A Single Amino Acid Mutation in Zebrafish (Danio rerio) Liver Bile Acid-binding Protein Can Change the Stoichiometry of Ligand Binding*

Received for publication, July 2, 2007, and in revised form, July 23, 2007 Published, JBC Papers in Press, August 1, 2007, DOI 10.1074/jbc.M705399200

Stefano Capaldi‡, Mara Guariento†, Gianmaria Saccomani§, Dimitrios Fessas†, Massimiliano Perduca‡, and Hugo L. Monaco†

From the ‡Biocrystallography Laboratory, Department of Science and Technology, University of Verona, Ca Vignal 1, Strada Le Grazie 15, 37134 Verona, Italy and the §Department of Food Science, Technology, and Microbiology, University of Milan, Via Celoria 2, 20133 Milan, Italy

In all of the liver bile acid-binding proteins (L-BABPs) studied so far, it has been found that the stoichiometry of binding is of two cholate molecules per internal binding site. In this paper, we describe the expression, purification, crystallization, and three-dimensional structure determination of zebrafish (Danio rerio) L-BABP to 1.5 Å resolution, which is currently the highest available for a protein of this family. Since we have found that in zebrafish, the stoichiometry of binding in the protein cavity is of only one cholate molecule per wild type L-BABP, we examined the role of two crucial amino acids present in the binding site. Using site-directed mutagenesis, we have prepared, crystallized, and determined the three-dimensional structure of co-crystals of two mutants. The mutant G55R has the same stoichiometry of binding as the wild type protein, whereas the C91T mutant changes the stoichiometry of binding from one to two ligand molecules in the cavity and therefore appears to be more similar to the other members of the L-BABP family. Based on the presence or absence of a single disulfide bridge, it can be postulated that fish should bind a single cholate molecule, whereas amphibians and higher vertebrates should bind two. Isothermal titration calorimetry has also revealed the presence in the wild type protein and the G55R mutant of an additional binding site, different from the first and probably located on the surface of the molecule.

The liver bile acid-binding proteins, formerly called liver “basic” fatty acid-binding proteins (FABPs),² belong to the conserved multigene family of the fatty acid-binding proteins (1–6). Originally, the different members of this family were named according to the tissue from which they were first isolated. More recently, an alternative nomenclature has been proposed, since it is well known that different FABP types can be present in the same tissue (7). In particular, in the liver of several vertebrates, two paralogous groups of FABPs have been described: the liver FABPs and another group that was initially called the liver “basic” FABPs and is currently referred to as the liver bile acid-binding proteins (L-BABPs) (8, 9). The reason for this unusual nomenclature was that the isoelectric point of the first protein of this group that was described, chicken liver BABP, is 9.0, and although it was evident that this protein was different from the canonical liver FABP, its specific ligand was unknown (10). When it was shown that this protein could specifically bind cholic acid, a change in its designation was suggested to avoid confusion (11). The cytosolic bile acid-binding proteins mediate the intracellular movement of bile acids to the membranes across which they exit or enter the cells via membrane-bound transporters (12). The L-BABPs have been shown to be present in birds, amphibians, reptiles, and fish but not in mammals that express the very similar ileal BABPs, formerly called gastrotropins. It is suspected, although not proved, that the function that they perform in the liver of the species in which they have been found may be carried out in mammals by the other paralogous protein in the liver, L-FABP. X-ray crystallography and NMR spectroscopy have been used to experimentally determine the three-dimensional structure of many members of the FABP family both in their apo form and complexed with lipophilic ligands. All of the members of the family share the same fold, a 10-stranded β barrel in which two short helices are inserted between the first and the second strand of antiparallel β sheet.

The first L-BABP three-dimensional structure determined by x-ray diffraction was that of chicken liver BABP (13), followed by toad (14) and axolotl L-BABP (15). NMR and other biophysical techniques have also been used to study chicken L-BABP, which remains the best characterized member of the L-BABP family (16–18). The most important result of these structural studies is that the stoichiometry of binding for cholate is of two ligands per protein molecule (11, 15), whereas that for oleate is of only one molecule per binding site (15, 17). It is an unusual characteristic of the paralogous L-FABPs that they bind two molecules of oleate in their binding sites (19, 20). Another important aspect described is the conformational
changes that take place upon ligand binding that have been studied in chicken L-BABP by both x-ray diffraction (11) and NMR spectroscopy that has highlighted the important role of a buried histidine (18).

Although L-BABPs have been studied in catfish (Rhantia sapo) (21), lungfish (Lepidosiren paradoxa) (22), Japanese sea perch (Lateolabrax japonicus) (23), gilthead sea bream (Sparus aurata) (24), salmon (Salmo salar L.) (25), shark (Halaieurus bivius) (26), and zebrafish (Danio rerio) (27), there is currently no three-dimensional structure available of an L-BABP belonging to this vertebrate group. The zebrafish (Danio rerio) is a powerful animal model that is widely used to study vertebrate development. The presence in this species of the gene coding for liver BABP has been reported, and the cDNA clone coding for the protein has been isolated (27).

In this paper, we report the expression, purification, crystallization, and three-dimensional structure determination of zebrafish liver BABP. We have prepared and examined co-crystals with cholate of the wild type protein and two mutants. The wild type protein and one of the mutants bind one cholate molecule per binding site, whereas the other mutant has the stoichiometry found in chicken and axolotl L-BABP of two ligands per binding site. Isothermal titration calorimetry has yielded results that are consistent with the x-ray data.

**MATERIALS AND METHODS**

**Protein Expression, Purification, Complex Formation, and Crystallization**—The cDNA coding for zebrafish liver BABP (IMAGE ID 5410136), obtained from RZPD (Deutsches Res sourcenzentrum fuer Genomforschung GmbH), was amplified by PCR using primers designed to introduce restriction sites for BamHI and HindIII endonucleases and a sequence coding for a digestion site for thrombin in the C-terminal end in the amplified fragment. After purification, the fragment and the expression vector pQE50 (Qiagen) were digested with the restriction enzymes mentioned above and incubated with ligase to insert the cDNA in the vector respecting the reading frame. SG13009 E. coli cells were transformed with the resulting vector and grown at 37 °C, and protein synthesis was induced overnight at 20 °C with 0.5 mM isopropyl β-d-thiogalactopyranoside. Under these conditions of subcloning in pQE50, the expressed intracellular domain is fused to a histidine tag through its C-terminus. The presence of the tag allowed the affinity purification of the fused protein by passing the bacterial extracts through a nickel-Sepharose column. The column was equilibrated with 20 mM Tris-HCl, pH 7.5, 0.5 M NaCl, 10 mM imidazole, and 0.02% NaN₃, and the bound protein was eluted with a linear gradient of imidazole from 10 to 500 mM. After the affinity column, the tag was removed by thrombin digestion, and the protein was further purified by gel filtration in a Superdex G-75 column equilibrated with 20 mM Tris-HCl, pH 7.5, 0.15 M NaCl, and 0.02% NaN₃. Complete removal of the tag was assessed by Western blot analysis using an anti-His-horseradish peroxidase-conjugated antibody (Sigma). The purified protein showed one band in SDS-PAGE. Five times the molar protein concentration of sodium cholate was added to the apoprotein at a concentration of about 30 mg/ml in 20 mM Tris-HCl buffer, pH 7.5, in order to prepare the complex with the ligand. The solution was stirred overnight at 20 °C and used at this concentration for the initial screen of crystallization conditions. Hampton Research Screens were used at 20 °C with the hanging drop method, mixing 1 μl of the protein solution with the same volume of the precipitating solution and equilibrating versus a volume of 0.3 ml of the latter in the reservoir. The conditions yielding very small crystals were later refined, and the sitting drop method with larger volumes was also tested until crystals that were large enough for data collection were obtained. The best crystals of the wild type z-L-BABP cholate complex grow by mixing equal volumes of the protein solution and 0.1 M sodium citrate, pH 5.6, 20% polyethylene glycol 4000, and 20% isopropyl alcohol. They are thin rods of about 2.5 mm × 30 μm × 30 μm and grow to their full size in very few days at 20 °C.

The mutants were generated using the commercial QuickChange kit (Stratagene) and primers designed to introduce the proper mutations. The presence of the desired mutations was confirmed by plasmid sequencing. Protein expression and purification followed the protocol used for the wild type protein. The co-crystals of both mutants were grown using a protocol identical to that of the wild type protein. In the case of the C91T mutant, the precipitant was a solution of 0.1 M sodium acetate, pH 4.6, 0.2 M ammonium acetate, and 30% polyethylene glycol 4000.

**Data Collection, Structure Solution, and Refinement**—The co-crystals of wild type z-L-BABP and cholate are tritonal, space group P3₂₁, with a = b = 43.4 Å and c = 67.5 Å and contain one molecule in the asymmetric unit (see Table 1). The co-crystals of the G55R mutant with cholate are isomorphous with this form, whereas those of the C91T mutant are monoclinic, space group P2₁, with a = 28.0 Å, b = 63.3 Å, c = 35.7 Å, and β = 105.5° (Table 1).

The data for the co-crystals with cholate of the wild type and C91T mutant of the protein were collected at the ID29 beam line of the European Synchrotron Radiation Facility in Grenoble (λ = 0.98 Å) at 100 K after a brief soaking in a mixture of 80% mother liquor and 20% glycerol. The data were indexed, integrated, and reduced using the programs MOSFLM and Scala (28, 29). The data for the co-crystals of the G55R mutant were collected on the home source at 100 K after a brief soaking in a mixture of 80% mother liquor and 20% glycerol. The detector was a MarResearch imaging plate mounted on a Rigaku RU-300 rotating anode x-ray generator. The source was operated at 50 kV and 100 mA, and monochromatic copper Kα radiation was obtained using Xenocs mirrors. The data were indexed, integrated, and reduced using the program AUTOMAR. The diffraction data statistics are summarized in Table 1.

The structure of the complex wild type zebrafish L-BABP-cholate was solved using the CCP4 suite of programs for crystallographic computing. The initial phases were calculated by the molecular replacement method as implemented in the program AMoRe (30), with the coordinates of chicken L-BABP (Protein data bank accession code 1TW4) as the search probe. When the rotation function was calculated with the data in the 8.0–3.0 Å resolution range, the two highest peaks had correlation coefficients of 20.2 and 11.5. The translation function, cal-
The crystals of the G55R mutant are isomorphous to the wild type liver BABP complexed with cholate, the highest resolution shells used in the refinements are: 1.54–1.50 Å for the wild type protein, 1.95–1.90 Å for the G55R mutant, and 1.58–1.50 Å for the C91T mutant. The three values listed in the first two columns as ligand atoms correspond to cholate, isopropyl alcohol, and glycerol, respectively. The average B factors are listed in the same order.

The values in parentheses refer to the highest resolution shells. For the wild type liver BABP complexed with cholate, the highest resolution interval is 1.58–1.50 Å, for the G55R mutant it is 1.94–1.87 Å, and for the C91T mutant it is 1.58–1.50 Å. The highest resolution shells used in the refinements are: 1.54–1.50 Å for the wild type protein, 1.95–1.90 Å for the G55R mutant, and 1.58–1.50 Å for the C91T mutant. The three values listed in the first two columns as ligand atoms correspond to cholate, isopropyl alcohol, and glycerol, respectively. The average B factors are listed in the same order.

The final refinement statistics for the models of the three crystal forms are summarized in Table 1.

Isothermal Titration Calorimetry (ITC)—The proteins and sodium cholate were dissolved in the following buffer: 20 mM Tris-HCl, pH 7.5, 0.15 mM NaCl, and 0.02% sodium azide. The sodium cholate concentrations in the titrating solution and the initial protein concentration in the measurement cell were 0.62 and 0.026 mM, respectively, for the wild type protein, 0.90 and 0.028 mM, respectively, for the G55R mutant, and 1.0 and 0.0136 mM, respectively, for the C91T mutant. The titrations were performed at 25 °C using a CSC 4200 isothermal titration calorimeter. A total of 25 injections of 10-μl aliquots of titrating solution were added to the 1.3-ml protein solution cell. The heat of the injections was corrected for the heat of dilution of the ligand into the buffer. Several binding models were tested to interpret the calorimetric data, and the fitting functions are described in the literature (34–36).

Briefly, the observable enthalpy is given by Equation 1,

\[
\Delta H(T,p,\mu_L) = -R [\delta \ln Q / \delta (1/T)]_{p,\mu_L} \tag{Eq. 1}
\]

and the degree of association (i.e. the concentration ratio \(\bar{x}\) = [bound ligand]/[total protein]) is given by Equation 2,

\[
\bar{x} = RT [\delta \ln Q / \delta \mu_L]_{T,p} = RT [\delta \ln Q / \delta \ln L]_{T,p} \tag{Eq. 2}
\]
where \( R \) is the universal gas constant, \( \mu_L \) is the chemical potential of the free ligand, \([L]\) is the concentration of the free ligand, and \( Q \) is the partition function of the system referred to the free protein state (35). Since we can approximate these systems as diluted solutions, the thermodynamic activities of the solutes may be replaced with their molar concentrations. Under this assumption, the partition function (Equation 3),

\[
Q = \sum_{j=0}^{n} \frac{[P_j]}{[P_0]}
\]

(Eq. 3)

is the sum of the concentrations of all protein species, \( P_j \), referred
Zebrafish Liver BABP

TABLE 2
Distances between the closest z-L-BABP residues and the cholate molecule bound in the wild type protein and the two ligands in the co-crystals of the C91T mutant

The co-crystal bound to the wild type protein is labeled 130, the second molecule bound only to the mutant, 131. This notation follows that used for chicken L-BABP.

| Cholate molecule | Main contacts between the cholic acid molecules and z-L-BABP residues | Atom | B Apost } | Distance | Atom | B Apost } | Distance |
|-----------------|-------------------------------------------------------------------|------|------------|----------|------|------------|----------|
| 130 C18 | Tyr\textsuperscript{14} | OH | 3.55 | 130 C19 | Leu\textsuperscript{18} | CD2 | 3.87 | 130 O25 | Lys\textsuperscript{56} | NZ | 3.39 | 130 O26 | Lys\textsuperscript{56} | NZ | 2.96 | 130 O12 | Met\textsuperscript{73} | CG | 3.34 | 130 O26 | Met\textsuperscript{73} | O | 3.81 | 130 O26 | Asp\textsuperscript{74} | CB | 3.47 | 130 C4 | Phe\textsuperscript{96} | CZ | 3.92 | 130 O3 | His\textsuperscript{98} | ND1 | 3.55 | 130 C19 | Met\textsuperscript{118} | SD | 3.83 | 130 C1 | Arg\textsuperscript{120} | NH1 | 3.50 | 130 C2 | Arg\textsuperscript{120} | NH2 | 3.74 |

| Cholate molecule | Contacts between the two cholic acid molecules in the C91T mutant | Atom | Cholate molecule | Atom | Distance | Atom | Cholate molecule | Atom | Distance |
|-----------------|-------------------------------------------------|------|-----------------|------|----------|------|-----------------|------|----------|
| 130 O3 | 131 C31 | 3.88 | 130 O3 | 131 C16 | 3.84 | 130 C4 | 131 C17 | 3.81 |

TABLE 3
Distances between the single cholate bound in the wild type protein and the two cholates in the C91T mutant and solvent molecules present in the crystals

The notation is the same as in Table 2. IPA, isopropyl alcohol; GOL, glycerol.

| Cholate molecule | Main contacts between the cholic acid molecules and solvent molecules in the crystals | Atom | Solvent molecule in wild type protein | Atom | Distance | Atom | Solvent molecule in C91T mutant | Atom | Distance |
|-----------------|----------------------------------------------------------------------------------|------|--------------------------------------|------|----------|------|----------------------------------|------|----------|
| 130 O3 | IPA | 140 | O | 2.66 | 130 O3 | HOH | 1 | O | 2.77 | 130 O7 | HOH | 2 | O | 2.79 | 130 O7 | GOL | 145 | O | 2.78 | 130 O12 | HOH | 2 | O | 2.95 | 130 O12 | HOH | 3 | O | 2.85 | 130 O25 | HOH | 4 | O | 2.75 | 130 O26 | HOH | 5 | O | 2.91 |

| Cholate molecule | Contacts between the two cholic acid molecules in the C91T mutant | Atom | Cholate molecule | Atom | Distance | Atom | Cholate molecule | Atom | Distance |
|-----------------|-----------------------------------------------------------------|------|-----------------|------|----------|------|-----------------|------|----------|
| 130 O3 | 131 C13 | 2.80 | 130 O3 | 131 O12 | 3.96 | 130 C4 | 131 C16 | 3.84 | 130 C4 | 131 C17 | 3.81 |

protein atoms, one cholate, 3 isopropyl alcohol, one glycerol, and 135 water molecules. The conventional R factor is 18.8%, and the free R factor is 20.9% (Table 1). The R factors and r.m.s. of Table 1 were calculated with the program REFMAC (31). The stereochemical quality of the protein model was assessed with the program PROCHECK (33). 94.5% of the residues are in the most favorable region of the Ramachandran plot, and the remaining 5.5% are in the additionally allowed region. The overall fold consists of the canonical β barrel with 10 strands of anti-parallel β chain and the two α helices inserted between the first and the second strand. The final secondary structure assignments are, for the β strands, the following: strand A, residues 4–12; B, residues 37–43; C, residues 46–53; D, residues 56–63; E, residues 67–71; F, residues 77–85; G, residues 88–92; H, residues 97–103; I, residues 106–113; and J, residues 116–124. The two α helices span residues 14–20 and 25–30.

Fig. 1a is a schematic diagram of the zebrafish L-BABP molecule with the experimental electron density for the single cholate ligand present in the binding site. Resolution and quality of the maps are such that this stoichiometry of binding is completely unambiguous. This result is totally unexpected, since in the other two L-BABP species of this family studied by x-ray diffraction of co-crystals, chicken (11) and axolotl (15), the stoichiometry found was unequivocally of two ligand molecules per binding site. Furthermore, although there is currently no x-ray structure available for the closely related mammalian ileal BABP, the same stoichiometry of two ligands is supported by NMR experiments (38, 39). Fig. 1b represents schematically hydrogen bonds and hydrophobic interactions between the protein and the ligand. Using the program LSQKAB (40), the model of zebrafish L-BABP was superim-

to the free protein, \( P_c \), \( Q \) depends on the assumptions made on the association (binding) mechanism and is the key function used to simulate the enthalpy so as to check the model with the experimental data and to obtain the association (or binding) constant, \( K_c \), and the binding enthalpy \( \Delta H_b \). The binding constant is a dimensionless quantity by definition. However, in order to stress the approximation of the thermodynamic activities with the molar concentrations, the use of m\(^{-1}\) units for this parameter is widely diffused and will be adopted in this paper.

The fit based on the binding models was accomplished using the nonlinear Levenberg-Marquardt method (37). The errors of each fitting parameter were calculated with a 95.4% confidence limit by the Monte Carlo simulation method.

RESULTS AND DISCUSSION

Structure and Ligand-binding Stoichiometry of the Wild Type Protein—The final model of the wild type co-crystals of z-L-BABP corresponds to the full-length 125-amino acid chain, 972
posed to that of chicken L-BABP (11) (Protein Data Bank entry 1TW4). A stereo diagram of the two models with the ligands is represented in Fig. 1c. Interestingly, the single cholate molecule present in the zebrafish L-BABP binding site superimposes very well with one of the two molecules bound in the chicken liver protein, the one whose carboxylate is closer to the end of the second of the two α helices that are believed to control access to the binding cavity. Fig. 1d represents the r.m.s. deviation between equivalent α carbons of the two models as a function of the amino acid number. The three peaks where the largest differences are found correspond to residues 90–97, 101–105, and 111–117 (i.e. the last part of strand G, strands H and I, and the connection with strand J). In chicken L-BABP, there are four residues in this area that have been identified as interacting with the same cholate molecule, which is the one absent in the zebrafish protein: Thr91, Phe96, His98, and Gln100 (11). Three of these residues are located in the region between the first and the second peak, and the first is at the beginning of the first peak (i.e. where the differences between the two structures are very small). It is

FIGURE 2. Sequence comparison of zebrafish L-BABP and other members of this protein family. The sequences were aligned using the program ClustalW (42). A white dot indicates the residues involved in cholate binding in chicken L-BABP, and a black dot indicates those that bind to cholate in axolotl L-BABP. The red and green arrows identify the residues involved in ligand binding of the wild type and C91T mutant, respectively. The disulfide bridge in the wild type protein is shown in blue, and the mutated glycine is shown in red. The last column on the right-hand side at the bottom gives the percentage identity of each sequence and that of zebrafish L-BABP. Zf, zebrafish; To, toad; Ax, axolotl; Lz, lizard; Ch, chicken.

FIGURE 3. Cholate binding to the G55R mutant of zebrafish L-BABP. a, left, electron density in the proximity of residue 55 in wild type zebrafish L-BABP. Right, electron density in the proximity of residue 55 of the G55R mutant. 

b, stereo view of the Cα chain trace of the G55R mutant of z-BABP with its single cholate molecule bound. The figure represents the side chains of selected amino acids interacting with the ligand, the three isopropyl alcohol, the single glycerol, and three water molecules present in the binding site. The mutated arginine is represented as green.
also worth noticing that for one of them, His98, an active role, upon protonation and deprotonation, has been assigned in the dynamic behavior of the protein (18). Table 2 lists the main contacts between the cholate molecule (labeled 130, following the notation used for chicken L-BABP) and the protein side chains in contact with it.

The main nonhydrophobic contacts are established between the cholate O25 and O26 and the NZ of Lys56 and between O3 and His98. Important hydrophobic contacts involve Tyr14, Leu18, Ile21, Ile70, Met73, Phe96, and Met118 as shown in Fig. 1b. In general, the residues involved agree quite well with the equivalent ones identified in chicken and axolotl L-BABP, but a few small variations are also observed.

The solvent-accessible volume of the ligand-binding cavity of wild type zebrafish L-BABP, calculated with the program CAST (41), is 513.6 Å³ (i.e. somewhat smaller than the equivalent value for chicken L-BABP, 627.4 Å³ (11), and axolotl L-BABP, 685.5 Å³ (15)). All of these values are significantly larger that the value observed for the apo form of chicken L-BABP, 143.7 Å³ (11).

Since the resolution of this crystal form, the best obtained so far for this family, is adequate for the purpose (see Table 1), it was possible to examine the structure of the solvent molecules within the ligand binding cavity of the protein. The wild type crystal form contains eight very well defined solvent molecules in the cavity. They are listed in Table 3. An isopropyl alcohol molecule, a compound present in the crystallization solution of this crystal form, binds to the cholate O3, a glycerol molecule binds to O7, and five water molecules bind to the ligand oxygens indicated in the table. An identical situation is observed for the isomorphous G55R mutant, and a very similar one is observed in the C91T mutant in which the glycerol and isopropyl alcohol, which is not present in the mother liquor, are replaced by the second cholate and by a water molecule.

The G55R Mutant of z-L-BABP—An examination of the information summarized in Fig. 2 prompted us to design the G55R mutant of zebrafish L-BABP. The sequences of five L-BABPs are aligned in the figure, and the amino acids involved in ligand binding in the only two structures of co-crystals that were known, chicken (11) and axolotl (15), are indicated with white and black dots. Six amino acids participate in the binding in both species, and five (Tyr14, Thr53, Phe96, His98, and Gln100) are conserved in all of the sequences. There is more variability in the amino acid in position 55, an Asn, Lys, or Arg in four of the five species listed, but a strikingly different Gly in the case of zebrafish L-BABP. We therefore prepared the mutant, expressed the protein, and grew co-crystals in the presence of cholate. The crystals are isomorphous with the wild type co-
crystals (see Table 1). The model was refined to 1.9 Å resolution using data collected at the home source. The refinement statistics are listed in Table 1 and indicate that this model is also very good; its lower resolution can be ascribed to the fact that the data were collected on a conventional rather than a synchrotron source. A Ramachandran plot of this model reveals that 94.6% of the residues are in the most favorable region of the plot, and the remaining 5.4% are in the additionally allowed region. Although the presence of the mutation has been controlled by plasmid sequencing, it is worthwhile examining the electron density of the mutated residue. It is displayed in Fig. 3a, which shows very clearly the presence of an Arg in position 55. Fig. 3b is a stereo diagram of the mutant with the side chains of the amino acids interacting with the single cholate molecule bound. Side chains involved and solvent structure are virtually identical to those of the wild type protein. Note the positions of Arg55 (green) and Lys56 (blue). Clearly, the position of either of the two residues is adequate to counterbalance the charge of the carboxylate of the single cholate molecule, which appears to be the only role played by residues 55–56 in that area.

The C91T Mutant of z-L-BABP—Another amino acid found in the internal binding cavity, in contact with the cholate molecule absent in wild type z-L-BABP and present in chicken L-BABP, is Thr91. In the case of zebrafish L-BABP, this amino acid is substituted by a Cys, which forms a disulfide bridge with Cys80. We therefore decided to probe the role of this disulfide bridge on the stoichiometry of ligand binding. Fig. 4a displays the electron density of the only disulfide bridge present in the wild type protein (left) and the density after the mutation (right). The co-crystals of the C91T mutant are monoclinic, space group P21, with \( a = 28.01 \) Å, \( b = 63.28 \) Å, \( c = 35.67 \) Å, and \( \beta = 105.5^\circ \). Although the crystals grow under rather different conditions from the wild type protein, they diffract to approximately the same resolution. The data used to solve this structure were collected at the ID29 beam line of the European Synchrotron Radiation Facility in Grenoble. Data collection statistics are summarized in Table 1. The structure was solved by molecular replacement and refined as described above. The refinement statistics are given in Table 1. In this case, the Ramachandran plot has 94.7% of the residues in the most favorable region of the plot and the remaining 5.3% in the additionally allowed region. The conventional R factor of the model is 22.0%, and the free R factor is 24.9%.

Two cholate molecules are present in the ligand binding site: one that overlaps very well with the molecule in the wild type co-crystals and a second molecule that is in the approximate position of the second molecule present in chicken and axolotl L-BABPs. Fig. 4b is a stereo diagram with the wild type and mutant superimposed. Fig. 4c shows the electron density of the two ligand molecules in the binding site. As the figure shows, there can be no doubt about the stoichiometry of ligand binding. The right-hand side of the same figure is a diagram representing the interactions of the two ligands with the protein. Table 2 lists the distances between side chains and cholate molecules. By and large, the interactions of the cholate molecule whose carboxylate is close to the end of the second helix (number 130 in our notation) and the distances to the side chains are very similar in the wild type protein and the two mutants. The distances to the other molecule are also listed in the same table, as are the interactions of the two cholate molecules with one another. In the C91T mutant, O3 of cholate molecule number 130 is found at 2.8 Å from O7 of the other molecule and at a distance shorter than 4 Å from O12. This geometry is the same as that described for chicken L-BABP (11), but the shorter and longer distances are exchanged. Thus, the suggestion that binding of the first cholate molecule may facilitate binding of the second can be extended to this case also. Table 3 lists the distances of the 11 water molecules in contact with the bile acids and the oxygens of the two ligands that are close to them. In the case of this mutant, neither isopropyl alcohol nor glycerol molecules were identified in the cavity, although the latter was present in the cryoprotectant. Two water molecules participate in contacts with the two cholates: molecule number 1, which is in...
contact with the two O3s, and molecule number 5, which is in contact with O7 of one molecule and O26 of the other.

Fig. 4d is a stereo diagram that represents the two cholate and the water molecules of Table 3. Notice that the carboxylates of both bile acids are quite exposed to the solvent, and one of them is close to the loop connecting the two helices, whereas the other is close to the end of the second helix. For the FABP family, it has been suggested that the two helices move to permit access to the ligand binding cavity. As this particular orientation of the molecule shows, it would seem that in this case access may be possible through two different channels and that the second site to appear in evolution is not too different from the first in terms of ligand-solvent contacts.

Fig. 4e represents the r.m.s. deviation between the α-carbons of the models of the co-crystals of the wild type and C91T mutant of zebrafish L-BABP. Not unexpectedly, the highest peaks are found where the major differences between wild type and chicken L-BABP had been identified (see Fig. 1c). In general, the two patterns are similar although with different values, and the peak spanned by amino acids 101–105 is absent. The solvent-accessible volume of the ligand-binding cavity of the C91T mutant is 675.4 Å³ (i.e. it has increased from 513.6 in the wild type protein by about 30% of its initial value and is now very close to that observed in axolotl L-BABP (685.5 Å³).

Isothermal Titration Calorimetry Experiments—Fig. 5a presents the experimental ITC data for the three proteins. The plots are given as the cumulative enthalpy expressed per mole of protein versus the concentration ratio (r) = total titrated ligand/total protein. Several binding models were tested to interpret the calorimetric data.

The simplest model assumes one binding site. In this case, the protein can be found in either the free or bound state, and the partition function is \( Q = 1 + K_b [L] \). The results of the best fit trials do not produce a satisfactory fit for any of the three proteins considered (see Fig. 5a).

The second model tested assumed two independent (but not necessarily equivalent) binding sites. In this case, four states are permitted (namely free, either site bound, and both sites bound); the relevant partition function is \( Q = 1 + K_{b1} [L] + K_{b2} [L] + K_{b1} K_{b2} [L]^2 \). This model produces a satisfactory fit for the wild type protein and the G55R mutant but not for the C91T mutant (see Fig. 5b and Table 4 for the fitting parameters). Errors are within 10% of the reported values.

In both cases, we observe that the first binding site implies a binding enthalpy compatible with the presence of the ligand in a cavity (high enthalpic contribution), whereas the second site is compatible with an entropy-driven association process (low enthalpic contribution). This situation suggests that the second site is not in the cavity and is likely to be on the surface of the molecule. Furthermore, the entropic nature of the association suggests a strong contribution from the solvent. This type of interaction is unlikely to be revealed by an x-ray structure, given the extensive intermolecular contacts among the protein molecules in

![FIGURE 5. Isothermal titration calorimetry experiments.](image)

**TABLE 4**

Best fit parameters obtained using the two-independent site binding model and the two-consecutive site binding model of the ITC data

| Model                      | \( K_{b1} \)  | \( K_{b2} \)  | \( \Delta H_{b1} \) | \( \Delta H_{b2} \) | \( T \Delta S_{b1} \) | \( T \Delta S_{b2} \) | \( \Delta G_{b1} \) | \( \Delta G_{b2} \) |
|----------------------------|---------------|---------------|---------------------|---------------------|----------------------|----------------------|---------------------|---------------------|
| Two-independent site binding model |               |               |                     |                     |                      |                      |                      |                      |
| Wild type                  | 1.10 \times 10^5 | 8.00 \times 10^5 | -14.50              | -2.80               | 14.28                | 30.89                | -28.78              | -33.69              |
| G55R                      | 0.95 \times 10^5 | 8.3 \times 10^5 | -11.35              | -2.40               | 17.06                | 31.39                | -28.41              | -33.79              |
| Two-consecutive site binding model |              |               |                     |                     |                      |                      |                      |                      |
| C91T                      | 2.74 \times 10^4 | 8.03 \times 10^4 | -3.67               | -6.47               | 21.66                | 21.53                | -25.33              | -28.00              |
the crystals that can displace a bile acid bound on the protein surface to establish a protein-protein contact in the lattice.

A third model was tested to describe the binding process of the C91T mutant, namely the consecutive binding to two not necessarily equivalent sites. In this case, the protein states are three (free, first site occupied, and both sites occupied), and the relevant partition function is

\[
Q = \frac{1}{H_1} + \frac{1}{H_1^2} K_1 + \frac{1}{H_2} K_2 \frac{1}{H_1} [L] + \frac{1}{H_2} K_2 \frac{1}{H_2} [L]
\]

This model produces a reasonably satisfactory fit (see Fig. 5b and Table 4 for the fitting parameters).

Given the values obtained in this case, this type of model implies that both ligands should be located within the cavity. Therefore, this rather simple model accounts for the stoichiometry of binding observed in the crystals. Refinements of this model are possible and will require a more extensive set of experiments to discriminate the influence of the solvent in the binding process and to assess the thermal stability of the proteins in the presence/absence of ligands.

**CONCLUSIONS**

The sequences of 10 fish, three amphibian, and a reptile L-BABP were aligned using the program ClustalW (42) and are represented in Fig. 6. The black arrows indicate the main residues identified in ligand binding of the wild type protein, whereas the white arrows identify those of the C91T mutant. In general, the residues are strictly conserved in all fish. It is worthwhile mentioning also that the two members of the broken disulfide bridge are in contact with the second cholate molecule bound in the mutant. This disulfide bridge is likely to be present in all of the fish listed with the exception of catfish and shark, which appear to lack one of the two cysteines. As for the other two vertebrate groups examined, it would seem that the strategy of evolution to break the bridge can vary, since in both toad and frog, there is a Val in position 80, whereas in the lizard, there is an Ala and in axolotl a Ser in place of the second Cys in position 91. Based on our data, we can thus predict that most if not all fish should present a stoichiometry of a single cholate molecule per binding site, whereas in amphibians and higher vertebrates, the stoichiometry should be of two molecules instead.

**FIGURE 6. Sequence alignment of 10 fish, three amphibian, and one reptile L-BABP.** ZF, zebrafish; CF, catfish; JSF, Japanese sea perch; EF, European flounder; GSB, gilthead sea bream; BP, black porgy; GP, green puffer; KF, killifish; LF, lungfish; SH, shark; TO, toad; BF, bull frog; AX, axolotl; LZ, lizard. Note that with only two exceptions, catfish and shark liver BABP, all of the other fish BABPs conserve Cys80 and Cys91. A black arrow identifies the residues that bind to cholate in wild type zebrafish L-BABP, and a white arrow shows those involved in ligand binding to the C91T mutant.
Zebrafish Liver BABP

the surface of the protein molecules can be explained by the likely participation of the residues involved in ligand binding in protein–protein contacts in the lattice. This superficial site was not included in the model used to fit the ITC data of the C91T mutant to avoid overfitting with too many parameters. A refinement of this second model (e.g. including cooperative effects) will require a more extensive set of measurements but is amply justified by the observation of cooperativity in the very similar chicken L-BABP (18).

Although more cases will have to be examined to validate this hypothesis, our data support the proposal that the stoichiometry of ligand binding in the cavity is subtly related to the presence of the disulfide bridge and the possibility of the loop connecting strands H and I and their connections with strands G and J to move and make more room to accommodate the second ligand molecule. This straightforward road to change the efficiency of the transport protein by doubling its ligand binding capacity appears to have required a very simple mechanism: a single mutation but with far reaching consequences on the volume of the ligand binding cavity and on secondary structure mobility. However, since the partners that interact with this protein and take up and/or release the ligand are not completely known, it is not possible to ascertain whether the two sites are functionally equivalent or not. As for the superficial binding site revealed by ITC, in a future work, a series of ITC and DSC experiments are planned to discriminate the influence of the solvent in the binding process, to assess the thermal stability of the proteins in the presence/absence of ligands, and thus to achieve a better insight into the binding mechanisms.

Acknowledgments—We are grateful to the staff of the European Synchrotron Radiation Facility in Grenoble for assistance during data collection (Proposal MX 342).

REFERENCES

1. Ockner, R. K., Manning, J. A., Poppenhausen, R. B., and Ho, W. K. (1972) Science 177, 56–58
2. Banaszak, L., Winter, N., Xu, Z., Bernlohr, D. A., Cowan, S., and Jones, T. A. (1994) Adv. Protein Chem. 45, 89–151
3. Zimmerman, A. W., and Veerkamp, J. H. (2002) Cell. Mol. Life Sci. 59, 1096–1116
4. Coe, N. R., and Bernlohr, D. A. (1998) Biochim. Biophys. Acta 1391, 287–306
5. Schaap, F. G., van der Vusse, G. J., and Glatz, J. F. (2002) Mol. Cell. Biochem. 239, 69–77
6. Haunerland, N. H., and Spern, F. (2004) Prog. Lipid Res. 43, 328–349
7. Hertzel, A. V., and Bernlohr, D. A. (2000) Trends Endocrinol. Metab. 11, 175–180
8. Ceciliani, F., Monaco, H. L., Ronchi, S., Faotto, L., and Spadon, P. (1994) Comp. Biochem. Physiol. Part B 109, 261–271
9. Santomé, J. A., Di Pietro, S. M., Cavagnari, B. M., Cordoba, O. L., and Dell’Angelica, E. C. (1998) Trends Comp. Biochem. Physiol. 4, 23–38
10. Scapin, G., Spadon, P., Pengo, L., Mammì, M., Zanotti, G., and Monaco, H. L. (1988) FEBS Lett. 240, 196–200
11. Nichesola, D., Perduca, M., Capaldi, S., Carrizzo, M. E., Righetti, P. G., and Monaco, H. L. (2004) Biochemistry 43, 14072–14079
12. Alrefai, W. A., and Gill, R. K. (2007) Pharm. Res., in press
13. Scapin, G., Spadon, P., Mammì, M., Zanotti, G., and Monaco, H. L. (1990) Mol. Cell. Biochem. 98, 95–99
14. Di Pietro, S. M., Corsico, B., Perduca, M., Monaco, H. L., and Santomé, J. A. (2003) Biochemistry 42, 8192–8203
15. Capaldi, S., Guariento, M., Perduca, M., Di Pietro, S. M., Santomé, J. A., and Monaco, H. L. (2006) Proteins 64, 79–88
16. Scheienvao, E., Quarzago, D., Spadon, P., Monaco, H. L., Zanotti, G., and Peggion, E. (1994) Biopolymers 34, 879–887
17. Beringhelli, T., Goldoni, L., Capaldi, S., Bossi, A., Perduca, M., and Monaco, H. L. (2001) Biochemistry 40, 12604–12611
18. Ragona, L., Catalano, M., Luppi, M., Cicero, D., Eliseo, T., Foote, J., Fogo- lari, F., Zetta, L., and Molinari, H. (2006) J. Biol. Chem. 281, 9697–9709
19. Thompson, J., Winter, N., Tervey, D., Bratt, J., and Banaszak, L. (1997) J. Biol. Chem. 272, 7140–7150
20. Thompson, J., Reese-Wagoner, A., and Banaszak, L. (1999) Biochim. Biophys. Acta 1441, 117–130
21. Di Pietro, S. M., Dell’Angelica, E. C., Veerkamp, J. H., Sterin-Speziale, N., and Santomé, J. A. (1997) Eur. J. Biochem. 249, 510–517
22. Di Pietro, S. M., and Santomé, J. A. (2001) Arch. Biochem. Biophys. 388, 81–90
23. Odani, S., Baba, K., Tsachida, Y., Aoyagi, Y., Wakui, S., and Takahashi, Y. (2001) J. Biochem. (Tokyo) 129, 69–76
24. Sarropoulou, E., Power, D. M., Magoulas, A., Geisler, R., and Kotoulas, G. (2005) Aquaculture 243, 69–81
25. Jordal, A. E., Hordvik, L., Pelsers, M., Bernlohr, D. A., and Torstensen, B. E. (2006) Comp. Biochem. Physiol. B Biochem. Mol. Biol. 145, 147–158
26. Cordoba, O. L., Sanchez, E. I., and Santomé, J. A. (1999) Eur. J. Biochem. 265, 832–838
27. Donovan-Wright, E. M., Pierce, M., Sharma, M. K., and Wright, J. M. (2000) Biochim. Biophys. Acta 1492, 227–232
28. Leslie, A. G. W. (1992) Int. CCP4/ESF-EACMB Newslett. Protein Crystallogr. 26, 27–33
29. Collaborative Computational Project Number 4 (1994) Acta Crystallogr. Sect. D 50, 760–767
30. Navaza, J. (1994) Acta Crystallogr. Sect. A 50, 157–163
31. Murshedov, G. N., Vagin, A. A., and Dodson, E. J. (1997) Acta Crystallogr. Sect. D Biol. Crystallogr. 53, 240–255
32. McRee, D. E. (1999) J. Struct. Biol. 125, 156–165
33. Laskowski, R. A., MacArthur, M. W., Moss, D. S., and Thornton, J. M. (1993) J. Appl. Crystallogr. 26, 283–291
34. Castronuovo, G., Elia, V., Fessas, D., Velleca, M., and Viscardi, G. (1996) Carbohydr. Res. 287, 127–138
35. Wyman, J., and Gill, S. J. (1990) Binding and Linkage, pp. 49–59, University Science Books, Mill Valley, CA
36. Gill, S. J., Robert, C. H., and Wyman, J. (1988) in Biochemical Thermodynamics (Jones, M. N., ed) Elsevier Science Publishers B.V., Amsterdam
37. Press, W. H., Flannery, B. P., Teukolsky, S. A., and Vetterling, W. T. (1989) Numerical Recipes: The Art of Scientific Computing, pp. 521–538, Cambridge University Press, Cambridge, UK
38. Tochtrop, G. P., Richter, K., Tang, C., Toner, J. J., Covey, D. F., and Cistola, D. P. (2002) Proc. Natl. Acad. Sci. U. S. A. 99, 1847–1852
39. Tochtrop, G. P., Bruns, J. L., Tang, C., Covey, D. F., and Cistola, D. P. (2003) Biochemistry 42, 11561–11567
40. Kabash, W. (1978) Acta Crystallogr. Sect. A 32, 922–923
41. Liang, J., Edelsbrunner, H., and Woodward, C. (1998) Protein Sci. 7, 1884–1897
42. Thompson, J. D., Higgins, D. G., and Gibson, T. J. (1994) Nucleic Acids Res. 22, 4673–4680
43. Kramer, W., Sauber, K., Baringhaus, K. H., Kurz, M., Stengelin, S., Lange, G., Corsiero, D., Giribeg, F., König, W., and Weyland, C. (2001) J. Biol. Chem. 276, 7791–7801
44. Potterton, E., McNicholas, S., Krissinel, E., Cowtan, K., and Noble, M. (2002) Acta Crystallogr. Sect. D 58, 1955–1957
45. Wallace, A. C., Laskowski, R. A., and Thornton, J. M. (1995) Protein Eng. 8, 127–134