as calculated from the respective CD spectra

	α -Helix	β -Sheet	Coil	Turn
HSA (defatted)	66 $\pm$ 0.6	6 $\pm$ 0.8	26 $\pm$ 1.1	2 $\pm$ 0.9
HSA-DOM I	46 $\pm$ 0.5	22 $\pm$ 0.7	20 $\pm$ 1.0	12 $\pm$ 0.6
HSA-DOM II	37 $\pm$ 0.5	34 $\pm$ 0.7	14 $\pm$ 0.8	15 $\pm$ 0.9
HSA-DOM III	53 $\pm$ 1.5	0	37 $\pm$ 3.0	10 $\pm$ 1.6
Human mercaptalbumin ^a	66	11	23	

^a Data are from Ref. 22.

three domains and with HSA. A ligand/protein molar ratio of 0.5 was selected to ensure that all of the ligand is bound to the primary binding site, avoiding interactions with secondary, weaker binding sites. Hemin complexed with HSA-DOM I and HSA, respectively, shows very similar spectra sharing the same minimum at 397 nm, indicating the location of the primary hemin binding site to be on HSA-DOM I. With this ligand, HSA-DOM II and HSA-DOM III display ICD spectra with different shapes, suggesting the presence of one or more secondary, weaker binding sites. However, it cannot be fully excluded that new binding sites, which are not present on native HSA, were created by exposing the domains as isolated proteins.

The results of the titration of HSA and HSA-DOM I with hemin are presented in Fig. 5B. No saturation of the ICD signal was achieved under our experimental conditions; however, a shift in the induced wavelength minimum was observed, which is in accordance with the experiments conducted with HSA by Beaven *et al.* (26). Starting at a molar ratio of hemin to HSA of 2, the negative extrinsic Cotton effect shifts from 397 nm to 400 nm with HSA, while under the same conditions HSA-DOM I causes a shift from 397 nm to 402 nm, which is already finished at a molar ratio of 2. The shift in the induced ellipticity is interpreted by the binding of hemin to one or more secondary binding sites with different binding properties and optical characteristics. This notion is further supported by similar observations in titration of HSA-DOM II and HSA-DOM III, which also show wavelength shifts upon titration with hemin (data not shown).

Warfarin—Fig. 6A shows the ICD spectra of warfarin complexed with the three domains and with HSA at a molar ratio of warfarin/protein of 0.5. The binding of warfarin to HSA generates extrinsic Cotton effects with a positive band at 308

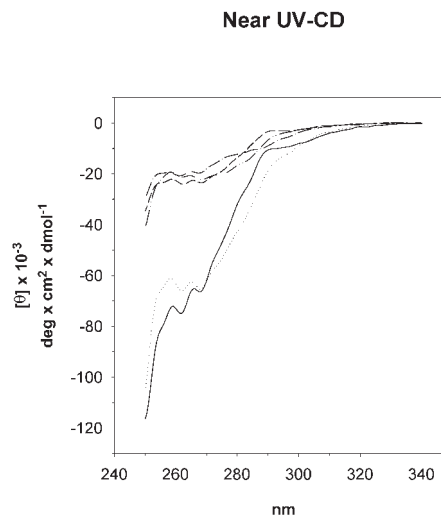


FIG. 4. Near UV-CD signals of HSA (—), HSA-DOM I (---), HSA-DOM II (·····), and HSA-DOM III (- · - · -), and calculated signal of the three domains combined (·····). Proteins were dissolved in 0.1 M sodium phosphate buffer, pH 7.4, at a concentration of 20 μ M.

nm, a negative one at 278 nm, and a positive shoulder near 340 nm, as previously shown by Fehske *et al.* (27) and Otagiri *et al.* (28). Warfarin binding to HSA-DOM II shows a similar spectrum with respect to positive and negative Cotton effects. However, the intensity of the bands is reduced, the wavelength of the positive and the negative bands is slightly shifted to 310 nm and 281 nm, respectively, and no shoulder is observed in the region of 340 nm. HSA-DOM I shows significant ellipticity with a broad band in the positive region (maximum at 306 nm), whereas no significant signal is observed for HSA-DOM III.

The results of the titration of HSA-DOM I and HSA-DOM II with warfarin are shown in Fig. 6 (B and C, respectively). A concentration-dependent signal is observed with both domains tested. HSA-DOM II yields a signal that is qualitatively comparable to that of HSA. However, it is clear that the level of the ICD signal obtained with this domain is lower than that obtained with HSA. Based on the concentration-dependent signal observed with HSA-DOM I, it can be concluded either that secondary binding site(s) are located here or, more likely, that a part of the primary binding site is present on this domain, as was previously suggested by Bos *et al.* (12, 13). An unambiguous interpretation of the interaction between warfarin and HSA-DOM III is not possible based on the data presented here, since the absence of an ICD signal with this domain could mean either the absence of binding or a binding event that does not produce an ICD signal.

Diazepam—The ICD spectrum of diazepam bound to HSA shows a strong positive band at 260 nm and a negative one at 317 nm, which is in accordance with the findings of Wilting *et al.* (29). Similarly, but with reduced intensity, binding of diazepam to HSA-DOM III shows a positive band near 260 nm and a negative one near 315 nm. In contrast, the induced ellipticity of diazepam bound to HSA-DOM I is negative at 260 nm and positive at 315 nm, with considerably lower intensity. Diazepam with HSA-DOM II shows no significant ICD spectrum in the observed wavelength range (Fig. 7A).

In the titration experiment (Fig. 7B), saturation of the signal obtained with HSA was observed at molar ratios higher than 2. However, no such saturation was obtained using HSA-DOM III, indicating that binding of diazepam to this domain, although qualitatively very similar, is probably taking place with reduced affinity as compared with HSA.

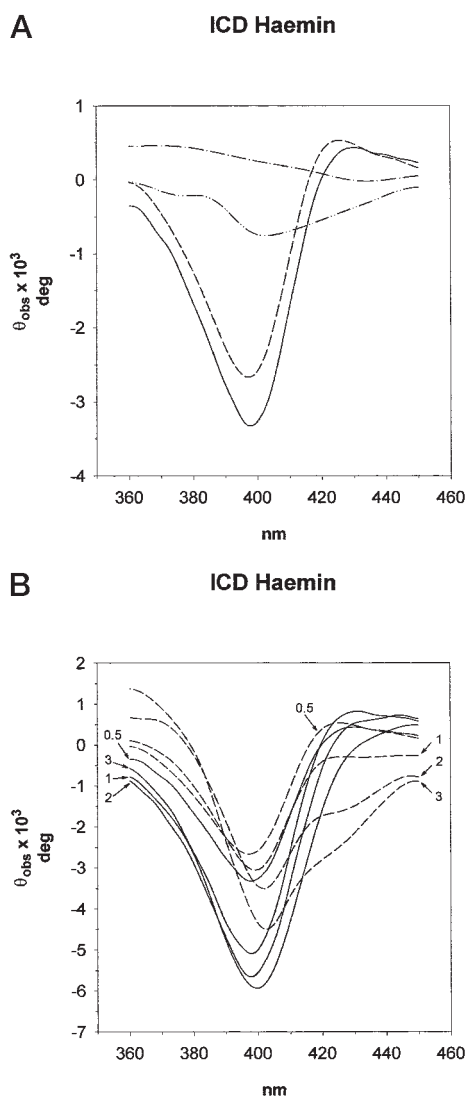


FIG. 5. A, ICD (expressed in degrees) of hemin bound to HSA-DOM I, HSA-DOM II, HSA-DOM III, and HSA at a molar ratio of 0.5. Line symbols are as in Fig. 3. Protein concentrations were 10 μM in 0.1 M sodium phosphate buffer, pH 7.4. B, ICD (expressed in degrees) of hemin bound to HSA-DOM I and HSA at molar ratios ranging from 0.5 to 3 as indicated by the arrows. Protein concentrations were 10 μM in 0.1 M sodium phosphate buffer, pH 7.4.

DISCUSSION

Numerous reports on the fragmentation of albumins, predominantly of bovine or human origin, have been published in the literature, all of which have been aimed at gaining insight into the structure/function relationships of the different regions of the albumin molecule (e.g. Refs. 8, 10, 12, 15, 25, and 30–33). The production of fragments employed standard methods, including cyanogen bromide cleavage and proteolytic digestion by pepsin or trypsin.

Drug binding studies have for a long time relied on the use of albumin fragments, which proved to be valuable research tools, allowing the dissection and separation of the complex array of different binding sites present on HSA from each other, thereby facilitating the study of the binding of ligands that have more than one binding site on albumin (10, 12, 13, 25, 33, 34).

However, the majority of the albumin fragments reported so far do not have boundaries that can be justified on grounds of structure or function, since they have been dictated by the cleavage specificity of the reagents used for their production, increasing the probability for impaired structural integrity and

potentially leading to a loss of functionality.

A recent report by Kjeldsen *et al.* (17) describes the cloning and expression of HSA domains I and III in *S. cerevisiae*. Attempts to produce domain II have not been successful. This is probably due to specific properties of the expression system used by these authors, which made use of a synthetic LA19 preproleader sequence for secretion of the recombinant protein from the host cell, since we were able to show here that domain II can in fact be produced in high yield with our expression system. Interestingly, Kjeldsen *et al.* found that domain I did not compete with immobilized HSA for binding to myristate and they concluded that domain I, when expressed as a separate entity, does not possess a myristate binding site. It should be noted that the domain preparations used in these experiments have not been subjected to any defatting or other delipidization procedure during their purification. In contrast, recent experiments in our laboratory² showed that palmitate bound to domains I and III and, to a lesser extent, to domain II, based on competition studies with the fluorescent long-chain fatty acid *cis*-parinaric acid. These findings are in accordance with the crystal structure of HSA complexed with myristate described by Curry *et al.* (7).

Based on the primary, secondary, as well as tertiary structure of albumin, natural borders between the domains can be defined. We propose here to set the boundaries as follows: HSA-DOM I, residues 1–197; HSA-DOM II, residues 189–385; HSA-DOM III, residues 381–585. These partially overlapping regions of the domains have been suggested by the atomic structures of various ligand complexes with albumin (1, 35),³ and each contains the major conserved portions of the relevant interdomain helix (h10_{DOM I}–h1_{DOM II} for HSA-DOM I, h10_{DOM II}–h1_{DOM III} for HSA-DOM III, and both for HSA-DOM II). They include cysteine residues only as needed for the formation of the intradomain disulfide bridges and have previously formed the basis for selecting essential portions of recombinant subdomains (IIA and IIIA) (36).

To overcome the serious limitations imposed by the cleavage reagents available, we have undertaken to produce custom designed HSA domains by gene cloning and expression techniques, allowing the preparation of all three domains of HSA as stand-alone unit molecules. As an expression system, we chose to use the methylotrophic yeast *P. pastoris*, which has previously been shown to be an excellent host for the production of various recombinant proteins, including HSA (37). Under the shake flask conditions used, the HSA domains were expressed and secreted at levels similar to HSA,⁴ ranging from 65 to 300 mg/liter. From the crude culture supernatant, the proteins were purified by a uniform procedure, employing Cibacron Blue affinity chromatography as a central step. After the final steps of the purification procedure, which consisted of preparative gel filtration and defatting, the purity of the preparations was estimated to be higher than 98% and therefore suitable for all further studies discussed below.

All three domains of HSA were purified using this affinity matrix, indicating the presence of Cibacron Blue binding sites on each of the domains. Compagnini *et al.* (16, 38) examined the binding behavior of cyanogen bromide fragments of HSA to immobilized Cibacron Blue F3G-A and found that a fragment containing residues 1–123 did not interact with the dye, while a fragment spanning residues 124–298 did. This fragment was interpreted by these authors to harbor two interaction sites, both located in the major hydrophobic binding pocket corresponding to subdomain IIA. Since we found that HSA-DOM I

² A. Schweinberger, unpublished data.

³ D. C. Carter, unpublished data.

⁴ M. Dockal, unpublished data.

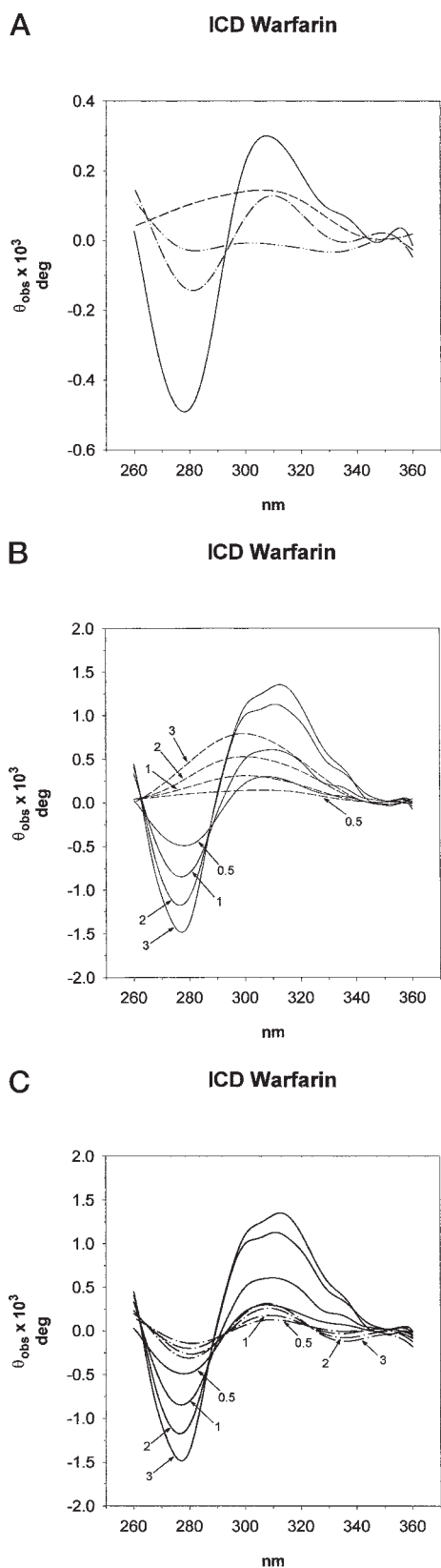


FIG. 6. A, ICD (expressed in degrees) of warfarin bound to HSA-DOM I, HSA-DOM II, HSA-DOM III, and HSA at a molar ratio of 0.5. Line symbols are as in Fig. 3. Protein concentrations were 20 μ M in 0.1 M sodium phosphate buffer, pH 7.4. B, ICD (expressed in degrees) of warfarin bound to HSA-DOM I and HSA at molar ratios ranging from 0.5 to 3 as indicated by the arrows. Protein concentrations were 20 μ M in 0.1 M sodium phosphate buffer, pH 7.4. C, ICD (expressed in degrees) of warfarin bound to HSA-DOM II and HSA at molar ratios ranging from 0.5 to 3 as indicated by the arrows. Protein concentrations were 20 μ M in 0.1 M sodium phosphate buffer, pH 7.4.

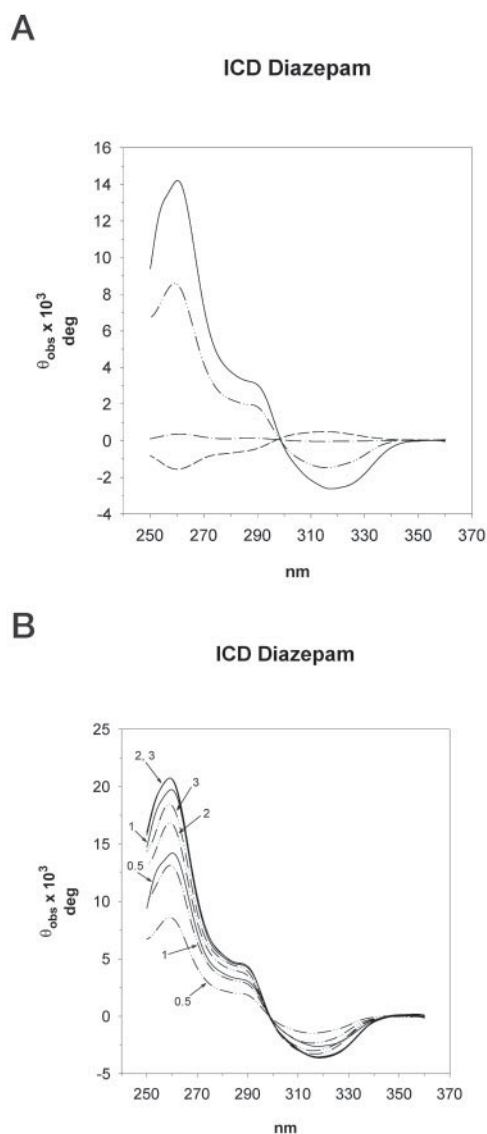


FIG. 7. A, ICD (expressed in degrees) of diazepam bound to HSA-DOM I, HSA-DOM II, HSA-DOM III, and HSA at a molar ratio of 0.5. Line symbols are as in Fig. 3. Protein concentrations were 20 μ M in 0.1 M sodium phosphate buffer, pH 7.4. B, ICD (expressed in degrees) of diazepam bound to HSA-DOM III and HSA at molar ratios ranging from 0.5 to 3 as indicated by the arrows. Protein concentrations were 20 μ M in 0.1 M sodium phosphate buffer, pH 7.4.

(residues 1–197) bound to Cibacron Blue under similar conditions, we predict that one of the two binding sites postulated by Compagnini in subdomain IIA is actually located in the C-terminal part of domain I.

A number of biochemical and biophysical techniques were employed for characterization of the domain preparations. The results of these measurements, which are presented in Table I, show that the proteins are present as monomers with a molecular weight and isoelectric point as expected, according to MALDI-TOF analysis, analytical gel filtration, as well as isoelectric focusing, indicating not only the absence of major unexpected posttranslational modifications, including glycosylation, but also correct processing of the α -factor leader peptide. This processing event was further corroborated by determination of the N-terminal sequence by Edman degradation.

Far UV-CD—To obtain an insight into the structure of the domains as compared with HSA, CD spectra were studied for each of the separate domains. In the far ultraviolet region, such spectra relate to the polypeptide backbone structures. A com-

parison of the spectra of HSA domains with the parent molecule on a mean residue weight basis is shown in Fig. 3, and a quantitative evaluation is given in Table II. HSA gives a result as expected (22–24). HSA-DOM I and HSA-DOM III show a similar signal typical for a predominantly α -helical protein, although with slightly reduced intensity as compared with HSA. The spectrum observed for HSA-DOM II is characterized by a different shape, indicating significant α -helical as well as β -sheet structure as further supported by the quantitative analysis. Since the domain boundaries selected by us are located centrally in the long interdomain helices (h10_{DOM I}–h1_{DOM II}, h10_{DOM II}–h1_{DOM III}), this cleavage could give rise in these regions to some loss of α -helical conformation for β -sheet and coil structure. As HSA-DOM II contains two such unnatural termini, the relative contribution of this effect could be more pronounced than for HSA-DOM I and HSA-DOM III, each of which contains only one such site. The C-terminal helix (h10_{DOM I}) of HSA-DOM I is partly buried in the native HSA molecule. If exposed, as is the case in a single domain, this could give rise to a destabilization of this region, manifesting itself by reduced ellipticity in the far UV-CD. In the case of HSA-DOM III, the interdomain helix connecting to HSA-DOM II is located on a more solvent-exposed patch, thereby maybe causing less perturbation to the molecule when exposed.

Furthermore, the isolated domains contain surface areas that are not solvent-exposed in the native protein. This newly created surface, which is predominantly hydrophobic in character, equals $\sim 7.0\%$ of the total molecular surface for HSA-DOM I and 5.4% for HSA-DOM III, whereas HSA-DOM II contains an extra 11.4% previously unexposed surface. This can be taken as another explanation for the specific structural behavior of the three domains.

Near UV-CD—Conformational changes around the aromatic amino acid residues and S-S bridges of the domains can conveniently be studied by CD measurements in the near ultraviolet region above 250 nm. The main contributions to the ellipticity from tryptophan and tyrosine groups are seen above 265 nm generally, with the largest maxima at 279, 284, and 291 nm for tryptophan (39) and at about 277 nm for tyrosine residues (40). The contribution from phenylalanine is most evident at about 255, 261, and 268 nm (41). When the CD spectra of the three recombinant domains were added and compared with that of native HSA, only small differences were noted above 265 nm, while the changes below 265 nm were somewhat larger. Thus, the structure around the lone tryptophan and the tyrosines largely remains intact in the fragments. However, the changes below 265 nm can be due to conformational changes around the S-S bridges and/or phenylalanines. In addition, due to the cloning procedure, there is an additional phenylalanine residue at the N terminus of each domain, which in sum may contribute to the ellipticity in this region providing further explanation for our findings.

In conclusion, we assume that the primary peptide structure contains sufficient information to fold the recombinant domains into secondary and tertiary structures closely approximating the structure occurring in native HSA, as was further corroborated by the ligand binding studies using ICD, which are discussed in the following sections.

Hemin Binding—We chose to compare the shape, minima, and maxima of the ICD spectra at a low molar ligand/HSA ratio, since under these conditions, predominantly the main interaction site on the protein is populated by the ligand, with only minor contributions from secondary sites. Based on this assumption, we clearly show here experimental evidence that HSA-DOM I harbors the primary binding site for hemin, as shown by spectra which are qualitatively very similar to those

obtained with native HSA, albeit with reduced intensity (Fig. 5A). This decrease in signal strength could indicate that the absence of the neighboring parts of the albumin molecule gives rise to a reduced binding affinity. Hrkal *et al.* (15) used cyanogen bromide fragments of HSA to study the interaction with hemin. In contrast to our domains, however, none of these fragments, which spanned residues 1–123, 124–298, and 299–585, respectively, yielded ICD spectra or affinity constants comparable to those of native HSA, indicating that neither of these fragments carries a fully functional binding site as presented on our HSA-DOM I. Unpublished crystallographic results³ support this view.

In the titration experiments with HSA (Fig. 5B), we observed a shift of the minimum to higher wavelength, in accordance with Beaven *et al.* (26). This is a clear indication of the presence of more than one binding site, since at higher concentrations of the ligand, the first binding site is saturated and secondary binding sites that exhibit altered optical properties become populated. Indeed, we found that hemin shows induced ellipticity not only upon binding to HSA-DOM I, but also to HSA-DOM II and HSA-DOM III, as indicated in Fig. 5A. These signals differ clearly from that of the postulated primary binding site. Furthermore, we were able to demonstrate that hemin, when added in increasing amounts to either of the three domains, exhibits continuously changing ICD spectra, suggesting the presence of additional secondary binding sites that are still not saturated under our experimental conditions (Fig. 5B; data not shown for HSA-DOM II and HSA-DOM III). Although the possibility cannot be completely excluded that by exposing the domains as isolated proteins, new binding sites were created that are not exposed in native HSA, we believe that the clear, concentration-dependent ICD signals that were obtained here and that are in accordance with the results obtained for native HSA by Beaven *et al.* (26) argue in favor of the applicability of the HSA domains as partial model proteins for native HSA.

Warfarin Binding—The spectra shown in Fig. 6A demonstrate the binding of warfarin to HSA, HSA-DOM I, HSA-DOM II, and HSA-DOM III. Since the shape of the spectrum of HSA-DOM II closely resembles that of native HSA, we conclude that the major part of the primary binding site is located here, in accordance with the crystallographic observations by Carter and Ho (1, 35). This finding also supports the results found by Fehske *et al.* (27), who described the participation of Trp-214 (located in domain II) in the non-overlapping part of the warfarin binding site. However, the titration of HSA-DOM II (Fig. 6C) with warfarin yields a much lower ellipticity than in the case of HSA, indicating distinct differences in the binding behavior of these two proteins. In view of the fact that warfarin binding to HSA-DOM I is also yielding a response (Fig. 6B), these two titrations can be seen as an indication for the presence of a secondary binding site in this domain, and/or for a considerable contribution of parts of HSA-DOM I to the primary binding site.

Bos *et al.* (12, 13) studied the binding of warfarin to a large peptic (domain I + II, residues 1–387) and a large tryptic (domain II + III, residues 198–585) fragment of HSA. They showed binding affinity as well as an induced CD similar to that of HSA for warfarin bound to the peptic fragment. The tryptic fragment, however, showed no induced ellipticity and a binding affinity 1 order of magnitude lower than HSA. This was interpreted such that the major binding site is located in domain II, with important contributions from domain I. HSA-DOM II as used in our experiments contains an extra 9 N-terminal residues compared with the tryptic fragment used by Bos and colleagues. Since, in contrast to these authors, we see a clear ICD signal with HSA-DOM II and warfarin, our find-

ings provide additional support for the importance of the N-terminal portion of HSA-DOM II and of the border region between subdomains IB and IIA for the warfarin binding site.

Diazepam Binding—Here we provide further experimental evidence that the primary diazepam binding site is located in HSA-DOM III, since the ICD signal obtained with this domain closely resembles that of HSA upon binding to diazepam (Fig. 7A). This is in accordance with Bos *et al.* (10), who used the peptic (residues 1–387) and the tryptic (residues 198–585) HSA fragments (see “Warfarin Binding” above) corresponding to domains I + II and domains II + III, respectively, to study the binding of diazepam. Crystallographic proof of the primary diazepam binding site was provided by Carter and Ho (1, 35). Our titration experiments showed that, under identical conditions, the binding of diazepam to HSA becomes saturated at a ligand/protein ratio of 2, which is not the case with HSA-DOM III, where a signal increase can still be detected at higher ligand concentration (Fig. 7B). These observations may be explained by a reduced binding affinity of HSA-DOM III as compared with HSA, in analogy to the equilibrium dialysis results obtained by Bos *et al.* (10) employing the proteolytic fragments.

In addition, we detected an ICD signal when diazepam was bound to HSA-DOM I. However, this signal is opposite in sign compared with HSA and HSA-DOM III, suggesting that the binding pocket in this domain is inducing different conformational changes in diazepam upon binding as compared with the binding site in HSA-DOM III. Indeed, Bos *et al.* (10) reported similar observations with their peptic fragment corresponding to domains I + II. The binding affinity of diazepam to this fragment was, however, reported to be up to 2000-fold lower than for the highest binding affinity found with HSA. Since we could not detect any ICD signal with HSA-DOM II, we therefore conclude that this secondary binding site described by Bos *et al.* was further pinpointed by our experiments to HSA-DOM I.

In conclusion, we were able to show that the three recombinant domains of HSA exhibit secondary and tertiary structure comparable to HSA, suggesting that their conformation is similar to their native structure in the context of HSA. From the evolutionary point of view, the domains as produced here can be seen as potential models of the primordial albumin, which is assumed to have consisted of only a single domain (2). The validity of the domains as stand-alone proteins is further underlined by the qualitative ligand binding studies presented here utilizing induced circular dichroism, with which we were able to provide experimental evidence that the primary binding site for hemin is located in HSA-DOM I, that for diazepam in HSA-DOM III, and that a major part of the warfarin binding site is located in HSA-DOM II, complemented by parts of HSA-DOM I. The potential for the use of the domains in drug binding and displacement studies is clearly demonstrated. We are presently extending the structural and functional characterization of the domains by a range of quantitative ligand binding studies, employing quenching of the intrinsic fluorescence of Trp-214, binding of fluorescent probes, difference spectroscopy, capillary electrophoresis, and chromatographic ap-

proaches. Crystallographic studies aimed at the elucidation of the atomic structures of the HSA domains are under way.

Acknowledgments—We thank Dr. Tilman Voss (Boehringer Ingelheim Austria GmbH) for use of their CD equipment, Dr. Friedrich Altmann for performing the MALDI-TOF analysis, Drs. Joseph Ho and John Ruble for help in preparing Fig. 1, Eva Obermayr for expert technical help, and Sabine Nussbaum for the gel filtration analysis.

REFERENCES

- Carter, D. C. & Ho, J. X. (1994) *Adv. Protein Chem.* **45**, 153–203
- Peters, T., Jr. (1996) *All about Albumin: Biochemistry, Genetics and Medical Applications*, Academic Press, Inc., Orlando, FL
- Carter, D. C., He, X. M., Munson, S. H., Twigg, P. D., Gernert, K. M., Broom, M. B. & Miller, T. Y. (1989) *Science* **244**, 1195–1198
- Carter, D. C. & He, X. M. (1990) *Science* **249**, 302–303
- Minghetti, P. P., Ruffner, D. E., Kuang, W. J., Dennison, O. E., Hawkins, J. W., Beattie, W. G. & Dugaiczak, A. (1986) *J. Biol. Chem.* **261**, 6747–6757
- Sudlow, G., Birkett, D. J. & Wade, D. N. (1975) *Mol. Pharmacol.* **11**, 824–832
- Curry, S., Mandelkow, H., Brick, P. & Franks, N. (1998) *Nat. Struct. Biol.* **5**, 827–835
- Reed, R. G., Feldhoff, R. C., Clute, O. L. & Peters, T., Jr. (1975) *Biochemistry* **14**, 4578–4583
- Ledden, D. J., Feldhoff, R. C. & Chan, S. K. (1982) *Biochem. J.* **205**, 331–337
- Bos, O. J., Fischer, M. J., Wilting, J. & Janssen, L. H. (1988) *Biochim. Biophys. Acta* **953**, 37–47
- Peters, T., Jr. (1985) *Adv. Protein Chem.* **37**, 161–245
- Bos, O. J., Remijn, J. P., Fischer, M. J., Wilting, J. & Janssen, L. H. (1988) *Biochem. Pharmacol.* **37**, 3905–3909
- Bos, O. J., Fischer, M. J., Wilting, J. & Janssen, L. H. (1989) *Biochem. Pharmacol.* **38**, 1979–1984
- Bos, O. J., Labro, J. F., Fischer, M. J., Wilting, J. & Janssen, L. H. (1989) *J. Biol. Chem.* **264**, 953–959
- Hrkal, Z., Kodicek, M., Vodrazka, Z., Meloun, B. & Moravek, L. (1978) *Int. J. Biochem.* **9**, 349–355
- Compagnini, A., Fisichella, S., Foti, S., Maccarrone, G. & Saletti, R. (1996) *J. Chromatogr. A* **736**, 115–123
- Kjeldsen, T., Pettersson, A. F., Drube, L., Kurtzhals, P., Jonassen, I., Havelund, S., Hansen, P. H. & Markussen, J. (1998) *Protein Exp. Purif.* **13**, 163–169
- Glatz, J. F. & Veerkamp, J. H. (1983) *J. Biochem. Biophys. Methods* **8**, 57–61
- Gill, S. C. & von Hippel, P. H. (1989) *Anal. Biochem.* **182**, 319–326
- Provencher, S. W. & Glöckner, J. (1981) *Biochemistry* **20**, 33–37
- Provencher, S. W. (1982) *Comput. Phys. Commun.* **27**, 229–242
- Wetzel, R., Becker, M., Behlke, J., Billwitz, H., Böhm, S., Ebert, B., Hamann, H., Krumbiegel, J. & Lassmann, G. (1980) *Eur. J. Biochem.* **104**, 469–478
- Lee, J. Y. & Hirose, M. (1992) *J. Biol. Chem.* **267**, 14753–14758
- Uversky, V. N., Narizhneva, N. V., Ivanova, T. V. & Tomashevski, A. Y. (1997) *Biochemistry* **36**, 13638–13645
- Sjöholm, I. & Ljungstedt, I. (1973) *J. Biol. Chem.* **248**, 8434–8441
- Beaven, G. H., Chen, S. H., d'Albis, A. & Gratzer, W. B. (1974) *Eur. J. Biochem.* **41**, 539–546
- Fehske, K. J., Müller, W. E., Wollert, U. & Velden, L. M. (1979) *Mol. Pharmacol.* **16**, 778–789
- Otagiri, M., Maruyama, T., Imai, T., Suenaga, A. & Imamura, Y. (1987) *J. Pharm. Pharmacol.* **39**, 416–420
- Wilting, J., Hart, B. J. & De Gier, J. J. (1980) *Biochim. Biophys. Acta* **626**, 291–298
- Peters, T., Jr. & Hawn, C. (1967) *J. Biol. Chem.* **242**, 1566–1573
- Peters, T., Jr. & Blumenstock, F. A. (1967) *J. Biol. Chem.* **242**, 1574–1578
- Markus, G., McClintock, D. K. & Castellani, B. A. (1967) *J. Biol. Chem.* **242**, 4395–4401
- Sjodin, T., Hansson, R. & Sjöholm, I. (1977) *Biochim. Biophys. Acta* **494**, 61–75
- Bos, O. J., Fischer, M. J., Wilting, J. & Janssen, L. H. (1988) *J. Chromatogr.* **424**, 13–21
- He, X. M. & Carter, D. C. (1992) *Nature* **358**, 209–215
- Carter, D. C. (July 14, 1998) *U. S. Patent* 5,780,594
- Ohi, H., Miura, M., Hiramatsu, R. & Ohmura, T. (1994) *Mol. Gen. Genet.* **243**, 489–499
- Compagnini, A., Fichera, M., Fisichella, S., Foti, S. & Saletti, R. (1993) *J. Chromatogr. A* **639**, 341–345
- Strickland, E. H., Horwitz, J. & Billups, C. (1969) *Biochemistry* **8**, 3205–3213
- Horwitz, J., Strickland, E. H. & Billups, C. (1970) *J. Am. Chem. Soc.* **92**, 2119–2129
- Horwitz, J., Strickland, E. H. & Billups, C. (1969) *J. Am. Chem. Soc.* **91**, 184–190
- Carson, M. (1987) *J. Mol. Graph.* **5**, 103–106
- Laemmli, U. K. (1970) *Nature* **227**, 680–685