

A Disulfide Bridge Allows for Site-Selective Binding in Liver Bile Acid Binding Protein Thereby Stabilising the Orientation of Key Amino Acid Side Chains

Simona Tomaselli,^[a] Michael Assfalg,^[b] Katuscia Pagano,^[a] Clelia Cogliati,^[a] Serena Zanzoni,^[b] Henriette Molinari,^{*,[b]} and Laura Ragona^{*,[a]}

Dedicated to Dr. Lucia Zetta, who introduced us to the beauty of biomolecular NMR

Abstract: The presence of a disulfide bridge in liver bile acid binding protein (L-BABP/S-S) allows for site-selective binding of two bile acids, glycochenodeoxycholic (GCDA) and glycocholic acid (GCA), differing only in the presence of a hydroxyl group. The protein form devoid of the disulfide bridge (L-BABP) binds both bile salts without discriminating ability. We investigate the determinants of the molecular recognition process in the formation of the heterotypic L-BABP/S-S complex

with GCA and GCDA located in the superficial and inner protein sites, respectively. The comparison of the NMR spectroscopy structure of heterotypic holo L-BABP/S-S, the first reported for this protein family, with that of the homotypic L-BABP complex demonstrates that the introduction of a

S-S link between adjacent strands changes the conformation of three key residues, which function as hot-spot mediators of molecular discrimination. The favoured χ_1 rotameric states (t , g^+ and g^- for E99, Q100 and E109 residues, respectively) allow the onset of an extended intramolecular hydrogen-bond network and the consequent stabilisation of the side-chain orientation of a buried histidine, which is capable of anchoring a specific ligand.

Keywords: heterotypic complexes • molecular recognition • NMR spectroscopy • proteins • site selectivity

Introduction

We address here the description of selective binding interactions between bile acids, the major metabolites of cholesterol, and their intracellular carriers, the bile acid binding proteins (BABPs). BABPs are small molecular mass soluble proteins, found in the ileum and in the liver, displaying a large internal cavity able to host up to two bile salts.^[1–6] They belong to the fatty acid binding protein family, the members of which behave as intracellular lipid chaperones capable of binding a wide variety of exogenous ligands^[7] and play a role in the defence of the cell against different oxidative stresses.^[8,9] A detailed mechanistic investigation of the binding mode of BABPs is of interest not only for clarifying their function but also for the design of ligands with

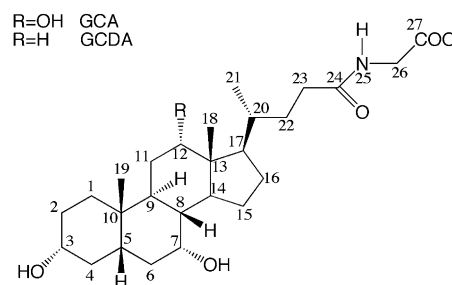
properties (for instance, fluorescence) that may arise from selective occupancy of different protein sites.^[10] BABPs may indeed behave as host systems to be exploited for the recognition of a variety of molecules in aqueous solution, with multiple potential applications in environmental and industrial chemistry.^[11]

The two most extensively characterised BABPs, namely, human ileal BABP and chicken liver BABP, share the common property of having two available binding sites for bile salts.^[2,12–14] Human ileal BABP was shown to display a nearly complete site selectivity for glycocholic (GCA) and glycochenodeoxycholic (GCDA) acids,^[15] the main components of the natural bile acid pool, differing only in the presence of a hydroxyl group at position 12 of the sterol moiety (Scheme 1). Chicken liver BABP instead exhibits a particu-

[a] Dr. S. Tomaselli, Dr. K. Pagano, Dr. C. Cogliati, Dr. L. Ragona
Laboratorio NMR, ISMAC-CNR
Via Bassini 15, 20133 Milano (Italy)
Fax: (+39) 223699620
E-mail: laura.ragona@ismac.cnr.it

[b] Dr. M. Assfalg, Dr. S. Zanzoni, Prof. H. Molinari
NMR Laboratory, Biotechnology Department
University of Verona
Strada Le Grazie 15, 37134 Verona (Italy)
Fax: (+39) 0458027929
E-mail: henriette.molinari@univr.it

Supporting information for this article is available on the WWW under <http://dx.doi.org/10.1002/chem.201102203>.



Scheme 1. Glycochenodeoxycholic and glycocholic acid structures: carbon-atom numbering is shown.

larly intriguing binding behaviour.^[16] Two forms of liver BABP have been reported in the literature, in which residue 91 is a threonine (L-BABP) or a cysteine (L-BABP/S-S) that can form a disulfide bridge with the conserved cysteine 80.^[17] In the presence of either GCA or GCDA both protein forms are able to give rise to homotypic complexes, in which the two sites are fully occupied by the same ligand (Figure 1). However, when both bile salts are present, L-

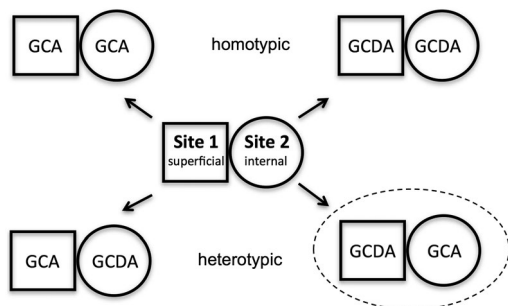


Figure 1. The four possible homotypic or heterotypic ligation states of BABPs in the presence of mixtures of GCDA and GCA. A dotted ellipse highlights the energetically favoured holo state observed for L-BABP/S-S.

BABP/S-S displays a nearly complete site selectivity towards GCA and GCDA,^[16] at variance with L-BABP in which the homotypic complex remains the most populated ligation state. The biological significance of site selectivity is related to the ability of the protein to sequester both dihydroxy and trihydroxy bile salts from the cytosol, thus protecting the cell against their cytotoxic effects.^[15]

NMR spectroscopy diffusion and ligand competition experiments, exploiting differential ¹⁵N labelling of the ligands, have been employed to characterise the four possible ligation states of L-BABP/S-S (Figure 1). The energetically favoured L-BABP/S-S heterotypic complex, hosting GCA and GCDA in the internal and superficial sites, respectively, was structurally characterised by employing 1) an iterative approach combining chemical-shift perturbation analysis and data-driven HADDOCK modelling for chemical shift and intermolecular NOE assignment and 2) XPLOR-NIH for structure determination. The structure of the heterotypic holo L-BABP/S-S, the first reported for this protein family, was used for a comparative analysis with the previously reported structure of the L-BABP homotypic complex. The obtained results demonstrate that the presence of a disulfide bridge, connecting two cysteine molecules in adjacent beta strands, affects the χ_1 rotameric state of a few key residues, favouring the formation of an extended hydrogen-bond network and allowing site discrimination between GCDA and GCA.

Results

Determination of the population of the four ligation states for L-BABP/S-S and L-BABP in the presence of GCDA and GCA:

Both NMR spectroscopy analysis and mass spectrometry data indicate the ability of both L-BABP/S-S and L-BABP to form ternary homotypic complexes with two molecules of GCDA or GCA. We have previously shown that, in the presence of both bile salts, L-BABP/S-S binds GCDA preferentially in the superficial site (site 1) and GCA in the internal one (site 2).^[16] Interestingly, site selectivity is observable only in the L-BABP/S-S variant. Indeed, competition experiments in which unlabelled GCDA (or GCA) was added to a solution containing a L-BABP/¹⁵N-GCA (or L-BABP/¹⁵N-GCDA) complex clearly indicated that L-BABP does not display any site preference (Figure 2). The power of these competition experiments resides, among others, in the property of the chemical shift of site 1 to be sensitive to the type of ligand (GCDA or GCA) occupying site 2. Thus the resonance of the ligand bound to the superficial site exhibits a different chemical shift (labelled as 1' in Figure 2) when internal site 2 is occupied by the competitor ligand and the corresponding peak volume is an indicator of the population of the heterotypic complex. The comparison of the relative intensities of peak 1 and 1', in competition experiments on L-BABP and L-BABP/S-S, suggested that the presence of the disulfide bridge could revert the relative populations of homotypic and heterotypic complexes (Figure 2b, d). Analysis of peak integrals from

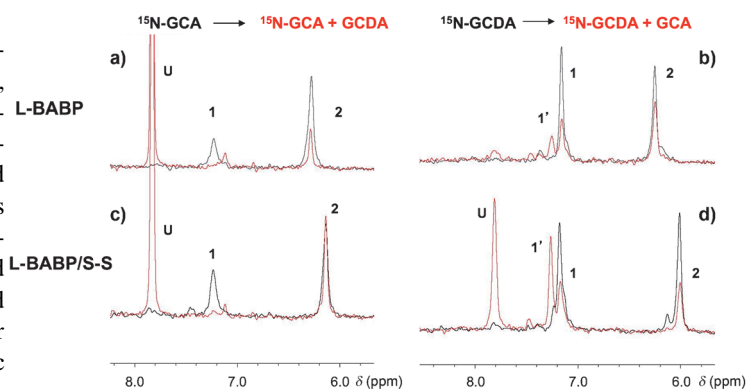


Figure 2. Bile salt competition experiments. The one-dimensional first increment of the 2D ¹H,¹⁵N-HSQC spectra collected on ¹⁵N GCA in a 1:2 protein/GCA molar ratio (L-BABP and L-BABP/S-S in (a) and (c), respectively; black lines); ¹⁵N GCA in the presence of equimolar amounts of unlabelled GCDA (protein/GCA/GCDA 1:2:2) (L-BABP and L-BABP/S-S in (a) and (c), respectively; red lines); ¹⁵N GCDA in a 1:2 protein/GCDA molar ratio (L-BABP and L-BABP/S-S in (b) and (d), respectively; black lines); ¹⁵N GCDA in the presence of equimolar amounts of unlabelled GCA (protein/GCA/GCDA 1:2:2) (L-BABP and L-BABP/S-S in (b) and (d); red lines). The resonances corresponding to the unbound ligand, the ligand bound to the superficial site and the ligand bound to the internal site are indicated as U, 1 and 2, respectively. See text for the definition of the 1' resonance. The different intensities exhibited by the unbound GCA and GCDA reflect both the lower affinity for GCA and the onset of different equilibria between monomeric and micellar bile salts.^[16]

NMR spectra acquired in the presence of either GCA or GCDA, or both, provided a rough estimate of the protein population of each variant in each of the four possible ligation states (Figure 1, Table 1). A higher population of the

Table 1. Population of the four possible ligation states estimated from NMR spectroscopy data recorded on protein/GCA/GCA 1:2:2 samples at 298 K.

Superficial site (site1)	Internal site (site 2)	L-BABP/S-S [% population]	L-BABP [% population]
GCDA	GCDA	≈35	≈45
GCA	GCDA	<5	≈25
GCA	GCA	<5	<5
GCDA	GCA	>55	≈25

GCDA homotypic complex compared to that of the GCA homotypic complex is observed for both protein variants, in agreement with the higher affinity observed for GCDA.^[1,16] In L-BABP the two heterotypic complexes are equally populated, whereas in L-BABP/S-S the population of the ligation state with GCDA in site 1 and GCA in site 2 is predominant (higher than 55%). This percentage further increases upon lowering the temperature and especially on changing the protein/ligand molar ratio. Different (L-BABP/S-S)/GCDA/GCA ratios were further characterised to identify the experimental conditions in which the heterotypic complex of interest is the only relevantly populated ligation state. The fractions of homotypic/heterotypic complexes were evaluated on the basis of the measured cross-peak volumes of selected residues in ¹⁵N-HSQC, CBCA(CO)NH and HNCOSY spectra. Indeed the amide resonances of three residues, namely, A20, T71 and I111, were easily recognisable and assigned in the spectra of both the heterotypic and homotypic complexes, thus acting as a probe of the population of the different species in solution. The population of the GCDA homotypic complex at a protein/GCDA/GCA 1:2:2 molar ratio was estimated around 35%. The population of the GCDA homotypic complex was significantly reduced (ca. 5–10%) in a sample of (L-BABP/S-S)/GCDA/GCA 1:1:1.5 at 298 K. The experimental conditions in which the L-BABP/S-S heterotypic complex is the largely predominant species in solution were employed for the collection of all NMR spectroscopy data used for structure determination. Self-diffusion coefficients (*D*) of bound ligands in homotypic and heterotypic complexes of L-BABP/S-S were also measured, as previously described,^[16] to monitor the presence of exchange between bound and free ligands. The obtained *D* values, summarised in Table 2, indicated a slower exchange process for GCA in the internal site when GCDA, rather than GCA, is bound at the superficial site. No significant difference could be inferred from *D* values measured for ligands bound to site 1 owing to high errors affecting the data. The diffusion measurements support the view, offered by competition data, that the ligation state showing GCDA in site 1 and GCA in site 2 is energetically favoured over the others. They also provide a preliminary indication that

Table 2. Diffusion coefficients of bile salt species in L-BABP/S-S homotypic and heterotypic complexes. *D* values measured for free chenodeoxycholate and holo protein are 3.97×10^{-10} and $1.04 \times 10^{-10} \text{ m}^2 \text{ s}^{-1}$, respectively.^[2]

Sample ^[a]	Superficial site (site1)	Internal site (site 2)	<i>D</i> _{site 1} [$\times 10^{-10} \text{ m}^2 \text{ s}^{-1}$]	<i>D</i> _{site 2} [$\times 10^{-10} \text{ m}^2 \text{ s}^{-1}$]
P/GCA/ GCDA	GCDA	GCA	–	1.24 ± 0.10
P/GCA/ P/GCDA/ GCA	GCA	GCA	–	1.55 ± 0.11
P/GCDA/ GCA	GCDA	GCA	2.58 ± 0.42	–
P/GCDA/ GCA	GCDA	GCDA	2.18 ± 0.15	–

[a] [¹⁵N]-enriched bile salts are indicated in bold.

one of the factors stabilising the heterotypic complex is an extended interaction network involving the inner ligand (see below).

Chemical-shift perturbation (CSP) data: Data derived from amide chemical-shift differences between L-BABP/S-S apo and holo states or between different holo states are of central importance in the iterative process of NOESY spectra assignment based on a data-driven docking approach.

Residues showing significant $\Delta\delta(\text{H}^{\text{N}})$ CSP on going from apo L-BABP/S-S to the heterotypic complex were located either at the level of the protein open end (R32, T57, D74, K76) or in a patch including the FGH strands (E67, T81, V82, V90, C91, K92, S93, F96, S97, H98, E99) (Figure 3a, b). It is, however, important to recall that amide chemical-shift changes, upon complexation, can be due either to direct interaction with the ligand or to indirect effects, such as allosteric changes or local structural rearrangement following the binding event. Independent analysis of ¹H and ¹⁵N chemical-shift changes has been shown to be of great help in distinguishing the two effects.^[18] ¹H chemical-shift changes ($\Delta\delta(\text{H}^{\text{N}})$) mainly reflect direct interactions with the ligands, whereas ¹⁵N chemical-shift changes ($\Delta\delta(\text{N})$) are influenced both by direct and indirect effects. Thus the identification of significant changes in the proton frequency together with negligible $\Delta\delta(\text{N})$, on going from an apo to a holo system (Figure 3c), allowed us to single out the subset of residues (CSP_{apo-holo}) whose chemical-shift changes are likely due to direct interactions with the ligand (D74, K76, V82, C91, K92, S93, F96, S97, H98). This information was used to define the ambiguous interaction restraints (AIR) to guide the first HADDOCK runs (see below). Residues showing significant changes in nitrogen frequency together with negligible $\Delta\delta(\text{H}^{\text{N}})$ (Q11, R32, K35, T57, E67, T81, L84, V90, E99) were considered affected mainly by structural changes.

To derive AIRs that could be selectively attributed to GCDA or GCA, ¹H,¹⁵N HSQC spectra acquired on the heterotypic complex (protein/GCDA/GCA 1:1:1.5) were compared with those of the homotypic complexes (with GCA or GCDA, protein/ligand ratio 1:3) (Figure 4a). Indeed, a chemical-shift comparison of proteins complexed with closely related ligands, such as GCDA and GCA (differing only

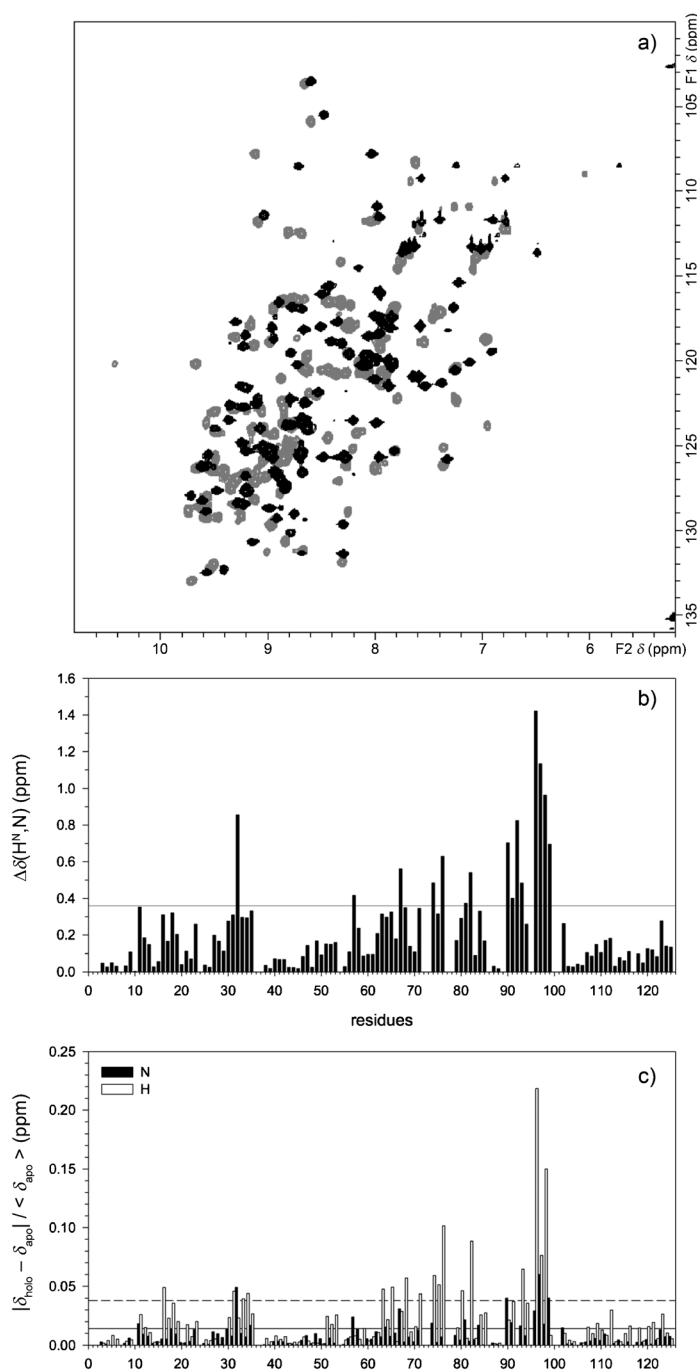


Figure 3. Chemical-shift changes upon heterotypic complex formation for (L-BABP/S-S)/GCDA/GCA 1:1:1.5 sample at pH 7 and 298 K. a) Superposition of apo (black) and holo $^1\text{H}, ^{15}\text{N}$ HSQC of the heterotypic complex (grey). b) Combined chemical-shift differences between apo and holo resonances (calculated as $\Delta\delta(\text{H}^{\text{N}}, \text{N}) = [(\Delta\delta(\text{H}^{\text{N}})(\text{apo}-\text{holo})^2 + \Delta\delta(\text{N})(\text{apo}-\text{holo})^2/25)^{1/2}]$) are plotted versus residue number. Missing data correspond to residues that could not be assigned in the apo state. The continuous line corresponds to the mean value plus one standard deviation. c) Plots for $\Delta\delta(\text{H}^{\text{N}})$ (white) and $\Delta\delta(\text{N})$ (black), calculated as $|\delta_{\text{holo}} - \delta_{\text{apo}}| / \langle \delta_{\text{apo}} \rangle$, versus residue number, in which $\langle \delta_{\text{apo}} \rangle$ is the average chemical shift calculated over all residues. The continuous and dashed lines correspond to the mean plus one standard deviation for N and H^{N} resonances, respectively.

in the presence of an OH group bound to C12), allowed the identification of the binding-site region proximal to the altered portion of the ligand, as elegantly shown for the FK506 binding protein upon addition of a series of FK506 analogues.^[19] Significant chemical-shift differences observed between heterotypic and homotypic GCDA complexes ($\text{CSP}_{\text{hetero-homo}}$) (Figure 4b) were attributed to the effect of the internal ligand and in particular of its OH-12. Large chemical-shift differences were observed for T91, H98, E99; however, only T91 and H98 were considered in direct contact with GCA, since E99 is expected to be affected by conformational changes rather than by a direct ligand effect, as discussed above. Similar arguments were employed to identify residue M73 as close to the superficial ligand (Figure 4a,b). The behaviour of glutamine and asparagine side chains could be analysed as well, as their NH_2 resonances were clearly detectable in the same $^1\text{H}, ^{15}\text{N}$ HSQC spectra. With similar arguments described for backbone amides, the side chain of Q100 was considered in contact with GCA and Q56 with GCDA.

Solution-structure determination of the heterotypic complex: The structure determination of the ternary complex represents a challenging task due to the extensive spectral overlap of GCDA and GCA resonances and to the related difficulties in the unambiguous assignments of intermolecular NOEs. The protocol employed to overcome these difficulties, summarised in Figure 5, consisted of 1) structure calculation of the holo scaffold (CYANA 2.1); 2) iterative docking procedure (HADDOCK 2.0) for the assignment of ligand resonances and of intermolecular NOEs; and 3) final structure calculation (XPLOR-NIH 2.26) and water refinement of the complex.

The backbone assignment of L-BABP/S-S in complex with GCDA and GCA was achieved by the analysis of standard triple-resonance spectra and was completed at 99.2%. Side-chain assignment was performed by analysing a set of 3D H(C)CH TOCSY and (H)CCH TOCSY experiments. Intramolecular NOEs were collected by analysing 3D NOESY- $^1\text{H}, ^{15}\text{N}$ -HSQC and 3D NOESY- $^1\text{H}, ^{13}\text{C}$ -HSQC spectra optimised for aliphatic and aromatic resonances. As previously shown,^[2] the protonation state of the buried histidine molecules plays an important role in the complex and its knowledge is of primary importance to set up structural calculations. Analysis of $^1\text{H}, ^{15}\text{N}$ long-range HSQC spectra, based on reported histidine cross-peak patterns,^[20] indicated that H83 is predominantly in the protonated form, while the $\text{N}_{\delta 1}$ tautomer seems to be the most abundant form for H98. This information was imposed in all the structural calculations. The solution structure of holo L-BABP/S-S was calculated by CYANA 2.1 from 2170 distance constraints and 228 angle constraints, collected by the analysis of NOESY spectra and of backbone chemical shifts employing TALOS+, respectively. The 20 lowest-energy conformers, over the 100 calculated, were then selected to represent the NMR spectroscopy solution structure. The obtained protein bundle is of high quality (root-mean-square deviation (RMSD)=

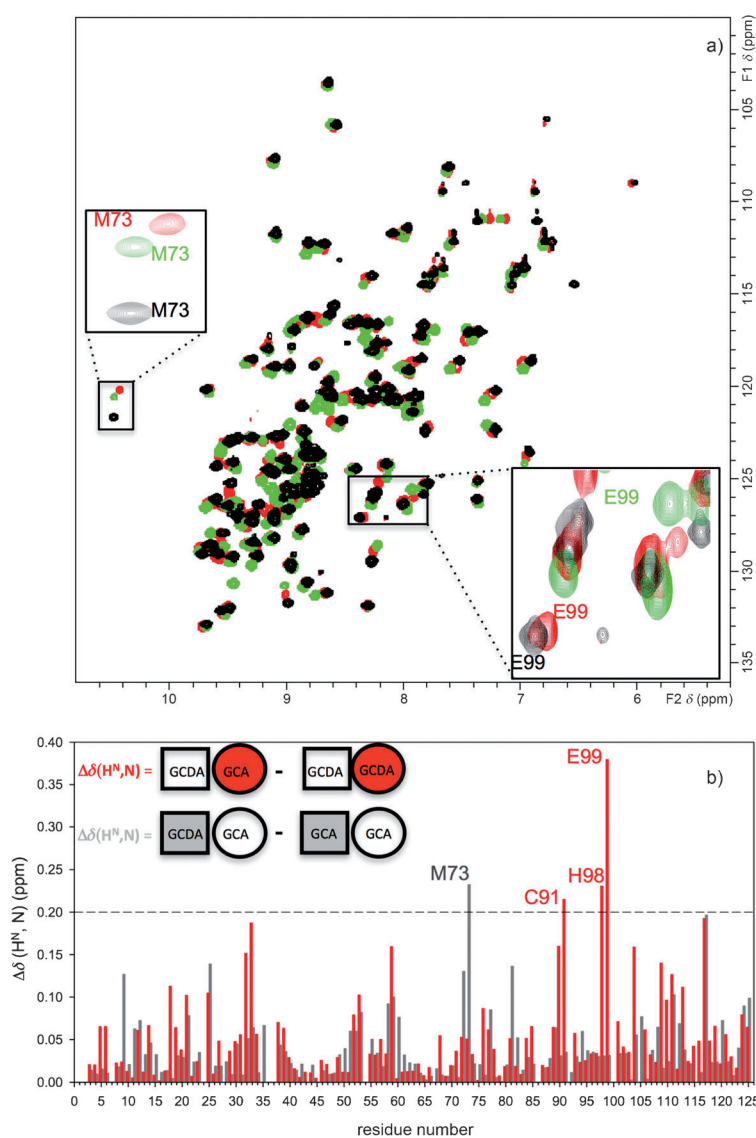


Figure 4. Comparison of chemical-shift changes of L-BABP/S-S in heterotypic and homotypic complexes. a) Superposition of ^1H , ^{15}N HSQC of heterotypic (red), GCA-homotypic (black) and GCDA-homotypic (green) complexes. Regions corresponding to residue M73 and E99 are magnified in the insets. b) Combined $\Delta\delta(\text{H}^{\text{N}}, \text{N})$ chemical-shift differences between resonances of heterotypic (L-BABP/S-S)/GCDA/GCA 1:1:1.5 sample and homotypic (L-BABP/S-S)/GCDA 1:3 samples (red histogram) or homotypic (L-BABP/S-S)/GCA 1:3 (grey histogram) are plotted versus residue number. Chemical-shift changes were calculated as $\Delta\delta(\text{H}^{\text{N}}, \text{N}) = [(\Delta\delta(\text{H}^{\text{N}})(\text{hetero-homo})^2 + \Delta\delta(\text{N})(\text{hetero-homo})^2/25)/2]^{1/2}$. The dashed line indicates the lower limit selected to define significant chemical-shift variations. In the graphical scheme a square and a circle represent superficial and inner binding sites, respectively. Graphical elements are colour coded according to the histogram.

(0.62 ± 0.11) Å for backbone and (1.27 ± 0.13) Å for all heavy atoms; 81.7% residues in the most favoured regions of the Ramachandran plot and 18.3% in the allowed regions). The statistics for the final bundle are reported in Table S1 in the Supporting Information.

Ligand resonances were assigned based on a F_1/F_2 - ^{13}C , ^{15}N -filtered NOESY spectra recorded at 900 MHz, showing the resonances of both the bound ligands and unbound GCA ligand. From the analysis of 3D F_1 - ^{13}C -filtered, F_2 - ^{13}C -separated, F_3 - ^{13}C -edited NOESY-HSQC spectra, displaying only intermolecular NOEs, it was possi-

ble to unambiguously assign only the resonance of GCA H12, showing two NOEs with H98 H_β atoms. Structural models were thus generated by HADDOCK to interpret the whole set of protein–ligand NOEs and to complete the assignment of ligand resonances. These first HADDOCK runs employed AIRs derived from 1) CSP and 2) ^{15}N relaxation data analysis.^[16] CSP_{apo-holo} were ambiguously attributed to both GCA and GCDA, while CSP_{hetero-homo} were unambiguously assigned to one of the two bile acids. The described iterative procedure required ten HADDOCK calculation steps and led to the assignment of the resonances of the two bound ligands (Table 3) and of 48 intermolecular NOEs (Table 4).

The structure determination of the heterotypic complex was then obtained through XPLOR-NIH, by employing all the constraints derived from the described iterative procedure. The final structural bundle is shown in Figure 6 and the relative statistics are reported in Table 5. GCA bound to the inner site is well defined with a RMSD of (0.85 ± 0.15) Å after protein backbone superposition (calculated with ProFit over the steroid moiety). The presence of a multiconformational state for GCDA in the superficial site is evident (Figure 6) and is reflected by a RMSD of (1.50 ± 0.66) Å after protein backbone superposition. The divergence

between the models is significantly higher when the glycine tail is included in the analysis (RMSD values up to 4 Å).

Structural determinants of site selectivity: The analysis of the structure of the L-BABP/S-S heterotypic complex sets the foundation for understanding the roles that specific interactions play in ligand binding and selectivity. Structural data shows that the superficial binding site is characterised by a highly hydrophobic surface defined by Y14, F17, L18, L21, L23, L27, A31, I34, P54, M73 and I111 side chains and is therefore suitable to host GCDA, which lacks the polar

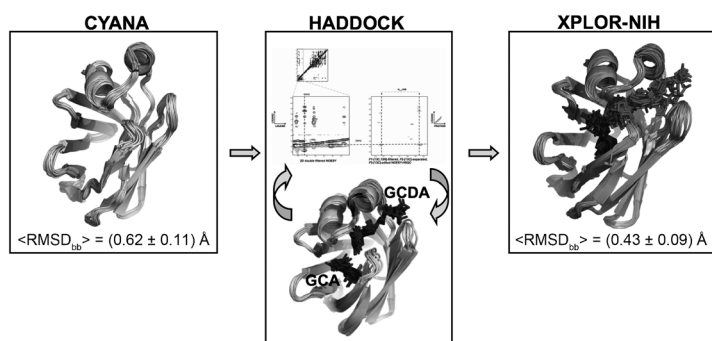


Figure 5. Summary of the structure calculation procedure and representation of the obtained bundles. Root-mean-square deviations of backbone atomic coordinates for the twenty structures of each CYANA and XPLOR-NIH bundles, calculated over residues 1–125 using CING, are reported in the boxes.

Table 3. Chemical shifts of bound GCA and GCDA ((L-BABP/S-S)/GCDA/GCA 1:1:1.5) measured at pH 7 and 298 K.

Atoms	GCA [ppm]	GCDA [ppm]
NH	6.20	7.30
Gly H α s	3.66, 3.30	3.71
H3	3.16	3.52
OH-3	3.53	–
H7	3.51	3.72
OH-7	3.74	–
H12	3.62	–
OH-12	4.19	–
CH ₃ -18	0.15	0.71
CH ₃ -19	0.12	0.99
CH ₃ -21	0.5	0.98

OH group bound to C12 (OH-12). The internal site, where GCA is preferentially bound, is more polar and is characterised by a hydrogen-bond cage anchoring the ligand to the protein. Indeed, LIGPLOT analysis revealed that GCA OH-3 is bound to Q100 NH ϵ_2 in most of the structures and, in a few structures, a further intermolecular hydrogen bond is observed between GCA OH-12 and H98 N δ_1 . It should be noted that GCA OH-3 and OH-12 labile protons are observable in the doubly filtered 2D-NOESY spectrum of the L-BABP/S-S heterotypic complex; their assignment is confirmed by NOEs with the proton bound to the same carbon (observed NOEs: OH-3/H3 and OH-12/H12).

Key intramolecular hydrogen bonds may play an important role in favouring the appropriate side-chain orientation for anchoring the ligand. This is the case of hydrogen bonds connecting the carboxyl group of E109 with both H98 NH ϵ_2 and Q100 NH ϵ_2 (Figure 6). The observation of these hydrogen bonds in most of the structures of the protein bundle is confirmed by the observation of the H98 NH ϵ_2 resonance, generally not visible due to its labile nature, at $\delta = 11.5$ ppm in the 2D NOESY spectra of L-BABP/S-S heterotypic complex. NOEs of H98 He $_2$ with the side chain of E109 confirm the presence of close contacts between these two residues (Table 6). Another hydrogen bond present in the majority of structures connects E99 carboxyl group with T110 OH γ

Table 4. Ambiguous intermolecular restraints used to calculate the HADDOCK structure of the heterotypic complex L-BABP/S-S with GCDA and GCA.

Active AIR	Ligand
Y14 (NOE) ^[a]	GCDA
F17 (NOE)	GCDA
L18 (NOE)	GCDA
L21 (NOE)	GCA
L23 (NOE)	GCDA
L27 (NOE)	GCDA
M30 (NOE)	GCDA
A31 (NOE)	GCDA
D33 (NOE)	GCDA
I34 (NOE)	GCDA
V49 (NOE)	GCA
R55 (NOE)	GCDA
Q56 side chain (CSP _{hetero-homo})	GCDA
N60 (NOE)	GCA
F62 (NOE)	GCA
I70 (NOE)	GCA
T71 (NOE, dynamics)	GCA
M73 (NOE, CSP _{hetero-homo})	GCDA
D74 (NOE, CSP, dynamics)	GCDA/GCA
K76 (CSP, NOE)	GCA
L78 (NOE)	GCA
V82 (CSP)	GCA
K89 (NOE)	GCA
C91 (CSP, CSP _{hetero-homo})	GCA
K92 (CSP, dynamics)	GCA
S93 (CSP)	GCA
F96 (CSP, NOE)	GCA
S97 (CSP, dynamics)	GCA
H98 (CSP _{hetero-homo} , NOE)	GCA
Q100 side chain (CSP _{hetero-homo})	GCA
I111 (NOE)	GCA
F113 (NOE)	GCA
L118 (NOE)	GCDA

[a] The experimental method used to define the residue as active is reported in parentheses. CSP: chemical shift perturbation upon binding; CSP_{hetero-homo}: chemical shift differences between heterotypic and homotypic complexes; dynamics: comparison between apo and holo ¹⁵N T₁, T₂ values;^[16] NOE: filtered-edited NOESY data.

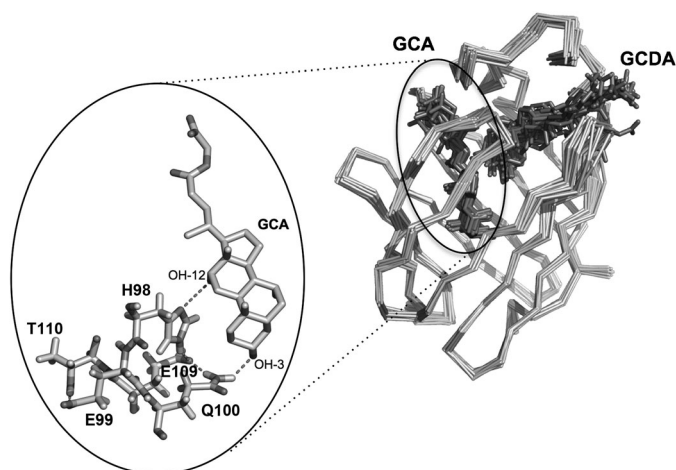


Figure 6. Superposition of the twenty structures of the XPLOR-NIH bundle. GCDA and GCA are represented as sticks and are labelled. The hydrogen atoms of the ligands are not shown for visual clarity purposes. Inset: polar interactions stabilising the inner ligand are shown. Hydrogen bonds are represented as dashed lines.

Table 5. Structural statistics of the twenty best structures calculated by XPLOR-NIH for L-BABP/S-S in complex with GCDA and GCA.

Average RMSD [\AA] ^[a]	
RMSD backbone	0.43 ± 0.09
RMSD heavy atoms	0.96 ± 0.06
RMSD GCA ^[b]	0.85 ± 0.15
RMSD GCDA ^[b]	1.50 ± 0.66
Number of interaction restraints	
TALOS+dihedral-angle restraints	228
intermolecular restraints ^[c]	48 ^[d]
intramolecular restraints	2168
intermolecular distance restraints violations	0
intramolecular distance restraints violations ^[d]	4
> 0.5 \AA	2
> 0.3 \AA	2
Intermolecular energies after water refinement	
E_{vdw} [kcal mol^{-1}]	-54.9 ± 4.8
E_{elec} [kcal mol^{-1}]	-42.8 ± 10.9
RMSD from idealised covalent geometry	
bonds [\AA]	0.004 ± 0.0002
angles [$^{\circ}$]	0.55 ± 0.01
improper dihedrals [$^{\circ}$]	0.51 ± 0.02
dihedrals [$^{\circ}$]	20.0 ± 0.70
Ramachandran plot [%]	
residues in most favoured regions	84
residues in allowed regions	15.5
residues in generously allowed regions	0.5
residues in disallowed regions	0

[a] Root mean square deviations of atomic coordinates were calculated over residues 1–125 using CING. [b] Calculated by the ProFit program by superimposing each structure to the lowest-energy structure and not considering the glycine moiety in the fit. [c] 24 restraints for GCDA and 24 for GCA. [d] Distance violated in more than 15% of the structures of the final bundle.

(vide infra). Its presence is supported by the observation of inter-side-chain NOEs. NMR spectroscopy experimental evidence supporting the hydrogen-bond network shown by structural results are summarised in Table 6. Additional hydrogen bonds were formed between the GCA carboxyl glycine moiety and T72 and/or K76 side chains.

Table 6. NMR spectroscopy experimental evidence supporting the hydrogen-bond network.

Hydrogen bonds	NOEs	Labile proton chemical shifts δ [ppm]
H98...C91	91HN...98HN	–
	91HN...98H β	–
	91HN...98H δ_2	–
H98...E109	98H ϵ_2 ...109H β	–
	98H ϵ_2 ...109H γ	NH ϵ_2 = 11.50
H98...GCA E99...T110	98H β ...GCA H12	OH-12 = 4.19
	99HN...110HN	–
	99HN...110H β	–
	99HN...110H γ	–
	99H β ...110H β	–
	99H γ ...110HN	–
99H γ ...110HN	–	
Q100...GCA	–	OH-3 = 3.53

Discussion

NMR spectroscopy competition experiments and self-diffusion coefficients of bound bile acids suggest that the asymmetric complex containing GCDA in the superficial site and GCA in the internal site is energetically favoured over the other ligand states for L-BABP/S-S. Asymmetry displayed by L-BABP/S-S in its interactions with GCA and GCDA has been reported for human ileal BABP,^[15] however, no structure of a heterotypic complex has been ever reported for this protein family. Here a combination of data-driven docking approaches for assignment purposes and restrained molecular dynamics led to the structural determination of the heterotypic complex of L-BABP/S-S with GCA and GCDA. Analysis of the obtained bundle reveals a well-defined orientation for GCA in the more internal site, whereas a higher variability is observed for the superficial site hosting GCDA. The conformational heterogeneity of the ligand at the superficial site and protein flexibility at the open end have been previously identified for both protein forms, with and without a disulfide bridge.^[2,16]

The superficial binding site displays a more hydrophobic character appropriate to host the dihydroxy bile salt (GCDA) rather than a trihydroxy bile salt (GCA). The inner ligand is part of a hydrogen-bond cage involving residues H98–Q100–E109, which bind GCA OH-3 and OH-12, thus determining site selectivity. The hydrogen-bond network identified in the presence of the disulfide bridge, connecting the carboxyl group of E109 with both H98 NH ϵ_2 and Q100 NH ϵ_2 , is not present in the homotypic complex of L-BABP (Protein Data Bank (PDB) id: 2JN3) (Figure 7), which does not discriminate between the two ligands. In L-BABP/S-S, H98 N δ_1 points towards the ligand, offering an anchoring point for the more hydrophilic GCA, and H98 NH ϵ_2 forms an intraprotein hydrogen bond with the carboxyl group of the E109 side chain. In L-BABP, the H98 side chain shows a reversed χ_2 orientation (Figure 8).

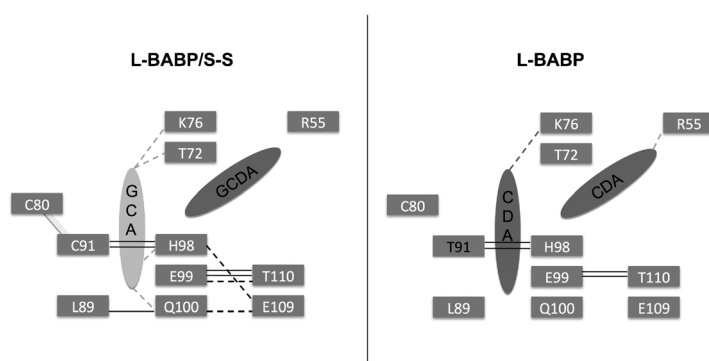


Figure 7. Hydrogen-bond interactions detected in the NMR spectroscopy structures of L-BABP/S-S in complex with GCDA and GCA (left) and L-BABP in complex with CDA (right). Black dashed lines represent side-chain/side-chain hydrogen bonds, continuous lines backbone hydrogen bonds. Grey dashed lines correspond to protein–ligand hydrogen bonds and the disulfide bridge between C80 and C91 is represented as a grey bar.

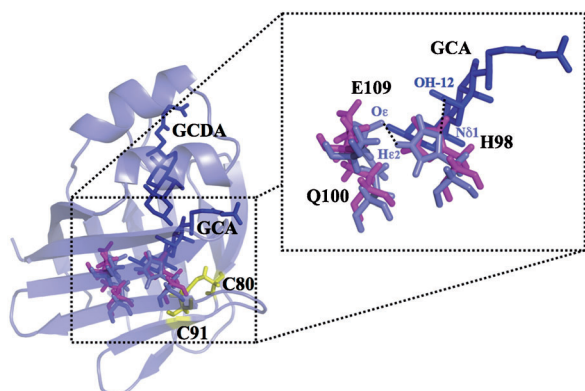


Figure 8. Rotameric state of H98 and the hydrogen-bond network with Q100 and E109. Superposition of H98, Q100 and E109 side chains in L-BABP/S-S (marine blue) and L-BABP (magenta) is shown. The L-BABP/S-S backbone cartoon is shown with transparency effects to allow visual inspection on the internal cavity. GCA and GCDA are coloured blue. The disulfide bridge C80–C91 is shown in yellow. In the magnified region hydrogen bonds between L-BABP/S-S and GCA are shown as black dotted lines and atoms involved are labelled and colour coded as protein or ligand. The hydrogen atoms of the ligands are not shown for visual clarity purposes.

The comparative analysis of the backbone (ϕ and ψ) and side-chain dihedral angles (χ_1) of the residues (H98, E99, Q100, E109 and T110) involved in the mentioned hydrogen-bond interactions in the two proteins offers an explanation for the modified hydrogen-bond network, in response to the presence of a disulfide bridge. Indeed, significant differences in backbone ϕ and ψ angles are observed for residues in the region 93–101 (close to the disulfide bridge) of the two proteins (Figure S1 in the Supporting Information). The preferred side-chain χ_1 rotameric state of a residue is strongly correlated to its ϕ and ψ dihedral-angle values.^[21] Analysis of χ_1 rotameric states for residues H98, E99, Q100, E109 and T110 in the twenty conformers of the L-BABP and L-BABP/S-S bundles highlighted differences only for E99,

Q100 and E109 (Figure 9). This has profound effects on the hydrogen-bond pattern permitted, as only the spatial arrangement of E99 t, Q100 g⁺, E109 g⁻ and T110 g⁺ rotamers favours the formation of the hydrogen bonds connecting the carboxyl group of E109 with Q100 NH₂ and the carboxyl group of E99 with T110 OH γ .

Conclusion

In summary, in the disulfide-bridged protein form, a highly hydrophobic surface characterises the superficial site that preferentially hosts the less hydrophobic ligand, while χ_1 rotameric states of key amino acids are determinant for the onset of a hydrogen-bond cage favouring GCA recognition. It should be mentioned that the disulfide bridge of L-BABP/S-S links two cysteine molecules located in adjacent antiparallel strands. These types of disulfides, which disobey rules expounded in classic structural studies by Richardson and Thornton^[22,23] and originally postulated to be forbidden, clearly exist in many proteins and appear to be associated with a functional role, rather than as structural stabilisers.^[24] The present results strongly support a functional role for the disulfide bridge, which confers site selectivity to the protein. Thus, thiol–disulfide interchange reactions, which could take place in response to modified cellular redox states, are capable of modulating ligand-binding properties. Preliminary data on intracellular L-BABP/S-S redox potentials, in progress in our laboratory, compare well with the liver cytoplasmic potential, mainly determined by the reduced-to-oxidised glutathione ratio (GSH/GSSG), thus confirming the functional relevance of the oxidised form of liver bile acid binding protein. The identification of a few key residues modulating L-BABP molecular recognition of physiological ligands opens the way to the engineering of the protein-binding pocket and/or to the rational design of functional molecules suitable for selective recognition processes.

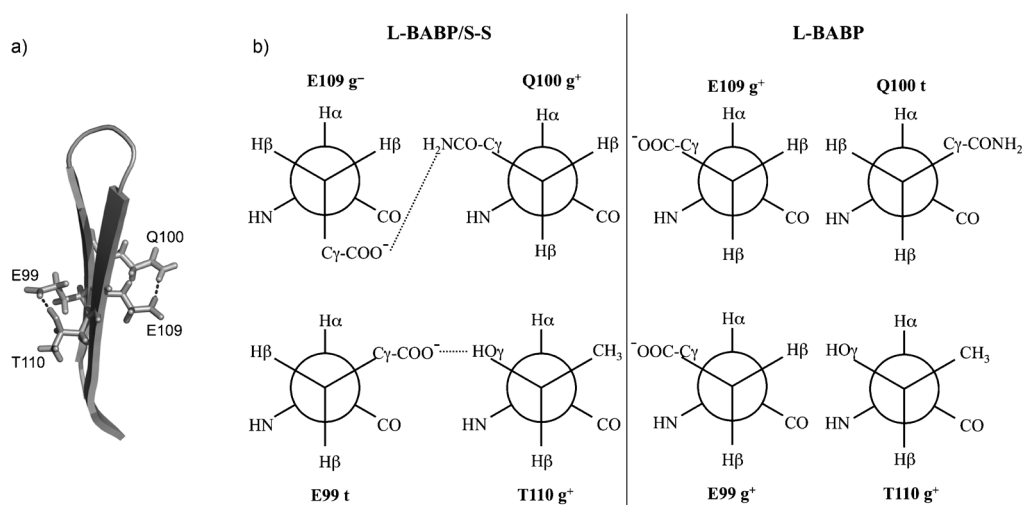


Figure 9. χ_1 rotameric states of E99, Q100, E109 and T110 in L-BABP/S-S (left panel) and L-BABP (right panel). Dotted lines represent hydrogen bonds. The structural details of side-chain orientation of E99, Q100, E109 and T110 in L-BABP/S-S are represented on the left.

Experimental Section

NMR spectroscopy sample preparation: $^{15}\text{N}/^{13}\text{C}$ -labeled L-BABP and L-BABP/S-S were expressed and purified as previously described.^[16,25] [^{15}N]Glycine conjugates of CDA and CA were synthesised as previously reported.^[1] Protein–ligand (P–L) complexes were prepared in 30 mM sodium phosphate buffer, pH 7.2, containing 0.03 % NaN_3 and 95 %/5 % $\text{H}_2\text{O}/\text{D}_2\text{O}$ or 99 % D_2O with a protein concentration of 0.5 mM. Diffusion experiments were performed on homotypic complexes with P/L ratio 1:4 and on heterotypic complexes with protein/GCDA/GCA ratio 1:2:2. Spectra recorded for chemical-shift assignment of homotypic complexes were run on samples with P/L ratio 1:3, whereas spectra employed for chemical-shift assignment and structural calculation of the L-BABP/S-S heterotypic complex were recorded on samples with protein/GCDA/GCA ratio 1:1:1.5 to reduce the content of the homotypic GCDA complex (ca. 5–10 %).

NMR spectroscopy, chemical-shift assignment and NOE analysis: NMR spectra were acquired at 298 K on Bruker spectrometers DMX 500, Avance III 600 and Avance 900, with the last two equipped with a 5 mm TCI cryoprobe and Z-field gradient. Data were processed with NMRPipe^[26] and visualised by using NMRView.^[27]

The sequential amino acid resonances assignment has been performed by analysing a complete set of triple-resonance experiments HNCACB, CBCA(CO)NH, HN(CA)CO and HNCO collected at 600 MHz. Side-chain carbon and proton resonances were assigned based on (H)CCH-TOCSY, H(C)CH-TOCSY acquired on a sample dissolved in D_2O . The resonance assignments were submitted to the BioMagResBank under accession number 17767. Intramolecular NOE-based distance restraints were extracted from 3D NOESY- $^1\text{H},^{15}\text{N}$ -HSQC, NOESY- $^1\text{H},^{13}\text{C}$ -HSQC in H_2O , as well as NOESY- $^1\text{H},^{13}\text{C}$ -HSQC spectra in D_2O optimised for the aromatic side chains, all acquired at 600 MHz. Intermolecular NOEs were obtained from NOESY- $^1\text{H},^{13}\text{C}$ -HSQC spectra, as well as from a 3D F1- ^{13}C -filtered, F2- ^{13}C -separated, F3- $^{15}\text{N},^{13}\text{C}$ -edited NOESY-HSQC spectrum acquired at 600 MHz. Mixing times of 120 ms were used in all NOESY spectroscopy experiments. $^1\text{H},^{15}\text{N}$ HSQC, CBCA(CO)NH and HNCO spectroscopy experiments performed at room temperature and at a protein/GCDA/GCA 1:2:2 ratio, employed to evaluate the relative fractions of homotypic and heterotypic complexes, afforded similar figures ($^1\text{H},^{15}\text{N}$ HSQC: 32 %, CBCA(CO)NH: 36 % and HNCO 37 %). Fractions of homotypic/heterotypic complexes at a protein/GCDA/GCA ratio 1:1:1.5 were evaluated from experiments based on both ligand ($^1\text{H},^{15}\text{N}$ HSQC, HNCA acquired on a heterotypic complex prepared with $^{15}\text{N}/^{13}\text{C}$ Gly-GCDA) and protein ($^1\text{H},^{15}\text{N}$ HSQC) observation. Similar figures, in the range 5–10 %, were obtained.

F1/F2- $^{15}\text{N},^{13}\text{C}$ -filtered NOESY spectrum was recorded at 900 MHz with a mixing time of 120 ms to assign ligand proton resonances of GCDA and GCA bound to the protein.

Diffusion experiments: [^{15}N]-edited diffusion experiments were performed using a pulse program obtained by combining the standard HSQC pulse scheme with a pulsed-field-gradient-stimulated echo module employing bipolar gradients under the same conditions as previously reported.^[16] The measured signal volumes (I), as a function of the applied gradient, were fitted to Equation (1) using a nonlinear least-squares minimisation:

$$I = I(0)\exp[-D\gamma^2G^2\delta^2(\Delta - \delta/3 - \tau/2)] \quad (1)$$

in which D is the translational diffusion coefficient, γ is ^1H gyromagnetic ratio, G is the gradient strength, δ is the duration of the gradient pulse, Δ is the delay between the focusing and refocusing gradients and τ is the gradient pulse separation.

Structure calculations: Structure calculations were performed in three main steps as follows.

Calculation of the protein scaffold: Simulated annealing in torsion-angle space was performed using the CYANA 2.1 package,^[28] which implements an efficient protocol for structure calculation/automated NOE assignment. The standard annealing protocol was used with 10000 steps of

torsion-angle dynamics; in each of the seven cycles, 100 structures were calculated and the 20 with the lowest target function were used in the next cycle. All the NOEs were automatically assigned, resulting in a total of 2170 upper distance bounds. A total number of 228 dihedral-angle restraints were derived from TALOS+.^[29] In the last cycle, the 20 structures with the lowest target function were selected as the final bundle.

Docking of ligands in the protein scaffold: An ensemble of 5 NMR spectroscopy structures of the holo protein calculated by CYANA was used as an input for the docking calculation performed with the software HADDOCK 2.0^[30] in combination with CNS.^[31] The starting coordinates of GCDA and GCA were obtained with the program SMILE.^[32] The structure of the complex was calculated by HADDOCK with 1) the final upper distance limits used for CYANA structure calculation; 2) the intermolecular NOEs and 3) TALOS+ restraints. The HADDOCK docking protocol consists of three steps: 1) randomisation of orientation and rigid-body minimisation, 2) semi-rigid simulated annealing in torsion-angle space, 3) final refinement in Cartesian space with explicit solvent. The rigid-body docking step was performed five times, with 4000 structures generated at each stage, the best 400 of which were refined in the semi-flexible stage and subsequently the best 300 in explicit water. Electrostatic and van der Waals terms were calculated with a 8.5 Å distance cutoff using the OPLS non-bonded parameters from the parallhdg5.3.pro parameter file.^[33] The Ne_2 protonated tautomer of H98 was selected in the input files. The topology and parameters of the ligand were calculated using the PRODRG server.^[34] The dielectric constant (ϵ) was set to 10.0 during rigid-body minimisation and semi-flexible simulated annealing and to 1.0 during explicit-solvent refinement. The ligand was kept semi-flexible during the semi-flexible stage. Ten HADDOCK calculation steps were required to assign the resonances of the bound ligand reported in Table 3. The final HADDOCK calculation was performed using 70 AIR restraints, 48 intermolecular and 2170 intraprotein distance constraints and 228 dihedral-angle constraints.

XPLOR-NIH structure calculation: The final structure of the heterotypic complex of L-BABP/S-S with GCDA and GCA was calculated with XPLOR-NIH 2.26^[35] and the PARALLHDG force field.^[33] The topology of H98 was modified to perform all the calculations with the Ne_2 protonated tautomer. Topology, force-field parameters and coordinates were calculated with the tool AnteChamber.^[36] Hydrogen atoms of both ligand molecules are included in the calculation. In the first part of the protocol, the L-BABP/S-S structure was generated in a completely extended conformation; in the second step the protein and the two ligand molecules were allowed to anneal during a high temperature phase (40000 steps of molecular dynamics at 4000 K, each step of 0.002 ps); in the third step the temperature was lowered to 100 K in 78 cycles of cooling phase with steps of 0.002 ps. Finally 2000 steps of Powell energy minimisation were performed. The simulated annealing was driven by NMR spectroscopy restraints: 2170 intraprotein-distance restraints, 228 dihedral-angle restraints derived from TALOS+ and 48 intermolecular-distance restraints (24 for each ligand) were used. Intraprotein-distance restraints were derived from the upper limit list calculated by CYANA ± 1.50 Å. Intermolecular-distance restraints were subdivided into three groups: strong (2.0–4.5 Å), medium (2.5–5.5 Å) and weak (3.0–7.0 Å).

Forty eight structures, over the 100 calculated by XPLOR-NIH, showed NOE violations or dihedral violations ≤ 1 , with an average total energy of (355.3 ± 16.4) kcal mol $^{-1}$. This subset was further refined in an explicit solvent with the CNS water-refinement protocol performed using all the constraints of the XPLOR-NIH calculation. The obtained structures display good structural statistics (Table 5). The protein–ligand contacts were analysed by using the software LIGPLOT^[37] and the RMSD, referred to ligand coordinates (carbon atoms), was calculated after global superposition of the protein backbone using the ProFit program.^[38] The final ensemble of structures has been deposited in the PDB (2lfo). Pymol was used for graphical representation of the results.^[39]

Acknowledgements

Authors are indebted to Lucia Zetta for helpful discussions. Renato Longhi is acknowledged for [¹⁵N]-enriched bile salts synthesis. Daniel Cicero and Andrea Giachetti are acknowledged for help with XPLOR-NIH. This research was possible thanks to the University of Verona for financial support in the acquisition of the NMR Bruker Avance 600 MHz spectrometer and to Cariverona Foundation for the acquisition of the cryoprobe. L.R. acknowledges Italian Ministry of Health (GR2007) and Fondazione De Marco for financial support. The work was supported by Ministero dell'Istruzione, dell'Università e della Ricerca (Fondo per gli Investimenti della Ricerca di Base, Futuro in Ricerca 2008, Grant RBF08R70U). The WeNMR project (European FP7 e-Infrastructure grant, contract no. 261572) is acknowledged for the use of web portals, computing and storage facilities. CIRMMMP (Consorzio Interuniversitario di Risonanze Magnetiche di Metalloproteine Paramagnetiche) is gratefully acknowledged for access to Bruker 900 MHz at CERM (University of Florence).

- [1] S. Tomaselli, L. Ragona, L. Zetta, M. Assfalg, P. Ferranti, R. Longhi, A. M. Bonvin, H. Molinari, *Proteins* **2007**, *69*, 177–191.
- [2] T. Eliseo, L. Ragona, M. Catalano, M. Assfalg, M. Paci, L. Zetta, H. Molinari, D. O. Cicero, *Biochemistry* **2007**, *46*, 12557–12567.
- [3] M. Guariento, D. Raimondo, M. Assfalg, S. Zanzoni, P. Pesente, L. Ragona, A. Tramontano, H. Molinari, *Proteins* **2008**, *70*, 462–472.
- [4] M. Guariento, M. Assfalg, S. Zanzoni, D. Fessas, R. Longhi, H. Molinari, *Biochem. J.* **2010**, *425*, 413–424.
- [5] G. Scapin, P. Spadon, L. Pengo, M. Mammi, G. Zanotti, H. L. Monaco, *FEBS Lett.* **1988**, *240*, 196–200.
- [6] E. Schievano, D. Quarzago, P. Spadon, H. L. Monaco, G. Zanotti and E. Peggion, *Biopolymers* **1994**, *34*, 879–887.
- [7] M. Furuhashi, G. S. Hotamisligil, *Nat. Rev. Drug Discovery* **2008**, *7*, 489–503.
- [8] J. Yan, Y. Gong, Y. M. She, G. Wang, M. S. Roberts, F. J. Burczynski, *J. Lipid Res.* **2009**, *50*, 2445–2454.
- [9] S. Odani, Y. Namba, A. Ishii, T. Ono, H. Fujii, *J. Biochem.* **2000**, *128*, 355–361.
- [10] C. Vasileiou, S. Vaezeslami, R. M. Crist, M. Rabago-Smith, J. H. Geiger, B. Borhan, *J. Am. Chem. Soc.* **2007**, *129*, 6140–6148.
- [11] G. Leone, U. Giovanella, W. Porzio, C. Botta, G. Ricci, *J. Mater. Chem.* **2011**, *21*, 12901–12909.
- [12] D. Nichesola, M. Perduca, S. Capaldi, M. E. Carrizo, P. G. Righetti, H. L. Monaco, *Biochemistry* **2004**, *43*, 14072–14079.
- [13] G. P. Tochtrop, K. Richter, C. Tang, J. J. Toner, D. F. Covey, D. P. Cistola, *Proc. Natl. Acad. Sci. USA* **2002**, *99*, 1847–1852.
- [14] C. Cogliati, L. Ragona, M. D'Onofrio, U. Gunther, S. Whittaker, C. Ludwig, S. Tomaselli, M. Assfalg, H. Molinari, *Chemistry* **2010**, *16*, 11300–11310.
- [15] G. P. Tochtrop, G. T. DeKoster, D. F. Covey, D. P. Cistola, *J. Am. Chem. Soc.* **2004**, *126*, 11024–11029.
- [16] C. Cogliati, S. Tomaselli, M. Assfalg, M. Pedo, P. Ferranti, L. Zetta, H. Molinari, L. Ragona, *FEBS J.* **2009**, *276*, 6011–6023.
- [17] A. Murai, M. Furuse, K. Kitaguchi, K. Kusumoto, Y. Nakanishi, M. Kobayashi, F. Horio, *Comp. Biochem. Physiol., Part A: Mol. Integr. Physiol.* **2009**, *154*, 216–223.
- [18] J. Krishnamoorthy, V. C. Yu, Y. K. Mok, *PLoS One* **2010**, *5*, e8943.
- [19] A. Medek, P. J. Hajduk, J. Mack, S. W. Fesik, *J. Am. Chem. Soc.* **2000**, *122*, 1241–1242.
- [20] J. L. Sudmeier, E. L. Ash, U. L. Gunther, X. Luo, P. A. Bullock, W. W. Bachovchin, *J. Magn. Reson., Ser. B* **1996**, *113*, 236–247.
- [21] G. G. Krivov, M. V. Shapovalov, R. L. Dunbrack, Jr., *Proteins* **2009**, *77*, 778–795.
- [22] J. S. Richardson, *Adv. Protein Chem.* **1981**, *34*, 167–339.
- [23] J. M. Thornton, *J. Mol. Biol.* **1981**, *151*, 261–287.
- [24] M. A. Wouters, S. W. Fan, N. L. Haworth, *Antioxid. Redox Signaling* **2010**, *12*, 53–91.
- [25] L. Ragona, M. Catalano, M. Luppi, D. Cicero, T. Eliseo, J. Foote, F. Fogolari, L. Zetta, H. Molinari, *J. Biol. Chem.* **2006**, *281*, 9697–9709.
- [26] F. Delaglio, S. Grzesiek, G. W. Vuister, G. Zhu, J. Pfeifer, A. Bax, *J. Biomol. NMR* **1995**, *6*, 277–293.
- [27] B. A. Johnson, *Methods Mol. Biol.* **2004**, *278*, 313–352.
- [28] P. Guntert, *Methods Mol. Biol.* **2004**, *278*, 353–378.
- [29] Y. Shen, F. Delaglio, G. Cornilescu, A. Bax, *J. Biomol. NMR* **2009**, *44*, 213–223.
- [30] C. Dominguez, R. Boelens, A. M. Bonvin, *J. Am. Chem. Soc.* **2003**, *125*, 1731–1737.
- [31] A. T. Brünger, P. D. Adams, G. M. Clore, W. L. DeLano, P. Gros, R. W. Grosse-Kunstleve, J. S. Jiang, J. Kuszewski, M. Nilges, N. S. Pannu, R. J. Read, L. M. Rice, T. Simonson, G. L. Warren, *Acta Crystallogr., Sect. D: Biol. Crystallogr.* **1998**, *54*, 905–921.
- [32] D. Eufri, A. Sironi, *J. Mol. Graphics* **1989**, *7*, 165–169, 158.
- [33] J. P. Linge, M. A. Williams, C. A. Spronk, A. M. Bonvin, M. Nilges, *Proteins* **2003**, *50*, 496–506.
- [34] A. W. Schüttelkopf, D. M. van Aalten, *Acta Crystallogr., Sect. D: Biol. Crystallogr.* **2004**, *60*, 1355–1363.
- [35] C. D. Schwieters, J. J. Kuszewski, N. Tjandra, G. M. Clore, *J. Magn. Reson.* **2003**, *160*, 65–73.
- [36] I. Bertini, D. A. Case, L. Ferella, A. Giachetti, A. Rosato, *Bioinformatics* **2011**, *27*, 2384–2390; AnteChamber available at <http://web-enmr.cerm.unifi.it:8085/load/index>.
- [37] A. C. Wallace, R. A. Laskowski, J. M. Thornton, *Protein Eng.* **1995**, *8*, 127–134.
- [38] Fitting was performed using the McLachlan algorithm, see: A. D. McLachlan, *Acta Crystallogr.* **1982**, *A38*, 871–873 as implemented in the program ProFit: A. C. R. Martin, C. T. Porter, <http://www.bioinf.org.uk/software/profit/>.
- [39] W. L. DeLano, *The PyMOL Molecular Graphics System*, DeLano Scientific, Palo Alto, CA, USA.

Received: July 18, 2011

Revised: December 5, 2011

Published online: February 1, 2012