BIOCHEMISTRY
including biophysical chemistry & molecular biology

Article

pubs.acs.org/biochemistry

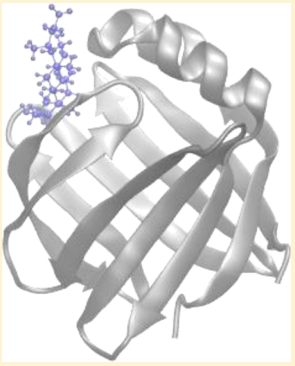
Cooperativity and Site Selectivity in the Ileal Lipid Binding Protein

Eleanor R. Turpin, Huey-Jen Fang, Neil R. Thomas, and Jonathan D. Hirst*

School of Chemistry, University of Nottingham, University Park, Nottingham NG7 2RD, United Kingdom

Supporting Information

ABSTRACT: The ileal lipid binding protein (ILBP or I-BABP) binds bile salts with positive cooperativity and has unusual site selectivity, whereby cholic acid binds preferentially in one site and chenodeoxycholic in another, despite both sites having an affinity for both ligands and the ligands only differing by a single hydroxyl group. Previous studies of the human variant have assumed that the ligand/protein binding ratio is 2:1, but we show, using electrospray ionization mass spectroscopy, that human ILBP binds bile acids with a 3:1 ratio, even at low protein and ligand concentrations. Docking calculations and molecular dynamics (MD) simulations identify an allosterically active binding site on the protein exterior that induces a change from a closed conformation to an open one, characterized by a movement of one of the α-helices by ~10° with respect to the β-clam shell. Additional independent MD simulations of several hundred nanoseconds implicate the change between conformations in the mechanisms of both cooperativity and ligand site selectivity.



Allostery is the term used to describe the process by which a binding event in one part of a macromolecule has an effect on another binding site in a different part of the macromolecule. Allostery is an important feature of many protein−ligand interactions. The binding of a ligand often induces a conformational change in the protein linked to changes in biological activity, as first recognized in 1963 by Monod, Changeux, and Jacob.[1] Cooperative binding is often a special case of allostery, where binding of a ligand promotes a conformation change that increases the affinity for binding another ligand. Recent reviews[2,3] provide a historical perspective to modern allostery models and comprehensive examples of applications to many biomolecular systems including regulatory enzymes, membrane proteins, and nuclear receptors. The ileal lipid binding protein (ILBP, also known as I-BABP) binds bile salt ligands cooperatively, on a level comparable to hemoglobin binding with oxygen,[4] and displays unusual selectivity of the ligands, despite these only differing by a single hydroxyl group.[5] These properties make ILBP an interesting candidate for the study of allosteric interactions by molecular dynamics (MD) simulations; one of the advantages of computational methods is that nonphysical protein−ligand combinations, where ligands bind in a different order to that observed by experiment, can be investigated. In this article, we show for the first time, using electrospray ionization mass spectroscopy (ESI MS), that human ILBP binds three ligands, even at low protein and ligand concentrations, and using docking and MD simulations propose a site on the protein exterior that leads to an allosterically induced change in the protein from a closed conformation to an open one. The open state is characterized by a movement of one of the α-helices by ~10° with respect to the β-clam shell, and further MD simulations implicate conformational shifts between the two states in the mechanisms of both cooperative binding and site selectivity.

The lock and key theory[6] was developed to describe enzyme−substrate binding, and it assumed that the enzyme binding site was rigid and was the geometric complement of its substrate. However, as proteins are flexible molecules, the induced fit model of Koshland[7] was developed to take account of conformation changes upon ligand binding. A current model of protein−ligand binding is population-shift theory, which takes into account the dynamic nature of proteins. Developed from the 1960s MWC model of Monod, Wyman, and Changeux,[8] the unbound native state of a protein is considered to be not a single conformation but an ensemble of low energy states with small energy barriers between them. The unbound ensemble samples some of the same conformations as the bound protein but with a lower probability because they have a higher energy. Upon binding, the energy of the bound conformations is lowered, so they are sampled more; hence, the population contains more bound conformations. Under the population shift paradigm, allosteric ligand binding induces changes that begin at the allosteric binding site and are propagated to other parts of the protein. Allostery is a thermodynamic effect,[9] and positive cooperative binding can be propagated by enthalpy through the formation of favorable interactions or by increasing the sampling of conformational states that contain a new binding site.

MD simulations have been used to determine the allosteric transitional pathway of cooperative binding systems. One important example is the protein GroEL, a chaperonin.[10] Its structure comprises 14-subunits, arranged into two rings. It cooperatively binds ATP and a second chaperonin with each ring but anticooperatively between the rings. The binding of

Received: February 14, 2013
Revised: June 7, 2013
Published: June 12, 2013

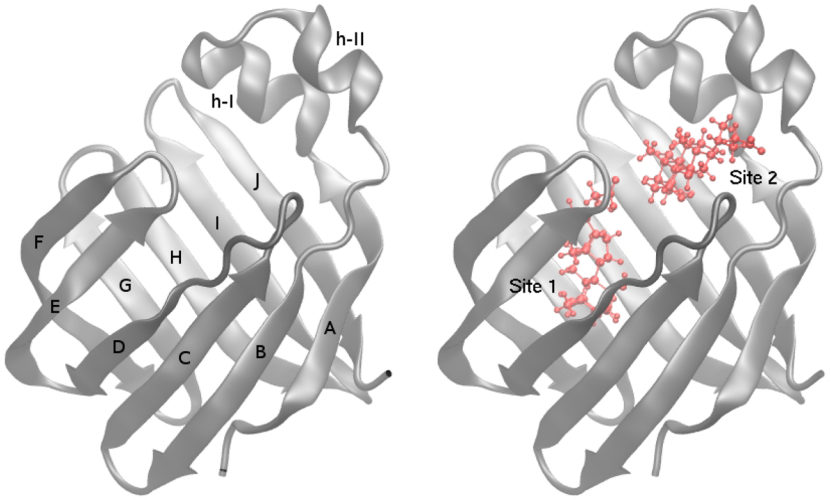


Figure 1. (Left) Convention for naming the β-strands and α-helices of ILBP. (Right) convention for naming the binding sites. Figure generated using the crystallographic structure of zebrafish ILBP bound to two ligands of cholic acid (PDB ID: 3ELZ, chain A).

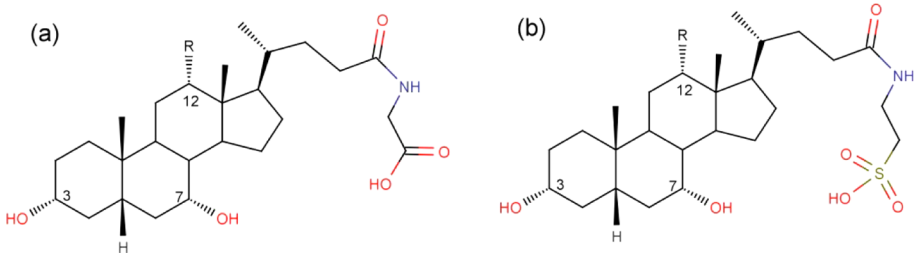


Figure 2. Primary bile salts. For cholic acid, R is OH, and for chenodeoxycholic acid, R is H; (a) is conjugated with glycine and (b) with taurine.

the ligands to one ring induces changes in all parts of the protein to move from a closed, unbound conformation to an open, bound conformation. The path of this transformation is impossible to measure experimentally,[11] but the transition states have been identified using targeted MD simulations.[12] A recent example of MD simulations being used to clarify the results of biophysics experiments of a system demonstrating unusual binding is the allosteric mechanism that inhibits 3-deoxy-D-arabino-heptulosonate 7-phosphate synthase (DAH7PS).[13] DAH7PS is the first enzyme on the shikimate pathway, which is the biosynthetic path for the production of aromatic compounds such as tyrosine, tryptophan, phenylalanine, and folic acid. Using results from X-ray scattering experiments, ITC, and MD simulations, it has been shown[13] that binding in the Trp/Phe site leads to cooperative binding in the Phe/Trp site by increasing flexibility in the upper part of the protein. Only when both sites are occupied (indicating an abundance of Trp and Phe) is the flexibility of the lower part of the protein reduced, inhibiting the catalytic site of the enzyme.

ILBP belongs to the intracellular lipid binding protein family.[14] All members have a 10-strand antiparallel β-sheet in a structure which looks like a clam shell, with two, short, nearly parallel α-helices covering the opening of the β-clam and one or two ligands bound inside (Figure 1). The ligands of ILBP are bile salts, which are amphiphilic with a hydrophobic steroid scaffold and a polar tail. The primary bile salts (Figure 2) cholic acid and chenodeoxycholic acid, and their derivatives, are biosynthesised in the liver in a ratio of 2:1, constituting 80% of the human bile acid pool.[5] The other 20% are secondary bile salts such as deoxycholic acid and lithocholic acid, which are modified by gut flora. When biosynthesized *in vivo*, bile salts are conjugated with glycine or taurine in a 3:1 ratio. Conjugation results in compounds that are fully ionized and soluble at physiological pH. This enhances their effectiveness as detergents,[5] which is important, as their biological role is to aid the digestion of fats. Bile salts are synthesized from cholesterol in the liver, and this synthesis accounts for about half of the catabolism of body cholesterol. When bile acids are reabsorbed in the ileum, they activate the nuclear farnesoid X receptor (FXR) in the enterocyte (an absorptive intestinal cell), which stimulates expression of ILBP.[15] This creates a positive feedback loop that leads to further bile acid absorption. Thus, ILBP is of interest as a pharmaceutical target. If its function were suppressed, the bile salts would be excreted from the body, resulting in more cholesterol being catabolized.

Tochtrop et al.[5] used NMR experiments with enriched bile salts to demonstrate site selectivity in human ILBP. In the two-dimensional heteronuclear NMR spectrum of ILBP and ^{15}N-glycocholate (GCA), three peaks are resolved, one for the unbound bile salts and one for each binding site. In a spectrum of ILBP, enriched ^{15}N-GCA, and unenriched glycochenodeoxycholate (GCDA) (invisible to the NMR experiment), only two peaks were resolved: one for unbound bile salts and one at binding site 2 (the sites are labeled in Figure 1); GCDA had completely displaced the labeled GCA from site 1. Similarly, in an experiment with ^{13}C enriched bile salts, unenriched GCDA displaced ^{13}C-GCA from binding site 1, and unenriched GCA displaced ^{13}C-GCDA from binding site 2. GCA and GCDA are identical except for a hydroxyl group at C12. In samples containing only one ligand species, both sites were occupied; so, selectivity cannot arise from steric exclusion or a lack of affinity for one site. ILBP shows positive cooperativity

dx.doi.org/10.1021/bi400192w | *Biochemistry* 2013, 52, 4723−4733

comparable with that of hemoglobin. The Hill coefficient is a macroscopic measurement of cooperativity; it is unity for noncooperative systems and 2 for extremely positively cooperative systems. Hemoglobin has a Hill coefficient of 2.8 to 3.0 and is greater than 2 because it has multiple binding steps. The Hill coefficient for ILBP was calculated as 1.94 by Tochtrop et al.[4] Toke et al.[16] showed that cooperative binding and site selectivity are independent in human ILBP, using isothermal titration calorimetry (ITC) and NMR experiments on mutated proteins; the choice of residue to mutate was determined from an unpublished NMR structure of ILBP in complex with GCA and GCDA, where GCA has the additional hydroxyl group. Zanzoni et al.[17] have introduced cooperative binding into the chicken variant of ILBP, which binds two bile salt ligands, but not cooperativity.[18] Following phylogenetic analysis of human ILBP and chicken ILBP, a double mutant, H99Q and A101S, was produced (of the chicken variant) which leads to a gain in cooperative binding. A change of the hydrogen bonding network in the bottom of the binding cavity appears to lead to cooperative binding between the ligands in the protein interior, as without Ser-101 to coordinate with the His-99 the binding cavity is less flexible.

The number of binding sites of human ILBP has not yet been conclusively determined experimentally. Tochtrop et al.[5] stated that ILBP binds to bile salts with a 1:2 ratio, but Lucke et al.[19] resolved only a single cholic acid in the binding cavity by NMR. Capaldi et al.[15] published the structure of zebrafish ILBP in complex with cholic acid. Two cholic acid molecules are present in the binding cavity of all five protein conformers produced from the two crystal forms; for four of the five conformers, the location of the binding site overlaps very well for both ligands; for the fifth conformer, the location of one binding site is the same as that of the other four, and the location of the second binding site is the same, but one cholic acid molecule has been rotated by 90° about its longest axis. Cholate molecules are found on the surface of all five of the models, although these have a much more variable alignment than those in the interior pocket of the protein. Two of the exterior cholic acids are present in all five models, but the precise location and orientation varies because of a lack of specific hydrogen bonds between the ligand and receptor. The region of the surface with the exterior cholic acids is almost entirely hydrophobic. Capaldi et al.[15] speculated that the physical role of exterior binding sites could be to help guide molecules into the interior or improve binding in the presence of excess ligand. ITC data confirmed the presence of bound ligands but could not discriminate between a model with (i) three independent binding sites, two in the interior and a third, entropically driven, adhesion site on the surface, or (ii) three states, free, first site occupied, and second site occupied. Experimental data to support specific binding sites on hydrophobic surfaces are difficult to acquire, and exterior ligands can disrupt crystallization. This might explain the difficulty in crystallizing rabbit ILBP described by Kouvatsos.[20,21] It has been shown using ITC and ESI MS that the rabbit variant of ILBP binds with a ligand/protein ratio of 3:1.[21] In the gain of function study by Zanzoni et al.,[17] only interior residues were considered when identifying mutants for introducing cooperative binding to chicken ILBP. The stoichiometry and the possibility of exterior binding sites were not discussed. Human ILBP may also have three binding sites; Toke et al.[16] reported the binding ratio determined by NMR of ILBP/GCDA/GCA as 1:1.5:1.5, but this is not discussed further, the authors explaining that the structure would be published subsequently. This binding ratio is also stated by Horvath,[22] but the accompanying discussion focuses on two binding sites only.

Using ESI-MS, we establish that the ligand/protein binding ratio of human ILBP with bile salts is 3:1. We then use the complementary, computational methods of docking to identify a possible location of the third binding site and use MD simulations to understand its role in cooperative binding and ligand site selectivity.

MATERIALS AND METHODS

Protein Expression and Purification. A synthetic gene for the human ILBP was synthesized by DNA2.0 (Menlo Park, CA) based on the reported sequence (PDB ID: 101U). The ILBP gene was inserted into the pET21a expression vector (Novagen) between the *NdeI* and *XhoI* restriction sites. The vector was transformed into $CaCl_2$ competent BL21 (DE3) (Strategene) *E. coli* for protein expression. This vector allows the expression of the human ILBP with a C-terminal addition of LEHHHHHH. An aliquot of 100 μL of the transformed culture was added into 100 mL of the sterile LB medium containing ampicillin (100 μg/mL), and the cell mixture was incubated at 37 °C with shaking (180 rpm) overnight. The overnight culture was added into the flask containing 1 L of the sterile LB medium with ampicillin (100 μg/mL), and this mixture was incubated for 1 h at 37 °C until OD_{600} reached 0.3. Ispropyl-β-D-thiogalactopyranoside (IPTG) (0.8 M, 1 mL) was added to the culture which was then incubated at 37 °C for another 4 h, and the cells were harvested by centrifugation at 5000 rpm for 15 min. The ILBP protein containing the C-terminal His_6 affinity tag was purified by immobilized metal affinity chromatography (IMAC), and 5 mL of HisPur Cobalt resin (Thermo Scientific) was used for protein purification. Before loading the protein lysate, the HisPur cobalt column was equilibrated with 20 mL of IMAC buffer. The purification process involved loading 20 mL of protein lysate, washing the column with 20 mL of IMAC buffer, eluting the bound protein with 20 mL of IMAC buffers containing 5, 50, and then 100 mM imidazole sequentially. All eluted fractions were analyzed by 15% SDS−PAGE to show which fraction contained the pure recombinant protein. The fractions containing the target protein were concentrated and buffer exchanged with IMAC buffer using Amicon Ultra-15 Ultracel-3K centricon spin concentrators (Millipore). Protein stocks (1 mM) were then stored at −80 °C and thawed prior to use.

Bile Acid Preparation. Bile acids were purchased from Sigma and dissolved in ddH_2O for preparing 100 mM liquid stocks. All liquid stocks were stored at −20 °C for future experiments.

Electrospray Ionization Mass Spectrometry. ESI-MS experiments were performed on a Synapt HDMS (Waters, Manchester, UK) equipped with the standard z-spray source. Samples were introduced into the source region using a Harvard (Hollistion, MA, USA) apparatus 22 dual syringe pump (model 55−2222) and a 100 μL Hamilton syringe (Bonaduz, Switzerland), at a 5 μL/min infusion rate. The typical setting for the ESI experiments was positive ion mode with the ESI capillary held at 2.8 kV; nitrogen desolvation gas with 200 L h^{-1}; 50 °C source temperature; cone gas flow with 30 L h^{-1}. The cone voltage was 40 V. Instrument control and data analysis were carried out using Masslynx software (Waters). The sample (100 μL) contained 10 μM protein

dx.doi.org/10.1021/bi400192w | *Biochemistry* 2013, 52, 4723−4733

and 40 μM ligand in 25 mM ammonium acetate (pH 7.0). Each protein solution for ESI-MS was prepared by the dilution of 100 μM protein stock which was desalted using a 3 kDa cutoff Vivaspin 30 centrifugal concentrator (Fish Scientific, UK) and buffer exchanged with 25 mM ammonium acetate (pH 7.0).

Initial Structures for Simulation. All simulations were performed using the human variant of ILBP. NMR structures are published of the apoprotein form of human ILBP and resolved with a single ligand of TCA.[23] These structures were accessed from the Protein Data Bank (PDB) using the codes PDB ID, 1O1U (apo), and PDB ID, 1O1V (single ligand). The first conformation in each PDB file was used as the starting coordinates. For the single ligand form, the TCA ligand was mutated in CHARMM[24,25] to the nonconjugated cholic acid ligand and energy minimized, with harmonic constraints applied to all atoms of the protein (force constant of 20 kcal mol^{-1} Å^{-2}), for 200 steps using the steepest descents algorithm.[25] The histidine residues, H52, H57, and H98, were examined in VMD[26] for hydrogen bonding with nearby neighbors to determine the protonation state of the side chain. No hydrogen bonds were present for the single ligand form, and the neutral form with the hydrogen on Nδ (HSD), as stated in the PDB file, was used. For the apo form, a hydrogen bond is formed between H98-Nδ and the backbone of Q99, justifying the use of the HSD form. The other histidines did not form any hydrogen bonds and were left as HSD, as stated in the PDB.

The models for the double ligand human ILBP were built from the crystallographic structure of zebrafish ILBP bound to two ligands of cholic acid (PDB ID: 3ELZ, chain A). The exterior ligands from 3ELZ were not included in the model. Zebrafish ILBP is three residues longer in the I-J loop region than the human variant. Sequence alignment of 1O1V and 3ELZ using ClustalW[27] shows that there are 72 conserved residues, 23 residues with strongly similar properties, 19 residues with weakly similar properties, and 13 nonconserved residues.

To build the human ILBP models with two ligands, residues 117−119 of the zebrafish ILBP structure were deleted. The nonconserved residues were changed in CHARMM by reading in the zebrafish ILBP primary sequence, renaming the residues to match the human variant in each position that needed altering, deleting the side chains of the modified residues, and rebuilding them from the internal coordinates in the parameter set.[28] The histidine residues were examined for hydrogen bonding using VMD,[26] and all were determined to be HSD. The structures were minimized using 2000 steps of steepest descents followed by 2000 steps of adopted-basis Newton−Raphson.

Three structures of human ILBP with two ligands were constructed. The first has a cholic acid molecule in both interior binding sites. The second has a cholic acid molecule in the first binding site (see Figure 1 for binding site naming convention) and a chenodeoxycholic acid molecule in the second site; this is termed mixed-CA. The third has a chenodeoxycholic acid molecule in the first binding site and cholic acid in the second site and will be named mixed-CDA. These mixed ligand structures have been constructed to study site selectivity and to confirm which binding site corresponds to which ligand, as no mixed ligand NMR or X-ray crystallographic structures have been published. Using MD simulations to study site selectivity behavior in this way makes it is possible to simulate the nonphysical protein−ligand combination where the unfavored ligands are in each binding site.

Molecular Dynamics Simulations. The initial apoprotein, single ligand, and double ligand complex structures were solvated in CHARMM in a box of 4500 to 6600 TIP3P water molecules.[29] The shape of the unit cell was a truncated octahedron with maximum length vectors of 68.7 Å (apo), 71.1 Å (single ligand), and 76.8 Å (double ligand). Each system was energy minimized in NAMD[30] using the standard conjugate gradient algorithm for 10,000 steps with the protein atoms fixed, 10,000 steps with the protein backbone atoms fixed, and a further 20,000 steps with no constraints.

All heating, equilibration, and production dynamics were performed using NAMD[30] with a time step of 2 fs, the CHARMM22 force field, and periodic boundary conditions. The parameters for the bile salts were taken from the CHARMM general force field (CGenFF) of drug-like small molecules.[31] The SHAKE algorithm[32] was applied to all bonds to hydrogen atoms. Long-range electrostatics were treated using the particle-mesh Ewald method.[33] Dynamics were performed using a Langevin dynamics integrator with a friction coefficient of 5 ps^{-1}. The solvent−protein system was heated from 0 to 298 K in increments of 30 K by temperature reassignment, where the velocities of all the atoms in the system are reassigned so that the entire system is set to the target temperature. The velocities were reassigned every 100 time steps for 10,000 time steps using the NVT ensemble. During the heating phase, the protein backbone atoms were initially constrained using harmonic constraints with a force constant of 25 kcal mol^{-1} Å^{-2}, and the force constant was gradually reduced by 2.5 kcal mol^{-1} Å^{-2} every 1000 steps to zero. The systems were equilibrated for a further 25,000 time steps using the NPT ensemble, with the pressure set to 1 atm. Production dynamics were run for 300 ns for the apo, single ligand, and double cholic acid ligand systems. For the mixed structures, the production dynamics length was 220 ns. The thermodynamic stability of all the MD simulations was confirmed by examining the kinetic energy, potential energy, and total energy (data not shown).

Docking Calculations. Docking was performed between ILBP and a CA ligand. The atomic positions were taken from 11 snapshots sampled uniformly across the apoprotein, single ligand, and double CA ligand complex trajectories to produce starting coordinates for the docking receptor. For the structures bound to a single ligand, two sets of starting coordinates were generated, one with the ligand present and one with it absent. In total, there were 44 sets of starting coordinates for the receptor. The ligand starting coordinates were taken from one of the interior ligands of the zebrafish X-ray crystallographic structure. The ligand has seven rotable bonds, C3−O3, C7−O7, and C12−O12 from the steroid rings and C17−C20, C20−C22, C22−C23, and C23-CD from the tail section, which were free to rotate during the docking calculations. The receptor structures did not have any rotable bonds.

Calculations were performed using AutoDock Vina.[34] In a systematic comparison[35] of widely used docking programs, AutoDock 4[36] performed reasonably compared with that of other docking programs and was able to identify over 93% of protein−ligand pairs from the PDBbind database.[37] All of the docking results reported here were produced using AutoDock Vina,[34] an update to AutoDock 4 that improves the accuracy of the predicted binding modes. The input files for the ligands and receptors were prepared using AutoDock Tools,[38] which

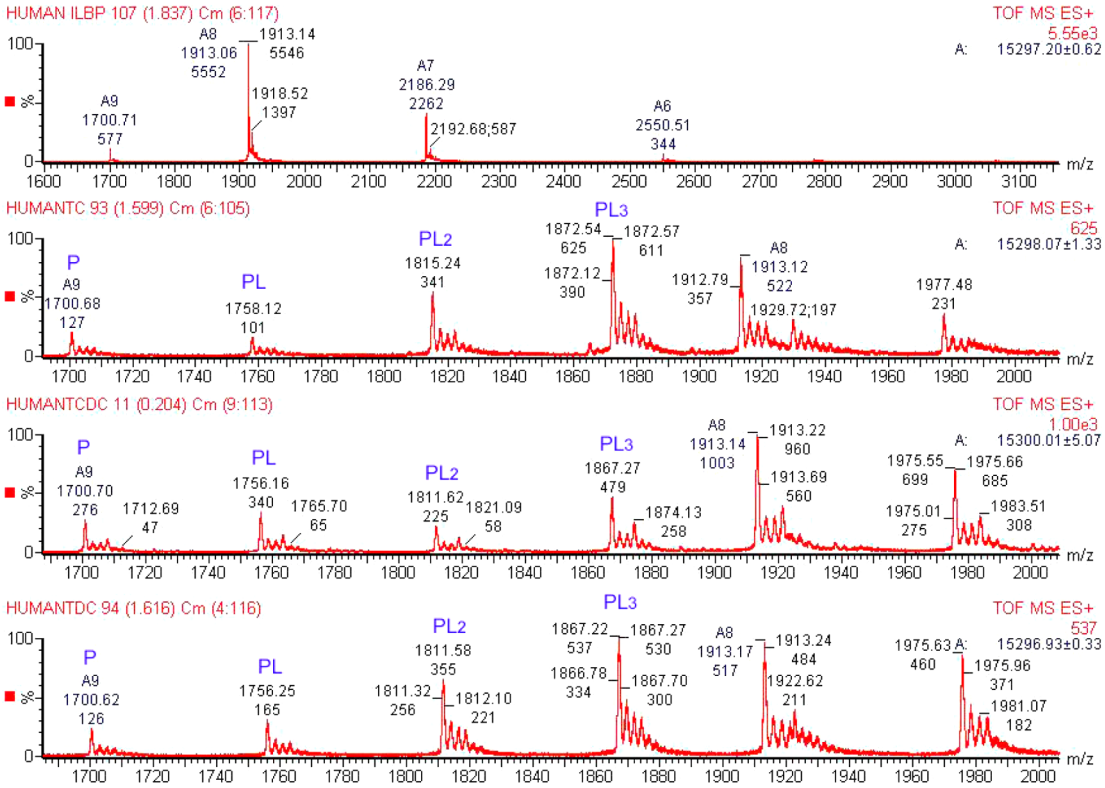


Figure 3. ESI-mass spectra of apo human ILBP (top) and, in descending order, taurocholic acid (TC), taurochenodeoxy cholic acid (TCDC), and taurodeoxycholic acid (TDC). Peaks corresponding to the apo protein (P) and protein/ligand complexes (PL_X) are labeled.

automatically merges nonpolar hydrogen atoms to form united atom groups and assigns the partial charges and atom types required by the AutoDock scoring function. The atom types assigned to the ligand atoms that were part of the receptor were manually checked, and they agreed with atoms types of protein side chains with the same or similar chemical groups. The origin of the search space was at the geometric center of each receptor. The edge lengths of the cuboid that define the search space were between 36 Å and 50 Å and chosen so that the search space covered the entire exterior of the protein receptor. The exhaustiveness parameter was set to 50 and the energy range parameter to 7 kcal mol^{-1}.

RESULTS

Bile Acid Binding Experiments. Three taurine-conjugated bile acids (L), taurocholic acid (TC), taurodeoxycholic acid (TDC), and taurochenodeoxycholic acid (TCDC), were used to study binding. The ligand stoichiometry of human ILBP was demonstrated using ESI-MS. The apoprotein was used to assess the mass-to-charge (m/z) fragments in the gas phase (Figure 3). Native ILBP apoprotein (A) produces multiply charged molecular ion peaks with 6^+, 7^+, 8^+, and 9^+ ions in the ESI-MS spectrum. The binding studies by ESI-MS were performed with the samples containing 10 μM protein and 40 μM taurine-conjugated bile acids. Protein–ligand complexes PL, PL_2, and PL_3 along with uncomplexed protein (P) appeared together in the ESI-MS spectrum, but their populations depend on the bile acids present. The ESI-MS spectra show the highest populations of the complexes (PL_X) with bile acids that can be formed if the bile acid is limiting rather than free protein (P). Within a given mass spectrum for human ILBP with TC or TCDC, the population (i.e., signal) due to the PL_3 complex is greater than that for PL_2, which in turn is greater than that for PL, which is greater than that for P.

In the gas phase, there is no mechanism for reorganization of the ligand/protein complexes, and as a result, it is a nonequilibrated system representing the complexes present in the solution before evaporation. To confirm that the non-specific binding peaks present in the ESI-MS spectra are not due to complexes that would be favored in the gas phase,[39] the experiments were repeated (data not shown) using 3 μM protein and 1.5 to 18 μM taurine conjugated bile acid. This corresponds to ratios from 2:1 up to 1:6 (protein/ligand). The ESI-MS shows a 1:3 complex for all concentrations from 9 to 18 μM, and there is no evidence for 1:4 or higher complexes.

Docking. The starting coordinates for the receptor were taken from snapshots from the MD trajectories with cholic acid as the ligand for all docking calculations (see the Materials and Methods section for details.) The docked ligand–protein complexes predicted by AutoDock Vina were compared with the X-ray crystallographic structure of zebrafish ILBP (PDB ID: 3ELZ), which has 89% sequence identity with human ILBP. To calculate the difference between ligands from a docked conformation and the experimental structure, the protein backbone atoms were aligned before the root-mean-square difference (RMSD) was calculated using the ligand heavy atoms. Results are reported for the first binding mode from each docking run, which is defined as the protein–ligand complex with the lowest binding affinity calculated by AutoDock Vina.

The first set of calculations was to redock a deleted ligand into receptor conformations taken from the single ligand trajectory. Nine of the 11 docking runs reproduce the original bound conformation with the same ligand orientation. The other two runs redock the ligand into the same site but with an

dx.doi.org/10.1021/bi400192w | Biochemistry 2013, 52, 4723–4733

Table 1. Results of Binding to Conformations from the Apo Trajectories[a]

		location (from visualization)		RMSD from zebrafish X-ray ligands (Å)				
receptor reference	binding affinity (kcal mol^{-1})			site 1	site 2	site 3	site 4	site 5
0	−6.3	exterior	rings near E-F loop, tail near α-II			4.7	6.0	4.6
1	−5.9	exterior	between strands E-F at bottom of β-clam shell			13.1	9.6	12.2
2	−8.7	interior	tail up	3.4	4.6			
3	−6.6	exterior	between strand A and α-I			8.6	11.7	9.8
4	−7.1	interior	tail down	4.3	5.2			
5	−7.9	interior	tail up	2.4	5.5			
6	−5.8	exterior	between strand A and α-I and α-II			10.0	12.1	10.5
7	−5.8	exterior	between strands F-G and α-I−α-II loop			9.5	3.3	6.5
8	−6.4	exterior	between strands F-G and α-I−α-II loop			8.4	2.9	5.8
9	−6.0	exterior	among strand A, α-I, and α-II			9.4	11.3	9.5
10	−6.1	exterior	among strand A, α-I, and α-II			9.5	11.6	9.9

[a]The receptor reference 0 refers to the initial conformation at the start of the production dynamics, 1 to the conformation after 15,000,000 steps, 2 to the conformation after 30,000,000 steps, etc.

incorrect orientation with the tail pointing toward the bottom of the binding cavity. The average RMSD between the docked ligand and the deleted ligand is 1.9 ± 0.9 Å.

Table 1 shows the docking results using the coordinates from the apoprotein MD trajectories. For three of the receptor structures (2, 4, and 5), the ligand is docked in the interior of the protein in site 1. For two of these, the ligand has the expected orientation with the tail pointing toward the helices; the third has the tail pointing toward the bottom of the binding cavity. The ligands docked into the protein interior have a greater binding affinity (−7.1 to −8.7 kcal mol^{-1}) compared with that of the exterior ligands (−5.8 to −6.6 kcal mol^{-1}). Two exterior binding sites were docked into more than one starting conformation. The first (referred to here as site D-1) is located between β-strands F and G and the loop region connecting the two α-helices (Figures 4a and 4b). Site D-1 is in a location similar to that of site 4 of the zebrafish crystal structure, and the docked ligands have an RMSD of 2.9 Å and 3.3 Å.

A second exterior binding site (referred to as site D-2), between β-strand A and the helices (Figure 4c), was predicted for four of the starting receptor conformations (3, 6, 9, and 10). This binding site does not correspond to any of the experimentally determined exterior sites. For the other two docking runs, the ligand was docked to the exterior on a site near the E-F loop and α-helix II and on a site against the E-F strand at the bottom of the β-clam shell.

The results of docking to receptor structures from the single ligand trajectories with a ligand already present in site 1 are shown in Table 2. Eight of the docking runs predict binding to site D1, and the average RMSD from experimental site 4 for these eight complexes is 3.4 ± 0.4 Å. Of the other three docking runs, one predicts binding to the D2 site, one to a site on the G-H β-strands, and one to a site on the B-C β-strands.

Table 3 reports the results of docking to structures taken from the double CA trajectories with two ligands already present. Four of the runs dock the ligand inside the binding cavity, three between the ligands in site 1 and site 2 and the fourth at the bottom of the cavity beneath the ligand in site 2. Five of the runs predict binding to site D-1 with an average RMSD from site 4 of zebrafish ILBP of 4.2 ± 0.8 Å. The other two runs predict binding to site D-2 and a site along the C-D, E-F turns and α-helix I.

MD Simulations from Docked Complexes. Three of the docked protein−ligand complexes were selected as starting structures for further MD simulations to study binding sites D-1 and D-2. Site D-1 is close to the experimentally identified exterior binding site 4 and 8 of the 11 docking runs using single ligand receptors docked into the position. Site D-2 is of interest, as 4 of the 11 docking runs using apo receptors docked to it. From docking to single ligand structures, the complexes corresponding to receptors 5 and 6 have the lowest binding affinity (−7.2 kcal mol^{-1}); model 6 was chosen for MD simulation of D-1 because it has a smaller RMSD with respect to experimental site 4. Models 3 and 8 were chosen from the receptors from the apoprotein trajectories for MD simulations of sites D-2 and D-1 respectively, as they have the lowest binding affinities (−6.6 kcal mol^{-1} and −6.4 kcal mol^{-1}). From analysis using the LigPlot program,[40] (data not shown) the binding site of apoprotein model 8 (D-1) is characterized by

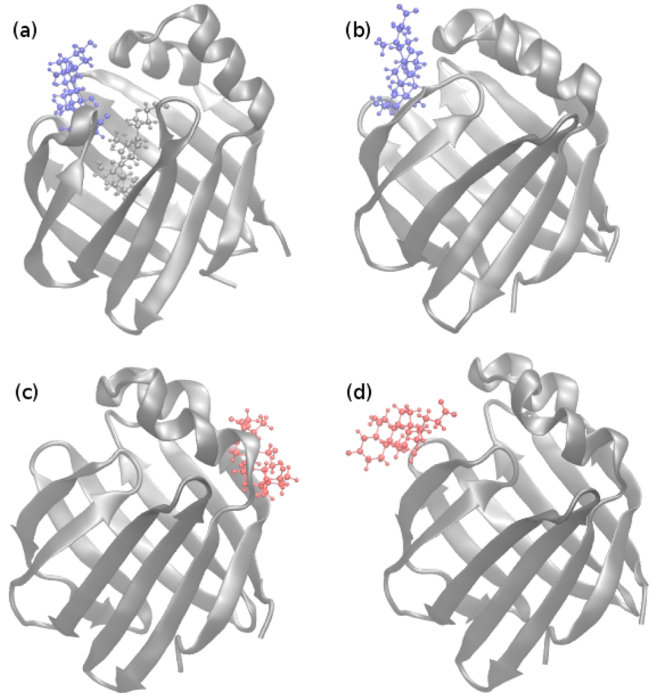


Figure 4. (a) Cholic acid docked into binding site D-1 with ligand present in interior site 1 (single ligand receptor model 6). (b) Cholic acid docked into binding site D-1 (apo receptor model 8). (c) Cholic acid docked into binding site D-2 at the start of the MD trajectory and having moved near to site D-1 after 63.3 ns of simulation (d).

dx.doi.org/10.1021/bi400192w | Biochemistry 2013, 52, 4723−4733

Table 2. Results of Binding to Conformations from the Single Ligand Trajectories

		location (from visualization)		RMSD from zebrafish X-ray ligands (Å)				
receptor reference	binding affinity (kcal mol^{-1})			site 1	site 2	site 3	site 4	site 5
0	−5.8	exterior	near G-H strand	7.5	11.0	14.1	8.2	11.5
1	−6.6	exterior	between F-G and α-I	6.7	8.8	10.0	3.9	7.0
2	−6.3	exterior	between F-G and α-I	6.0	7.2	8.0	3.6	5.5
3	−6.0	exterior	between F-G and α-I	7.1	8.1	9.1	3.2	5.7
4	−6.2	exterior	near B-C strand	8.4	7.5	8.5	11.1	10.2
5	−7.2	exterior	between F,G and α-I–α-II loop	6.1	7.7	9.2	3.5	6.3
6	−7.2	exterior	between F,G and α-I–α-II loop	6.7	6.6	8.0	3.3	4.7
7	−6.6	exterior	between F,G and α-I–α-II loop	6.0	7.5	8.8	2.7	5.8
8	−6.9	exterior	between F,G and α-I–α-II loop	6.3	8.0	9.5	3.6	6.6
9	−6.4	exterior	between F,G and α-I–α-II loop	7.4	7.2	7.6	3.7	4.2
10	−6.1	exterior	between A, α-I and α-II	10.4	6.9	9.2	12.5	10.7

Table 3. Results of Binding to Conformations from the Double Ligand Trajectories

		location (from visualization)		RMSD from zebrafish X-ray ligands (Å)				
receptor reference	binding affinity (kcal mol^{-1})			site 1	site 2	site 3	site 4	site 5
0	−7.3	exterior	between F-G and α-I	5.6	6.9	8.6	3.7	6.0
1	−8.3	interior	between site 1 and site 2	3.9	4.4	7.1	5.0	6.1
2	−7.9	interior	between site 1 and site 2	4.7	4.4	7.4	6.2	6.6
3	−6.8	exterior	between F, G and α-I	6.3	6.5	9.0	4.3	6.2
4	−6.9	exterior	between F, G and α-I–α-II loop	5.5	5.5	7.4	3.1	4.6
5	−8.2	exterior	between F, G and α-I–α-II loop	4.7	4.5	7.1	5.5	5.6
6	−9.1	interior	between site 1 and site 2	4.8	3.9	7.5	7.2	7.3
7	−8.7	interior	beneath site 2	4.7	4.8	8.5	8.5	7.8
8	−7.1	exterior	between C-D turn, E-F turn and α-I	7.0	4.5	4.2	5.2	3.7
9	−6.8	exterior	among A, α-I, and α-II	10.3	7.1	9.9	12.3	10.8
10	−6.7	exterior	between F, G and α-I	5.6	5.8	7.4	4.3	5.0

hydrophobic interactions with Gly22, Ile23, Gly76, Phe79, Phe94, and Pro95 and hydrogen bonds with Lys77, Tyr97, and Asn96; the binding site of the single ligand protein model 6 (D-1) shows hydrophobic interactions with Leu20, Leu21, Ile23, Phe79, Phe94, Asn96, Tyr97, Ile114, and the other bile salt ligand and has a hydrogen bond with Met74; the binding site of apoprotein model 3 (D-2) is characterized by hydrophobic interactions with Met8, Glu9, Lys12, Arg32, and Phe34 and hydrogen bonds with Asn13, Tyr14 and Lys35.

The MD simulations were set up and run following the same protocol as the initial simulations, as described in the Materials and Methods section. For the complex with a single ligand in site D-1, the duration of the production dynamics was 40 ns for four of the runs, and an extended run of 320 ns was produced for statistical analysis. An extended run was also performed with an interior ligand in site 1 and a ligand in site D-1 for 270 ns. Figure 5 shows the RMSD of the ligands from their starting position during these extended simulations. For the trajectory with only the exterior ligand, the average RMSD is 2.5 ± 1.1 Å, indicating that the ligand remains in position in the binding site throughout. Visualization of the four 40 ns runs shows that for three of the runs the ligand remains in the D-1 site and that during the other it leaves the site after 11 ns. For the simulation with a ligand in interior binding site 1, the average RMSD is 8.8 ± 2.8 Å, with the ligand initially moving away from its starting position before moving close to the D-1 site again for the last 70 ns. Visualization shows that although the ligand moves away from the binding site, it stays within the vicinity of D-1. During an extended simulation with the ligand in site D-2, the RMSD from the starting position was 19.7 ± 10.1 Å after 63 ns. Over the course of the trajectory, the ligand detached entirely from

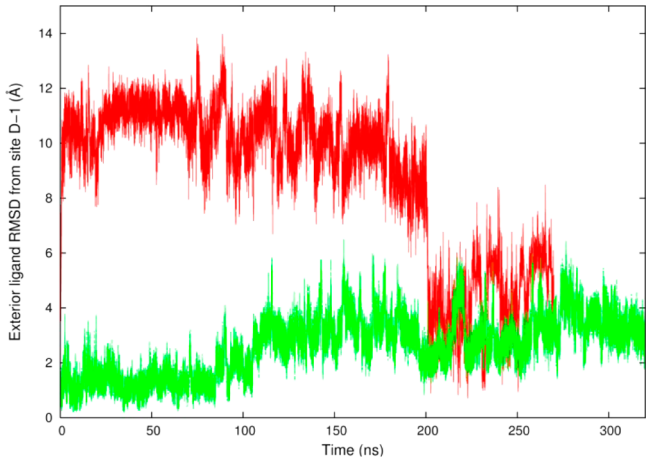


Figure 5. RMSD of the ligand from its initial position in site D-1. Red corresponds to cholic acid in site 1 and green to no interior ligand.

site D-2 and moved to site D-1. This simulation was repeated a further five times, with a production dynamics of 40 ns. During four of these runs, the ligand detached from the D-2 binding site and remains in the solvent, and for the fifth, the ligand leaves D-2 and goes into D-1.

Interaction between Gln51 and Ligands. Toke et al.[16] reported that the point mutation Gln51Ala disrupts the site selectivity mechanism of human ILBP. Gln51 is located in the center of β-strand C, and its location makes it a useful reference point within the binding cavity, to measure how deep the ligands are within the binding pocket. The distance between the center of mass (COM) of each ligand and Gln51 was calculated

dx.doi.org/10.1021/bi400192w | Biochemistry 2013, 52, 4723−4733

every 2 ps and plotted as a histogram with a bin size of 0.1 Å (site 2 is shown in Figure 6 as an example). The averages of

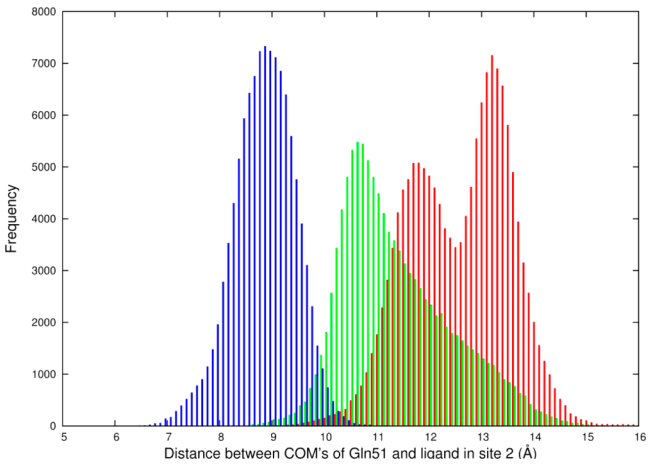


Figure 6. Histogram of the distance between the COM of the ligand in site 2 and Gln51 for the double (red) CA, mixed CA (green), and mixed CDA (blue) trajectories.

these distributions were considered in terms of the mode of distribution. Part of the motivation for calculating the mode of distribution, as opposed to the mean, is that a distribution is more suitable for characterizing an ensemble of conformations.

When chenodeoxycholic acid is in site 1 and cholic acid in site 2, the mode of the distance between the COMs is 8.1 Å; when cholic acid is in both sites, the mode is 8.6 Å. For the trajectory with cholic acid in site 1 and chenodeoxycholic acid in site 2, there are peaks at 8.6 Å and 11.3 Å. The distribution of the distances between the COMs as a function of time (data not shown) indicates that the first peak corresponds to the ligand's position for the first ~70 ns, after which the ligand moves away further from Gln51 for the remainder of the simulation, corresponding to the second peak. During the mixed CDA trajectory, the ligand in site 2 (cholic acid) is closer to Gln51, and deeper in the binding pocket than the ligand in site 2 for the double CA trajectory and the mixed CDA trajectory. For the mixed CDA simulation, the average distance (the mode of the distribution) between the COMs of the ligand and Gln51 is 8.8 Å, for mixed CA it is 10.6 Å, and for double CA there are two peaks at 11.8 Å and 13.2 Å.

Hydrogen bonding between the side chain of Gln51 and the hydroxyl groups of the steroid ring region of the ligands was identified using the CHARMM Hbond analysis module with a 2.8 Å cutoff for the hydrogen-acceptor distance, no cutoff for the bond angle, and sampling structures every 4 ps. During the double CA trajectory, a hydrogen bond is present between H_7 (the hydrogen in the OH group at steroid ring position 7) and O_ε (the oxygen in the Gln-51 side chain) for 3% of the trajectory frames. For the mixed CA trajectory, this hydrogen bond is present for 17% of the frames, but for the mixed CDA trajectory, the hydrogen bond is present for 50% of the frames. No other direct hydrogen bonds between the ligands or between the ligands and Gln51 are present.

Orientation of Ligands and Helices. The second α-helix is implicated in the mechanism of ligand binding for several members of the ILBP family.[41] To examine the position of the helices and ligands and to determine if the apo, single, and doubly complex proteins adopt similar conformations in this domain, three pairs of vectors were defined. To measure the angle of α-I with respect to the β-clam shell, a vector was defined along the length of the helix, from C_α-Glu9 (strand A) to C_α-Gly22 (end of α-I) and along the top of the β-clam shell from C_α-Glu9 to C_α-Thr78 (strand G). A pair of vectors from C_α-Val37 (strand B) to C_α-Ser25 (end of α-II) and C_α-Val37 to C_α-Gln72 (strand F) was used to measure the angle of α-II with respect to the β-clam shell. For the simulations with two ligands inside the binding cavity, vectors were defined along the length of the ligands from C_3 to C_{24} (i.e., a carbon in the first steroid ring to the tail carboxylate group). The angle between these three pairs of vectors was calculated every 2 ps and was plotted as histograms with a bin size of 1° (see Figure 7 for an

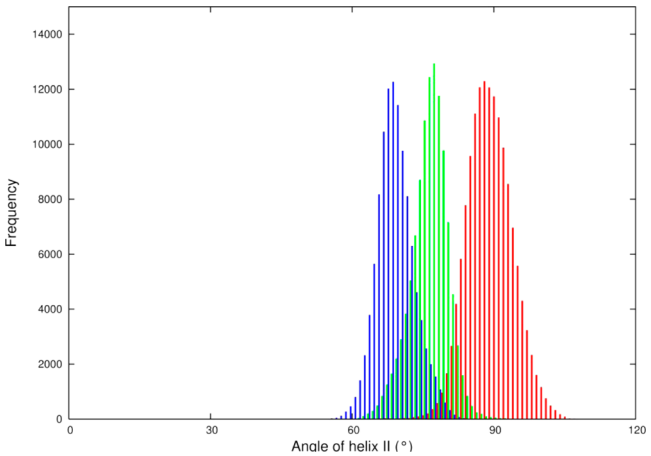


Figure 7. Histogram showing the angle of helix II with respect to the β-clam shell for the double CA (red), mixed CA (green), and mixed CDA (blue) simulations.

example). Table 4 reports the mean of the vectors, the standard deviation of the mean, and the mode of distribution. The significance of this result with respect to the mechanisms of cooperativity and site selectivity is discussed in the next section.

DISCUSSION

Human ILBP Has Three Binding Sites. The ESI-MS results show that human ILBP binds bile salts with a 3:1 ligand/protein ratio. This is in agreement with variants from other species, such as rabbit ILBP.[21] Docking to conformations from the MD trajectories identified three possible sites for the third ligand: site D-1, site D-2 both on the exterior of the protein, or inside the binding cavity between the interior ligands. During the MD simulations with the docked structures as starting structures, the ligand in D-2 dissociated from its initial position and for two of the production runs migrated toward the D-1 binding site. This suggests that the binding affinity predicted by AutoDock Vina for site D-2 was partly based on a steric interaction that is not maintained when the equilibrium conformational dynamics of the protein are considered. During four of the six simulations with the ligand starting in site D-1, the ligand remains in the binding site, including the extended production run of 320 ns. Site D-1 also corresponds to site 4 or site 153 of the X-ray crystallographic structure of zebrafish ILBP.[15] A ligand is resolved in this location in all models of both unit cells and is the only exterior ligand consistently identified in the same place. AutoDock Vina predicts four structures with three ligands inside the binding cavity due to the strong hydrophobicity and lack of steric clashes. However, the five X-ray crystallographic structures of zebrafish ILBP[15]

Table 4. Average Values of the Angles, and Their Standard Deviations, in Degrees, Defined to Examine the Orientation of the α-Helices with Respect to the β-Clam Shell and the Angle between the Ligands in Sites 1 and 2

vector	α-I w.r.t β-clam shell			α-II w.r.t β-clam shell			site 1 w.r.t site 2		
trajectory	mean	std dev	mode	mean	std dev	mode	mean	std dev	mode
apo	34.7	2.7	34	56.6	3.6	55			
single CA	41.5	4.8	43	75.8	7.1	78			
double CA	49.1	6.6	46	89.6	4.8	88	92.6	24.8	88
mixed CA	44.7	3.8	44	76.5	3.9	77	73.8	18.0	71
mixed CDA	40.1	3.2	40	69	3.9	68	54.5	11.0	56

indicate only two ligands in the interior and several on the exterior, including the D-1 position. Taken together, it is more probable that the third binding site is D-1 and not between the interior ligands.

D-1 Is an Allosteric Binding Site That Induces ILBP to Move from a Closed State to an Open State. Further evidence that D-1 is the third binding site is that our MD simulations show that binding to D-1 results in an allosteric interaction. The protein exhibits notable variation in the tertiary structure during the different simulations. The angle between the helices and the β-clam shell is smallest during the apoprotein simulation (α-I 34.7 ± 2.7° and α-II 56.6 ± 3.6°). When a single ligand is present, the helices move away from the β-clam shell to an angle of 41.5 ± 4.8° for α-I and 75.8 ± 7.1° for α-II. Figure 8 shows the angles of the helices with respect to

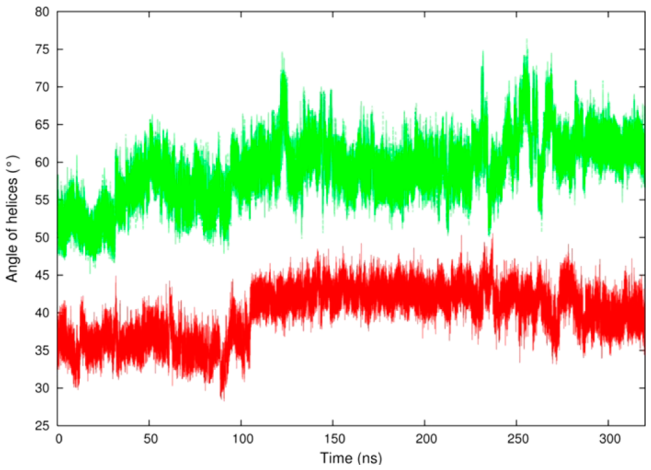


Figure 8. Orientation of α-I with respect to the β-clam shell (as defined in the Results section) as a function of time when a ligand is bound to exterior site D-1 and the interior sites are empty (red), and a ligand is bound to site 1 (green).

the β-clam shell during the simulations of the protein with cholic acid docked to the exterior binding site D-1 and no ligands inside the binding cavity. The angle of α-I increases from ∼36° to ∼41°, and the angle of α-II increases from ∼52° to ∼62°. During the simulation of the apoprotein structure, the helices do not show this increase in angle with respect to the β-clam shell (data not shown). This suggests that a ligand associated with binding site D-1 induces changes in the protein conformation to match the conformation when ligands are bound to the protein interior.

A recent study of the internal backbone motions of IBLP using NMR spectroscopy[22] supports the hypothesis that there is an allosteric mechanism of ligand binding and that the ensemble of protein states shifts from favoring a closed conformation to favoring an open conformation before interior ligand binding.[42] The simulation suggests that the hydrophobic adhesion of a ligand to the exterior of the protein into site D-1 initiates such a transition to an open state, where the α-helices move away from the β-clam shell to allow ligands into the binding cavity. A proposed external binding site does not contradict the gain of function study by Zanzoni et al.,[17] where an enthalpic mechanism of positive cooperative binding is introduced to the interior ligands of the chicken variant of ILBP by a double mutation. ITC data,[43] where binding constants for human ILBP have been calculated by fitting to a three step binding model, indicate that for taurocholic acid there is cooperative binding between both the first and second binding events and the second and third events.

Proposed Mechanism of Site Selectivity. Although no complete NMR or X-ray crystallographic structure of ILBP bound to both a conjugate of cholic acid and chenodeoxycholic acid is available, Toke et al.[16] propose that GCDA is bound to site 1 and GCA to site 2, based on homology modeling of porcine ILBP[19] and preliminary NMR results. A covariance analysis of the mixed CA and mixed CDA simulations (Figures S1d and S1e in Supporting Information) also supports this proposal. During the mixed CA simulation, the helices move away from the front region of the β-clam shell. The simulations with mixed CDA do not show this type of internal motion. This comparative structural stability indicates that chenodeoxycholic acid is the preferred ligand in site 1 and that cholic acid is the preferred ligand in site 2.

The increased stability of the mixed CDA simulation compared with the mixed CA simulation, together with differences in secondary structure, lead us to suggest a mechanism of site selectivity. Chenodeoxycholic acid bile salts have one less OH group than cholic acid and are therefore more hydrophobic. When chenodeoxycholic acid is in binding site 1, this increased hydrophobicity leads it to sit deeper in the binding pocket. This is illustrated by the distribution of distances between Gln51 and the site 1 ligand, which has a mode of 8.1 Å for mixed CDA and had two peaks at 8.6 Å and 11.3 Å for mixed CA. When chenodeoxycholic acid is in site 1, it is anchored in the binding pocket by a hydrogen bond between the OH group on carbon 7 and the side chain of Gln51. This hydrogen bond is present for 50% of the simulation when chenodeoxycholic acid is in site 1, but only 17% of the time during mixed CA simulation and for just 3% of the double CA trajectory. With the bile salt in site 1 deeper in the binding pocket, this increases the hydrophobicity of the binding cavity and encourages the ligand in site 2 to be deeper in the pocket also, as shown in Figure 6, where the mode of the distance between the ligands in site 2 and Gln51 is 8.8 Å for mixed CDA (with cholic acid in site 2) compared with 10.6 Å for mixed CA (with chenodeoxycholic acid in site 2) and two peaks for double CA at 11.8 Å and 13.2 Å. The mixed CDA configuration of ligands induces the protein to change to the

dx.doi.org/10.1021/bi400192w | Biochemistry 2013, 52, 4723−4733

closed conformation where the helices are closer to the β-clam shell compared with that of mixed CA and double CA (Table 4), and the angle between the vectors along the bile salts is smaller for the mixed CDA trajectory than the double CA or mixed CA. This proposed mechanism of site selectivity is energetic rather than steric, in agreement with the mechanism identified by Tochtrop et al.[5] using NMR methods.

Conclusions. In summary, using ESI-MS we show that human ILBP has a 3:1 ligand/protein binding ratio, and with a combined computational approach of docking and MD simulation, we identify a highly probable location for the third binding site that is involved in the allosteric mechanism of cooperative binding. An initial binding event to the exterior binding site D-1, located between β-strands F and G and the loop region connecting the two α-helices, induces the protein to change from a closed structure, to an open one, characterized by a movement of ~10° of α-II. This open structure then adopts a conformation that allows the second and third ligands to bind to the interior of the protein's β-clam shell. Changes between an open and closed structure also support a mechanism of site selectivity, where the higher hydrophobicity of chenodeoxycholic acid leads it to sit deeper in the binding cavity while in site 1 inducing the cholic acid in site 2 to sit deeper in the cavity. The helices then move closer to the β-clam shell, returning to the closed conformation, preventing further ligand exchange.

■ ASSOCIATED CONTENT

Supporting Information

Covariance analysis of the MD trajectories. This material is available free of charge via the Internet at http://pubs.acs.org.

■ AUTHOR INFORMATION

Corresponding Author

*Tel: +44 115 9513478. Fax: +44 115 9513562. E-mail: jonathan.hirst@nottingham.ac.uk.

Notes

The authors declare no competing financial interest.

■ ACKNOWLEDGMENTS

We are grateful for access to the University of Nottingham High Performance Computing facility. We thank Neil Oldham for useful discussions. E.R.T. is grateful for funding from an EPSRC Doctoral Prize.

■ REFERENCES

(1) Monod, J., Changeux, J. P., and Jacob, F. (1963) Allosteric proteins and cellular control systems. *J. Mol. Biol. 6*, 306−329.

(2) Changeux, J.-P. (2012) Allostery and the Monod-Wyman-Changeux model after 50 years. *Annu. Rev. Biophys. 41*, 103−133.

(3) Cui, Q., and Karplus, M. (2008) Allostery and cooperativity revisited. *Protein Sci. 17*, 1295−1307.

(4) Tochtrop, G. P., Bruns, J. L., Tang, C. G., Covey, D. F., and Cistola, D. P. (2003) Steroid ring hydroxylation patterns govern cooperativity in human bile acid binding protein. *Biochemistry 42*, 11561−11567.

(5) Tochtrop, G. P., DeKoster, G. T., Covey, D. F., and Cistola, D. P. (2004) A single hydroxyl group governs ligand site selectivity in human ileal bile acid binding protein. *J. Am. Chem. Soc. 126*, 11024−11029.

(6) Fischer, E. (1894) Einfluss der configuration auf die wirkung der enzyme. *Ber. Dtsch. Chem. Ges. 27*, 2985−2993.

(7) Koshland, D. E. (1958) Application of a theory of enzyme specificity to protein synthesis. *Proc. Natl. Acad. Sci. U.S.A. 44*, 98−104.

(8) Monod, J., Wyman, J., and Changeux, J.-P. (1965) On the nature of allosteric transitions: A plausible model. *J. Mol. Biol. 12*, 88−118.

(9) Tsai, C. J., del Sol, A., and Nussinov, R. (2009) Protein allostery, signal transmission and dynamics: a classification scheme of allosteric mechanisms. *Mol. Biosyst. 5*, 207−216.

(10) Adcock, S. A., and McCammon, J. A. (2006) Molecular dynamics: survey of methods for simulating the activity of proteins. *Chem. Rev. 106*, 1589−1615.

(11) Karplus, M., and McCammon, J. A. (2002) Molecular dynamics simulations of biomolecules. *Nat. Struct. Biol. 9*, 646−652.

(12) Ma, J., Sigler, P. B., Xu, Z., and Karplus, M. (2000) A dynamic model for the allosteric mechanism of GroEL. *J. Mol. Biol. 302*, 303−313.

(13) Jiao, W., Hutton, R. D., Cross, P. J., Jameson, G. B., and Parker, E. J. (2012) Dynamic cross-talk among remote binding sites: the molecular basis for unusual synergistic allostery. *J. Mol. Biol. 415*, 716−726.

(14) Hanhoff, T., Lücke, C., and Spener, F. (2002) Insights into binding of fatty acids by fatty acid binding proteins. *Mol. Cell. Biochem. 239*, 45−54.

(15) Capaldi, S., Saccomani, G., Fessas, D., Signorelli, M., Perduca, M., and Monaco, H. L. (2009) The X-ray structure of zebrafish (Danio rerio) ileal bile acid-binding protein reveals the presence of binding sites on the surface of the protein molecule. *J. Mol. Biol. 385*, 99−116.

(16) Toke, O., Monsey, J. D., DeKoster, G. T., Tochtrop, G. P., Tang, C. G., and Cistola, D. P. (2006) Determinants of cooperativity and site selectivity in human ileal bile acid binding protein. *Biochemistry 45*, 727−737.

(17) Zanzoni, S., Assfalg, M., Giorgetti, A., D'Onofrio, M., and Molinari, H. (2011) Structrual requirements for cooperativity in ileal bile acid binding proteins. *J. Biol. Chem. 286*, 39307−39317.

(18) Guariento, M., Assfalg, M., Zanzoni, S., Fessas, D., Longhi, R., and Molinari, H. (2009) Chicken ileal bile-acid-binding protein: a promising target of investigation to understand binding co-operativity across the protein family. *Biochem. J. 425*, 413−424.

(19) Lucke, C., Zhang, F. L., Hamilton, J. A., Sacchettini, J. C., and Ruterjans, H. (2000) Solution structure of ileal lipid binding protein in complex with glycocholate. *Eur. J. Biochem. 267*, 2929−2938.

(20) Kouvatsos, N. (2006) Characterisation of Rabbit Ileal Lipid Binding Protein and Design of New Beta-Scaffold Proteins, Ph.D. Thesis, University of Nottingham, Nottingham, U.K.

(21) Kouvatsos, N., Thurston, V., Ball, K., Oldham, N. J., Thomas, N. R., and Searle, M. S. (2007) Bile acid interactions with rabbit ileal lipid binding protein and an engineered helixless variant reveal novel ligand binding properties of a versatile beta-clam shell protein scaffold. *J. Mol. Biol. 371*, 1365−1377.

(22) Horvath, G., Kiraly, P., Tarkanyi, G., and Toke, O. (2012) Internal motions and exchange processes in human ileal bile acid binding protein as studied by backbone ^{15}N nuclear magnetic resonance spectroscopy. *Biochemistry 51*, 1848−1861.

(23) Kurz, M., Brachvogel, V., Matter, H., Stengelin, S., Thuring, H., and Kramer, W. (2003) Insight into the bile acid transportation system: The human ileal lipid-binding protein-cholyltaurine complex and its comparison with homologous structures. *Proteins 50*, 312−328.

(24) Brooks, B. R., Bruccoleri, R. E., Olafson, D. J., States, D. J., Swaminathan, S., and Karplus, M. (1983) CHARMM: A program for macromolecular energy, minimization, and dynamics calculations. *J. Comput. Chem. 4*, 187−217.

(25) Brooks, B. R., Brooks, C. L., Mackerell, A. D., Jr., Nilsson, L., Petrella, R. J., Roux, B., Won, Y., Archontis, G., Bartels, C., Boresch, S., Caflisch, A., Caves, L., Cui, Q., Dinner, A. R., Feig, M., Fischer, S., Gao, J., Hodoscek, M., Im, W., Kuczera, K., Lazaridis, T., Ma, J., Ovchinnikov, V., Paci, E., Pastor, R. W., Post, C. B., Pu, J. Z., Schaefer, M., Tidor, B., Venable, R. M., Woodcock, H. L., Wu, X., Yang, W., York, D. M., and Karplus, M. (2009) CHARMM: The biomolecular simulation program. *J. Comput. Chem. 30*, 1545−1614.

(26) Humphrey, W., Dalke, A., and Schulten, K. (1996) VMD: Visual molecular dynamics. *J. Mol. Graphics 14*, 33−38.

(27) Larkin, M. A., Blackshields, G., Brown, N. P., Chenna, R., McGettingan, P. A., McWilliam, H., Valentin, F., Wallace, I. M., Wilm, A., Lopez, R., Thompson, J. D., Gibson, T. J., and Higgins, D. G. (2007) ClustalW and ClustalX version 2. *Bioinformatics 23*, 2947−2984.
(28) Schleif, R. (2006) *Analysis of Protein Structure and Function: A Beginner's Guide to CHARMM*, Johns Hopkins University, Baltimore, MD.
(29) Jorgensen, W. L., Chandrasekhar, J., Madura, J. D., Impey, R. W., and Klein, M. L. (1983) Comparison of simple potential functions for simulating liquid water. *J. Chem. Phys. 79*, 926−936.
(30) Phillips, J. C., Braun, R., Wang, W., Gumbart, J., Tajkhorshid, E., Chipot, C., Skeel, R. D., Kale, L., and Schulten, K. (2005) Scalable molecular dynamics with NAMD. *J. Comput. Chem. 26*, 1781−1802.
(31) Vanommeslaeghe, K., Hatcher, E., Acharya, C., Kundu, S., Zhong, S., Shim, J., Darian, E., Guvench, O., Lopes, P., Vorobyov, I., and A. D., M., Jr. (2009) CHARMM general force field: A force field for drug-like molecules compatible with the CHARMM all-atom additive biological force fields. *J. Comput. Chem. 31*, 671−690.
(32) Ryckaert, J. P., Ciccotti, G., and Berendsen, H. J. C. (1977) Numerical integration of the Cartesian equations of motion of a system with constraints: molecular dynamics of n-alkanes. *J. Comput. Phys. 23*, 327−341.
(33) Darden, T. A., York, D., and Pedersen, L. (1993) Particle-mesh Ewald: An N.log(N) method for Ewald sums in large systems. *J. Chem. Phys. 98*, 10089−10092.
(34) Trott, O., and Olson, A. J. (2010) AutoDock Vina: Improving the speed and accuracy of docking with a new scoring function, efficient optimization and multithreading. *J. Comput. Chem. 31*, 455−461.
(35) Plewczynski, D., Lazniewski, M., Augustyniak, R., and Ginalski, K. (2011) Can we trust docking results? Evaluation of seven commonly used programs on PDBbind database. *J. Comput. Chem. 32*, 742−755.
(36) Morris, G. M., Goodsell, D. S., Halliday, R. S., Huey, R., Hart, W. E., Belew, R. K., and Olson, A. J. (1998) Automated docking using a Lamarckian genetic algorithm and an empirical binding free energy. *J. Comput. Chem. 19*, 1639−1662.
(37) Wang, R., Fang, X., Lu, Y., and Wang, S. (2004) The PDBbind database: Collection of binding affinities for protein−ligand complexes with known three-dimensional structures. *J. Med. Chem. 47*, 2977−2980.
(38) Sanner, M. F. (1999) Python: A programming language for software integration and development. *J. Mol. Graphics Modell. 17*, 57−61.
(39) Sun, J., Kitova, E. N., Wang, W., and Klassen, J. S. (2006) Method for distinguishing specific from nonspecific protein-ligand complexes in nanoelectrospray ionization mass spectrometry. *Anal. Chem. 78*, 3010−3018.
(40) Wallace, A. C., Laskowski, R. A., and Thornton, J. M. (1995) LIGPLOT: A program to generate schematic diagrams of protein-ligand interactions. *Protein Eng. 8*, 127−134.
(41) Eberini, I., Rocco, A. G., Ientile, A. R., Baptista, A. M., Gianazza, E., Tomaselli, S., Molinari, H., and Ragona, L. (2008) Conformational and dynamics changes induced by bile acids binding to chicken liver bile acid binding protein. *Proteins 71*, 1889−1898.
(42) Toke, O., Monsey, J. D., and Cistola, D. P. (2007) Kinetic mechanism of ligand binding in human ileal bile acid binding protein as determined by stopped-flow fluorescence analysis. *Biochemistry 46*, 5427−5436.
(43) Fang, H.-J. (2011) An Investigation of Ileal Bile Acid Binding Protein and Its Application as a Biosensor for Performance Enhancing Drugs, Ph.D. Thesis, University of Nottingham, Nottingham, U.K.