



UNIVERSITA' DEGLI STUDI DI VERONA

DIPARTIMENTO DI
BIOTECNOLOGIE

DOTTORATO DI RICERCA IN
BIOTECNOLOGIE MOLECOLARI, INDUSTRIALI ED AMBIENTALI

CICLO XXII

**Liver bile acid intracellular transporter:
NMR studies of unfolding, cooperative binding and
interaction with membrane mimetic.**

S.S.D. CHIM/06

Coordinatore: Prof. Roberto Bassi

Tutor: Prof.ssa Henriette Molinari

Dottorando: Dott. Massimo Pedò

Abstract

Bile acids are cholesterol derived molecules which, together with phospholipids, cholesterol and bilirubin, represent the principal constituents of bile. Bile acids are detergent molecules which, in the intestinal tract, help the emulsification of fat, thus improving its absorption by the intestinal epithelia. Recently another important bile acid function has emerged linked to their role in many signalling processes. Some bile acid receptors have been indeed identified and several metabolic disorders, related to impaired bile acid dependent signalling, have been reported. Bile acids undergo, in the body, a recycling pathway between the intestine and the liver, called “enterohepatic circulation”, which assure the recovery of these molecules and their subsequent reutilization. Bile acids transport is achieved, in the enterocytes and in the hepatocytes, mainly through a protein mediated mechanism that involves transmembrane and intracellular transporters.

In this study the intracellular bile acid transporter from chicken liver (cL-BABP) has been characterized through NMR spectroscopy.

cL-BABP is a soluble protein of about 14 KDa with the three-dimensional structure represented by a 10 stranded beta barrel with a helix-turn-helix motive that covers the open end of the barrel.

In the first part of the project the stability and unfolding mechanism of cL-BABP was investigated. The structural stability was analyzed through hydrogen-deuterium (HX) exchange experiments and the ΔG_{op} was estimated. The protein unfolding mechanism was also investigated in a residue specific manner through urea titration NMR experiments and the results compared with the data available for other members of the same protein family.

In a second part of the project a cL-BABP mutant (H98Q), in which the histidine in position 98 was substituted with a glutamine, was produced and analyzed to investigate the role of such residue in triggering the allosteric binding mechanism. The overall structure of the protein was substantially unaffected by the mutation, in spite of a slightly reduced stability. The analysis of the chemical shift changes between the WT and the mutant revealed the presence of long range effects in the 55-59 segment of the protein which makes contacts with one ligand. The analysis of the binding mode, investigated both through NMR and mass spectrometry (MS), indicated some disruption of binding cooperativity as evidenced by the reduced binding affinity for one ligand. The loss of affinity for one ligand was also confirmed by the comparison of the ligand diffusion coefficients (D) in the complexes of WT and H98Q. The protein mobility was also investigated and the NMR relaxation parameters, T1 and T2, were measured. The main effect of the mutation resulted in the quenching of the slow motions in the C-terminal region of the WT protein.

The last part of the project was focused on the NMR characterization of the interactions of cL-BABP with anionic vesicles in the presence of the physiological ligand, i.e. glycochenodeoxycholic acid (GCDA). Initially the association with anionic vesicles, i.e., 1,2-dimyristoyl-sn-glycero-3-phospho-(1'-rac-glycerol) (DMPG) liposomes, was analyzed in the absence of bile acids and the electrostatic nature of the interaction was confirmed. Then the ternary complex cL-BABP/DMPG/GCDA was studied. The analysis of several NMR titration experiments clearly indicated that cL-BABP associates with DMPG liposomes in the apo form, while such interaction is weakened in the holo form. Indeed the addition of the physiological ligand to the protein-liposome mixture modulated this interaction, shifting the equilibrium towards the free folded holo protein. Thus membrane and ligands were shown to establish competing binding equilibria for the interaction with cL-BABP. These results supported a mechanism of ligand binding and release controlled by the onset of a bile salt concentration gradient within the polarized hepatic cell.

The data obtained for the H98Q mutant allowed to clarify the key role of His98 in the modulation of the mobility observed in cL-BABP and its importance for an efficient cooperative bile acids binding. Finally, on the basis of the results of the interaction with the liposomes, it has been proposed model for the intracellular bile acids transport in the hepatocytes that takes into account the intracellular bile acid concentration gradient.

Riassunto

Gli acidi biliari sono molecole derivanti dal colesterolo le quali, assieme con fosfolipidi, colesterolo e bilirubina rappresentano i principali costituenti della bile. Gli acidi biliari sono molecole detergenti le quali, nel tratto intestinale, favoriscono l'emulsione dei grassi migliorandone così l'assorbimento da parte dell'epitelio intestinale. Recentemente è stata scoperta un'altra importante funzione degli acidi biliari legata al loro ruolo in molti processi di signalling. Alcuni recettori degli acidi biliari sono stati identificati e sono stati riportati diversi disordini metabolici legati a difetti negli eventi di segnalazione dipendente da acidi biliari. Gli acidi biliari sottostanno, nel corpo, ad un processo di ricircolo tra l'intestino ed il fegato chiamato "circolazione enteroepatica" il quale garantisce il recupero ed il successivo riutilizzo di queste molecole. Il trasporto degli acidi biliari è realizzato, negli enterociti e negli epatociti, principalmente attraverso un meccanismo mediato da proteine trasportatrici e coinvolge trasportatori trans-membrana e intracellulari.

In questo studio è stato caratterizzato, attraverso spettroscopia NMR, il trasportatore intracellulare di acidi biliari di fegato di pollo (cL-BABP).

cL-BABP è una proteina solubile di circa 14 KDa la cui struttura tridimensionale è costituita da un barile beta con 10 filamenti e un motivo helix-turn-helix che copre la regione aperta del barile.

Nella prima parte del progetto è stata studiata la stabilità ed il meccanismo di unfolding di cL-BABP. La stabilità strutturale è stata analizzata attraverso esperimenti di scambio idrogeno/deuterio (XH) ed è stato stimato il ΔG_{op} . Il meccanismo di unfolding è stato studiato in maniera residuo specifica attraverso esperimenti NMR di titolazione con urea e i risultati sono stati confrontati con i dati disponibili per altri membri della stessa famiglia proteica.

Nella seconda parte del progetto è stato prodotto e analizzato un mutante di cL-BABP (H98Q) in cui l'istidina in posizione 98 è stata sostituita con un residuo di glutammina allo scopo di studiare il ruolo di tale residuo nell'avviare il meccanismo di legame allosterico. La struttura globale della proteina è risultata sostanzialmente inalterata dalla mutazione, nonostante la stabilità risulti leggermente ridotta. L'analisi delle variazioni dei chemical shift tra la proteina WT ed il mutante ha rivelato la presenza di effetti a lunga distanza nel segmento 55-59 il quale risulta in contatto con un ligando. L'analisi della capacità di legame, studiato attraverso NMR e spettrometria di massa (MS), ha indicato una rottura del meccanismo di legame cooperativo evidente come una riduzione dell'affinità di legame per un ligando. La perdita di affinità per un ligando è stata confermata anche attraverso il confronto del coefficiente di diffusione (D) per i ligandi tra WT e H98Q. Nel corso del progetto è stata inoltre studiata la mobilità attraverso la misura NMR dei parametri di rilassamento T_1 e T_2 . L'effetto principale della sostituzione è stato lo smorzamento dei moti lenti osservati nella regione C-terminale nella proteina WT.

L'ultima parte del progetto è stata focalizzata sulla caratterizzazione, via NMR, dell'interazione di cL-BABP con vescicole anioniche in presenza del ligando fisiologico, cioè acido glicochenodeossicolico (GCDA). Inizialmente l'associazione con vescicole anioniche, cioè liposomi di 1,2-dimyristoyl-sn-glycero-3-phospho-(1'-rac-glycerol) (DMPG), è stata analizzata in assenza di acidi biliari ed è stata confermata la natura elettrostatica dell'interazione. Successivamente è stato studiato il complesso ternario cL-BABP/DMPG/GCDA. L'analisi di diversi esperimenti NMR di titolazione ha indicato che cL-BABP si associa con liposomi di DMPG nella condizione apo, mentre l'interazione è indebolita nella forma olo. Infatti l'aggiunta del ligando fisiologico alla miscela proteina-liposoma modula l'interazione, spostando l'equilibrio verso la forma olo libera della proteina. Quindi membrana e ligandi sono in competizione per l'interazione con cL-BABP. Questi risultati supportano un meccanismo di legame e rilascio dei ligandi controllato dal gradiente di concentrazione di sali biliari presente nelle cellule epatiche polarizzate.

I dati ottenuti dallo studio di H98Q hanno permesso di chiarire il ruolo chiave dell'istidina 98 nella modulazione della mobilità osservata in cL-BABP e la sua importanza nel determinare un efficiente legame cooperativo degli acidi biliari.

È stato infine proposto, sulla base dei dati di interazione con liposomi, un meccanismo per il trasporto intracellulare degli acidi biliari nelle cellule epatiche che tiene conto del gradiente intracellulare di acidi biliari.

Acknowledgment

Thanks to everyone I worked with during my PhD in the NMR group of the University of Verona: Prof.ssa Henriette Molinari, Michael Assfalg, Mariapina D'Onofrio, Alejandro Giorgetti, Serena Zanzoni, Mara Guariento, Alberto Ceccon, Silvia Verzini, Rudi Remelli. Thanks also to the people of ISMAC of Milano with which I collaborated: Lucia Zetta, Laura Ragona, Simona Tomaselli and Clelia Cogliati.

Thanks to Prof. Volker Dötsch and Frank Löhr and to all the people I met at the Institute of Biophysical Chemistry of the Wolfgang Goethe University of Frankfurt. A special thank to Solmaz, Michael, Mahdi and Lei.

Thanks to all the people of the Biocristallography group of the University of Verona with which I had a lot of funny discussions: Emanuele, Gianmaria, Michele, Beniamino, Laura, Manuel.

Thanks to my parents Giovanni and Maria Antonietta and to my brother Marco.

Finally the biggest thank to Sara.

Contents

Abstract/Riassunto	1
Acknowledgement	5
Contents	6
Abbreviations	10
CHAPTER 1: Introduction	12
1.1 Bile acid chemical structure and physical properties	13
1.2 Bile acid biosynthesis	14
1.2.1 The classical pathway	14
1.2.2 The alternative pathway	16
1.3 Bile acid biosynthesis regulation	16
1.3.1 Nuclear receptors	17
1.3.2 Regulation of key genes	17
1.3.2.1 CYP7A1	17
1.3.2.2 CYP8B1	17
1.3.2.3 CYP27A1	18
1.3.2.4 CYP7B1	18
1.4 Enterohepatic circulation and bile acid transport	18
1.4.1 Bile acid transport in the hepatocytes	18
1.4.1.1 Basolateral transport	18
1.4.1.1.1 Na ⁺ -dependent bile acid transport	19
1.4.1.1.2 Na ⁺ -independent bile acid transport	19
1.4.1.2 Bile acid efflux	19
1.4.1.3 Intracellular transport	21
1.4.1.4 Canalicular transport	21
1.4.1.4.1 Bile Salt Export Pump	21
1.4.1.4.2 Multidrug Resistance Protein	21
1.4.2 Bile acid transport in the enterocytes	21
1.4.2.1 Apical transport	22
1.4.2.1.1 Passive diffusion	22
1.4.2.1.2 Na ⁺ -dependent uptake	22
1.4.2.1.3 Na ⁺ -independent uptake	22
1.4.2.2 Intracellular transport	22

1.4.2.3	Basolateral efflux	22
1.5	Focus on intracellular bile acid transport proteins	23
1.5.1	The fatty acid-binding proteins (FABPs) family	23
1.5.2	The birth of Bile Acid-Binding Proteins	24
1.5.3	The liver Bile Acid-Binding Proteins	25
1.5.4	The ileal Bile Acid-Binding Proteins	27
1.6	Bile acid functions and regulation	28
1.6.1	Bile acid homeostasis	28
1.6.2	Bile acids modulate glucose homeostasis	29
1.6.3	Bile acid effects on lipid homeostasis	29
1.6.4	Bile acid increase energy expenditure	29
1.6.5	Further endocrine effects of bile acids	30
1.6.6	Paracrine role of bile acids	30
1.7	Bile acids related damage	30
1.7.1	Bile acid induced hepatocyte injury	30
1.7.2	Bile acid induced oxidative stress	30
1.7.3	Bile acid induced cell death pathways	31
1.7.3.1	Bile acid induced necrosis	31
1.7.3.2	Bile acid induced apoptosis	31
1.7.4	Bile acid induced damage in the plasma membrane	32
1.7.5	Bile acid induced injury in other cell types	32
CHAPTER 2:	Materials & methods	33
2.1	Materials	34
2.1.1	Chemicals	34
2.1.2	DNA coding sequence	34
2.1.3	Primers	34
2.1.4	Plasmid	34
2.1.5	Bacterial strains	35
2.1.6	Chromatography resins and columns	35
2.1.7	Instruments	36
2.2	Culture growing media	36
2.3	Competent cells	38
2.4	Heat shock transformation	39

2.5	Plasmid DNA extraction	39
2.6	Site-direct mutagenesis	39
2.7	DNA sequencing	40
2.8	Protein expression	40
2.8.1	Preliminary protein expression evaluation	40
2.8.2	Protein solubility test	40
2.8.3	Large-scale soluble expression	41
2.9	Protein purification	41
2.9.1	Anion exchange chromatography	42
2.9.2	Size exclusion chromatography	42
2.9.3	Hydrophobic interaction chromatography (delipidation)	42
2.9.4	SDS-PAGE	43
2.9.5	Estimation of protein concentration	44
2.9.6	Buffer exchange	44
2.9.7	Sample storage	44
2.9.8	Sample quality control	44
2.10	NMR analysis	45
2.10.1	Sample preparation	45
2.10.2	Data collection	45
2.10.3	Data acquisition, processing and analysis	45
CHAPTER 3:	Protein NMR spectroscopy	46
3.1	The NMR phenomenon	47
3.1.1	Magnetic dipole and energy levels	47
3.1.2	Population distribution	49
3.1.3	Bulk magnetization and Larmor precession	49
3.1.4	Radio frequency pulses and signal detection	50
3.2	Chemical shift	51
3.3	Scalar coupling	51
3.4	Relaxation	53
3.4.1	Time dependent field fluctuations	53
3.4.2	Longitudinal and transverse relaxation	55
3.4.3	Nuclear Overhauser Effect	57
3.4.4	Relaxation and protein mobility	60

3.5	NMR experiments	61
3.5.1	1D NMR experiments	61
3.5.2	Multi-dimensional NMR experiments	61
3.6	NMR experiments in protein studies	63
3.6.1	Two-dimensional homonuclear J-correlated spectroscopy	63
3.6.2	Two-dimensional heteronuclear J-correlated spectroscopy	64
3.6.3	Two-dimensional homonuclear dipolar correlated spectroscopy	65
3.6.4	Heteronuclear relaxation experiments	66
3.6.5	Transverse relaxation-optimized spectroscopy	67
CHAPTER 4:		69
	Short presentation of publication 1	70
	NMR unfolding studies on a liver bile acid binding protein reveal a global two-state unfolding and localized singular behaviours.	71
	Supplementary of publication 1	80
CHAPTER 5:		86
	Short presentation of publication 2	87
	Towards the elucidation of molecular determinants of cooperativity in the liver bile acid binding protein.	89
	Supplementary of publication 2	103
CHAPTER 6:		106
	Short presentation of publication 3	107
	Disulfide bridge regulates ligand-binding site selectivity in liver bile acid-binding proteins.	109
	Supplementary of publication 3	122
CHAPTER 7:		124
	Short presentation of publication 4	125
	NMR Studies Reveal the Role of Biomembranes in Modulating Ligand Binding and Release by Intracellular Bile Acid Binding Proteins	127
	Supplementary of publication 4	139
CHAPTER 8:		140
	Conclusions	141
	Bibliography	143
	Figure list	152
	Table list	154
	Appendices	155

Abbreviations

1D	Mono – dimensional
2D	Bi – dimensional
3D	Three – dimensional
3 α -HSD	3 α -hydroxysteroid dehydrogenase
ABC	ATP-binding cassette
AKR1C4	3 α -hydroxysteroid dehydrogenase
AKR1D1	Δ 4-3-oxosteroid-5 β -reductase
ApoB	Apolipoprotein B
ASBT	Apical sodium dependent bile acid transporter
BA	Bile acid
BAAT	bile acid CoA: amino acid N-acyltransferase
BABP	Bile acid-binding protein
BACS	Bile acid CoA synthetase
BARE-I	Bile acid response elements - I
BARE-II	Bile acid response elements - II
BSEP	Bile salt export pump
CA	Cholic acid
cAMP	Cyclic AMP
CDCA	Chenodeoxycholic acid
cL-BABP	Chicken liver bile acid-binding protein
CMC	Critical micellar concentration
COSY	Correlated spectroscopy
CPMG	Carr-Purcell-Meiboom-Gill
CRABP	Cellular retinoic acid binding proteins
CRBP	Cellular retinol binding protein
CSA	Chemical shift anisotropy
CYP27A1	sterol 27-hydroxylase
CYP7A1	cholesterol 7 α -hydroxylase
CYP8B1	sterol 12 α -hydroxylase

CYP7B1	oxysterol 7 α -hydroxylase
D	Diffusion coefficient
DEAE	Diethylaminoethyl cellulose
DHCA	3 α ,7 α -dihydroxy-5 β -cholestanoic acid
DQF-COSY	Double quantum filtered COSY
ER	Endoplasmic reticulum
FABP	Fatty acid binding protein
FID	Free induction decay
FT	Fourier transformation
FXR	Farnesoid X receptor
GCDA	Glycochenodeoxycholic acid
GSH	Glutathione (reduced state)
HepG2	Human hepatocellular liver carcinoma cell line
hI-BABP	Human ileal bile acid-binding protein
HSD3B7	3 β -hydroxy- Δ 5-C27-steroid dehydrogenase/isomerase
HSQC	Heteronuclear single quantum correlation
I-BABP	Intestinal bile acid-binding protein
IBAT	Ileal bile acids transporter
IL-6	Interleukine – 6
iLBP	Intracellular lipid binding protein
INEPT	Insensitive nuclei enhanced by polarization transfer
IPTG	Isopropyl- β -D-thiogalactopyranosid
Kan	Kanamycin
LB	Luria Bertani medium
L-BABP	Liver bile acid binding protein
Lb-FABP	Liver basic fatty acid binding protein
LCFA	Long chain fatty acids

L-FABP	liver fatty acid binding protein
LUV	Large unilamellar vesicles
LXR	Liver X receptor
mEH	Microsomal epoxide hydrolase
MM	Minimal medium
MPT	Mitochondrial permeability transition
MRP	Multidrug resistance protein
MTP	Microsomal triglyceride transfer protein
MS	Mass spectrometry
MW	Molecular weight
NMR	Nuclear magnetic resonance
NOE	Nuclear Overhauser effect
NOESY	Nuclear Overhauser effect spectroscopy
NTCP	Na ⁺ -dependent taurocholic cotransporting polypeptide
OATP	Organic anion transporting polypeptides
OD	Optical density
Ost α/β	Organic solute transporter α/β
PCR	Polymerase chain reaction
PDB	Protein data bank
PDK4	Pyruvate dehydrogenase kinase-4
PEPCK	phosphoenolpyruvate carboxykinase

PMSF	phenylmethanesulphonyl fluoride
PPAR α	Peroxisome proliferators activated receptor α
ppm	Parts per million
RF	Radio frequency
ROS	Reactive oxygen species
SDS-PAGE	Sodium dodecyl sulphate polyacrilamide gel electrophoresis
SOFAST	Band-selective optimized flip-angle short transient
SREBP	sterol-regulatory-element binding protein
SXR	Steroid and xenobiotic receptor
TEMED	tetramethylethylenediamine
TGR5	G-protein coupled bile acid receptor
THCA	3 α ,7 α ,12 α -trihydroxy-5 β -cholestanoic acid
TOCSY	Total correlated spectroscopy
TROSY	Transverse relaxation optimized spectroscopy
VDR	Vitamin D receptor
VLCS	Very long chain acyl CoA synthetase
VLDL	Very low density lipoproteins
WT	Wild type

CHAPTER 1: Introduction

1.1 Bile acid chemical structure and physical properties

Bile acids are acidic steroids synthesized in the liver starting from cholesterol. Bile acids have a molecular structure based on a saturated tetracyclic hydrocarbon perhydrocyclopentanophenanthrene system, usually known as the steroid nucleus. The scaffold with a C₅ side chain on the carbon 17 is termed the *cholane* nucleus, and the bile acid without any hydroxylic substituents is termed cholanic acid⁴. In higher vertebrates bile acids have 24 carbon atoms in contraposition to “primitive” bile acids which have 25-27 carbon atoms which are present in primitive vertebrates, reptiles and amphibians. The structures of some of the most abundant bile acids in humans are depicted in Figure 1.1A⁵.

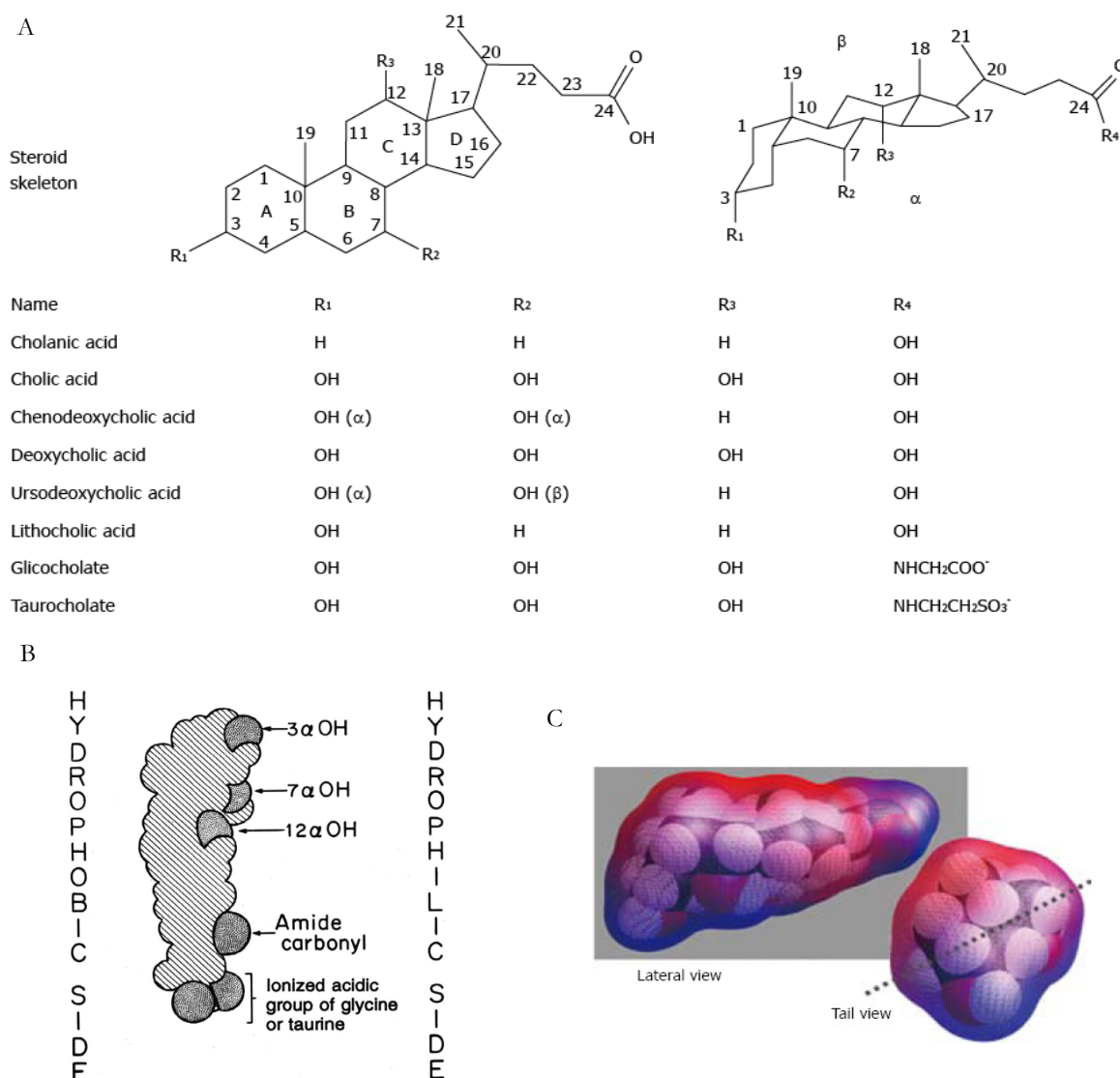


Figure 1.1 A: Structure of the most abundant bile acids in humans and their glycine and taurine conjugates⁵; B: Space filling model of the taurine conjugate of cholic acid showing its planar amphipathic structure with the hydrophilic side and the hydrophobic side¹⁰. C: Calculated molecular lipophilic potential⁵. Blue colour shows polar surface and red colour shows apolar surface.

The steroid nucleus consists of three six-member rings (A, B and C) and a five-member ring (D), with a curved (beaked) or flat structure (depending on a *cis*- or *trans*-fused configuration between the A and B rings). The common bile acids in higher vertebrates are all 5β i.e. A/B ring junction is in *cis* configuration⁴. Bile acid molecules are approximately 20 Å long, with an average radius of about 3.5 Å. Bile acids are rigid molecules with a hydrophilic side (α -side) and a hydrophobic side (β -side), and a negative charge at the end of the molecule (Figure 1.1B)¹⁰. The negative charge derives from the carboxylic group present on the C₂₄ carbon atom which has a pKa of about 5. The ionization of this group is responsible for the water solubility of bile acids which are poorly soluble in the protonated state⁴. In Figure 1.1C is shown the calculated lipophylic potential pointing out the anphipathic nature of bile acids⁵ which derives from the presence of hydrohyl groups (in mammals commonly in positions 3, 7 and/or 12) oriented toward the α -side and methyl groups (at C-18 and C-19) oriented towards the β -side. Bile acids can form small aggregates or micelles (usually with less than 10 monomers) in aqueous solution when their concentration is above the so called critical micellar concentration (CMC)¹¹.

1.2 Bile acid biosynthesis

Bile acids are synthesized in the liver cells (hepatocytes) through an enzymatic process that involves 17 different enzymes¹². The bile acid biosynthesis process starts from cholesterol and represents the major process by which cholesterol is consumed, indeed approximately 500 mg of cholesterol is converted into bile acids each day in the adult human liver. The direct products of this process are the so called “primary bile acids”. Primary bile acids can be converted into secondary and tertiary bile acids by anaerobic bacteria present in the gut¹³. Two main biosynthetic pathways have been described accounting for bile acid biosynthesis: the “classical pathway” and the “alternative pathway”, summarized in Figure 1.2. The classical pathway accounts for at least 75% of the total bile acid pool, therefore it represents the major bile acid biosynthetic process. Moreover some other routes have been recognized which can be important in some species and conditions⁵.

1.2.1 The classical pathway

The classical pathway, present exclusively in the liver, is also known as the “neutral” pathway because the intermediate metabolites are neutral sterols. The enzymes involved are located in the cytosol, microsomes, mitochondria and peroxisomes⁵. The process, shown in the left panel of Figure 1.2 starts in the microsomes with the conversion of cholesterol into 7α -hydroxycholesterol by cholesterol 7α -hydroxylase (CYP7A1). 7α -hydroxycholesterol is converted to 7α -hydroxy-4 cholesten-3-one by the 3β -hydroxy- Δ^5 -C27-steroid dehydrogenase/isomerase (HSD3B7), which is also microsomal. After this the enzyme sterol 12 α -hydroxylase (CYP8B1), another microsomal enzyme, catalyzes the hydroxylation of 7α -hydroxy-4 cholesten-3-one at the C-12 position. The next step is the reduction of the double bond

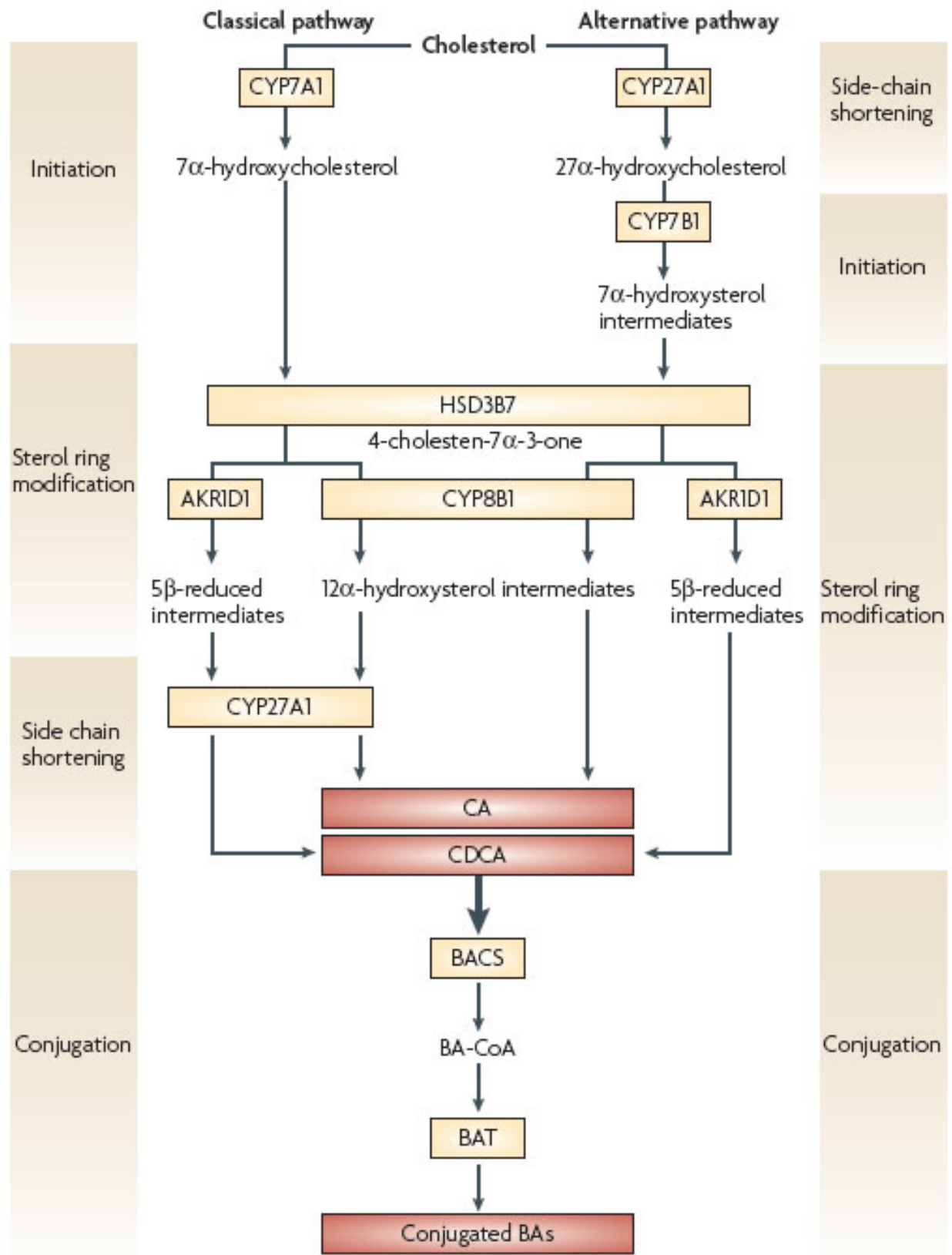


Figure 1.2 Bile acids biosynthesis.

The 4 principal steps involved in the bile acids biosynthesis: initiation, sterol ring modification, side chain shortening and conjugation are highlighted. The principal enzymes and the intermediate metabolites are also indicated³. The main products of these pathways are cholic acid (CA) and chenodeoxycholic acid (CDCA) which are finally extensively conjugated (98%) with glycine or taurine amino acids.

carried out by two cytosolic enzymes: Δ^4 -3-oxosteroid-5 β -reductase (AKR1D1) and 3 α -hydroxysteroid dehydrogenase (AKR1C4). The resulting molecules are 5 β -cholestan-3 α ,7 α -diol or 5 β -cholestan-3 α ,7 α ,12 α -triol, the precursors of chenodeoxycholic acid (CDCA) and cholic acid (CA), respectively. In the mitochondrion the enzyme sterol 27-hydroxylase (CYP27A1) oxidizes the side-chain of these precursors by introducing a hydroxyl group to the C-27 position. This group will be subsequently oxidized to an aldehyde and then to a carboxylic acid. The resulting acids 3 α ,7 α -dihydroxy-5 β -cholestanoic acid (DHCA) and 3 α ,7 α ,12 α -trihydroxy-5 β -cholestanoic acid (THCA) are then activated by conjugation with coenzyme A by bile acid CoA synthetase (BACS) or very long chain acyl CoA synthetase (VLCS), both localized at the endoplasmic reticulum. The resulting cholestanoyl-CoAs are then transported into peroxisomes where the side-chain is shortened by β -oxidation, a process that involves the action of four peroxisomal enzymes. The final step in bile acid biosynthesis is the conjugation of the carboxylic group with the amino acids glycine or taurine. This reaction is catalyzed by the enzyme bile acid CoA: amino acid N-acyltransferase (BAAT)¹² and increases the water solubility of bile acids. C₂₄ bile acids constitute a major part of the bile in higher vertebrates and in humans and these compounds are almost completely in conjugated form with glycine (75%) or taurine (25%)¹⁴.

1.2.2 The alternative pathway

In the “alternative pathway” side-chain oxidation of cholesterol precedes steroid ring modification as depicted in Figure 1.2 right panel. Acidic intermediate metabolites are formed during the process, therefore this route is also known as “acidic” pathway. Initially cholesterol is oxidized to 27-hydroxycholesterol and 3 β -hydroxy-5-cholestenoic acid by sterol 27-hydroxylase (CYP27A1). These two compounds are then converted to 7 α ,27-dihydroxycholesterol and 3 β ,7 β -dihydroxy-5-cholestenoic acid by oxysterol 7 α -hydroxylase (CYP7B1). CYP27A1 and CYP7B1 are microsomal enzymes and CYP7B1 is acidic pathway specific¹⁵. Subsequently the intermediate compounds follow the same route as in the classical pathway with the involvement of the enzymes HSD3B7, CYP8B1 and AKR1D1 who catalyze the formation of CDCA and CA. Finally the enzymes BACS or VLCS catalyze the activation of bile acids with coenzyme A and the enzyme BAAT conjugates the activated bile acids with the amino acid glycine or taurine.

1.3 Bile acid biosynthesis regulation

The rate of bile acid synthesis, bile acid composition, and bile acid pool size vary significantly depending on the species, sex, genetic, pathophysiological conditions and environmental factors such as diets and drugs¹⁶. Many nuclear receptors have been found to play pivotal roles in regulating transcription of the genes involved in bile acid synthesis¹⁷. In general hydrophobic bile acids (CA and

CDCA) are potent inhibitors of bile acid biosynthesis, whereas hydrophilic bile acids, such as 7 β -bile acids, e.g. ursodeoxycholic acid and β -muricholic acids are not.

1.3.1 Nuclear receptors

Nuclear receptors are essential proteins crucial for the regulation of biologically important processes in development and homeostasis. They act as transcription factors that can regulate the expression of many genes upon activation by interaction with specific ligands¹⁸.

In the bile acid related metabolism we can find three nuclear receptors who can be activated by bile acids. Farnesoid X receptor (FXR) regulates bile acid synthesis, transport and absorption¹⁴. Steroid and xenobiotic receptor (SXR) regulates lithocholic acid and drug metabolism¹⁴. Vitamin D receptor (VDR) regulates calcium and phosphate homeostasis¹⁶. Bile acid metabolites can also activate liver X receptor α (LXR α) which is an oxysterol receptor that plays a central role in lipid metabolism¹⁹. These nuclear receptors regulate key bile acid biosynthetic genes and have been referred to as the 'metabolic receptors' that coordinately regulate a network of genes involved in integrated control of energy metabolism, bile acid metabolism, lipoprotein metabolism and triglyceride metabolism²⁰.

1.3.2 Regulation of key genes

1.3.2.1 CYP7A1

In the CYP7A1 promoter two potential transcription factor-binding regions, named bile acid response elements (BARE-I and BARE-II), have been identified. These two conserved sequences are essential for basal transcription and for bile acid inhibition¹⁹. Bile acid activated FXR inhibits CYP7A1 gene transcription. Although FXR response element has been mapped onto the BARE-II element, FXR does not bind to this element and apparently inhibits the CYP7A1 gene by an indirect mechanism²¹. CYP7A1 activity and mRNA expression show a regulation on circadian rhythm. The expression is highest at midnight and decreases during the day¹⁸. Insulin is also known to inhibit bile acid synthesis by inhibiting CYP7A1 gene transcription¹⁹. The identification of a mutation in the CYP7A1 gene in patients with hypercholesterolemia, premature gallstone disease and premature atherosclerosis supports the hypothesis that CYP7A1 plays a critical role in maintaining bile acid and cholesterol homeostasis²⁰.

1.3.2.2 CYP8B1

CYP8B1 is required for synthesis of cholic acid and determines the ratio of CA to CDCA. CYP8B1 gene transcription is also strongly inhibited by bile acids and insulin can inhibit CYP8B1 activity and mRNA levels²⁰. CYP8B1 activity and mRNA show a circadian rhythm. The rhythmicity of CYP8B1 and CYP7A1 together can alter the CA/CDCA ratio and regulate bile acid synthesis.

1.3.2.3 CYP27A1

A defect in CYP27A1 leads to excessive accumulation of 7α -hydroxycholesterol, 7α -hydroxy-4-cholesten-3-one, 5β -cholestane- $3\alpha,7\alpha,12\alpha$ -triol, cholesterol, and cholestanol²¹. The Cyp27a1 $-/-$ mice have enlarged livers and kidneys, and have increased triglyceride levels, SREBP expression, fatty acid synthesis, cholesterol absorption, and cholesterol synthesis²³. These phenotypes suggest that CYP27A1 may play a major role in cholesterol and triglyceride metabolism. In addition insulin suppresses CYP27A1 gene transcription to reduce bile acid synthesis and pool size¹⁹.

1.3.2.4 CYP7B1

A CYP7B1 mutation has been identified and patients with this mutation show a severe defect in bile acid synthesis, neonatal cholestasis, cirrhosis, high level of 24-, 25- and 27-hydroxycholesterol accumulation in the serum and 3β -hydroxy-5-cholestenoic and 3β -hydroxy-5- cholenoic acids in the urine²⁴. This finding supports the hypothesis that a major function of CYP7B1 may be to detoxify oxysterols. It is possible that CYP7B1 is expressed very early in life, when CYP7A1 is not expressed, to produce bile acids via the alternative pathways.

1.4 Enterohepatic circulation and bile acid transport

Bile acids derive from the enzymatic catabolism of cholesterol in the liver. Upon synthesis bile salts are secreted into the gallbladder where they compose the bile together with bilirubin, phospholipids, cholesterol, fatty acids, proteins and several ions. After food ingestion gallbladder content flush into the small intestine and here bile salts act as detergent molecules emulsifying dietary fats and lipid-soluble vitamins therefore enhancing the absorption of several water insoluble molecules. Bile salts are efficiently reabsorbed in the ileum and returned to the liver in the portal circulation through the cycle of enterohepatic circulation. Back in the liver bile salts are secreted again in the bile.

After biosynthesis bile acids are conjugated with amino acids glycine or taurine. Conjugated bile acids have a pKa of approximately 4 (for glycoconjugates) and 2 (for tauroconjugates) and at physiological pH they are present as membrane impermeable anionic salts. From this derives the common name of “bile salts”. Because of their membrane impermeability the transcellular movement of bile salts is accomplished by specialized transporters localized in a highly polarized manner both in hepatocytes and enterocytes²².

1.4.1 Bile acid transport in the hepatocytes

1.4.1.1 Basolateral transport

Bile acids are circulated in the blood bound to albumin (Figure 1.3). Upon reaching the space of Disse and contacting the hepatocytes basolateral membrane albumin undergoes a conformational change that

facilitate the bile acid dissociation²³. The transport of bile acids from the portal blood to the hepatocytes involves two main processes: Na⁺-dependent and Na⁺-independent uptake.

1.4.1.1.1 Na⁺-dependent bile acid transport

This transport process is driven by an inwardly Na⁺ gradient generated by the activity of Na⁺/K⁺ ATPase²⁴.

The most important Na⁺-dependent transporter is the Na⁺-dependent Taurocholic Cotransporting Polypeptide (NTCP) which has been shown to localize exclusively to the basolateral membrane of the hepatocytes²⁵ (Figure 1.3). NTCP is a 56 KDa transmembrane protein with seven predicted transmembrane domains, an intracellular C-terminal region and an extracellular N-terminal region²⁵.

It has been suggested a role in the Na⁺-dependent bile acid uptake for Microsomal Epoxide Hydrolase (mEH)²⁴. This enzyme is present in two topological forms and one of this, the type II form, has been shown to be targeted to the plasma membrane. Since mEH seems to transport glyco-conjugates more efficiently than NTCP and the circulating bile acids are conjugated with glycine, it is possible that mEH is responsible for a higher uptake of bile acids if compared to NTCP²⁶.

1.4.1.1.2 Na⁺-independent bile acid transport

The Na⁺-independent bile acid transport in hepatocytes is mediated by several members of the Organic Anion Transporting Polypeptides (OATP) superfamily of transporters²⁴ (Figure 1.3). OATPs members involved in bile acid transport are transmembrane glycoproteins with predicted 12 transmembrane domains with both C- and N-terminal region protruding in the cytosol²⁷.

OATP proteins mediate the exchange of extracellular organic anions or bile acids with intracellular HCO₃⁻ or glutathione (GSH)²⁸.

The main human OATP proteins involved in bile acid uptake are OATP1A2, OATP1B1 and OATP1B3. OATP1A2 is mainly expressed in the brain and in the liver and in addition to bile acids is capable of transporting a variety of amphipathic organic compounds including steroid conjugates, thyroid hormones, prostaglandins, peptides and drugs²⁸. OATP1B1 expression is strictly limited to the liver and its transport activity has been suggested to be charge-selective and restricted to organic anions²⁴. OATP1B3 is also expressed specifically in the liver and is able to transport a wide spectrum of substrates²⁸.

1.4.1.2 Bile acid efflux

In physiological condition bile acid uptake (influx) represents the major transport process across the basolateral membrane whereas their efflux is negligible. In some pathological states, such as cholestatic conditions, bile acid efflux can be upregulated to enhance the excretion of bile acids from

hepatocytes²⁹. The process of bile acid efflux across the basolateral membrane of hepatocytes is mediated by members of the multidrug resistance proteins (MRPs) subfamily. Among MRPs subfamily four members are expressed in the liver. MRP3 and MRP4 are localized to the basolateral membrane of the hepatocytes and MRP2 is expressed on the canalicular membrane (Figure 1.3). These transporters function as ATP-dependent pumps with a range of substrates including glucuronide and glutathione conjugates of endogenous and exogenous compounds²⁴.

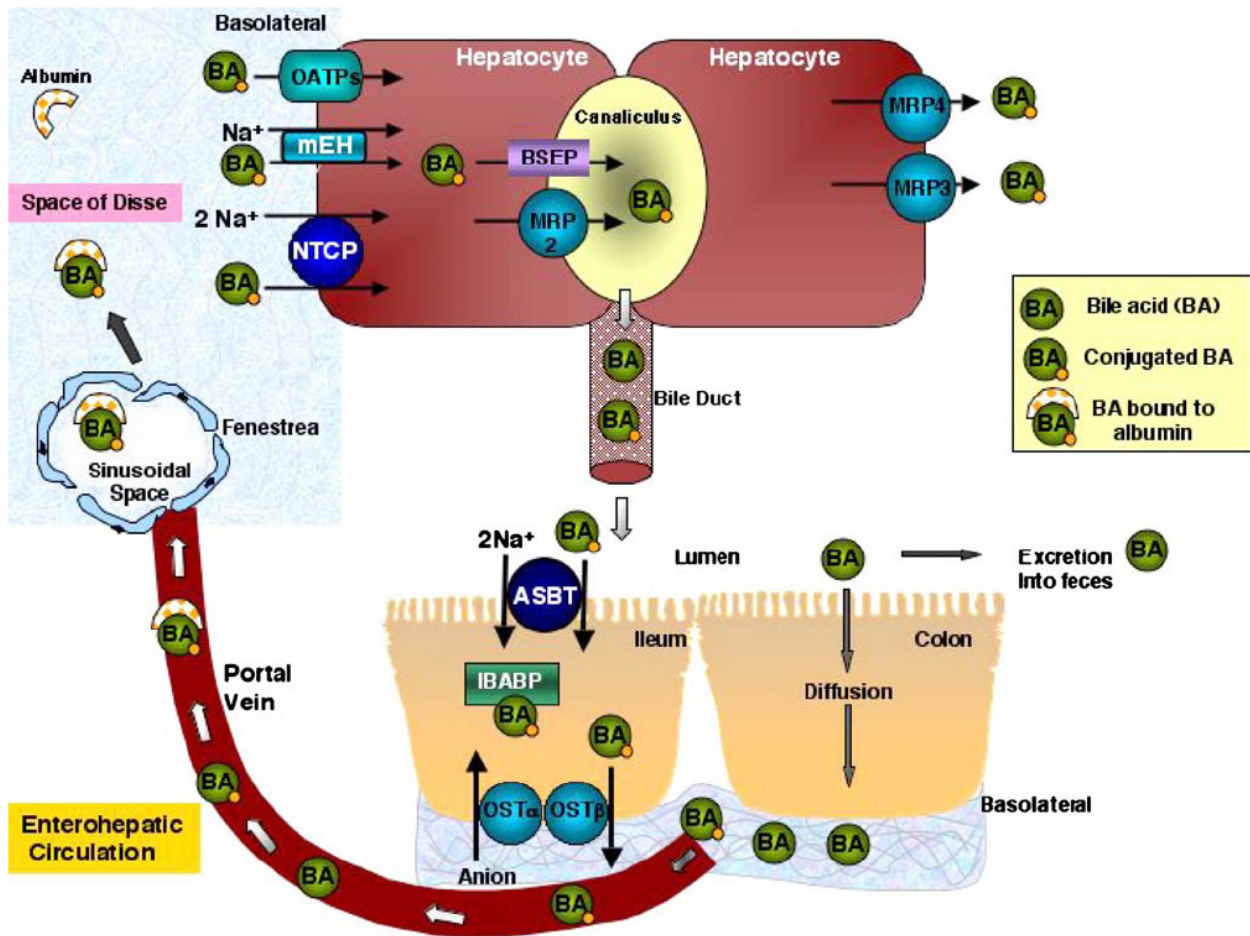


Figure 1.3 Enterohepatic circulation of bile acids. In the picture the principal cytosolic and transmembrane transporters, along with their localization, involved in the bile acid transport mechanism across the body are shown. The conjugation state of bile acids is also indicated¹.

1.4.1.3 Intracellular transport

Movement of bile acids from sinusoidal to canalicular membrane have been proposed to be mediated by two distinct processes: intracellular trafficking and vesicle-mediated transport. In physiological conditions the major intracellular transport process is represented by intracellular trafficking whereas the vesicle-mediated pathway is involved when bile acid load is imposed on the liver²⁹.

The intracellular trafficking is driven by basolateral to canalicular bile acid concentration gradient and is proposed to be mediated by intracellular binding proteins. Several hepatic intracellular bile acid-binding proteins were identified in rat liver including the 3 α -hydroxysteroid dehydrogenase (3 α -HSD), glutathione S-transferase and liver fatty acid-binding protein (L-FABP)^{45, 46}.

1.4.1.4 Canalicular transport

Canalicular excretion of bile acids is mediated by transmembrane transporters whose function depends on ATP hydrolysis²⁴. The principal transporters involved in hepatic bile acid excretion are Bile Salt Export Pump (BSEP) and Multidrug Resistance Protein (MRP2).

1.4.1.4.1 Bile Salt Export Pump

BSEP is mainly responsible for transport of conjugated monovalent bile acids into bile with high affinity for taurochenodeoxycholate > taurocholate > tauroursodeoxycholate > glycocholate⁴². This transmembrane protein is exclusively expressed in the liver and is localized to the canalicular membrane of hepatocytes⁴⁸ (Figure 1.3). BSEP is a member of the multidrug resistance P-glycoprotein that belongs to the ATP-binding cassette (ABC) superfamily of transporters. Similarly to other ABC transporters, BSEP consists of 12 transmembrane domains with two intracellular nucleotide-binding domains for binding and hydrolysis of ATP^{30, 31}.

1.4.1.4.2 Multidrug Resistance Protein

MRP2 is another ABC transporter and is localized in the canalicular membrane of the hepatocytes and in the enterocytes of the duodenum and jejunum³⁰ (Figure 1.3). MRP2 mediates the export of bilirubin conjugates and its substrate profile includes a wide range of organic substrates such as glutathione, glucuronide, sulphate conjugates and some unconjugated drugs³⁰. MRP2 transports divalent bile acids such as sulphated tauro and glycolithocholate but does not transport monovalent bile acids^{40, 51}.

1.4.2 Bile acid transport in the enterocytes

The bile acid transport in the enterocytes consists of three steps: i) apical uptake of conjugated bile acids in the terminal ileum, ii) intracellular bile acid transport mediated by cytosolic intestinal bile acid-

binding protein (I-BABP) and iii) anion exchange mechanism for basolateral efflux of bile acids from enterocytes.

1.4.2.1 Apical transport

1.4.2.1.1 Passive diffusion

Bile acids conjugated with glycine or taurine are impermeable to cell membranes because of the low pKa value. This allows high concentrations to persist in bile and intestinal lumen. Bile acid deconjugation by the intestinal microflora increases the pKa and allows the passive diffusion transport of deconjugated bile acids into the enterocytes³¹ (Figure 1.3).

1.4.2.1.2 Na⁺-dependent uptake

Apical Sodium Dependent Bile Acid Transporter (ASBT) transports efficiently conjugated and unconjugated bile acids with a preference for the taurine and glycine conjugated forms. Moreover it exhibits a higher affinity for dihydroxy bile acids (chenodeoxycholate, CDCA) and deoxycholate compared to trihydroxy bile acids such as cholate, taurocholate and glycocholate⁵³. ASBT mediated bile acid transport is coupled to sodium ion transport, indeed for each bile acid molecule two sodium ions are cotransported³² (Figure 1.3).

1.4.2.1.3 Na⁺-independent uptake

The sodium independent transport of bile acids in the intestine is mediated by the Oatp1a5 transporter for which the human ortholog has been suggested to be the OATP1A2 protein⁴¹. The uptake of bile acids into jejunum brush border membrane vesicles has been characterized and occurs via outward HCO₃⁻ gradient, however is not clear whether this mechanism pertains to Oatp1a5³³.

1.4.2.2 Intracellular transport

The intracellular bile acid transport in the enterocytes is mediated by the Intestinal Bile Acid-Binding Protein (I-BABP) (Figure 1.3). I-BABP belongs to a family of intracellular lipid-binding proteins that bind fatty acids, retinoids, cholesterol and bile acids²⁴. I-BABP is a cytoplasmatic protein of 14-15 KDa that reversibly binds bile acids. Human I-BABP binds two molecules of glycocholate (GCA) with low intrinsic affinity but high degree of positive cooperativity³⁵.

1.4.2.3 Basolateral efflux

The bile acid transport in ileal basolateral membrane was demonstrated to occur via a sodium-independent exchange mechanism. Recently, a transmembrane transporter, the Organic Solute Transporter Ostα-ostβ⁴⁴, responsible for basolateral efflux of bile acids from enterocytes, (Figure 1.3)

has been identified. The transport activity requires the co-expression of two gene products: Ost α , that is a seven transmembrane domain protein and Ost β , a putative single transmembrane domain ancillary polypeptide³⁶.

1.5 Focus on intracellular bile acid transport proteins

It is now well accepted that the main intracellular transport mechanism of bile acids is mediated by cytosolic proteins able to bind bile acids. Several Bile Acid-Binding Proteins (BABP) have been identified in the liver of non mammalian organisms all of which belongs to the intracellular Lipid Binding Protein (iLBPs) family³⁴. Among the putative mammalian hepatic bile acid transport proteins, the liver fatty acid-binding protein (L-FABP) is the most prominent cytoplasmic protein that is labelled by photoreactive bile acid probes³⁸. A recent work confirmed the ability of human L-FABP to bind bile acids supporting therefore the role of L-FABP as bile acid transporter in the human liver³⁵.

For the cytosolic transport of bile acids in the intestine, the role of I-BABP is well established. I-BABP has been shown to be functionally associated with the nuclear receptor FXR and the ileal bile acid transporter IBAT³⁶.

1.5.1 The fatty acid-binding proteins (FABPs) family

Both L-FABP and I-BABP belong to the Fatty Acid-Binding Proteins (FABPs) family. FABPs, together with cellular retinol- and retinoic acid-binding proteins (CRBPs, CRABPs), constitute the iLBP family. iLBPs are cytoplasmic proteins thought to be involved in cellular uptake and intracellular transport of various hydrophobic compounds, such as long-chain fatty acids (LCFAs) and retinoids³⁷. It is important to mention that iLBPs, along with avidins and lipocalins, constitute the calycins superfamily. All the calycins members, although having a low primary sequence homology, share a similar structural fold represented by an antiparallel β -barrel³⁸. iLBPs in particular have a β -barrel structure formed from two five-stranded β -sheets arranged almost perpendicular to each other. The strands are linked by β -turns with

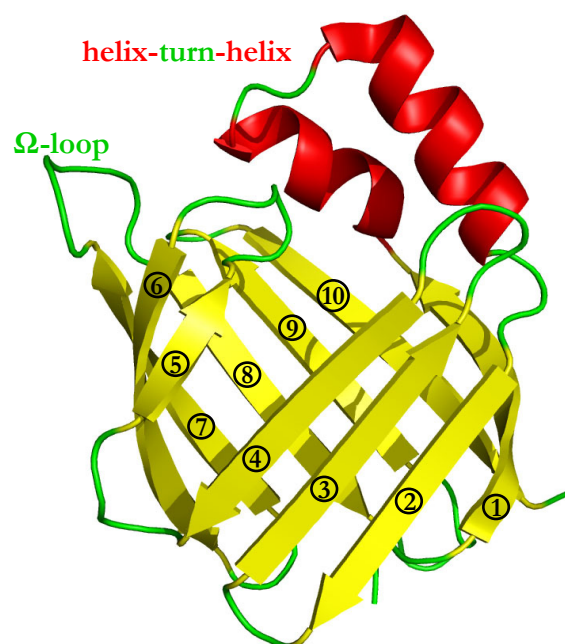


Figure 1.4 Three-dimensional structure of a typical intracellular lipid-binding protein (PDB code 2FS6). The secondary structure elements are in different colour: sheets in yellow, helices in red and loops in green. Sheets are numbered in a progressive way. The peculiar characteristics in the structure are indicated².

the exception of strands 1 and 2, which are connected by an helix-turn-helix motif and strands 7 and 8, which are joined by an Ω -loop of variable length³⁹ (Figure 1.4).

The first FABPs was discovered in 1972⁴⁰ and after this the number of proteins of the FABPs family grew up to reach at least nine members. The members of FABPs family exhibit unique patterns of tissue expression and the name derives from the organ in which the protein has been originally found. The family is composed of fatty acid-binding proteins isolated from liver (L-), intestine (I-), heart (H-), adipocyte (A-), epidermis (E), ileum (Il-), brain (B-), myelin (M-) and testis (T-). However it has to be stressed that FABPs are not expressed specifically in just one tissue since most tissues express several FABP isoforms⁶. Table 1.5 lists the known FABPs and the tissues in which they are present.

Gene	Common name	Alternative names	Expression	Chromosomal location		
				<i>Homo sapiens</i>	<i>Mus musculus</i>	<i>Rattus norvegicus</i>
<i>Fabp1</i>	Liver FABP	L-FABP	Liver, intestine, pancreas, kidney, lung, stomach	2p11	6 C1	4q32
<i>Fabp2</i>	Intestinal FABP	I-FABP	Intestine, liver	4q28–q31	3 G1	2q42
<i>Fabp3</i>	Heart FABP	H-FABP, MDGI	Heart, skeletal muscle, brain, kidney, lung, stomach, testis, aorta, adrenal gland, mammary gland, placenta, ovary, brown adipose tissue	1p32–p33	4 D2.2	5q36
<i>Fabp4</i>	Adipocyte FABP	A-FABP, aP2	Adipocyte, macrophage, dendritic cell	8q21	3 A1	2q23
<i>Fabp5</i>	Epidermal FABP	E-FABP, PA-FABP, mal1	Skin, tongue, adipocyte, macrophage, dendritic cell, mammary gland, brain, intestine, kidney, liver, lung, heart, skeletal muscle, testis, retina, lens, spleen	8q21.13	3 A1-3	2
<i>Fabp6</i>	Ileal FABP	Il-FABP, I-BABP, gastrotropin	Ileum, ovary, adrenal gland, stomach	5q33.3–q34	11 B1.1	10q21
<i>Fabp7</i>	Brain FABP	B-FABP, MRG	Brain, glia cell, retina, mammary gland	6q22–q23	10 B4	20q11
<i>Fabp8</i>	Myelin FABP	M-FABP, PMP2	Peripheral nervous system, Schwann cell	8q21.3–q22.1	3 A1	2q23
<i>Fabp9</i>	Testis FABP	T-FABP	Testis, salivary gland, mammary gland	8q21.13	3 A2	2q23

Table 1.5 Family of fatty acid-binding proteins. The nine members of FABPs family are indicated with the gene name annotation, the common name and eventual alternative nomenclature. Is also specified the organs in which the protein is expressed along with the chromosomal localization of the corresponding genes in three model organisms⁶.

(aP2, adipocyte P2; I-BABP, ileal bile acid-binding protein; MDGI, mammary derived growth inhibitor; MRG, MDGI-related gene; PA-FABP, psoriasis-associated FABP; PMP2, peripheral myelin

1.5.2 The birth of Bile Acid-Binding Proteins

About 20 years ago a protein was isolated and crystallized from chicken liver. It was very similar to another fatty acid-binding protein present in the same tissue⁴¹. This protein had an isoelectric point of about 9.0 and was named liver “basic” Fatty Acid-Binding Protein (Ib-FABP). In 1989 it was

discovered that gastrotropins, proteins believed to be enterooxyntins, i.e., stimulators of gastric acid secretion, belonged to the FABP family⁴². Few years later it was shown that these proteins are bile acid transporters and were renamed Bile Acid-Binding Proteins (BABPs)⁴³.

By comparing the primary sequence of lb-FABPs with BABPs and FABPs it turned out that there is a higher sequence similarity between lb-FABPs and BABPs than between lb-FABPs and FABPs. This result suggested the hypothesis that lb-FABPs could be bile acid transporters, as shown by the structure determination of the chicken lb-FABP complexed with two molecules of cholic acid⁴⁴.

1.5.3 The liver Bile Acid-Binding Proteins

Liver Bile Acid-Binding Proteins (L-BABPs) have been identified and characterized in some vertebrates but not in mammals. Currently 19 amino acid sequences of L-BABPs from several non mammal vertebrates have been determined. These sequences have about 22% of sequence identity indicating that they form a distinct subfamily³⁴. In Figure 1.5 a multiple alignment of 14 sequences of L-BABPs from several non mammalian organism such as birds, reptiles, amphibians and fishes is shown.

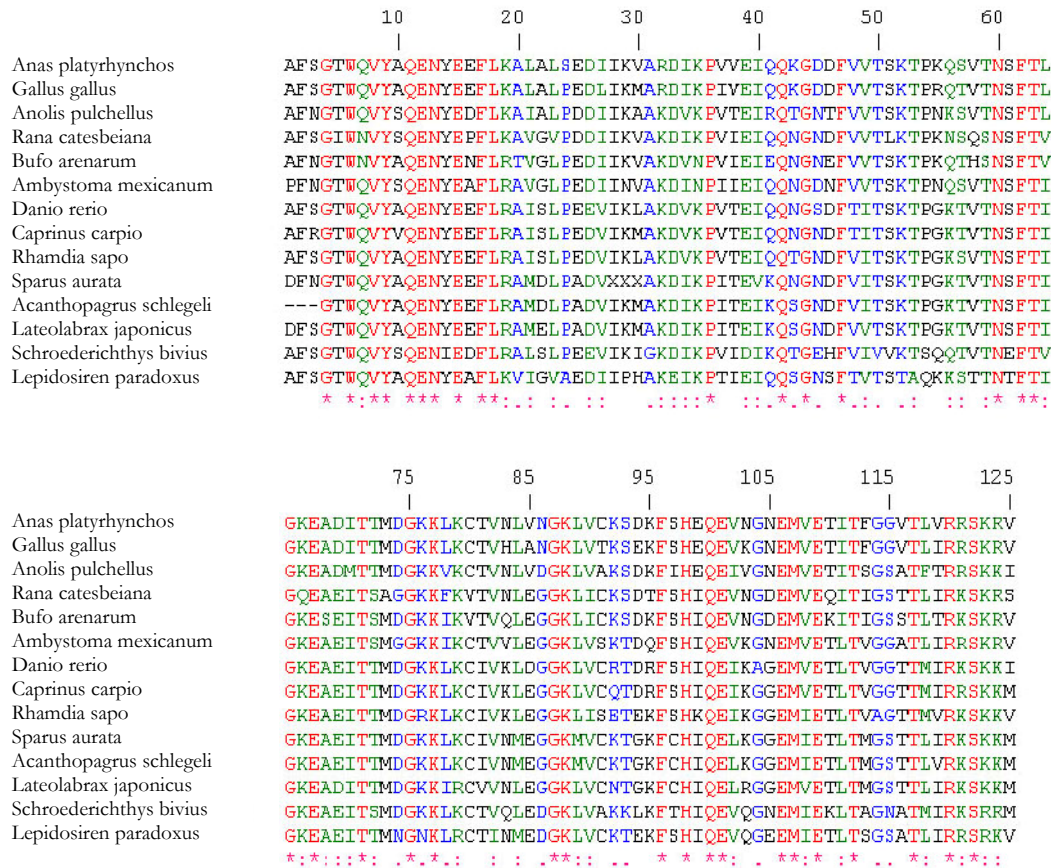


Figure 1.5 Sequence alignment of some (14) known L-BABPs sequences from birds, reptiles, amphibians and fishes. The scientific names of the organisms at which each sequence refers are indicated. The positions with conserved residues are marked with a star (*) at the bottom of the alignment. Positions with strong similarity and weak similarity are also marked with the signs (:) and (.), respectively. The alignment was performed with clustalW program⁸.

At present, four three-dimensional structures of L-BABPs resolved by X-Ray and NMR spectroscopy are available. All these structures display a canonical FABPs fold characterized by 10 antiparallel β -strands and two α -helices, as described previously. One of the most studied L-BABP is that from chicken (cL-BABP).

The binding capability of cL-BABPs has been initially investigated in the presence of fatty acids, as the bile acid binding properties of L-BABPs were not known. The fatty acid binding ability of cL-BABP was investigated by fluorescence spectroscopy and ^{13}C NMR spectroscopy using ^{13}C -enriched fatty acids⁴⁵.

The results indicated that cL-BABP binds palmitic acid and oleic acid with a 1:1 stoichiometry, as confirmed by the structure determination of the complex between the L-BABP from axolotl and oleic acid⁴⁶. The three-dimensional structure of cL-BABP, complexed with cholic acid, was

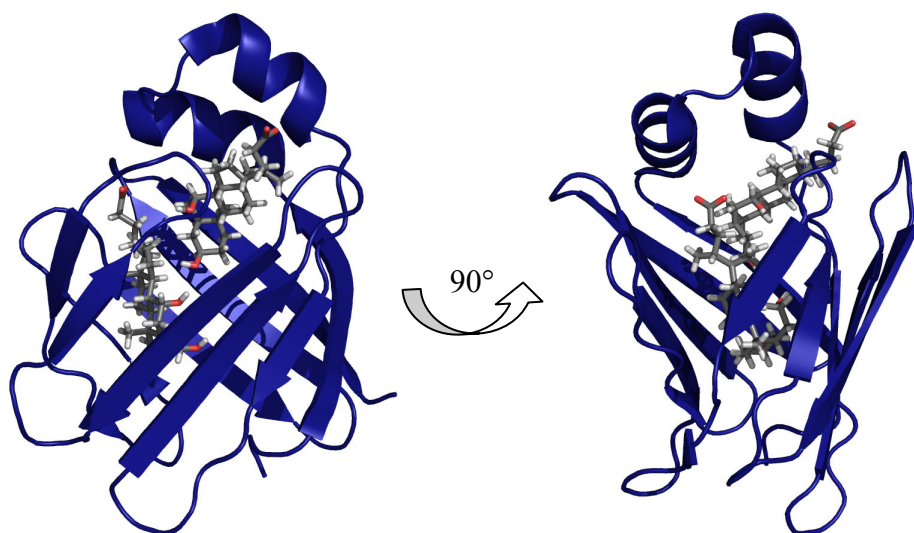


Figure 1.6 Three-dimensional structure of cL-BABP in complex with two molecules of chenodeoxycholic acid (PDB code: 2JN3)⁹.

subsequently resolved by X-Ray⁴⁴ proving the ability of the protein to bind two molecules of bile acids. Recently, the ternary complex between cL-BABP and two molecules of chenodeoxycholic acid has been determined by NMR spectroscopy⁹ as shown in Figure 1.6.

cL-BABP has been also used as a model for the study of bile acids binding mechanism to L-BABPs. It was proposed that the dynamic properties of this protein can regulate the allosteric activation functional to the ligand binding event⁴⁷. cL-BABP was reported to exhibit a highly dynamic behaviour mainly localised on specific protein regions, including the helix loop helix and C-D and E-F loops⁴⁷.

It is well known that proteins belonging to the iLBP family carry water molecules in their central cavity. The behaviour of these water molecules have been studied for several FABPs and related to the ligand binding mechanism and protein stability^{48, 49}. Several water molecules were shown to be present in the internal binding pocket of cL-BABP, both in the apo and holo form⁵⁰. It has been recently shown that *in silico* molecular docking methods reproduced the position of the ligands inside the binding site only taking into account the resident water molecules, thus suggesting that water plays a relevant role in the bile acid binding processes⁵⁰.

cL-BABP has been also shown to interact with membrane mimetics such as Large Unilamellar Vesicles (LUVs). cL-BABP can interact with negatively charged LUVs through electrostatic interaction undergoing a partial unfolding⁵¹. The protein-membrane interaction has been further investigated using multiple-run molecular dynamics simulations, showing that cL-BABP can interact with anionic lipid membranes through the region opposite to the portal area⁵².

1.5.4 The ileal Bile Acid-Binding Proteins

Ileal Bile Acid-Binding Proteins (I-BABPs), previously called gastrotropins, are mainly involved in bile acid transport in enterocytes. These proteins have been found in mammalian organisms and only recently their presence has been identified in non mammalian species³⁵. Several papers were published on the bile acid binding properties of human I-BABP (hI-BABP) making it the most studied ileal bile acid-binding protein. hI-BABP binds two molecules of bile acids with low intrinsic affinity but high cooperativity⁵³. It has been shown that hI-BABP can bind glycocholic acid (GCA) and glycochenodeoxycholic acid (GCDA) in a site selective manner. When hI-BABP is incubated with GCA or GCDA both binding

sites are occupied; while when the protein is incubated with a mixture of the two bile salts GCA binds nearly exclusively to one site and GCDA binds nearly exclusively to the other site. In Figure 1.7A is shown the heterotypic model of hI-BABP with a molecule of GCA bound in one site and a molecule of GCDA bound in the other site. The binding

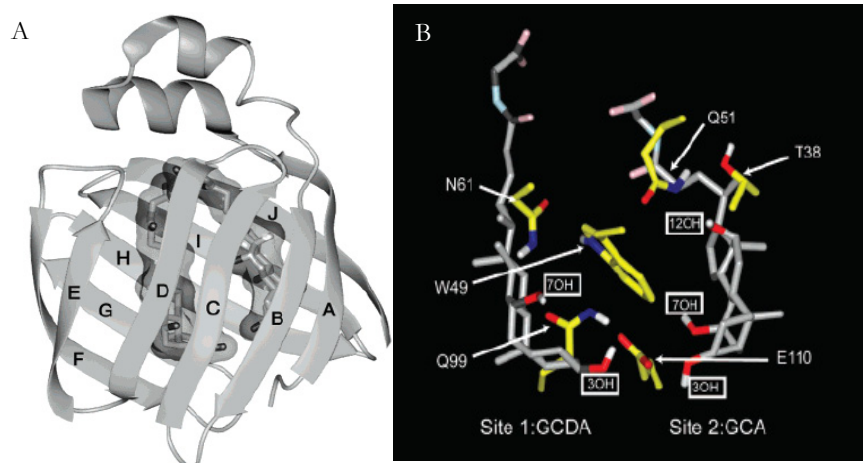


Figure 1.7 Homology model of human I-BABP with GCA and GCDA docked by restrained energy minimization (A). Focus on the human I-BABP binding pocket. The bound molecules of GCA and GCDA are shown along with the lateral chains of key residues involved in binding mechanism (B)⁷.

The binding and cooperativity determinants have been studied and several key residues involved in both processes have been identified (Figure 1.7B). The macroscopic binding parameters have been determined through isothermal titration calorimetry measurements on the wt protein and several mutants and an estimate of the cooperativity factor was obtained^{7, 54}. The calculated Hill coefficient, a parameter used to estimate positive cooperativity, for GCA binding was 1.94 and the cooperativity factor exceed 1000. The Hill coefficient adopts a value of 1 for a non-cooperative system and a maximum value of 2 for an extremely positively cooperative two-site system. The cooperativity parameters observed in this system have been reported to be the largest ever measured for a ligand–protein interaction⁵⁴.

Recently the three-dimensional structure of the I-BABP from zebrafish has been resolved in both its apo and holo-form, complexed with cholic acid. This work reported an unexpected result represented by the ability of the zebrafish I-BABP to bind bile acid molecules on the surface of the protein¹ (Figure 1.8). This was the first report in which bile acid molecules were found on the external surface of BABPs.

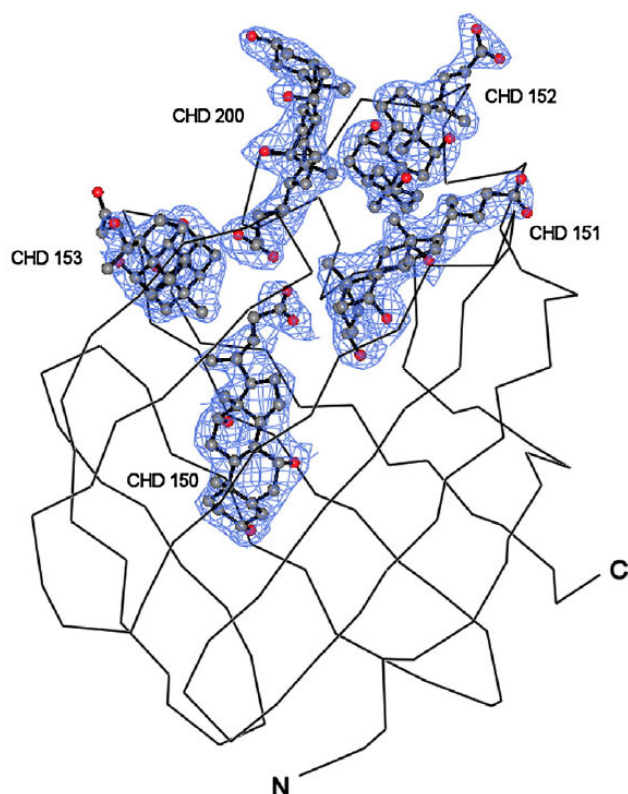


Figure 1.8 Three-dimensional structure of ileal BABP from zebrafish (*Danio rerio*) complexed with cholic acid; pdb code: 3MLZ¹. The α -carbon chain trace is shown with a straight black line. The structures of cholic acid molecules bound to the protein are shown along with the experimental electron density map. In the structure are clearly visible bile acids molecules (CHD 152, CHD 153 and CHD 200) bound on the surface of the protein in addition of two bile acids molecules (CHD 150 and CHD 151) bound in the internal binding pocket.

1.6 Bile acid functions and regulation

Bile acids were traditionally considered as detergent molecules with the main function to enhance the absorption of dietary fats and liposoluble vitamins. Due to their physical-chemical properties bile acids are able to form mixed micelles together with biliary phospholipids, which allow the solubilization in bile of cholesterol and other lipophilic compounds. Mixed micelles also account for the emulsion of dietary fat and liposoluble vitamins in the gut, thus helping their absorption⁵⁵.

1.6.1 Bile acid homeostasis

It is well established that bile acids regulate their own biosynthesis, detoxification and transport both in the liver and the intestine. Bile acid accumulation leads to liver damage and may promote the development of liver tumors⁵⁶. The identification of bile acids as natural ligands for FXR⁵⁷ led to the discovery that FXR is essential for the maintenance of bile acid homeostasis and protects the organism from the accumulation of potentially toxic bile acids¹⁷.

1.6.2 Bile acids modulate glucose homeostasis

Several recent studies indicate that bile acids regulate hepatic gluconeogenesis. Bile acids have been reported to inhibit gluconeogenesis via the downregulation of phosphoenolpyruvate carboxykinase (PEPCK) mRNA levels both in an FXR-dependent and FXR-independent manner^{58, 59}. In other studies, instead, FXR activation was shown to activate PEPCK⁶⁰. Activation of FXR leads also to an increase of glycogen synthesis in the liver, thereby lowering blood glucose levels⁶¹.

In addition bile acids have been shown to modulate glucose homeostasis through alteration of the cellular hydration state. Hypo-osmotic hepatocyte swelling, which can be induced by taurocholate⁶², stimulates glycogen synthesis⁶³ as well as flux through the pentose phosphate pathway⁶⁴ and inhibits glycolysis and glycogenolysis⁶⁵.

1.6.3 Bile acid effects on lipid homeostasis

An inverse relationship between bile acids and triglyceride levels has been recognized through the effects of the administration in patients of bile acids and bile acids binding resins. Bile acids administration lead to a reduction of plasma triglyceride levels⁶⁶, whereas bile acids binding resins induced an increased production of very low density lipoproteins (VLDL) and an elevation of serum triglyceride levels⁶⁷.

Bile acids can also stimulate the triglyceride clearance from serum through activation of lipoprotein lipase and subsequent hydrolysis of triglycerides in VLDL and chylomicrons⁶⁸.

Bile acids, such as CDCA, repressed the expression of Microsomal triglyceride Transfer Protein (MTP) and ApoB, which are essential for the assembly of chylomicrons and VLDL particles⁶⁹.

Bile acids may also influence lipid metabolism through cross-talk with peroxisome proliferator-activated receptor α (PPAR α)-dependent pathways. PPAR α is a nuclear receptor which plays an important role in lipid and lipoprotein metabolism and controls several enzymes critical for fatty acid oxidation⁷⁰. Expression of human PPAR α is directly regulated by bile acids⁷¹. Furthermore, expression of pyruvate dehydrogenase kinase-4 (PDK4) is upregulated by bile acids *via* FXR, which leads to inactivation of the pyruvate dehydrogenase complex with subsequent suppression of glycolysis and increased fatty acid oxidation⁷².

1.6.4 Bile acids increase energy expenditure

Bile acids modulate energy metabolism through the membrane bile acid receptor TGR5. TGR5 is coupled to stimulatory G-protein and activation of the receptor by bile acids increases intracellular cyclic AMP (cAMP) levels⁷³. Administration of bile acids to mice increased energy expenditure in brown adipose tissue and prevented development of obesity and insulin resistance⁷⁴.

1.6.5 Further endocrine effects of bile acids

Bile acids show immunomodulatory functions as oral administration of bile acids reduced endotoxin-related complications following surgery in patients with obstructive cholestasis⁷⁵.

Bile acids can alter macrophage function by affecting phagocytic activity as well as cytokine production in a TGR5-cAMP-dependent manner supporting the hypothesis that TGR5 plays an important role for macrophage function^{73, 76}.

Bile acids can prevent bacterial overgrowth and mucosal injury in the small intestine of mice revealing a novel mechanism of bile acid dependent enteroprotection⁷⁷.

Bile acids may also interfere with interleukin-6 (IL-6) signalling in hepatocytes. IL-6 has hepatoprotective properties and can protect liver from injury induced by obstructive cholestasis⁷⁸.

Recently a role of bile acids in liver regeneration has been identified. Increased bile acid levels after partial hepatectomy promoted liver regeneration which was attenuated in FXR knockout mice, suggesting that these effects are mediated by FXR⁷⁹.

1.6.6 Paracrine role of bile acids

Bile acids secreted into the bile are in close contact with cholangiocytes, the epithelial cells forming the bile duct. Bile acids modulate cholangiocyte secretion, proliferation and survival in paracrine manner. Taurocholic acid and tauroolithocholic acid increased secretin stimulated intracellular cAMP levels and Cl⁻/HCO₃⁻ exchanger activity thus promoting ductal secretion and bile flow⁸⁰.

1.7 Bile acid related damage

1.7.1 Bile acids induced hepatocyte injury

In healthy hepatocytes uptake of bile acids across the basolateral membrane and export via the canalicular export pump are tightly coupled. Bile acids accumulate within the hepatocyte when canalicular export is defective or when there is a physical obstruction to bile flow. When bile acid concentration exceeds the binding capacity of the binding proteins, bile acids could induce disruption of cell membranes through a detergent action on lipid components⁸¹ and promote the generation of reactive oxygen species (ROS) that, in turn, oxidatively modify lipids, proteins and nucleic acids and eventually cause hepatocyte apoptosis⁸².

1.7.2 Bile acids induced oxidative stress

Several studies have suggested that oxidative stress may play an important role in the pathogenesis of hepatic injury and hepatic mitochondria have been proposed as a major source of the oxidative stress induced by bile acids. This hypothesis is supported by the observation that hepatic mitochondria undergo lipid peroxidation during experimental cholestasis, condition where bile can not flow from the

liver to the duodenum⁸³. It has been reported also that hydrophobic bile acids can damage respiration and electron transport in hepatic mitochondria and decrease the activity of several enzyme complexes involved in the electron transport chain⁸⁴. Bile acids seems to stimulate ROS generation in human hepatic mitochondria⁸⁵, process that is considered to be an early event in hydrophobic bile acids-induced hepatocyte toxicity⁸⁶.

Hydrophobic bile acids can also induce the mitochondrial permeability transition (MPT)⁸⁵, a crucial intracellular event that triggers the apoptotic and necrotic cell death of hepatocytes⁸⁷. Induction of MPT is associated with mitochondrial swelling, collapse of the mitochondrial membrane potential, reduced oxidative phosphorylation, rupture of the outer mitochondrial membrane, cytochrome c release and generation of ROS⁸⁷.

1.7.3 Bile acids induced cell death pathways

Bile acids accumulation within hepatocytes can result in cell injury and death. Two mechanisms are thought to be involved in these cell death pathways: hepatocellular apoptosis can be induced by low bile acids concentrations while high bile acids concentrations can induce cellular necrosis.

1.7.3.1 Bile acids induced necrosis

Cellular necrosis in rat hepatocytes is characterized by cell swelling and disruption of the intracellular and plasma membranes. This effects have been attributed to direct membrane damage due to the detergent properties of hydrophobic bile acids⁸¹ as well as to depletion of ATP, ion disregulation, mitochondrial and cellular swelling, plasma membrane failure and cell lysis⁸⁸. It has been proposed that MPT induction by ROS generated in hepatocyte mitochondria⁸⁹ is a critical event promoting bile acid-induced hepatocytes necrosis.

1.7.3.2 Bile acids induced apoptosis

Bile acids can induce apoptosis in three different ways.

Bile acids can activate the death receptor starting the so called extrinsic pathway in which a caspase cascade is induced that leads to cell death⁹⁰. Alternatively they can generate intracellular stress which causes mitochondrial disfunction and subsequent release of proapoptotic factors⁹¹. Finally bile acids can induce apoptosis in hepatocytes by causing endoplasmic reticulum (ER) stress. In this case bile acids determine Ca^{2+} release from ER which induces extracellular Ca^{2+} influx followed by the activation of the caspase cascade⁹².

1.7.4 Bile acids induced damage in the plasma membrane

The efficiency of bile acids to solubilise membrane lipids, such as phospholipids, cholesterol or fatty acids, is generally enhanced with increasing bile acids hydrophobicity⁹³. Conjugated bile acids are more hydrophilic and not cytotoxic until their concentration approach their critical micellar concentration¹⁶. In the hepatocytes, when the concentration of hydrophobic bile acids exceeds the binding capacity of the cytosolic proteins, these compounds can damage organelle membranes, especially in the case of mitochondria, leading to mitochondrial damage and ultimately to apoptosis or necrosis¹⁶.

1.7.5 Bile acids induced injury in other cell types

Gastrointestinal cells

High levels of hydrophobic bile acids in humans is associated with an increased incidence of cancer of the laryngopharyngeal tract, esophagus, stomach, pancreas, small intestine and colon⁹⁴. Bile acids cause DNA damage in several human colon cancer cell lines, probably inducing ROS production⁹⁵. The DNA damage increase the mutation rate in several genes, including those coding for tumour suppressors and oncogenes. Hydrophobic bile acids may also behave as selection agents for apoptosis-resistant cells⁹⁶.

Kidney and lung cells

Bile acid accumulation in the systemic circulation may contribute to endothelial injury in the kidney and lungs⁹⁷. This endothelial damage is due to ROS which are generated by hydrophobic bile acids from mitochondria as well as released from neutrophils and macrophages in the presence of bile acids. In addition oxidative stress can promote the formation of a variety of vasoactive mediators which can affect renal function directly by causing renal vasoconstriction⁹⁸.

CHAPTER 2: Material & Methods

2.1 Materials

2.1.1 Chemicals

All chemicals and reagents used were of analytical grade purity or higher and, if not otherwise specified, were purchased from Sigma-Aldrich, Carlo Erba, Merck and Pharmacia Biotech. ^{15}N enriched ammonium chloride, ^{13}C enriched glucose and D_2O were purchased from Cambridge Isotope Laboratories.

2.1.2 DNA coding sequence

The DNA coding for cL-BABP, already cloned in the pET24d(+) vector, has been originally provided by Dr. J. Foote (Fred Hutchinson Cancer Research, Seattle, USA).

2.1.3 Primers

All primers used in the cloning protocols were purchased from Eurofins MWG Operon and are listed in Table 2.1.

primer name	primer sequence
H98Q-for	CCAAATCTGAAAAGTTCTCTCA G GAACAGGAAGTTAAAGGTAAC
H98Q-rev	GTTACCTTTAACTTCCTGTT C CTGAGAGAACTTTTCAGATT T TGG
T91C-for	CCTGGCTAACGGTAAACTGGTT T T GTAAATCTGAAAAGTTCTCTCACGAA
T91C-rev	TTCGTGAGAGAACTTTTCAGATT T T ACAACCAGTTTACCGTTAGCCAGG

Table 2.1 Primer sequences used in mutagenesis procedures. Annealing sequences are reported in black. The nucleotide substitutions for the generation of mutant proteins using the QuikChange™ site-directed mutagenesis kit are shown in red.

2.1.4 Plasmid

pET24d(+) pET24d(+) (Novagen) is a 5.3 Kb expression vector based on the strong T7 promoter. T7 promoter is recognized by T7 RNA polymerase, therefore T7 transcription/expression is induced by providing a source of T7 RNA polymerase in the host cell. T7 RNA polymerase is so selective and active that, when fully induced, almost all of resources of the cell are converted to target gene expression.

A schematic representation of this plasmid is shown in Figure 2.1.

The elements characterising this vector are:

- Optimized promoter/operator element: consists of phage T7 *lac* promoter and *lacI* operator sequence. The presence of *lac* repressor bound on T7 *lac* promoter repress the production of T7 RNA polymerase. In presence of IPTG *lac* repressor detaches from the T7 *lac* promoter de-repressing T7 RNA polymerase gene expression and in turn the expression of genes under its control.
- Multiple cloning site (shown in Figure 2.2).
- N-terminal 6xHis-tag coding sequence.
- Origin of replication from pBR322 plasmid.
- Kan gene: confers kanamycin resistance.

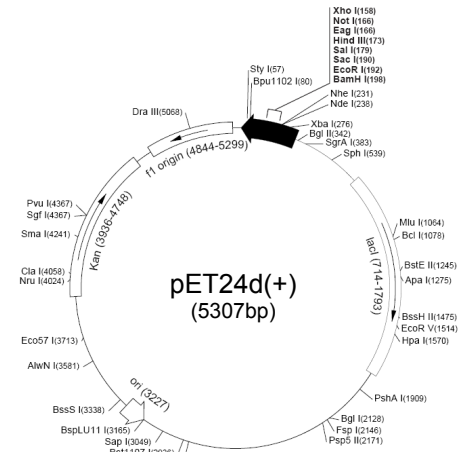


Figure 2.1 Schematic representation of pET24d(+) plasmid (from Novagen website).

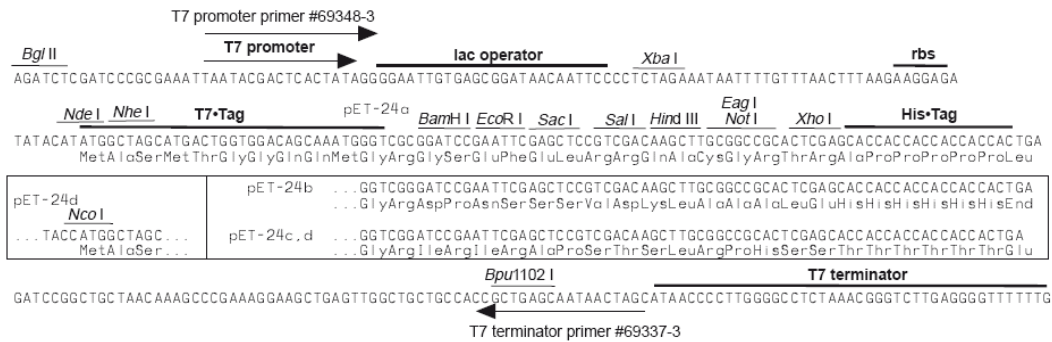


Figure 2.2 Cloning/expression region of pET24d(+) plasmid (from Novagen website).

2.1.5 Bacterial strains

XL1 Blue (Stratagene): *E. coli* strain particularly indicated for cloning and plasmidic DNA replication procedures. XL1 Blue strain is endonuclease deficient, which greatly improves the quality of plasmidic DNA obtained from it. The genotype is: *recA1*, *endA1*, *gyrA96*, *thi*, *hsdR17* (r^-_K , m^+_K), *supE44*, *relA1*, *lac*, [*F'*, *proAB*⁺, *lacI*^fZΔ*M15*::*Tn10*(*tetR*)].

BL21(DE3) (Stratagene): *E. coli* strain for protein production in T7 promoter-driven vectors. The BL21(DE3) strain has a T7 phage RNA polymerase gene inserted in its genome inducible with IPTG. The genotype is: *F*⁻, *ompT*, *hsdSB*, (*rB*⁻, *mB*), *dcm*, *gal*, λ (DE3).

2.1.6 Chromatography resins and columns

Ion exchange chromatography resin

It has been used the weak anion exchanger DE52 preswollen microgranular diethylaminoethyl (DEAE) cellulose (Whatman).

Size exclusion chromatography resin

It has been used the Sephacryl S-100 HR (GE Healthcare) resin – MW fractionation range 10^3 - 10^5 .

Hydrophobic interaction chromatography resin

It has been used the Lipidex 1000 (Packard Bioscience) resin.

Chromatography columns

XK26/100, XK26/40 and XK16/20 columns (GE Healthcare) were used.

2.1.7 Instruments

For protein concentration the Amicon[®] concentrator (10 ml, 50 ml and 400 ml) with regenerated cellulose membranes (YM serie) of 3000 and 5000 Da cutoff was used.

Liquid chromatography was carried out using ÄKTA Prime Automated Liquid Chromatography System (GE Healthcare).

For gel electrophoresis the electrophoresis power supply EPS 301 (Pharmacia Biotech) and the Mini-PROTEAN[®] 3 cell (Biorad) was used.

Media sterilization was achieved using an Alfa 10 plus (PBI) autoclave.

Bacterial growth were performed using a FOC 225E refrigerated incubator (VELP scientifica) and bacterial lysis were carried out with the sonicator Sonoplus HD 2070 (Bandelin Electronic).

Site-directed mutagenesis was performed using a Mastercycler personal PCR machine (Eppendorf).

Centrifugation steps were carried out using centrifuges J2-HS (Beckman), Centrifuge 5804R and MiniSpin[®] plus (Eppendorf) and a Biofuge Fresco (Heraeus instruments).

2.2 Culture growing media

Selective media were prepared adding kanamycin, from a sterile stock solution of 1 M, at a final concentration of 50 µg/ml. Sterilisation was achieved in autoclave if not otherwise specified. Media for bacterial growth were prepared as follows.

Luria-Bertani (LB)

compound	quantity [^]
triptone	10g
yeast extract	5g
NaCl	10g

[^] For a final volume of 1L.

pH was adjusted to 7.0. For the preparation of solid LB media 15g of bacto agar were added and after sterilization the solution was poured into sterile petri dishes.

Minimal medium (MM)

compound	quantity [^]
sterile water	400 ml
M9 salts	100 ml
MgSO ₄ (1 M stock solution)*	1 ml
ZnSO ₄ (50 mM stock solution)*	1 ml
CaCl ₂ (0.1 M stock solution)*	500 µl
FeCl ₃ (10 mM stock solution)*	50 µl
glucose (40% stock solution)*	10 ml
vitamin stock solution*	3 ml

[^] For a final volume of 500ml.

* Sterilized by filtration through a 0.22 µm filter.

M9 salts

compound	quantity [^]
Na ₂ HPO ₄	33.9 g
KH ₂ PO ₄	15 g
NaCl	2.5 g
NH ₄ Cl	5 g

[^] For a final volume of 1L.**Vitamin stock solution**

compound	quantity [^]
thiamine	100 mg
d-biotin	20 mg
choline chloride	20 mg
folic acid	20 mg
niacinamide	20 mg
d-panthothenate	20 mg
pyridoxal hydrochloride	20 mg
riboflavin	2 mg

[^] For a final volume of 200ml. Sterilization was achieved by filtration through a 0.22 µm filter.

2.3 Competent cells

The following protocol has been used to make *E. coli* cells competent for transformation through heat shock:

- 100 ml of sterile LB were inoculated with the *E. coli* strain of interest. The culture was then incubated in a shaker at 37°C until OD₆₀₀ reached 0.35-0.4 value.
- The culture was transferred into sterile tubes and centrifuged at 3000g for 5 minutes at 4°C.
- The pellet obtained was gently resuspended in 16 ml of solution 1 and left in ice for 15 minutes.
- After putting together the content of the tubes a second centrifugation step was performed at 3000g for 5 minutes at 4°C.
- 8 ml of solution 2 were used to resuspend the pellet.
- The culture was then aliquoted in sterile 1.5 ml tubes, frozen in liquid nitrogen and stored at –80°C.

Solution 1

compound	quantity [^]
RbCl	0.6 g
MnCl ₂ ·4H ₂ O	0.5 g
potassium acetate pH 7.5 (1 M stock solution)	1.5 ml
CaCl ₂ ·2H ₂ O	75 mg
glycerol	7.5 g

[^] For a final volume of 50ml. pH was adjusted to 5.8 with acetic acid and sterilization was achieved by filtration through a 0.22 µm filter.

Solution 2

compound	quantity [^]
MOPS pH 6.8 (0.5 M stock solution)	0.4 ml
RbCl	24 mg
CaCl ₂ ·2H ₂ O	0.22 g
glycerol	3 g

[^] For a final volume of 20ml. Sterilization was achieved by filtration through a 0.22 µm filter.

2.4 Heat shock transformation

The following protocol has been used to transform *E. coli* competent cells through heat shock:

- 🖼 A 100 μ l aliquot of competent cells was defrosted in ice.
- 🖼 30-50 ng of DNA were added and the culture was incubated in ice for 30 minutes.
- 🖼 The heat shock was achieved by placing the culture in a water bath at 42°C for 90 seconds.
- 🖼 After heat shock cells were kept in ice for 2 minutes.
- 🖼 800 μ l of non selective, sterile LB media were added directly to the test tube and the cells were incubated at 37°C for 45 minutes.
- 🖼 The treated cells were then plated on selective LB agar and incubated overnight at 37°C.

2.5 Plasmid DNA extraction

Cells from an overnight grown bacterial culture (1-5 ml volume) were collected by centrifugation at 10000g for 1 minutes. Plasmid DNA was extracted using the GenElute Plasmid Miniprep kit (Sigma) according to the manufacturer's instructions.

2.6 Site-direct mutagenesis

QuikChange II Site-Directed Mutagenesis kit (Stratagene) has been used to generate different plasmids coding for mutant proteins. According to the manufacturer's instructions the mutagenesis was performed using the reaction mixture reported in Table 2.2.

Site-directed mutagenesis was performed on a Mastercycler personal PCR machine (Eppendorf) using the cycling parameters presented in Table 2.3.

reagent	volume (μ l)	quantity
template DNA vector	1.0	5-50 ng
dNTPs	1.0	10 mM stock
forward primer	1.0	125 ng
reverse primer	1.0	125 ng
QuickChange reaction buffer	5.0	10X stock
<i>Pfu</i> Ultra HF DNA polymerase	1.0	2.5 units/ μ l
ddH ₂ O	40.0	-
total volume	50.0	

Table 2.2 Site-directed mutagenesis volumes and stock concentrations.

step	temperature (C°)	time (min)	number of cycles
initial denaturation	95	5	1
denaturation	95	0.5	
annealing	55	1	16
elongation	68	1/Kbp template DNA	

Table 2.3 Site-directed mutagenesis cycling parameters.

Following the temperature cycling the reaction product was treated with 1 μ L of *Dpn*I endonuclease for one hour at 37°C. After digestion 2 μ L of the *Dpn*I-treated DNA reaction mix were used to transform 100 μ L of XL1Blue competent cells.

2.7 DNA sequencing

The sequences of the mutated plasmids were verified through the sequencing service provided by Eurofins MWG Operon.

2.8 Protein expression

2.8.1 Preliminary protein expression evaluation

Several transformed bacterial colonies were tested for recombinant protein expression and colonies showing higher protein expression were selected. Protein expression was assessed on a small-scale by growing single colonies in 5 ml LB or M9 medium in test tubes. Colonies were grown at 37°C until OD₆₀₀ reached the 0.5-0.6 value and induction was achieved by addition of 0.5 mM IPTG. After induction colonies were incubated overnight at 22°C. 1 ml of each bacterial culture was collected by centrifugation at 10000g for 5 minutes and each pellet was resuspended in 500 μ L of distilled water. Protein content was then analysed by SDS-PAGE, together with samples collected before induction for the evaluation of basal protein expression.

2.8.2 Protein solubility test

The colonies showing higher protein expression were then tested for the soluble expression level in different conditions of induction temperature and time. Five colonies for each protein were grown in 5 ml of LB or M9 media in test tubes. Colonies were grown at 37°C until OD₆₀₀ reached the 0.5-0.6 value and induction was achieved by addition of 0.5 mM IPTG. Test tubes were incubated at 22°, 28° and 37°C and 1 ml sample was collected for each culture at 4, 8 and 16 hours after induction. The bacterial cells were harvested by centrifugation at 10000g for 5 minutes and each pellet was resuspended in 500 μ L of lysis buffer. Each sample was then sonicated with several 1 minute pulses and between each pulse

samples tubes were kept in ice to avoid sample overheating. After cells lysis 20 μ l lysate was collected that represent the total protein fraction. Each lysate was treated by centrifugation at 10000g for 10 minutes to pellet cellular debris and after this 20 μ l of supernatant was collected that represent the soluble protein fraction. Both total and soluble protein contents were then analysed by SDS-PAGE.

2.8.3 Large-scale soluble expression

For unlabelled proteins a 10 ml starter culture of the best expression colony was grown in LB medium shaking overnight at 37°C. The culture was used to inoculate 1L of sterile LB media and then incubated in a shaker at 37°C until OD₆₀₀ value of 0.6 was reached. Induction was achieved by addition of 0.5 mM IPTG and the culture was incubated at the best temperature and time found with the solubility test. Bacterial cells were harvested by centrifugation at 8000g for 10 minutes. The pellet was resuspended in 20 ml of lysis buffer and PMSF (1 mM final concentration) and β -mercaptoethanol (0.7 μ l/ml) were added. Cells lysis was achieved by sonication with 4 pulses of 1 minute each keeping the sample in ice. The lysate was then centrifuged at 15000g for 10 minutes. The supernatant, containing the soluble protein fraction, was then loaded on the first purification column. Before induction a 1 ml sample was collected and centrifuged at 10000g for 5 minutes and the pellet was stored at -20°C for the evaluation of basal protein expression.

The protocol for ¹⁵N or ¹⁵N-¹³C labelled proteins was the same as for unlabelled proteins except that a 30 ml starter culture and M9 minimal medium was used.

Lysis buffer
50 mM Tris-HCl pH 8.0
10% sucrose
1 mM EDTA
1 mM PMSF freshly added
10 mM β -mercaptoethanol

2.9 Protein purification

For all proteins, wild type and mutants, the purification protocol was the same. The purification started with an anion exchange chromatography followed by a size exclusion chromatography with a final delipidation step. The isolated proteins were transferred, by dialysis, in a buffer solution optimal for NMR measurements and finally processed by lyophilization and stored at 4°C.

2.9.1 Anion exchange chromatography

The weak anion exchanger DE52 preswollen microgranular DEAE cellulose (Whatman) resin was used. 80g of pre-swollen resin was resuspended in 200 ml of 0.2 M Tris-acetate pH 7.8. After sedimentation the supernatant was removed and the resin washed twice with 200 ml of 50 mM Tris-acetate pH 7.8. The resin was then poured into a XK 26/40 (GE Healthcare) column and extensively washed with 50 mM Tris-acetate pH 7.8, 0.02% NaN₃ (running buffer). The cleared cellular lysate was loaded in the column and the protein was eluted with several column volumes of running buffer. The chromatography was performed at a constant flow rate of 1 ml/min using ÄKTAprime (GE Healthcare) and the proteins elution profile was followed monitoring the absorbance at 280nm (Abs₂₈₀). Fractions of 10 ml were collected and analyzed by SDS-PAGE.

After use the resin was regenerated according to the manufacturer's instructions.

2.9.2 Size exclusion chromatography

500 ml of Sephacryl S-100 HR (GE Healthcare) resin were poured into a XK26/100 (GE Healthcare) column and extensively washed with 50 mM Tris-HCl pH 7.2, 0.2 M NaCl, 0.02% NaN₃ (running buffer). Fractions containing the protein of interest, resolved with the previous chromatographic step, were concentrated to 5 ml with an Amicon[®] concentrator using N₂ as gas. The concentrated protein sample was injected in the column and the chromatography was performed at a constant flow rate of 1 ml/min using ÄKTAprime (GE Healthcare). The proteins elution profile was followed monitoring the absorbance at 280nm (Abs₂₈₀). Fractions of 10 ml were collected and analyzed by SDS-PAGE.

After use the resin was regenerated according to the manufacturer's instructions.

2.9.3 Hydrophobic interaction chromatography (delipidation)

In order to remove endogenous hydrophobic ligands bound to the purified protein a delipidation procedure was required. 30 ml of Lipidex 1000 (Packard Bioscience) resin were washed twice with 5 volumes of 10 mM sodium phosphate pH 7.4, 0.02% NaN₃ (running buffer) at 4°C. The resin was then poured into a XK16/20 (GE Healthcare) column and extensively washed with the running buffer. Before sample loading the resin was equilibrated for an hour at 37°C by connecting the column to a preheated water bath. Fractions containing the protein of interest, resolved with the previous chromatographic step, were concentrated at 1 mg/ml and applied to the column. The chromatography was performed at a constant flow rate of 12 ml/hour using a peristaltic pump system and fractions of 2 ml were collected and analyzed by SDS-PAGE.

After use the resin was regenerated according to the manufacturer's instructions.

2.9.4 SDS-PAGE

For protein analysis with Sodium Dodecyl Sulphate PolyAcrylamide Gel Electrophoresis (SDS-PAGE) the samples were added with a reducing loading buffer and heated at 95°C for 5 minutes. 15 µl of each sample was loaded into a separate well of a discontinuous 15% polyacrylamide gel and 5 µL of molecular weight marker (GE Healthcare) was also loaded in a separate well for internal referencing. Gels were run at 120 V for approximately one hour. After running gels were immersed in the staining solution for 10 minutes and destained with bidistilled hot water.

Reducing loading buffer

50 mM Tris-HCl pH 6.8

100 mM DTT^{*}

2% SDS

10% glycerol

0.1% bromophenol blue

Polyacrylamide gel

- Stacking gel solution

compound	quantity [^]
acrylamide mix (30% stock solution - BioRad)	330 µl
Tris-HCl pH 6.8 (1 M stock solution)	250 µl
SDS (10% stock solution)	20 µl
ammonium persulfate (10% stock solution)	20 µl
TEMED	2 µl

[^] For a final volume of 2ml.

- Running gel solution

compound	quantity [^]
acrylamide mix (30% stock solution - BioRad)	2 ml
Tris-HCl pH 8.8 (1.5 M stock solution)	1 ml
SDS (10% stock solution)	40 µl
ammonium persulfate (10% stock solution)	40 µl
TEMED	16 µl

[^] For a final volume of 4ml at 15% acrylamide.

Running buffer

25 mM Tris-HCl pH 8.3

192 mM glycine

0.1% (w/v) SDS

Staining solution

compound	quantity [^]
methanol	400 ml
acetic acid	100 ml
Comassie Brilliant Blue	1.5 g

[^] For a final volume of 1L.**2.9.5 Estimation of protein concentration**

Protein concentration (c) was calculated from the Beer-Lambert equation:

$$Abs_{280} = \epsilon_{280} \cdot l \cdot c$$

The absorbance at a wavelength of 280 nm (Abs_{280}) was measured using a Unicam UV2 UV/VIS spectrometer. The pathlength (l) of the sample in a measurement cuvette was 1 cm. The theoretical extinction coefficient at a wavelength of 280 nm measured in water (ϵ_{280}) was estimated for each construct using the ProtParam Tool from the Expasy Tools server (<http://www.expasy.ch/tools/protparam.html>).

2.9.6 Buffer exchange

After purification proteins were transferred in a buffer solution suitable for NMR measurements. The final buffer was 30 mM Na_2HPO_4/NaH_2PO_4 , 0.02% NaN_3 , pH 7.0. Buffer exchange was performed at 4°C by 3 dialysis steps with “buffer : protein solution” ratio of 100 each.

2.9.7 Sample storage

After buffer exchange in the desired NMR buffer the purified protein samples were concentrated to 0.5 mM, lyophilized and stored at 4°C.

2.9.8 Sample quality control

Every purified protein was checked by mass spectrometry to evaluate the correspondence of the measured molecular mass with the theoretical value. In the case of labelled protein samples mass spectrometry analysis were performed to estimate the ^{15}N and/or ^{13}C isotope inclusion percentage. In addition every unlabelled protein was checked by 1H mono-dimensional NMR experiment. For ^{15}N

and/or ^{13}C isotope labelled proteins basic bi-dimensional heteronuclear NMR experiments were performed to confirm the isotope inclusion and to evaluate the goodness of the protein preparation.

2.10 NMR analysis

2.10.1 Sample preparation

For NMR analysis lyophilized samples were hydrated in 90% double distilled water and 10% D_2O (v/v) or 100% D_2O . After hydration each protein sample was centrifugated at 10000g for 1 minute to remove any insoluble particle and the protein concentration was checked by Abs_{280} measurement.

2.10.2 Data collection

The majority of NMR experiments were recorded on a Bruker DRX 500 spectrometer, operating at 500.13 MHz, equipped with a triple resonance probehead, and on a Bruker Avance III 600 spectrometer, operating at 600.13 MHz, equipped with a triple resonance TCI cryoprobe, incorporating gradients in the z -axis. Both these instruments are present at the University of Verona.

Experiments were also recorded using the Bruker Avance spectrometers operating at 600.13 MHz, 700.13 MHz and 800.13 MHz all equipped with cryogenic probe present in the NMR facility at the Center for Biomolecular Magnetic Resonance of the University of Frankfurt (Germany).

The experimental temperature was set to 298 K, unless otherwise specified.

2.10.3 Data acquisition, processing and analysis

Acquisition and processing were carried out by using the software Topspin (Bruker) version 2.0 and 2.1. Data visualisation and analysis were carried out using the software Topspin and the software Sparky⁹⁹.

CHAPTER 3: Protein NMR spectroscopy

NMR spectroscopy, similarly to other forms of spectroscopy, deals with transitions from a ground state to an excited state. In the case of NMR spectroscopy these transitions affect nuclear spins. It is important to specify that NMR spectroscopy differs from other spectroscopy in the fact that the generation of the ground and excited states requires the existence of an external magnetic field. Because of this it is possible to change the transition frequencies by changing the applied magnetic field strength. In addition the NMR excited state has a lifetime that is of the order of 10^9 times longer than the lifetime of excited electronic states. This property allows to attain very narrow spectral lines and facilitates the performance of multi-dimensional experiments.

3.1 The NMR phenomenon

3.1.1 Magnetic dipole and energy levels

The ground and excited states in NMR spectroscopy originate from the interaction of a *nuclear magnetic dipole moment* with an intense external magnetic field. Transitions between these states are stimulated using radio-frequency (RF) electromagnetic radiation.

The nuclear magnetic dipole moment arises from the *spin angular momentum* of the nucleus. All nuclei with an odd mass number (e.g. ^1H , ^{13}C , ^{15}N) and nuclei with an even mass number and an odd charge (e.g. ^2H , ^{14}N) have spin angular momentum. The spin angular momentum, \vec{S} , is quantized and the total angular momentum of a nuclear spin is:

$$\vec{S} = \hbar\sqrt{I(I+1)} \quad (3.1)$$

where I represent the spin quantum number.

We are interested in the z-component of the angular moment, S_z , which is restricted to integral steps of \hbar ranging from $-I$ to $+I$. For a nucleus with $I = 1/2$ we will have two possible values of S_z : $+1/2\hbar$ and $-1/2\hbar$, corresponding to spin quantum numbers $m_z = +1/2$ and $m_z = -1/2$, respectively. The magnetic moment of a nuclear spin, $\vec{\mu}$, is proportional to its spin angular momentum, $\hbar\vec{I}$, by a factor, γ , which depends on the type of nuclei.

$$\vec{\mu}_n = \gamma_n \hbar \vec{I} \quad (3.2)$$

In Table 3.1 are reported some useful NMR properties of the most important nuclear spins for protein NMR studies.

In the absence of a magnetic field a collection of nuclear spins can adopt all possible spatial orientations. When placed in a magnetic field magnetic moments of spin $1/2$ nuclei assume two orientations (Figure 3.1), either in the same direction or in the opposed direction with respect to the magnetic field.

nuclei	γ (rad \cdot sec ⁻¹ \cdot gauss ⁻¹)	I	Natural abundance (%)
¹ H	26.753	1/2	99.98
² H	4.106	1	0.016
¹⁹ F	25.179	1/2	100
¹³ C	6.728	1/2	1.108
¹⁵ N	-2.712	1/2	0.37
³¹ P	10.841	1/2	100

Table 3.1 Several properties of NMR active nuclei are reported such as giromagnetic ratio (γ), spin quantum number (I) and natural abundance¹⁰⁰.

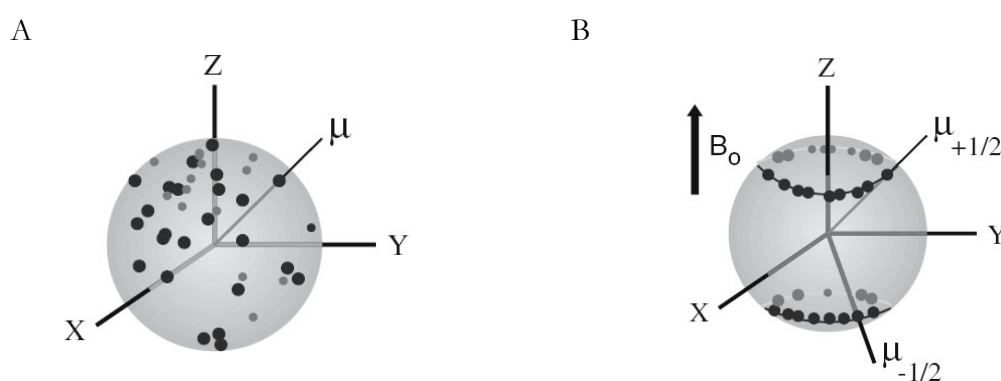


Figure 3.1 A) Orientation of nuclear magnetic dipoles in the absence of magnetic field. The orientation of one dipole μ is indicated by a line starting from the center of axis and the orientations of dipoles are indicated as dots. B) Orientation of nuclear spin dipoles in a magnetic field (B_0)¹⁰⁰.

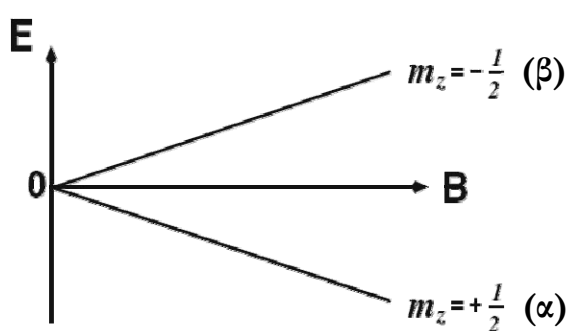


Figure 3.2 The energy separation between the ground and excited states increases with the magnetic field. The ground (lower energy, $m_z = +1/2$) state is referred also as α and the excited (higher energy, $m_z = -1/2$) state is referred also as β ¹⁰⁰.

The possible orientations of nuclear spins depends on the spin quantum numbers m_z which are in turn restricted in the range $\pm \hbar$. The magnetic moments cannot orient parallel to the magnetic field because of the restrictions placed on the value of μ_z by the quantum mechanical properties of the system.

The energy of a state depends on the interaction of the aligned magnetic dipole with an externally applied magnetic field. The calculation of the energy is expressed in the form:

$$E = -\vec{\mu} \cdot \vec{B}_o \quad (3.3)$$

where \vec{B}_o indicates the applied magnetic field.

Considering a nucleus with spin quantum number $I = 1/2$ ($m_z = \pm 1/2$) it is possible to draw an energy diagram as shown in Figure 3.2 which shows the dependence of the energy levels on the magnetic field. Calculating the energy associated with each of the energy levels is possible to compute the energy difference between the ground and excited states.

$$E_\alpha = -\frac{\gamma\hbar B_o}{2} \quad E_\beta = +\frac{\gamma\hbar B_o}{2} \quad (3.4)$$

$$\Delta E = E_\beta - E_\alpha = \gamma\hbar B_o \quad (3.5)$$

Finally the application of the relationship $E = \hbar\omega$ gives the Larmor equation:

$$\omega_o = \gamma B_o \quad (3.6)$$

where ω_o refers to the frequency (in rad s^{-1}) in the applied magnetic field B_o .

3.1.2 Population distribution

To achieve the transition from a ground state to an excited state it is necessary to have a population difference between the two energetic states. For any system of energy levels at thermal equilibrium there will be more particles in the lower energy state(s) than in the higher energy state(s). The population distribution, at thermal equilibrium, can be calculated by using the Boltzmann distribution relationship:

$$\frac{P_{(m=-1/2)}}{P_{(m=+1/2)}} = e^{-\Delta E/KT} \approx 1 - \frac{\gamma\hbar B}{kT} \quad (3.7)$$

where P is the population in each state, T is the absolute temperature in Kelvin and k is the Boltzmann constant with value of $1.381 \cdot 10^{-23} \text{ JK}^{-1}$.

The population difference between the two levels is very small, on the order of 1 part in 10^6 . This small population difference makes NMR spectroscopy a relatively insensitive technique.

3.1.3 Bulk magnetization and Larmor precession

As described in the previous section, at thermal equilibrium the nuclear magnetic dipole moments are asymmetrically distributed on the energy levels with a slight preference for the lower energy state. By adding all the magnetic moments there will be a net magnetic field along the direction of the applied field (B_o). This is called the *bulk magnetization* of the sample and can be represented by a vector, called

the *magnetization vector*.

If we perturb the equilibrium (e. g., applying a RF pulse) the magnetization vector will shift from the equilibrium position and will start to rotate about the direction of B_o (Figure 3.3A). The vector rotates with a constant angle β in respect to the B_o direction and this particular motion is called *Larmor precession*. The Larmor precession is described by the relationship in equation 3.6 in which the frequency can be expressed in Hertz (Hz):

$$\nu_o = -\frac{1}{2\pi} \gamma B_o \quad (3.8)$$

3.1.4 Radio frequency pulses and signal detection

The perturbation of the equilibrium condition of the magnetization vector is achieved by applying an additional magnetic field (B_1) along the x or y axis oscillating at or near the Larmor frequency, i.e., *resonant* with the Larmor frequency. This second oscillating magnetic field is generated by applying some radiofrequency power through a coil which is conveniently placed near the sample. By appropriately setting the frequency of the oscillating field it is possible to reach an *on resonance* condition where the magnetization vector is forced to move away from the z axis, usually coincident with the B_o direction, and to place itself and precess in the xy plane (Figure 3.3B).

The precession of the magnetization vector in the xy plane is detected in the NMR experiment. The signal detection is achieved by using a coil in which precession of the magnetization vector induces a current that is amplified and recorded (Figure 3.3C). The return to the equilibrium state determines a

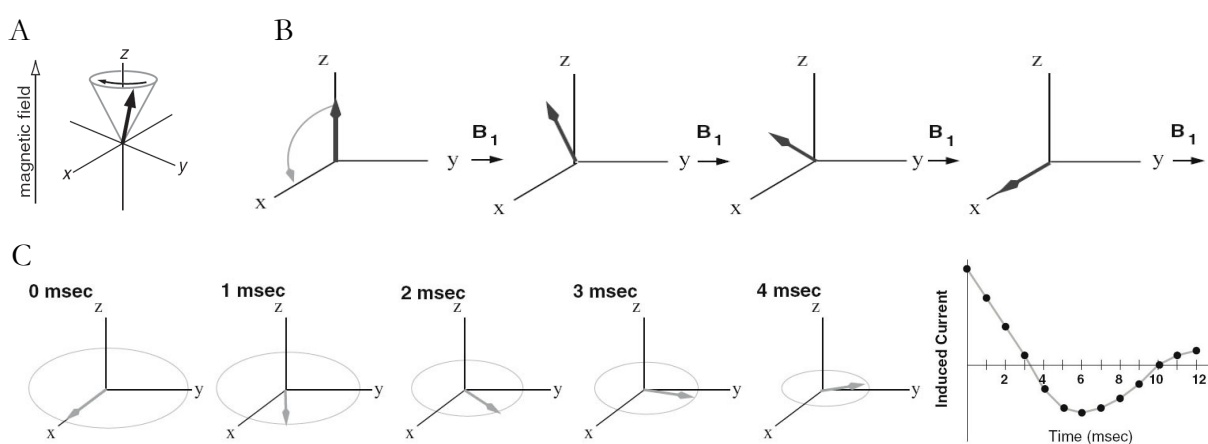


Figure 3.3 A) The magnetization vector, once away from the equilibrium position, executes a precessional motion in which the vector sweeps out a cone of constant angle to the magnetic field direction. B) Effect of a RF magnetic field (B_1) applied along the y axis on the magnetization vector. The magnetization moves towards the transverse plane along the x - y plane. C) Transverse magnetization precession after a 90° pulse and FID with the points indicating the data sampled during acquisition¹⁰⁰.

decrease of the acquired signal during time: this is the so-called *free induction decay* (FID).

3.2 Chemical shift

The NMR frequency of a nucleus in a molecule is determined by its gyromagnetic ratio (γ) and the strength of the applied magnetic field (B_0) (equations 3.6 and 3.8). For example proton and ^{13}C nuclei resonate respectively at 500 and 125 MHz in a 11.7 Tesla field. However not all protons, nor all ^{13}C nuclei, have identical resonance frequencies since ν depends on the local electron distribution. This effect is called *chemical shift* (δ) and is generally defined as the difference between the resonance frequency of the nucleus of interest (ν) and that of a reference nucleus (ν_{ref}):

$$\delta = \frac{\nu - \nu_{ref}}{\nu_{ref}} \quad (3.9)$$

The frequency difference $\nu - \nu_{ref}$ is divided by ν_{ref} so that δ is a molecular property, independent of the magnetic field used to measure it. Chemical shift values are quoted as parts per millions, or ppm. In the NMR terminology a resonance is at high field (or up-field) relatively to another resonance if it has a smaller δ in ppm values.

The chemical shift is a fundamental NMR parameter as it gives detectable and separate signals for the hundreds of protons of a protein and these signals can be therefore distinguished and assigned. Chemical shift dispersion arises because different residues experience different microenvironments; e.g., internal residues in globular proteins are shielded from the solvent and more in contact with other residues. Chemical shift dispersion is also observed among protons within the same residue, due to different local electron distribution.

The characteristic ^1H chemical shifts for amino acid residues of extended random chain structures are reported in Table 3.2. The data were obtained from NMR measurements in aqueous solution of 20 tetrapeptides H-Gly-Gly-Xxx-Ala-OH, where Xxx represents one of the 20 different amino acids¹⁰². It is well established that chemical shifts deviations from the random coil values are strongly correlated with protein secondary structures¹⁰³. In particular the chemical shifts of NH and H α can be indicative of regular secondary structures. In helices NH and H α have chemical shifts up-field respect to the average for extended chain conformations (≈ 0.1 ppm for NH; ≈ 0.39 ppm for H α), while in β -sheets they resonate at lower field.

3.3 Scalar coupling

Scalar coupling, known also as spin-spin or J-coupling, arises from the interaction of the nuclear magnetic moment with the electrons involved in the chemical bond. Because of its nature scalar

coupling is the basis of the resonance assignment methods. The nuclear magnetic dipole of one nucleus can affect the polarization of surrounding electrons and in turn this produces a change in the magnetic field at the coupled spin. For example we can consider a C-H group in a molecule, as illustrated in Figure 3.4. The proton nuclear spin polarizes the electrons in the bonding orbitals and this alters the magnetic field at the carbon nucleus. Protons have two possible spin states, one will increase the

residue	NH	H α	H β	others
Gly	8.39	3.97		
Ala	8.25	4.35	1.39	
Val	8.44	4.18	2.13	γ CH ₃ 0.97, 0.94 γ CH ₂ 1.48, 1.19
Ile	8.19	4.23	1.90	γ CH ₃ 0.95 δ CH ₃ 0.89 γ H 1.64
Leu	8.42	4.38	1.65, 1.65	δ CH ₃ 0.94, 0.90 γ CH ₂ 2.03, 2.03 δ CH ₂ 3.68, 3.65
Pro		4.44	2.28, 2.02	
Ser	8.38	4.50	3.88, 3.88	
Thr	8.24	4.35	4.22	γ CH ₃ 1.23
Asp	8.41	4.76	2.84, 2.75	
Glu	8.37	4.29	2.09, 1.97	γ CH ₂ 2.31, 2.28 γ CH ₂ 1.45, 1.45 δ CH ₂ 1.70, 1.70 ϵ CH ₂ 3.02, 3.02 ϵ NH ₃ 7.52
Lys	8.41	4.36	1.85, 1.76	γ CH ₂ 1.70, 1.70 δ CH ₂ 3.32, 3.32 NH 7.17, 6.62
Arg	8.27	4.38	1.89, 1.79	γ NH ₂ 7.59, 6.91 γ CH ₂ 2.38, 2.38 δ NH ₂ 6.87, 7.59
Asn	8.75	4.75	2.83, 2.75	γ CH ₂ 2.64, 2.64 ϵ CH ₃ 2.13
Gln	8.41	4.37	2.13, 2.01	
Met	8.42	4.52	2.15, 2.01	
Cys	8.31	4.69	3.28, 2.96	2H 7.24 4H 7.65
Trp	8.09	4.70	3.32, 3.19	5H 7.17 6H 7.24 7H 7.50 NH 10.22 2,6H 7.30
Phe	8.23	4.66	3.22, 2.99	3,5H 7.39 4H 7.34
Tyr	8.18	4.60	3.13, 2.92	2,6H 7.15 3,5H 6.86
His	8.41	4.63	3.26, 3.20	2H 8.12 4H 7.14

Table 3.2 Random coil ¹H chemical shifts for the 20 amino acid residues.

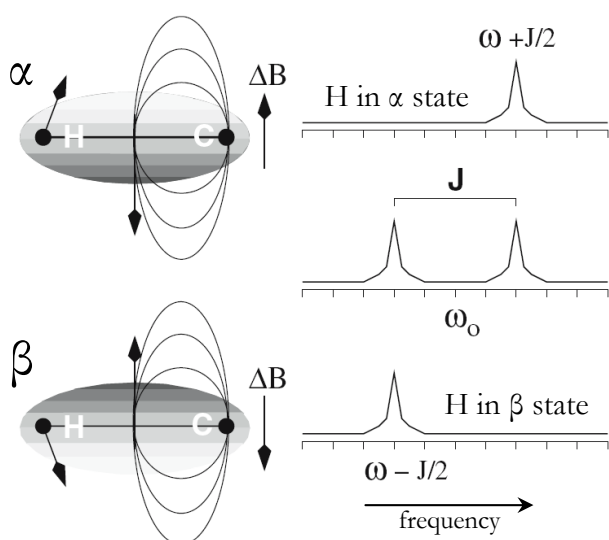


Figure 3.4 Nuclear spin coupling in a ^{13}C -H group. Two C-H groups with different proton spin states are shown. The oval black lines at the carbon nucleus indicate the polarized bonding electrons. The observed carbon spectrum is shown in the middle and is the sum of the carbon spectra shown in the upper and lower left panels. The sign of the difference in the magnetic field is also indicated with a black arrow on the right of each C-H group¹⁰⁰.

effective field at the carbon nucleus (Figure 3.4 upper molecule) and the other will decrease it (Figure 3.4 lower molecule). This alteration in the effective magnetic field at the carbon nucleus will lead to a shift of the carbon resonance frequency. In addition there will be a splitting into two signals of the carbon spectrum due to the two spin states possible for the proton. The separation between the two signals is the J-coupling. The proton spin will experience a shift in frequency in the same way as its coupled partner, giving rise to exactly the same splitting of the proton resonance signal as the attached carbon.

The strength of the J-coupling depends on several factors, including the gyromagnetic ratio of the coupled spins, the number of bonds connecting the coupled spins, and the conformation of the intervening bonds in the case of multiple bond couplings. In Table 3.3 are reported some values

of J-coupling constants that are important in biomolecular NMR.

Couplings involving heteronuclear (^{13}C or ^{15}N) spins		Proton-proton couplings	
C-N	14 Hz	H-C-H	-12 to -15 Hz
C-C	35 Hz	H-C-C-H	2-14 Hz
H-N	92 Hz	H-C=C-H	10 (cis)/17 (trans)
H-C	130 Hz	H-N-C-H	1-10 Hz
H-C-C	5 Hz (two bond coupling)		(3 Hz for α -helix β) (10 Hz for β -strand)

Table 3.3 Homonuclear and heteronuclear coupling constants. Homonuclear (proton-proton) and heteronuclear coupling constants commonly found in biopolymers are listed¹⁰⁰.

3.4 Relaxation

3.4.1 Time dependent field fluctuations

All excited systems eventually return to the ground state by relaxation. In the case of NMR, relaxation is the process by which the spins in the sample come to equilibrium with the surroundings. There are

three main relaxation processes involved in NMR magnetization relaxation: spin-lattice relaxation (T_1), spin-spin relaxation (T_2) and the Nuclear Overhauser Effect (NOE).

The relaxation process in NMR is due to the presence of oscillating electromagnetic fields whose frequencies match the absorption frequencies of the NMR transitions. In the case of T_1 relaxation these field fluctuations are orthogonal to the B_0 while for T_2 the fluctuations are in the B_0 direction. The oscillating fields are created by random rotational motion of the molecules or by internal motions within the molecule. These oscillating fields are produced by two principal mechanisms: dipolar coupling between nearby spins and chemical shift anisotropy (CSA) associated with a single spin.

CSA

The electron shielding is anisotropic therefore the effective field at the nucleus will vary depending on the orientation of the electrons with respect to the external field. The effective field B is reduced by the surrounding electron density, by a shielding factor σ , giving an observed chemical shift of ω_s as shown in the following equations:

$$\omega_s = \gamma B \quad B = (1 - \sigma)B_0 \quad (3.10)$$

Dipolar coupling

Dipolar coupling arises when the magnetic field of one nuclear spin affects the local magnetic field of another spin. The intensity of the dipole field depends on both the orientation of the two spins and the distance between them. If either of these properties are time dependent then the magnetic field will vary with time.

Both dipolar coupling and chemical shift anisotropy generate magnetic field fluctuations at the nucleus. The random molecular motion must contain field fluctuations at the appropriate frequencies to stimulate transitions. The intensity of the magnetic field fluctuations as a function of the frequency is represented by the spectral density function ($J(\omega)$) which is dependent on the rotational correlation time (τ_c) (Figure 3.5). For random isotropic tumbling due to Brownian motion it is expressed with the following equation:

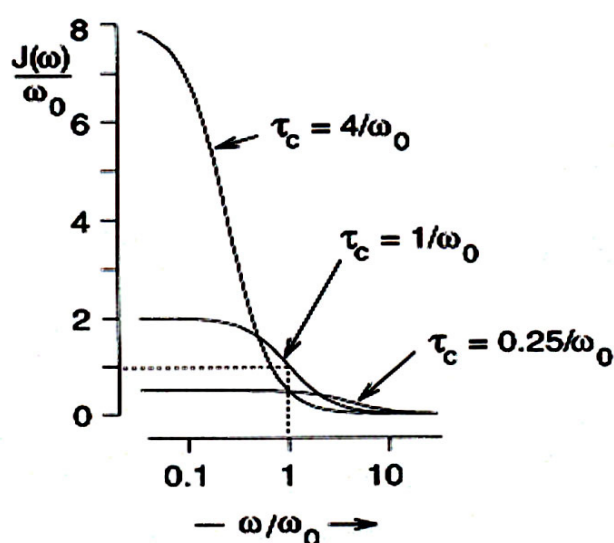


Figure 3.5 The spectral density function $J(\omega)$ drawn for three values of the correlation time τ_c ¹⁰¹.

$$J_{(\omega)} = \frac{2\tau_c}{1 + \omega^2\tau_c^2} \quad (3.11)$$

The rotational correlation time is related to the isotropic rotational diffusion constant (D_{iso}) and proportional to the size of the protein and the viscosity of the solution:

$$\tau_c = \frac{1}{6D_{iso}} \quad \tau_c = \frac{4\pi\eta a^3}{3kT} \quad (3.12)$$

where a is the radius of the protein and η is the viscosity.

3.4.2 Longitudinal and transverse relaxation

Spin-lattice relaxation

Spin-lattice relaxation, or longitudinal relaxation, arises from a re-alignment of the bulk magnetic moment along the static field to regenerate the original Boltzmann population difference (Figure 3.6 left panel). It is referred to as spin-lattice relaxation because it involves the transfer of energy from the excited state to the surroundings or lattice. The rate of this process is characterized by a relaxation time T_1 or by a relaxation rate R_1 , ($R_1 = T_1^{-1}$). Typical T_1 values range from 1 sec for proteins to tens of seconds for small molecules.

T_1^{-1} is proportional to the spectral density ($J_{(\omega_o)}$) as follows:

$$\frac{1}{T_1} = \gamma^2 \langle B_\mu^2 \rangle J_{(\omega_o)} \quad (3.13)$$

where $\langle B_\mu^2 \rangle$ is the mean square value of the local field. Figure 3.5 shows that $J_{(\omega_o)}$ is small for $\tau_c^{-1} \ll \omega_o$ or $\tau_c^{-1} \gg \omega_o$ and reaches a maximum when τ_c^{-1} matches the Larmor frequency ($\omega_o\tau_c = 1$). This behaviour is summarised in Figure 3.6 right panel. For rapidly tumbling molecules with $\omega_o\tau_c \ll 1$, $J_{(\omega_o)} \approx 2\tau_c$ and the relaxation gets slower as the mean tumbling rate is increased. Conversely, slowly tumbling molecules have $\omega_o\tau_c \gg 1$ and $J_{(\omega_o)} \approx \frac{2}{\omega_o^2\tau_c}$, so that the relaxation accelerates as the tumbling speeds up. The maximum relaxation rate occurs for $\omega_o\tau_c = 1$ (minimum T_1), at which point $J_{(\omega_o)} = \frac{1}{\omega_o}$. For macromolecules that rotate slowly relative to ω_o , so that $\omega_o\tau_c \gg 1$, T_1 increases proportionally to τ_c .

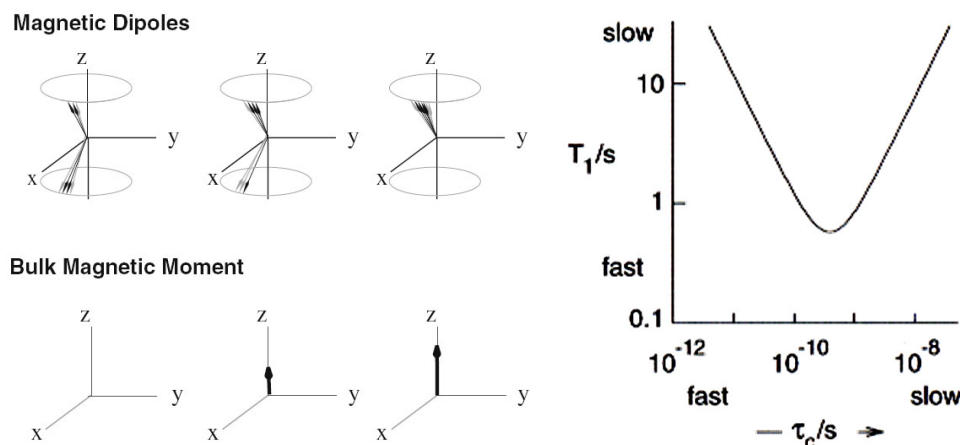


Figure 3.6 The left panel shows, from top to bottom, the individual nuclear moments as a function of time and the bulk magnetization as a function of time. The dependence of the spin-lattice relaxation time (T_1) on the correlation time (τ_c) is shown (right panel). Regions with fast and slow tumbling and relaxation are indicated^{100, 101}.

Spin-spin relaxation

The spin-spin relaxation, or transverse relaxation, process causes dephasing of the coherent transverse magnetization as it precesses around B_o . This process is illustrated in Figure 3.7 left panel. Since the dephasing does not affect the net population of excited spins, but simply their phase coherence, the system does not reach thermal equilibrium by this relaxation process. The time constant for this process is T_2 or in terms of rate constant R_2 , ($R_2 = T_2^{-1}$).

The relation between T_2 and the spectral density ($J_{(\omega_o)}$) is:

$$\frac{1}{T_2} = \frac{1}{2} \gamma^2 \langle B_\mu^2 \rangle J_{(\omega_o)} + \frac{1}{2} \gamma^2 \langle B_\mu^2 \rangle J_{(o)} \quad (3.14)$$

The dependence of the spin-spin relaxation time T_2 on the correlation time τ_c is shown in Figure 3.7 right panel. T_2 increases as the tumbling gets faster, while in the slow motion limit $\omega_o \tau_c \gg 1$, typical of macromolecules, it is simply inversely proportional to the correlation time. The two relaxation times, T_1 and T_2 , are identical in the extreme narrowing conditions $\omega_o \tau_c \gg 1$.

In general T_1 is greater than T_2 because there are additional causes for loss of transverse magnetization. The main reason for this loss is that the static magnetic field B_o is not uniform throughout the sample. If we divide the sample in small regions where the field is uniform (isochromats), then the total magnetization is the sum of all these regions. Each isochromat contributes a precessing vector which differs slightly in frequency. Consequently the transverse magnetization of the whole sample is reduced and the transverse relaxation is faster (i.e., T_2 is shorter).

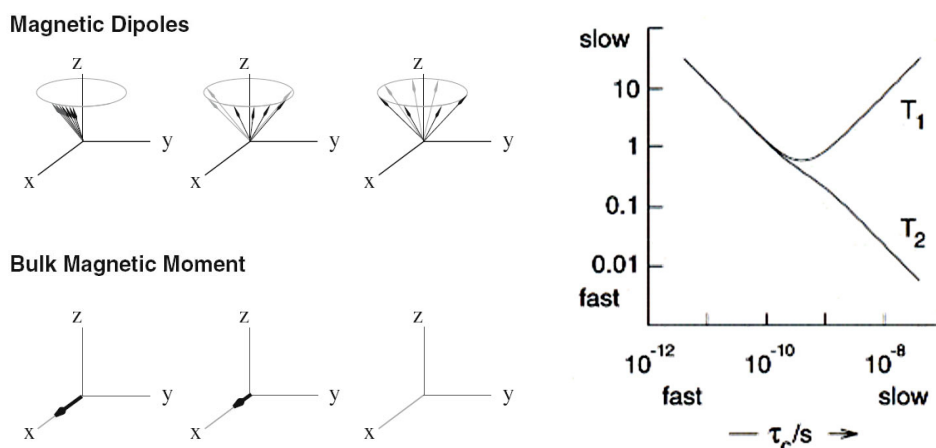


Figure 3.7 The left panel shows, from top to bottom, the individual nuclear moments as a function of time and the bulk magnetization as a function of time. The dependence of the spin-lattice relaxation time (T_1) and spin-spin relaxation time (T_2) on the correlation time (τ_c) is shown (right panel). Regions with fast and slow tumbling and relaxation are indicated^{100, 101}.

Both longitudinal and transverse relaxation, characterized by the relaxation time constants T_1 and T_2 respectively, are processes that follow a first-order exponential decay. The estimate of the T_1 and T_2 rate constants is usually achieved by fitting the magnetization decay with the following function:

$$I(t) = I_0 e^{-t/T} = I_0 e^{-Rt} \quad (3.15)$$

3.4.3 Nuclear Overhauser Effect

In addition to scalar coupling, i.e., interaction of the nuclear magnetic moment with the electrons involved in the chemical bond, nuclear spins can interact through space via dipolar coupling. Dipolar coupling occurs when the magnetic field generated by one nuclear dipole affects the magnetic field at another nucleus. The magnitude of the dipolar coupling depends on the strength of the magnetic field generated by one spin and the size of the magnetic moment of the recipient spin. This form of coupling does not require the presence of connecting bonds; it is possible for spins on entirely different molecules to show dipolar coupling. Dipolar coupling can be used to determine the distance between two spins, and in certain cases, the orientation of the inter-atomic vector relative to the applied B_0 field. Consequently dipolar coupling is a major source of information for structure determination by NMR spectroscopy.

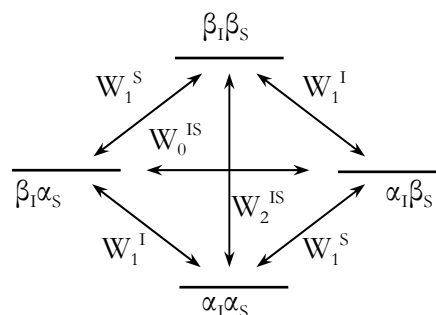


Figure 3.8 Energy levels for a couple of $1/2$ spins nuclei I and S, showing the six possible relaxation pathways¹⁰¹.

The Nuclear Overhauser Effect (*NOE*) is defined as the fractional change in intensity of one NMR line when another resonance is selectively perturbed.

Consider a pair of dipolarly coupled $\frac{1}{2}$ spins, I and S . Denoting the $+\frac{1}{2}$ and $-\frac{1}{2}$ spin state of each spin with α and β respectively, the four possible combination of the two spins I and S are usually represented as $\alpha_I\alpha_S$, $\alpha_I\beta_S$, $\beta_I\alpha_S$ and $\beta_I\beta_S$ (see Figure 3.8).

In Figure 3.8 six relaxation pathways are indicated:

- W_1^I and W_1^S , also called single quantum transitions, correspond to the single spin flipping $\alpha_I \leftrightarrow \beta_I$ and $\alpha_S \leftrightarrow \beta_S$ and are the spin-lattice relaxation processes.
- W_o^{IS} and W_2^{IS} , also called zero quantum and double quantum transition, respectively, correspond to the simultaneous flipping of both I and S : $\alpha_I\alpha_S \leftrightarrow \beta_I\beta_S$ (both spins flipping in the same direction) and $\alpha_I\beta_S \leftrightarrow \beta_I\alpha_S$ (I and S flipping in opposite directions). These relaxation processes are known as cross relaxation and allow the spin state populations to return to equilibrium following some disturbance.

Cross relaxation happens because the chaotic molecular motion, combined with the mutual dipolar interaction, causes the fluctuating local fields experienced by I and S to be correlated, with the results that the nuclei can undergo simultaneous spin-flips.

The NOE effect can be explained through the experiment described in Figure 3.9. In a homonuclear system, at thermal equilibrium, level $\alpha_I\alpha_S$ is the most populated, followed by $\alpha_I\beta_S$ and $\beta_I\alpha_S$ with essentially identical populations, followed by $\beta_I\beta_S$. The same population difference exists across all four single quantum transitions ($\alpha_I\alpha_S \leftrightarrow \beta_I\alpha_S$, $\alpha_I\alpha_S \leftrightarrow \alpha_I\beta_S$, $\alpha_I\beta_S \leftrightarrow \beta_I\beta_S$ and $\beta_I\alpha_S \leftrightarrow \beta_I\beta_S$) and so the equilibrium intensities are identical (Figure 3.9a). When the S transitions are saturated by an appropriate radiofrequency field, this equalizes the populations of levels $\alpha_I\alpha_S$ and $\alpha_I\beta_S$, and similarly those of levels $\beta_I\alpha_S$ and $\beta_I\beta_S$ (Figure 3.9b). This has no effect on the population difference across the I transitions, indeed the population differences $P\alpha_I\alpha_S - P\beta_I\alpha_S$ and $P\alpha_I\beta_S - P\beta_I\beta_S$ are, at this stage, unchanged. If the transition W_2^{IS} occurs, it will transfer population between $\alpha_I\alpha_S$ and $\beta_I\beta_S$ and restores the equilibrium population of these two states $1 + 2\Delta$ and $1 - 2\Delta$. The population difference across the I transition is now 3Δ (Figure 3.9c), and therefore the intensity of the I signal has increased by 50%. Cross relaxation has transferred magnetization from the saturated spin S to its dipolar-coupled spin I . With the same logic, if the W_o^{IS} transition takes place, it will transfer population between $\beta_I\alpha_S$ and $\alpha_I\beta_S$ and restore the equilibrium population of the two states. The population

difference across the I transition is now Δ (Figure 3.9d), and therefore the intensity of the I signal has decreased by 50%.

The NOE can be quantified by the parameter η :

$$\eta = \frac{i - i_o}{i_o} \quad (3.16)$$

where i is the perturbed NMR intensity of the spin I and i_o its normal intensity.

The maximum homonuclear NOE can be $\frac{1}{2}$ and the minimum -1 . In reality, neither W_2^{IS} nor W_o^{IS} dominates the other relaxation pathways and η is somewhere between the extremes. The parameter η has the same sign as $W_2^{IS} - W_o^{IS}$.

Both the relaxation rates W_2^{IS} and W_o^{IS} are related to the spectral density $J(\omega)$ as follows:

$$W_2^{IS} \approx J(2\omega_o) \quad \text{and} \quad W_o^{IS} \approx J(\omega_o) \quad (3.17)$$

Without giving all the mathematical details, it occurs that the proton-proton NOE η should be positive for fast motions ($\omega_o\tau_c \ll 1$) and negative for slow motions ($\omega_o\tau_c \gg 1$). The change of sign occurs when $W_2^{IS} = W_o^{IS}$ so that the effect of the two cross relaxation pathways cancel: this happens when $\omega_o\tau_c \approx 1$.

NOE are also observable for heteronuclear pairs of spins.

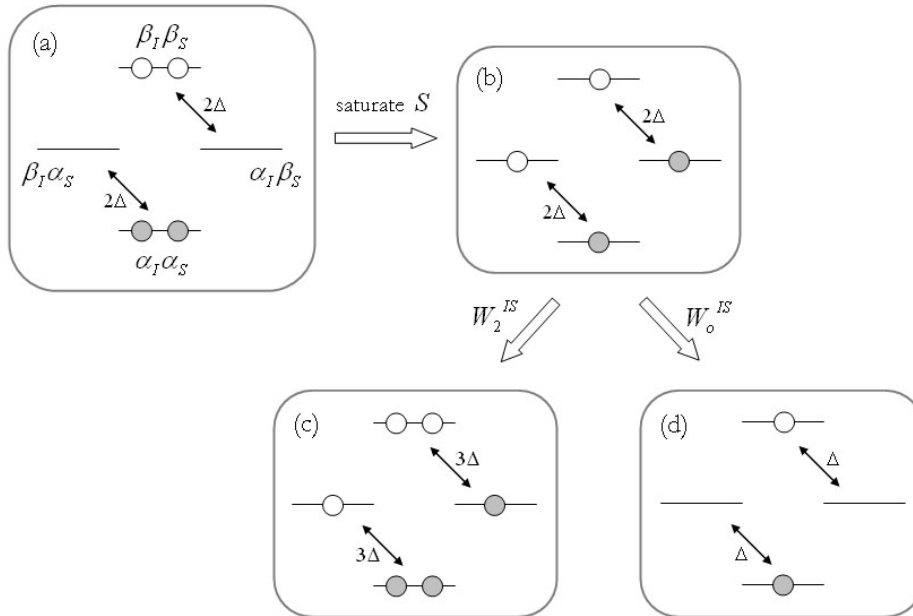


Figure 3.9 The origin of the NOE in a homonuclear $\frac{1}{2}$ spin system (I and S). Shaded circles indicate a population excess of Δ ; open circles indicate a population deficit of Δ . (a) Thermal equilibrium. (b) Effect of saturating both transitions of spin S . (c) Effect of W_2^{IS} cross relaxation. (d) Effect of W_o^{IS} cross relaxation¹⁰¹.

3.4.4 Relaxation and protein mobility

Protein backbone and side chains display varying degrees of flexibility, which allows many slightly different but related conformational sub-states to occur. Such fluctuations are known to differ in both timescale and magnitude, from rotation of methyl group (nanoseconds) to the flipping of buried tyrosine rings (seconds). Movement at molecular level is essential for many biological processes, including protein folding, protein-protein interaction and protein-ligand interaction.

The NMR relaxation is strictly related to the internal motions of a molecule and is sensitive to motions over a wide range of timescale. In the last 10 years NMR relaxation experiments have become an essential instrument in the study of protein dynamics as it can provide information about the internal motions at many sites throughout the protein. The development of heteronuclear and multidimensional NMR and the isotopic labelling methods, were fundamental in the growth of the relaxation studies applied to proteins. In particular, ^{15}N and ^{13}C relaxation data are widely used for studying backbone and side chains dynamics, respectively. The relaxation properties of protonated heteronuclei such as ^{15}N and ^{13}C are typically dominated by the dipole-dipole interaction with the attached proton(s); therefore, the relaxation data can be interpreted in terms of the motions of the ^{15}N - ^1H or ^{13}C - ^1H bond vector. Heteronuclear relaxation experiments typically provide heteroatom T_1 , T_2 and heteronuclear steady-states NOE . These data are subsequently interpreted in the context of motional models. The most used approach is the “model-free” analysis described by Lipari & Szabo¹⁰⁴, which in its simplest form allows extraction, from NMR relaxation measurements, of an overall rotational correlation time characterizing internal dynamics and of an order parameter, describing the amplitude of the internal motions.

Heteronuclear relaxation experiments can be used to study several biological processes.

In the characterization of ligand binding properties of a protein, the analysis of how the relaxation parameters vary upon binding can highlight the regions of the protein that show different flexibility in the apo and holo forms and are more affected by the presence of the ligand. Moreover, under some assumptions, changes in the NMR relaxation parameters can be linked to changes in thermodynamic quantities so that the entropic contribution to the site-specific binding can be delineated. Therefore, the information obtained from relaxation parameters, combined with structural data, can help in the comprehension of the mechanism of binding giving also an estimate of the role of the dynamics in driving molecular interaction.

In protein folding the NMR relaxation experiments are usually applied to unfolded or partially folded states. The variations of the NMR parameters are measured between the active and the denatured or partially folded states. This can highlight the regions of the protein that have a different backbone flexibility in the different stability states. The regions that are motionally more restricted in the unfolded or partially folded state are probably involved in some kind of structural interaction and can therefore

play an important role in the stability of the protein. Restriction of backbone motions can occur through formation of local hydrophobic clusters, through formation of elements of secondary structure or through long-range interactions in compact folding intermediates.

3.5 NMR experiments

3.5.1 1D NMR experiments

The simplest pulsed NMR experiment consist in a short RF-pulse followed by detection of the signal as depicted in Figure 3.10. The recorded signal, also called Free Induction Decay (FID), is then Fourier transformed, affording the normal NMR spectrum with absorption peaks at frequencies that represent the energy difference between the ground and excited states. A typical 1D NMR pulse sequence is formed by three sequential steps:

Preparation: Prior to the excitation pulse the spins are at thermal equilibrium and are subject to only the static B_0 field.

Excitation: During the excitation pulse the spins are subject to the static field B_0 plus the oscillatory excitation field B_1 .

Detection: The excited spins precess under the static field generating the free induction decay.

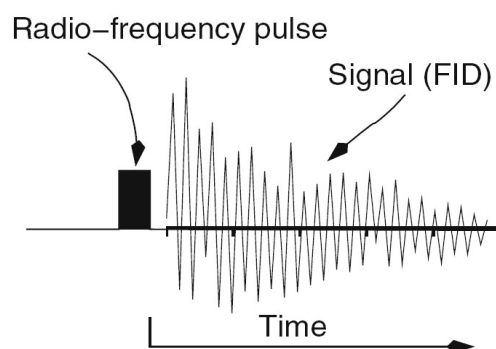


Figure 3.10 A simple one-pulse NMR experiment. The experiment begins with a short RF pulse and the induced signal (FID) is sampled as it evolves over time¹⁰⁰.

Therefore in one-dimensional pulsed NMR experiment the signal is recorded as a function of one time interval (t_1) and then Fourier transformed to give a spectrum which is in function of one frequency variable (f_1).

3.5.2 Multi-dimensional NMR experiments

In multi-dimensional NMR the signal is recorded as a function of more than one time variables; (t_1), (t_2), ... (t_n) and the resulting data are Fourier transformed a number of times correspondent to the number of time variables employed. The final spectrum has multiple frequency dimensions; (f_1), (f_2), ... (f_n), the number of which depends on the number of time variables included in the experiment. A general scheme for two-dimensional and three-dimensional experiments is given in Figure 3.11. In Figure 3.12 a general multi-dimensional NMR experiment is summarized. Any multi-dimensional NMR experiment can be divided into four basic elements: preparation, evolution, mixing and detection. For

clarity, each element will be described in the following for a general two-dimensional experiment (Figure 3.11 upper scheme):

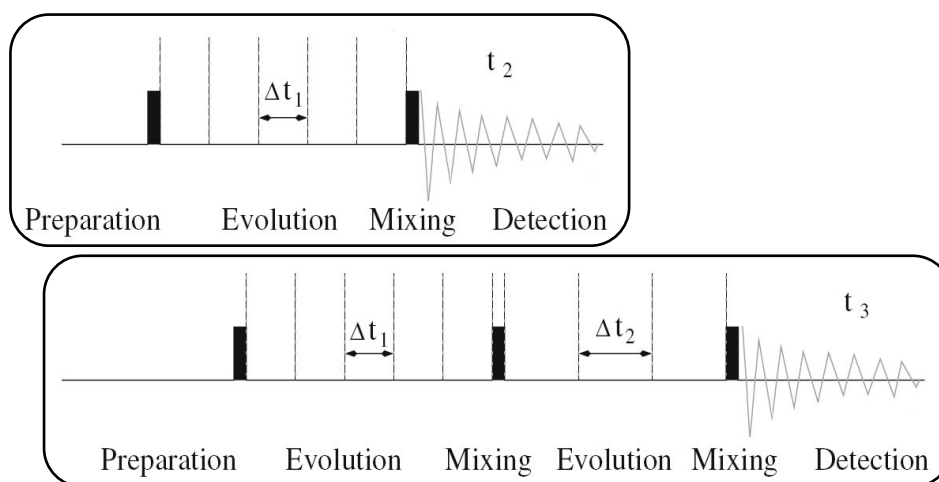


Figure 3.11 Two-dimensional and three-dimensional NMR experiments. The four central steps of multi-dimensional NMR experiments are shown: preparation, evolution, mixing and detection. Upper scheme: 2D experiment. Lower scheme: 3D experiment¹⁰⁰.

Preparation: This period has a fixed length and usually is employed to allow the spins to return to, or near, thermodynamic equilibrium.

Evolution: This period is used for the evolution of the magnetization that will generate the second dimension. It is referred to as the *indirect* detection domain, or dimension, because the excited state of the spin evolving under this period is not directly detected by the receiver coil. It is indeed sampled digitally, i.e., t_1 begins at zero and then is incremented by a constant amount, (Δt_1) .

Mixing: This event causes the magnetization that is associated with spin evolving under the t_1 period to become associated with spin evolving under the t_2 period. In the mixing period the chemical shift information of the spin evolving in t_1 is transferred to the spin evolving in t_2 through J-coupling or dipolar coupling.

Detection: During this time the magnetization precessing in the xy plane is detected in the normal fashion. It is referred to as the *direct* detection domain, or dimension, because the magnetization present in this period is directly detected by the receiver coil.

In the case of three- or more dimensional experiments, additional evolution and mixing periods will be inserted in the sequence thus introducing additional indirect dimensions (Figure 3.11 lower scheme).

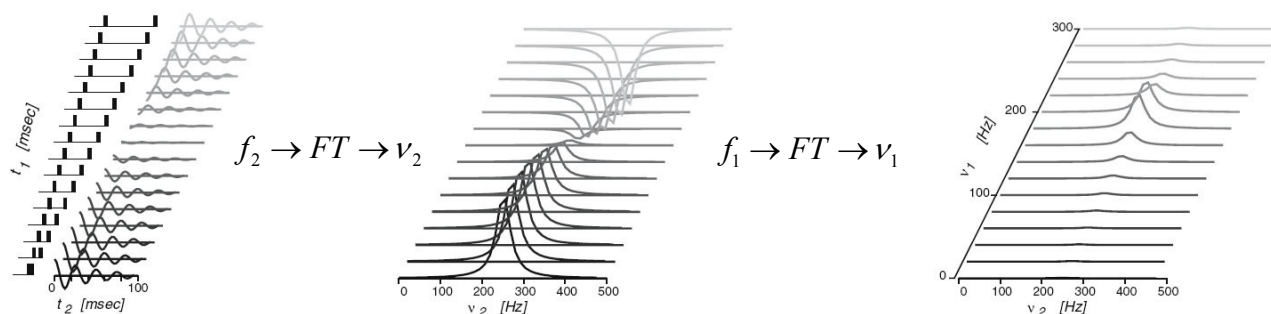


Figure 3.12 Steps of a general 2D NMR experiment. At the extreme left are shown the series of t_1 increments that generate the indirect dimension and the FIDs obtained for each of these increments. In the middle is shown the result of the first Fourier transform (FT) that converts the horizontal axis from the time domain t_2 to the frequency domain ν_2 . In the right panel the result of the second Fourier transform along t_1^{100} is shown.

3.6 NMR experiments in protein studies

3.6.1 Two-dimensional homonuclear J-correlated spectroscopy

The most common 2D homonuclear J-correlated experiments used in protein studies are the CORrelated SpectroscopY (COSY) and the TOrtal Correlation SpectroscopY (TOCSY).

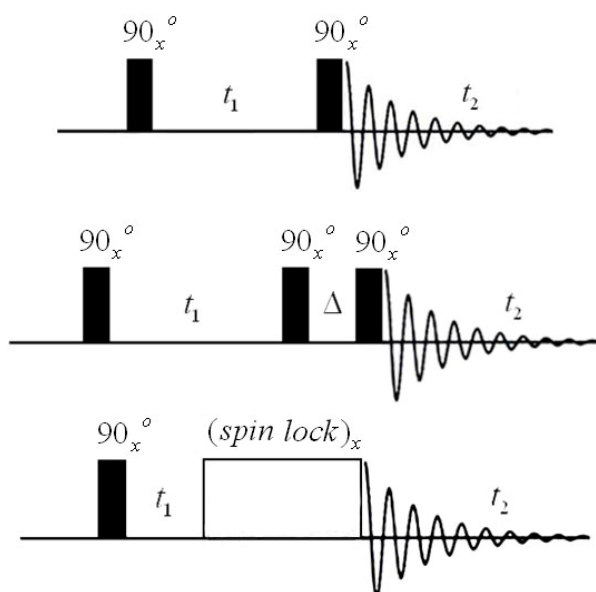


Figure 3.13 Pulse sequence diagram for COSY, DQF-COSY and TOCSY experiments. Upper panel: COSY; middle: DQF-COSY; lower panel: TOCSY¹⁰⁰.

The COSY experiment is the simplest and oldest 2D experiment and is used to detect scalar coupling interactions in spin systems (spin system = group of spins that are connected by scalar coupling). The experiment, as shown in Figure 3.13 upper panel, consists of a 90° excitation pulse, an evolution period, a second pulse that serves to transfer the magnetization from one coupled spin to another and a detection period. In a COSY experiment only J-coupling interactions between spins connected by three or less covalent bonds are detected.

The COSY experiment has been improved during the time and one of the most used variant is the double quantum filtered COSY (DQF-COSY). This experiment filters out any signal that do not arise from coupled spins giving a much cleaner spectrum.

The DQF-COSY sequence (Figure 3.13 middle) is very similar to the COSY sequence with the only difference that the single mixing pulse in the COSY has been replaced by two 90° pulses in the DQF-COSY.

Another homonuclear J-correlated NMR experiment largely used in protein NMR spectroscopy is the TOCSY. In TOCSY experiment the chemical shift information of one proton within a spin-system is relayed to all other protons within the spin system. The relay occurs by the sequential transfer of magnetization through the coupled network of spins. A schematic pulse sequence is reported in Figure 3.13 lower panel. The experiment consists of an initial 90° pulse, an evolution period and a long series of 180° (π) pulses that are applied to all of the coupled spins. The 180° pulses prevent evolution of the magnetization by chemical shift during the mixing time and is called spin lock because has the effect of keeping the magnetization in the xy plane. Under spin lock the system therefore only evolves under J-coupling causing transfer of magnetization between coupled spins. In a polypeptide chain the peptide bond interrupts the scalar coupling between the protons, and each amino acid residue constitutes an independent spin system. In the TOCSY spectrum scalar connections between protons within a single amino acid are highlighted.

3.6.2 Two-dimensional heteronuclear J-correlated spectroscopy

Proteins contain, in addition to proton nuclei, other nuclei species, i.e., nitrogen, carbon and oxygen. Among these, nitrogen and carbon are the most helpful in protein NMR spectroscopy. Heteronuclear experiments are designed to detect correlations between protons and their attached heteronuclear spins, such as nitrogen and carbon, using heteronuclear J-correlated spectroscopy. Several experiments involving magnetization transfer between hetero-nuclei has been developed that exploit heteronuclear scalar coupling. One of the most popular is the Heteronuclear Single Quantum Coherence (HSQC) experiment which, since its development, had a wide diffusion in protein NMR.

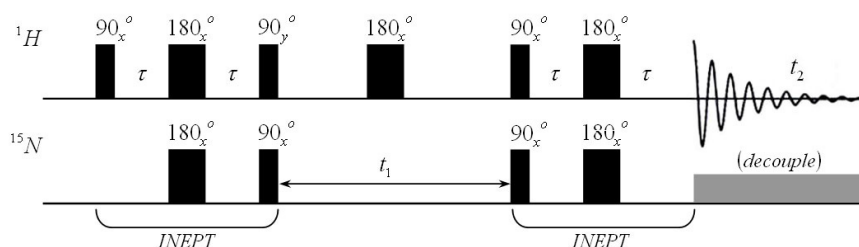


Figure 3.14 Pulse sequence diagram for ^1H - ^{15}N HSQC experiment¹⁰⁰.

The HSQC experiment, as

reported in the diagram in Figure 3.14 for a ^{15}N heteronucleus, contains two INEPT (Insensitive Nuclei Enhanced by Polarization Transfer) sequences which have the aim to transfer magnetization from the proton nuclei to the heteronuclei and viceversa. This step is very important to enhance the sensitivity of the experiment, indeed ^{15}N and ^{13}C heteronuclei have 300 and 32 fold less sensitivity, respectively, than protons. The HSQC experiment starts with an INEPT sequence that transfers the magnetization from the proton to the J-coupled insensitive spin. This first step is followed by the evolution period during which the magnetization evolves under the heteronucleus chemical shift. In the middle of the evolution period a 180° pulse is applied at the proton frequency that refocus the evolution due to J-coupling. The

second INEPT step, also called reverse INEPT, transfers back to the proton the magnetization which carries now the modulation due the heteronucleus chemical shift. At the end of this second INEPT the signal is acquired and recorded. During the acquisition time a decoupling pulse sequence is applied to the heteronucleus that remove the coupling and allows to obtain a single signal for every J-coupled pair of spins. The global appearance of an ^1H - ^{15}N HSQC spectrum is a series of dot signals characterized by

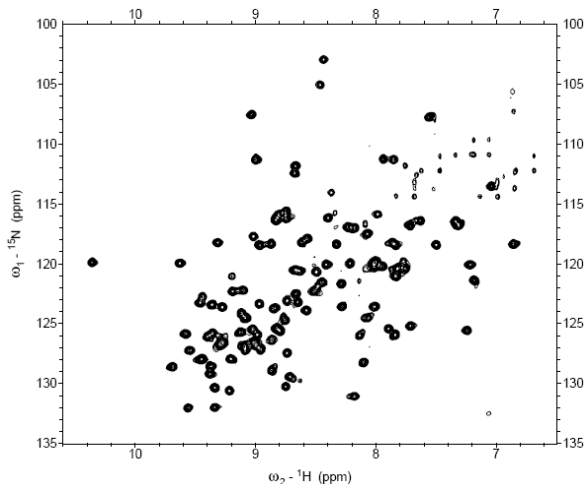


Figure 3.15 Example of an ^1H - ^{15}N HSQC spectrum.

two frequency coordinates, i.e., proton and heteronucleus (Figure 3.15). The ^1H - ^{15}N HSQC spectrum represents the fingerprint of a protein because the number and the position of each signal is representative for each single protein sample. Indeed a protein sequence contains one NH group for each amino acid residue, without considering the side chains, which will give a well distinguishable signal. The amide chemical shifts depend on the protein tertiary structure and are therefore specific of each protein.

3.6.3 Two-dimensional homonuclear dipolar correlated spectroscopy

One of the most popular 2D homonuclear experiments involving dipolar coupling is the Nuclear Overhauser Effect Spectroscopy (NOESY).

NOESY is one of the most useful experiments as it allows to correlate nuclei through space, provided that the two dipolar coupled spins are closer than 5\AA . In addition the measurement of signal intensity can afford distance information.

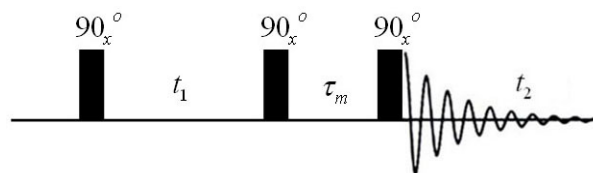


Figure 3.16 Pulse sequence diagram for NOESY experiment¹⁰⁰.

The pulse sequence starts with a 90° pulse followed by the evolution time. After the evolution time there are two 90° pulses separated by a mixing time (τ_m) after which the FID is recorded (Figure 3.16). During the mixing time magnetization is exchanged between dipolarly coupled spins, through the dipole-dipolar relaxation mechanism (see section 3.4.3). Since it is during the mixing period that magnetization is transferred between coupled spins, the mixing time length has to be properly set to allow the magnetization exchange.

3.6.4 Heteronuclear relaxation experiments

The measure of T_1 and T_2 relaxation times gives information on the motion rates associated with the observed nucleus (see section 3.4.4). In protein NMR, T_1 and T_2 relaxation rates are very important parameters used to identify and characterize mobility properties that can present functional relevance.

Measuring heteronuclear T_1

The measure of T_1 relaxation constant is based on the so called *inversion recovery* experiment. A schematic picture of inversion recovery pulse sequence is shown in Figure 3.17. The sequence starts

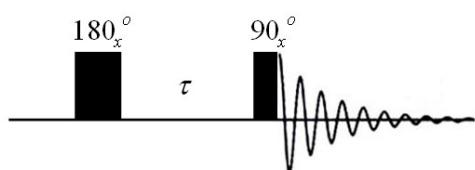


Figure 3.17 Pulse sequence diagram for T_1 inversion recovery experiment¹⁰⁰.

with a 180° pulse that inverts the spin populations and move the magnetization from the equilibrium $+z$ position to the anti-parallel $-z$ position. After a τ time period, during which the magnetization approaches the equilibrium population distribution, a 90° pulse is applied to move the magnetization in the transverse plane. Finally the signal is

recorded. The second pulse is essential because detection is possible only in the transverse plane. The spin-lattice relaxation process is hence sampled by recording several spectra with increasing τ time delays and the T_1 value is extracted by fitting the magnetization reduction trend with an exponential decay function as a function of the τ delay (see section 3.4.2).

In proteins, T_1 relaxation measurement is commonly performed on the ^{15}N heteronucleus. An HSQC-based pulse sequence is used in which, during the evolution period, an inversion recovery module is applied at the nitrogen frequency. The global appearance of the final spectrum is like an HSQC and the T_1 value is extracted by fitting the volume of the signals with an exponential decay function as a function of the τ delay.

Measuring heteronuclear T_2

Spin-spin relaxation refers to the dephasing of magnetization in the xy plane (see section 3.4.2).

The measure of T_2 constant is based on the Carr-Purcell-Meiboom-Gill (CPMG) pulse sequence.

In Figure 3.18 the basic pulse sequence of the CPMG experiment is shown. The CPMG sequence starts with a 90° pulse that move the magnetization in the xy plane. The spins start now to dephase and the magnetization decays. After a

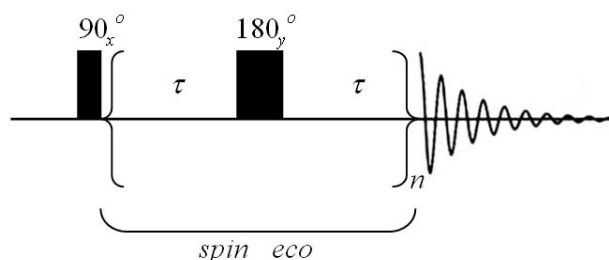


Figure 3.18 CPMG basic pulse sequence¹⁰⁰.

time τ a 180° pulse is applied that invert the spins precession direction. After another τ time period all the spins will be refocused in the same position as just after the first 90° pulse. The $\tau - 180_y^\circ - \tau$ is reported also as spin echo sequence because the 180_y° pulse, applied in the middle of two equal time periods, has the property to reflect the precessing spins in the yz plane and to refocus them after the second τ period. During the spin echo step, lost of spins coherence and refocusing takes place, and the spins undergoes spin-spin relaxation determining a reduction of magnetization. The spin echo module is repeated several times in the CPMG pulse sequence and the residual magnetization after each spin echo step is measured. Subsequent fitting of the magnetization decay leads to the extraction of the T_2 time constant (see section 3.4.2).

As in the case of T_1 , also T_2 relaxation parameter estimate, in protein samples, is performed on the ^{15}N heteronucleus. The pulse sequence for heteronuclear T_2 measurement is based on the HSQC pulse sequence in which, during the relaxation period, a CPMG module is applied at the nitrogen frequency. The global appearance of the final spectrum is like an HSQC and the T_2 value is extracted by fitting the volume of the signals with an exponential decay function as a function of the CPMG module length.

3.6.5 Transverse relaxation-optimized spectroscopy

Transverse Relaxation-Optimized Spectroscopy (TROSY) is a technique that increases the resolution and sensitivity of heteronuclear NMR experiments on larger proteins at high magnetic field strengths. This technique can be applied to 2D proton-nitrogen and proton-carbon correlated spectroscopy as well as to several triple resonance experiments. To explain how the TROSY technique works we will consider its application to a 2D proton-nitrogen HSQC experiment. If decoupling of the spectrum during t_1 and t_2 is not performed then a quartet will be found for each NH group, as illustrated in Figure 3.19B left panel, where each signal corresponds to one energetic level associated with the coupled nitrogen and proton spins. The four signals of the quartet show different linewidths due to the different relaxation rate of each energy level. The relaxation transitions associated with the NH coupled spins are reported in Figure 3.19A. The transitions between levels 3 and 4 ($3 \leftrightarrow 4$) and levels 2 and 4 ($2 \leftrightarrow 4$) have smaller relaxation rates than the transitions between levels 1 and 3 ($1 \leftrightarrow 3$) and levels 1 and 2 ($1 \leftrightarrow 2$). For this reason the signal associated with the $3 \leftrightarrow 4$ and $2 \leftrightarrow 4$ transitions shows a narrow linewidth in both dimensions and the signal associated with $1 \leftrightarrow 3$ and $1 \leftrightarrow 2$ transitions shows a broad linewidth (Figure 3.19B left panel). If proton decoupling is applied during t_1 , and nitrogen decoupling during t_2 , then spin states of the coupled spins are mixed leading to an averaging of relaxation rates. This results in a single signal, as in the case of a common HSQC spectrum, with an average linewidth (Figure 3.19B middle panel). The TROSY technique allows to select the magnetization correspondent

to the $3 \leftrightarrow 4$ and $2 \leftrightarrow 4$ transitions affording signals with narrower linewidth (Figure 3.19B right panel). Since the TROSY removes three signal components it suffers of a reduced sensitivity, around one-fourth, with respect to a normal HSQC. However, in the case of proteins larger than 40-50 KDa, the observed TROSY signal-to-noise results higher than that found in a standard HSQC.

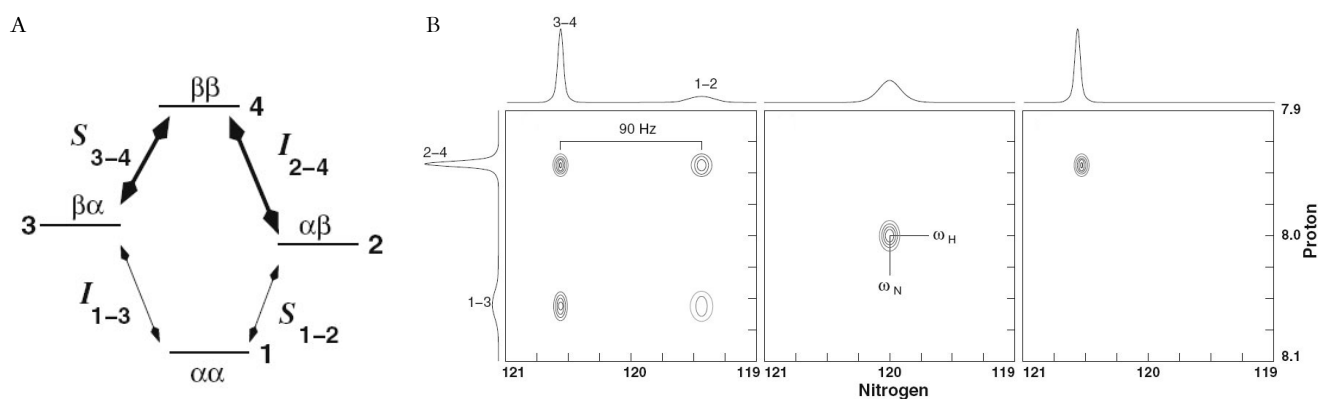


Figure 3.19 A) N-H coupled spin pair relaxation rates. Energy levels diagram is shown along with transitions between levels. The thicker arrows represent transitions that have narrow resonance lines. I refers to the proton and S to the nitrogen. B) Left panel shows the appearance of a signal in an HSQC spectrum in which decoupling was omitted both during t_1 and t_2 . A 90Hz splitting, equal to $J_{\text{N-H}}$, is observed in each dimension. Middle panel shows the effect of decoupling in both t_1 and t_2 dimensions. Right panel shows the TROSY spectrum of the same N-H group. For each signal cross section and associated transition are shown in either the nitrogen dimension (top) or proton dimension (side).¹⁰⁰

CHAPTER 4: Publication 1

Title: NMR unfolding studies on a liver bile acid binding protein reveal a global two-state unfolding and localized singular behaviors.

Short presentation of publication 1:

Folding studies on some members of iLBP family, i.e., cellular retinoic binding protein I (CRABI), intestinal fatty acid-binding protein (IFABP) and ileal lipid binding protein (iLBP) have been reported and the folding mechanism of these proteins have been characterized.

In the following paper the urea unfolding and the structural stability of chicken liver bile acid-binding protein (cL-BABP) is studied by nuclear magnetic resonance (NMR) spectroscopy and differential scanning calorimetry (DSC). Comparison of the folding properties of closely related proteins is also reported.

Hydrogen-deuterium (HX) exchange experiments were carried out to estimate the protection factors and the protein stability. The band-selective optimized flip-angle short transient (SOFAST) experiment has been used that allows very short repetition rates and fast acquisition, thus extending the analysis to residues exhibiting fast exchange. The protection factors indicated that the most stable region of cL-BABP involves A-C, I and J strands. In addition, ΔG_{op} value was determined for a large group of residues and the three largest values (detected for T110, I111 and R120) were used to estimate the conformational stability of the protein, affording a value of ΔG_{op} of 7.78 ± 0.03 kcal mol⁻¹.

The unfolding of cL-BABP was also analysed in a residue specific manner by urea unfolding experiments in equilibrium conditions. ¹H, ¹⁵N-HSQC experiments were collected at different urea concentrations and the volume of signals was monitored. The majority of the residues displayed a similar behaviour consistent with a one step cooperative unfolding transition. A minor group did not fit the one step unfolding model and exhibited a very fast decrease of signal intensity at low urea concentration. This residues, located in the G and H strands and in DE, EF and FG loops, were suggested to play some role as unfolding initiation sites. The interested regions are indeed characterized, in the folded structure, by a high mobility possibly responsible for a higher urea susceptibility. The singular behaviour of this group of residues could depend on the onset of intermolecular interactions. DSC measurements indeed identified strong exothermic phenomena that prevented a thermodynamic evaluation of the unfolding enthalpy. The strong exothermic behaviour observed was related to an aggregation processes involving those residues exhibiting a singular behaviour.

cL-BABP shows different stability if compared with proteins of the same family. In IFABP native-like structures were detected, at very high urea concentrations, in or near DE, IJ and helixI-helixII turns, while in cL-BABP structured regions at high urea concentrations are present in A-D strands.

In this work I contributed in the data analysis and in the writing of the scientific paper.



NMR unfolding studies on a liver bile acid binding protein reveal a global two-state unfolding and localized singular behaviors

Mariapina D'Onofrio^a, Laura Ragona^b, Dimitrios Fessas^c, Marco Signorelli^c,
Raffaella Ugolini^a, Massimo Pedò^a, Michael Assfalg^a, Henriette Molinari^{a,*}

^a Dipartimento Scientifico e Tecnologico, Strada le Grazie 15, Università degli Studi di Verona, 37134 Verona, Italy

^b Laboratorio NMR-ISMAC, CNR, Via Bassini 15, 20133 Milano, Italy

^c Università di Milano DiSTAM, Via Celoria 2, 20133 Milano, Italy

ARTICLE INFO

Article history:

Received 23 July 2008

and in revised form 7 October 2008

Available online 21 October 2008

Keywords:

Urea unfolding

Nuclear magnetic resonance

Chicken liver bile acid binding protein

Hydrogen exchange

Differential scanning calorimetry

Transient aggregation

Residual structure

ABSTRACT

The folding properties of a bile acid binding protein, belonging to a subfamily of the fatty acid binding proteins, have been here investigated both by hydrogen exchange measurements, using the SOFAST NMR approach, and urea denaturation experiments. The urea unfolding profiles of individual residues, acting as single probes, were simultaneously analyzed through a global fit, according to a two-state unfolding model. The resulting conformational stability $\Delta G_U(H_2O) = 7.2 \pm 0.25 \text{ kcal mol}^{-1}$ is in good agreement with hydrogen exchange stability ΔG_{op} . While the majority of protein residues satisfy this model, few amino-acids display a singular behavior, not directly amenable to the presence of a folding intermediate, as reported for other fatty acid binding proteins. These residues are part of a protein patch characterized by enhanced plasticity. To explain this singular behavior a tentative model has been proposed which takes into account the interplay between the dynamic features and the formation of transient aggregates. A functional role for this plasticity, related to translocation across the nuclear membrane, is discussed.

© 2008 Published by Elsevier Inc.

Introduction

Bile Acid Binding Proteins (BABPs)¹ are members of the large Fatty Acid Binding Proteins (FABPs) family, which includes small proteins (15 kDa) consisting of two β -sheets, surrounding a large interior cavity which can host a diverse set of ligands (fatty acids, retinoids and bile salts), and of a small helical region acting as a portal for the cavity [1]. All members of this family share this common fold in spite of a poor sequence identity.

FABPs have been considered good models for folding studies of β -sheet systems and many investigations have been reported [2–7]. Specifically, the unfolding behavior of one member of this family, the rat intestinal fatty acid binding protein (rat IFABP), has been investigated, using different spectroscopic and biochemical approaches, thus offering a thorough description of the steps involved [8–14]. BABPs represent a very interesting FABP subfamily, from both the folding and the functional point of view. These proteins are indeed characterized by a high plasticity, possibly related to their capability of hosting two bulky and rigid ligand molecules, such as steroid derivatives [15]. They play a functional role as bile acid transporters, within the framework of the enterohepatic circulation [16],

as well as signaling proteins, capable of translocation to the cellular nucleus for the delivery of bile acids which regulate the transcriptional activity of hormone nuclear receptors [17,18]. In the context of these multiple functionalities, the elucidation of the determinants of stability and flexibility of BABPs becomes important to understand the global action of these proteins within the cell. In this work, the stability and unfolding mechanism of chicken liver bile acid binding protein (cL-BABP) have been investigated by NMR spectroscopy. The analysis of the protection factors and protein stability, as derived from NMR hydrogen exchange (HX) and urea unfolding experiments, is consistent with a global cooperative $N \rightleftharpoons U$ unfolding mechanism. However, few residues, located in the C-terminal region of the protein, exhibited a singular behavior which did not fit with the presence of a folding intermediate and was tentatively analyzed using different models, keeping into account an interplay between flexibility and aggregation. A possible functional role of the detected plasticity, on one face of the protein, could be related to protein translocation across the nuclear membrane.

Materials and methods

Protein expression and purification

The wt cL-BABP was expressed and purified as soluble protein in *Escherichia coli*, in rich and ¹⁵N enriched minimal media, as

* Corresponding author. Fax: +39 0458027929.

E-mail address: henriette.molinari@univr.it (H. Molinari).

¹ Abbreviations used: BABPs, bile acid binding proteins; FABPs, fatty acid binding proteins; NMR, nuclear magnetic resonance.

previously described [15]. The expression plasmid for H98YcL-BABP was obtained from that of wt cL-BABP using the Quickchange (Stratagene) mutagenesis kit. The expression and purification protocol of the mutant is the same followed for the wt cL-BABP. The molecular weight of the two proteins and the extent of labelling was verified by MALDI mass analysis. All NMR samples were dissolved in $\text{Na}_2\text{HPO}_4/\text{NaH}_2\text{PO}_4$ 30mM buffer at pH 7.4.

The delipidation of the purified proteins was achieved by Lipidex chromatography. Because the presence of bound lipids may affect the quality of the NMR and calorimetric measurements, the general efficiency of the separation procedure was carefully checked by 1D ^{13}C NMR analysis of a cL-BABP protein sample incubated with $^{13}\text{C}_1$ enriched palmitic acid (PA). The spectrum revealed the presence of a single sharp resonance for the bound carboxylate at 184.5 ppm (free palmitic acid is insoluble in aqueous solution at pH 7), which disappeared after delipidation. At the same time, the ^1H NMR profile of the delipidated protein (displaying marked differences of protein resonances in the high-field region, compared to the lipid-bound protein spectrum) was taken as the reference for the apo-protein and used in every fresh preparation as a check for delipidation. Finally, also mass spectrometry analysis was used to verify the absence of lipids in the analyzed samples.

NMR spectroscopy

Measurements were performed at 25 °C on a Bruker DRX 500 MHz and an Avance 600 MHz spectrometer with cryoprobe. Two-dimensional ^1H - ^{15}N HSQC spectra were collected with 2.0 and 6.5 kHz spectral widths, 256 and 1024 complex points along the F_1 (^{15}N) and F_2 (^1H) frequency dimensions, respectively. The standard gradient- and sensitivity-enhanced pulse sequences were used. Each FID consisted of 8 signal-averaged scans separated by a 1.7 s relaxation delay. Spectra were processed using gaussian enhancement in both dimensions and zero-filled once in the F_1 dimension. SOFAST-HSQC experiments [19] were collected with 700 points in F_2 and 192 complex data points in F_1 , using a relaxation delay of 0.5 s with 2 scans and a duty cycle of 7.6%. The band-selective ^1H excitation (PC9) and refocusing (RSNOB) pulses were characterized by pulse lengths of 2.33 and 0.817 ms, respectively. ^{13}C decoupling was achieved through an adiabatic pulse (smoothed chirp) employing a 500 μs pulse, where the carbon hard pulse corresponded to 11.8 μs . An adiabatic pulse of 2 ms was employed for ^{15}N decoupling, where the nitrogen hard pulse was 45 μs . Processing of spectra was carried out using TOPSPIN (Bruker). Cross-peak volumes were calculated using both TOPSPIN and CARRA [20] softwares.

^1H 2D TOCSY spectra were recorded for the analysis of the mutant. Isotropic mixing times varying from 20 to 80 ms were employed. A total of 470 increments were collected in t_1 with 1908 data points and 64 scans/FID in t_2 , over a spectral width of 6 KHz in both dimensions. Prior to 2D FT the experimental array was zero-filled to a final matrix of 2048×1024 data points.

^1H chemical shifts were referenced versus external sodium 3-trimethylsilyl-propionate-2,2,3,3- $^2\text{H}_4$ (TSP), and ^{15}N chemical shifts were referenced indirectly using a reported protocol [21].

Hydrogen exchange experiments

HX experiments were performed at p^2H 7.4 ($\text{p}^2\text{H} = \text{pH}_{\text{read}} + 0.4$). The HX sample of cL-BABP was prepared by dissolving 5 mg of delipidated and lyophilized protein in 600 μl of 30 mM deuterated $\text{Na}_2\text{HPO}_4/\text{NaH}_2\text{PO}_4$ buffer. The final protein concentration was about 0.4 mM. The solution was centrifuged briefly at 4 °C to remove insoluble protein and transferred to a 5 mm NMR tube.

The exchange behavior of cL-BABP was initially followed by the acquisition of standard ^1H - ^{15}N HSQC spectra. The time point for

experiments was taken to be the time elapsed from the addition of D_2O buffer to the beginning of that experiment plus half the duration of the acquisition (33 min). For the SOFAST NMR experiments the time required to dissolve the protein, transfer it to the NMR tube, setting the homogeneity and spectrometer parameters, and acquiring the data went down to 11 min comprehensive of 1 min acquisition on a 600 MHz spectrometer equipped with the cryoprobe. Experiments were performed at p^2H 7.4 and 6.5. The folding reaction was followed by a series of SOFAST-HSQC experiments. The reproducibility of the SOFAST experiments has been evaluated by repeating them twice.

HX rate constants were determined by fitting cross-peak volumes to a first-order exponential decay:

$$I(t) = I(0) \exp(-k_{\text{ex}}t)$$

where $I(t)$ represents the volume of the cross-peaks at the time t , $I(0)$ the volume of the cross peaks at $t = 0$; k_{ex} is the observed rate of hydrogen exchange and t is the time in minutes. In the SOFAST approach, $t = 0$ corresponded to 11 min and k_{ex} greater than 10^{-1} min^{-1} could not be determined.

The free energy of opening, ΔG_{op} , was derived from the equation:

$$\Delta G_{\text{op}} = -RT \ln K_{\text{op}},$$

where $K_{\text{op}} = k_{\text{ex}}/k_{\text{ch}}$ and k_{ch} for model random peptides was derived as previously reported [22].

Data were fitted with the program Sigmaplot (Jandel Scientific).

Urea equilibrium unfolding

cL-BABP was unfolded at pH 7 in ten additions of urea in the concentration range from 0 to 4.6 M. The urea concentration was determined by measuring the refractive index at 25 °C. The volumes of the signals, at each denaturant concentration, were normalized to those measured in spectrum recorded at 0 M urea. Peak volumes were measured using two different software packages, TOPSPIN and CARRA [20]. For many resolved resonances, peak heights were also quantified and compared to peak volumes. The figures obtained with all the mentioned approaches were fully consistent and the same trends were observed. Experiments at selected urea concentrations (2.2 M and 3.6 M) were repeated twice, and the volumes calculated in the repeated experiments differed less than 5%.

The reversibility of the unfolding was assessed by diluting the 4.6 M urea cL-BABP sample down to 0.1 M urea by addition of the buffer solution. 1D ^1H experiments recorded on the subsequently concentrated sample appeared fully superimposable to the spectrum of the native protein.

Thermodynamic models were tested to describe the experimental data. The fit based on these models was accomplished using the non-linear Levenberg–Marquardt method [23]. The errors of each fitting parameter were calculated with a 95% confidence limit by the Monte Carlo simulation method.

For the one-step unfolding model, $N \rightleftharpoons U$, the relevant fraction of the unfolded state is given by $f_U = 1/Q$ where $Q = 1 + K_{\text{NU}}$ is the partition function of the system referred to the native protein state with $K_{\text{NU}} = \exp(-\Delta G_U/RT)$. We assumed that the drop of the Gibbs energy related to the above one-step transition follows the simple linear relationship $\Delta G_U = \Delta G_U(\text{H}_2\text{O}) - m [\text{U}]$.

In all our analyses, calculations were performed using thermodynamic activity evaluated according to the reported polynomial equation [24]. However, in order to allow a direct comparison with literature data about the same or similar subjects, the values of the parameters and the data reported in the figures were back converted and referred to urea molar concentration units.

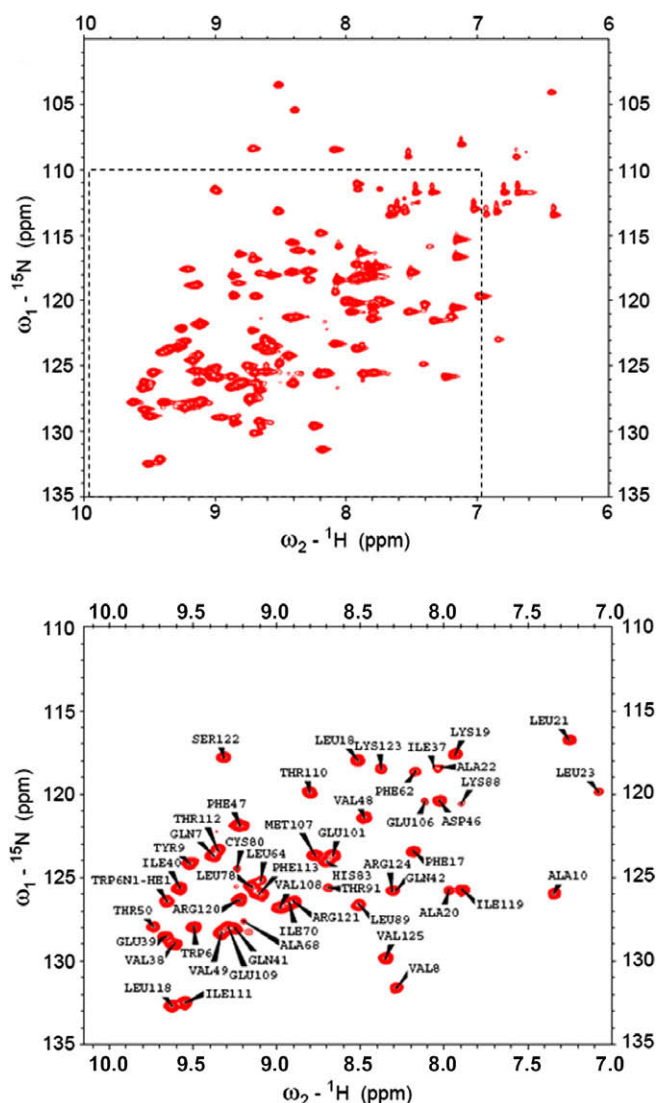
$$f_D = \frac{1}{O}, f_N = \frac{K_{DN}[N]^{-1}}{O}, f_U = \frac{K_{DN} \cdot K_{NU}[N]^{-1}}{O}$$
$$Q = 1 + K_{\text{DN}}[N]^{-1} + K_{\text{DN}} \cdot K_{\text{NU}}[N]^{-1}$$

Calorimetry measurements

A Setaram Micro-DSC III apparatus that is suitable for dilute solutions of biological macromolecules in the temperature range from -20 to 120°C was used with a $0.5^{\circ}\text{C}/\text{min}$ scan rate.

Hydrogen exchange measurements

Measured exchange rates differed by more than three orders of magnitude along the sequence and were slower by six orders of magnitude, compared to the intrinsic rates. Exchange rates derived from standard and SOFAST HSQC experiments were in good agreement, wherever the comparison was possible. For an evaluation of K_{op} , protection factors P and ΔG_{op} , based on the exchange rates, it is important, as clearly stated by many authors [22], to elucidate the mechanism by which exchange occurs in the conditions employed for the kinetic measurements [22,26]. Under unimolecular exchange (EX₁) there is no pH dependence of k_{ex} , whereas in the bimolecular exchange (EX₂) limit, k_{ex} will be proportional to the concentration of H_2O^+ or OH^- [22]. Measurement of exchange rates



at two different pH values and subsequent plotting of $\log k_{\text{ex}}$ for individual residues at one pH (pH_1) against the value at another pH (pH_2) can thus provide insight into the mechanism of exchange. We therefore performed an additional set of kinetic experiments at p^2H 6.5 (Table 1, Supplementary Material). Representative exchange curves measured for selected residues of cL-BABP at p^2H 6.5 and 7.4 at 25 °C are shown in Fig. 2A and B. The results of the test, reported in Fig. 2C, show that a straight line fits the experimental data with an $R^2 = 0.86$, a slope close to one, within experimental error, and an intercept equal to the pH difference, thus confirming the presence of an EX_2 mechanism. The exchange rate values, obtained for each residue at pH 7.4, were used to evaluate protection factors ($P = k_{\text{ch}}/k_{\text{ex}}$) and ΔG_{op} , which are reported in Table 1 and mapped onto the protein structure in Fig. 3.

The denaturation of cL-BABP was followed in a residue-specific manner collecting a series of eleven ^1H - ^{15}N HSQC spectra at increasing urea concentration in equilibrium conditions and mon-

Table 1

Summary of k_{ex} , protection factors (P) and ΔG_{op} derived from hydrogen exchange SOFAST NMR experiments at p²H 7.4^a. Hydrogen bonds involving the listed residues are reported.

Residue	k_{ex}^b (min ⁻¹)	$P \times 10^3$	ΔG_{op} (kcal mol ⁻¹)	H-bonds (NH-OC) ^c
W6	8.70 ± 0.2E-03	29.06	6.09 ± 0.01	W6-V38
Q7	5.50 ± 0.2E-03	66.45	6.58 ± 0.02	Q7-K123
V8	5.98 ± 0.29E-02	2.16	4.55 ± 0.03	V8-P36
Y9	1.36 ± 0.12E-03	117.3	6.91 ± 0.05	Y9-R121
A10	4.89 ± 0.21E-02	9.41	5.42 ± 0.03	A10-R121
F17	2.93 ± 0.06E-02	5.8	5.13 ± 0.01	F17-N13
L18	2.23 ± 0.04E-02	5.64	5.11 ± 0.01	L18-Y14
K19	5.20 ± 0.16E-02	4.43	4.97 ± 0.02	K19-E15
A20	1.17 ± 0.15E-01	4.61	4.99 ± 0.08	A20-E16
L21	6.01 ± 0.19E-02	1.81	4.44 ± 0.02	L21-F17
V38	2.80 ± 0.1E-03	17.13	5.77 ± 0.02	V38-W6
E39	6.70 ± 0.2E-03	13.89	5.65 ± 0.02	E39-T50
I40	7.60 ± 0.2E-03	7.16	5.26 ± 0.02	I40-G4
D46	2.11 ± 0.4E-02	6.53	5.20 ± 0.11	D46-K43
F47	1.77 ± 0.3E-02	8.94	5.39 ± 0.10	F47-F62
V48	8.4 ± 0.2E-03	11.21	5.52 ± 0.01	V48-Q41
V49	8.60 ± 0.2E-03	6.86	5.23 ± 0.01	V49-N60
T50	3.63 ± 0.18E-02	7.03	5.24 ± 0.03	T50-E39
F62	4.65 ± 0.37E-02	10.21	5.47 ± 0.05	F62-F47
L64	3.43 ± 0.2E-02	5.03	5.05 ± 0.03	L64-D45
L78	2.71 ± 0.06E-02	5.29	5.08 ± 0.01	L78-I70
H83	2.35 ± 0.09E-02	30.69	6.12 ± 0.02	H83-V90
L89	2.43 ± 0.1E-02	5.9	5.14 ± 0.02	L89-Q100
T91	3.73 ± 0.29E-02	6.84	5.23 ± 0.04	T91-H98
E101	1.45 ± 0.04E-02	13.79	5.65 ± 0.02	E101-V108
M107	4.60 ± 0.2E-03	62.27	6.54 ± 0.03	M107-S122
V108	1.10 ± 0.1E-03	95.39	6.79 ± 0.05	V108-E101
T110	3.48 ± 0.1E-04	721.37	7.99 ± 0.02	T110-E99
I111	1.53 ± 0.1E-04	790.94	8.08 ± 0.04	I111-L118
T112	7.50 ± 0.2E-03	27.64	6.06 ± 0.02	T112-S97
F113	1.88 ± 0.06E-02	20.08	5.87 ± 0.02	F113-V116
L118	6.03 ± 1.92E-04	285.93	7.44 ± 0.19	L118-I111
I119	4.33 ± 0.17E-02	1.09	4.14 ± 0.02	I119-E12
R120	5.19 ± 0.25E-04	555.11	7.83 ± 0.03	R120-E109
S122	1.69 ± 0.04E-02	94.39	6.78 ± 0.01	S122-M107
K123	2.75 ± 0.14E-02	27.11	6.0 ± 0.034	K123-Q7
V125	3.17 ± 0.06E-02	0.07	2.49 ± 0.01	V125-T5

^a Final measured p²H of sample: p²H = pH_{read} + 0.4.

^b Reported errors refer to the errors in the least squares fitting.

^c The presence of a hydrogen bond was assessed from the analysis of 1ZRY and 1TVQ PDB structures.

itoring peak volumes. Resonances corresponding to the fully denatured protein became measurable at 2.2 M urea and their intensities reached a plateau at 4.3 M urea. Most of the peaks displayed minor chemical shift variations upon urea titration. Inspection of cross-peaks sections indicated that the line widths of the native resonances did not vary significantly during the titration, thus allowing reliable intensity measurements. The behavior of sixty-nine resonances could be followed upon unfolding. The amide resonances of five residues (V90, S93, K95, E99 and Q100) were exchange broadened at pH 7.4 and the corresponding peaks were missing in the spectra, while the intensities of the remaining resonances could not be correctly evaluated due to overlap.

The resonances of sixty amides, evenly distributed along the protein sequence, displayed a similar behavior, as shown by the corresponding NMR peak volumes, reported in Fig. 4 as a function of urea concentration. In general, the decay trend of the NMR signal corresponding to a given amino-acid may not be representative of the whole biopolymer, nor may it be used to assess whether the conformational transition is a one- or a multi-step process. However, since the changes occurring to a given amino-acid must likely perturb its close neighbors, it seems reasonable to cluster residues that undergo transitions according to a similar trend.

To this aim, the intensity decay of each amino-acid was scaled with respect to a descendent straight line, observed at low urea

concentration, treated as a mean pre-denaturation baseline, and normalized to a unity drop (Fig. 4). The normalized data, shown in Fig. 5, followed a similar pattern, suggesting the occurrence of a cooperative unfolding with a sharp volume decrease around 2.2–3.6 M urea, with a complete loss of intensity at 4.3 M urea. It was found that the sigmoid decay of the NMR signals from most of the residues could be separately fit to a one-step model, $N \rightleftharpoons U$. For evaluation of the thermodynamic parameters, we assumed, as in most analyses of this kind [27], that the drop of the Gibbs energy related to the above one-step transition, follows a simple linear relationship

$$\Delta G_U = \Delta G_U(H_2O) - m[U].$$

It was therefore possible to define the overall data sets $\{\Delta G_U^{H_2O}\}$ and $\{m_i\}$ that correspond to the best fit of the data of each residue. By reporting these data in the $(\Delta G(H_2O), m)$ plane, two distinct clusters appeared. Most of the points were indeed positioned in a restricted region of the plot and were distributed along one straight line trend (Fig. S1), while the data corresponding to a minor fraction of residues, which could not be fit with a one-step model (see Discussion Section), clustered separately. The linear relationship between ΔG and $[U]$ allows one to immediately realize that the equilibrium condition, i.e., $\Delta G = 0$, is attained at about the same denaturant concentration, $[U]_d$, for all the residues in this major group. This observation supports the hypothesis that they cooperatively undergo a one-step denaturation. The residues belonging to this cooperative unfolding unit are mapped in blue onto protein ribbon structure in Fig. 6.

Given the established mechanism, common thermodynamic parameters may apply to the whole body of data relevant to this major fraction. Taking into account the $[U]_d$ boundary condition, a 3D fit (residue, NMR signal, $[U]$) was essayed for this fraction of amino-acids employing the one-step denaturation model (Fig. 5). The best fit parameters obtained were $\Delta G_U(H_2O) = 7.20 \pm 0.25$ kcal mol⁻¹ and $m = 2.15 \pm 0.05$ kcal mol⁻¹ M⁻¹, implying the mid-point transition denaturant concentration, $[U]_d = 3.35$ M (data derived from this thermodynamic analysis are expressed in kcal instead of Joule to compare the results with data reported in the literature).

The above described analysis is based on criteria similar to statistical methods such as the singular value decomposition, a kind of analysis that has been employed in a number of cases to analyze multivariate unfolding data [28,29]. The two methods were found to give the same results in this study.

The overall analysis of the separately observed NMR signals from each residue allows the assessment of reliably averaged parameters that account for the entire structure and are suitable to support a thermodynamic analysis of the unfolding transition. The global energetics derived from such a high-resolution analysis has already been found to be fully consistent with the information drawn from other experimental approaches, like CD, and fluorescence [30].

A smaller subset of residues (D33, L64, G65 (DE loop), D74, G75 (EF loop), T81, A85 (β-F), L89 (β-G), H98 (β-H), referred to as subset A and mapped in red onto the protein structure (Fig. 6), displayed a singular behavior, characterised by a smooth decrease of the NMR volume on adding urea even after application of the pre-transition baseline described before (Fig. 7). These residues are mainly located in the C-terminal face of the protein barrel, where the presence of conformational averaging was previously assessed [15]. Indeed ¹⁵N relaxation measurements revealed that seven residues located in DE, EF loops and in β-F, β-G and β-H strands exhibited an R_{ex} contributions while other six residues located as well in EF loop and β-G, β-H strands were broadened beyond detection.

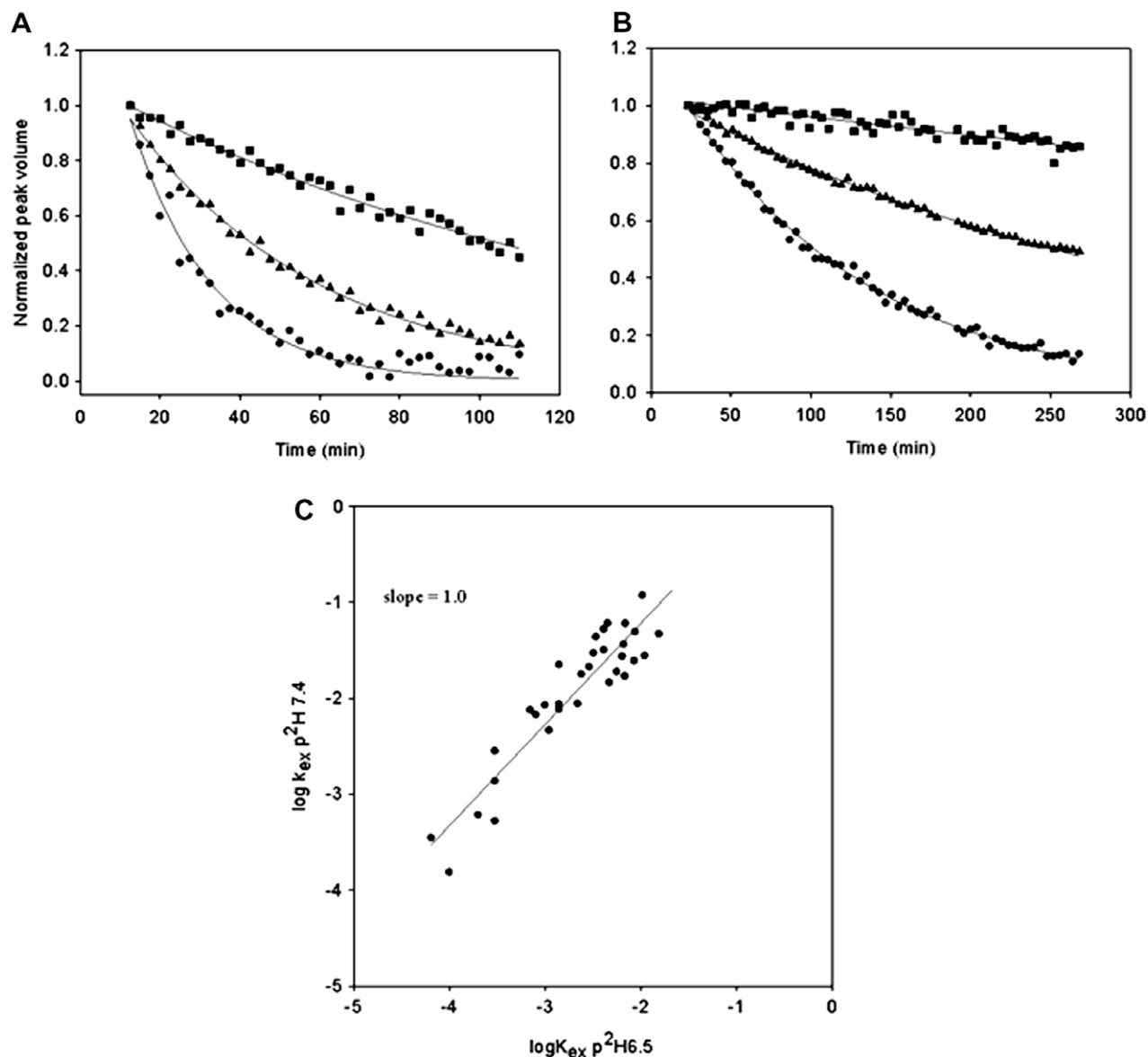


Fig. 2. Representative examples of exchange kinetics of three residues (T_{112} ●, D_{46} ■, K_{19} ▲) obtained from the analysis of HX from SOFAST NMR experiments performed at 600.13 MHz on a 0.4 mM sample of cL-BABP, p²H 7.4 (panel A) and p²H 6.5 (panel B). Plot of pH dependence of exchange (panel C): the value of log k_{ex} at p²H 7.4 is plotted against the log k_{ex} at p²H 6.5 for the same residue. The data relative to 25 residues could be compared. The experimental points were fitted with a linear regression: a slope of 1.06 ± 0.07 and an intercept of 0.9 ± 0.2 with $R^2 = 0.86$ were obtained (B).

DSC measurements and analysis

To confirm the thermodynamic parameterization obtained from the NMR data, a DSC analysis was undertaken. Unfortunately, strong exothermic phenomena were observed, related to aggregation processes that are typical for this class of proteins [31] and possibly in line with the presence of soluble aggregates between native monomers. Exploited changes of pH (from 5 to 8.5) and ionic strength (see Materials and methods) did not allow a shift of the exotherm to either lower or higher temperature, preventing any reliable evaluation of the unfolding enthalpy and any tentative thermodynamic treatment, including a simple van't Hoff analysis (Fig. S2). In an attempt to overcome this problem, and assuming the possibility of aggregation promoted by the formation of intermolecular disulfide bridges, the free cysteine present in the protein was methylated, following standard procedures. However, the obtained data were affected by the same problems.

The only conclusion drawn from these investigations is that the unfolding temperature of this protein ranges between 60 and

65 °C, which is indeed common for many molecules of similar molecular mass.

The substantial failure of these experiments is here reported to emphasize the problems met when approaching this protein family.

Discussion

Hydrogen exchange analysis

Information on amides protected from exchange (Table 1) was derived from HX analysis, indicating that 50% of the residues predicted (on the basis of available structures) to be involved in backbone hydrogen bonds, are protected. The mapping of protection factors onto the protein structure (Fig. 3) suggested that the most stable core involves A–C, I and J strands, while most of the amides in D–H strands and all those of helix II are solvent accessible. A gradual increase in the rate of opening fluctuations across one face of the β -sheet, from strand D to H, was also observed, suggesting

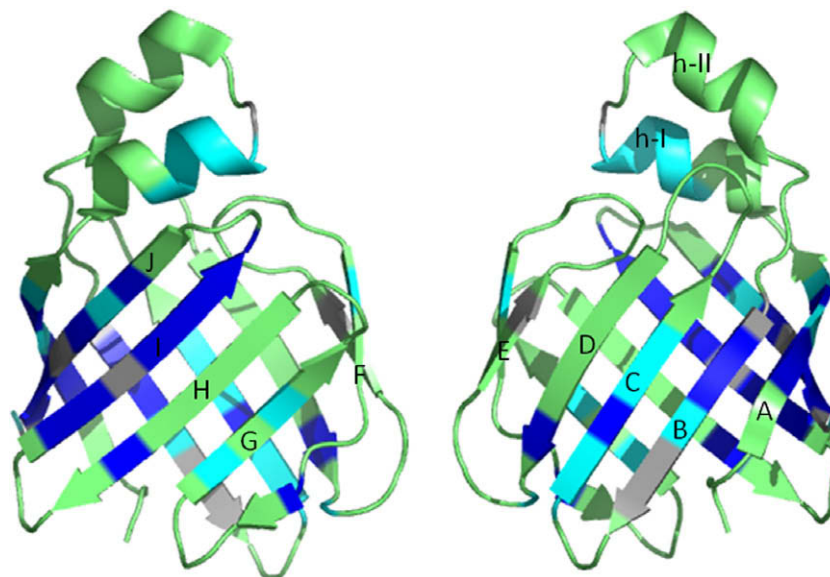


Fig. 3. Protection factors are mapped onto the protein structure with the following color code: green $P < 10^3$, cyan $10^3 < P < 10^4$, blue $10^4 < P < 10^6$. Residues exhibiting overlap which could not be analyzed are indicated in grey. Secondary structure elements are labeled. Left and right panels correspond to the same view of the protein rotated by 180° around the y axis. (For interpretation of the references to colour in this figure legend, the reader is referred to the web version of this paper.)

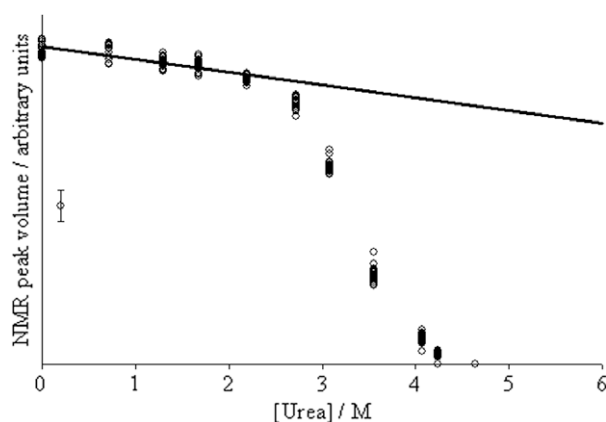


Fig. 4. Selected peak volumes of cL-BABP ($\text{Na}_2\text{HPO}_4/\text{NaH}_2\text{PO}_4$ 30 mM buffer, pH 7, 25°C) measured from ^1H - ^{15}N HSQC experiments recorded at 500.13 MHz are reported as a function of urea concentration (circles). No discrimination of single amino-acid was applied for visual reasons. The mean pre-denaturation baseline is also shown as a continuous line.

the presence of a highly flexible region in this element of secondary structure. Although it is clear that slow HX exchange in the native state does not necessarily reflect regions of stability in the urea denatured state [6], it is interesting to note here that the HX pattern exhibited by cL-BABP seems to be correlated to the conservation of the native structure at high urea concentration. At 4.3 M urea, where the native tertiary structure of the protein is not populated to any significant extent, few peaks, mostly localized in strands A–D, i.e. on one face of the protein, exhibit chemical shifts similar to their values in the absence of urea and are well-resolved from the crowded region of the HSQC spectrum corresponding to the unfolded state (Fig. 8). We interpret the presence of these amide resonances as indicative of some residual localized structure in the conformational ensemble of the denatured proteins.

The analysis of HX exchange allowed the determination of ΔG_{op} values. It is expected that a subset of these values generally represent the complete unfolding reaction and indeed it has been proposed that the average of the three largest ΔG_{op} values, calculated

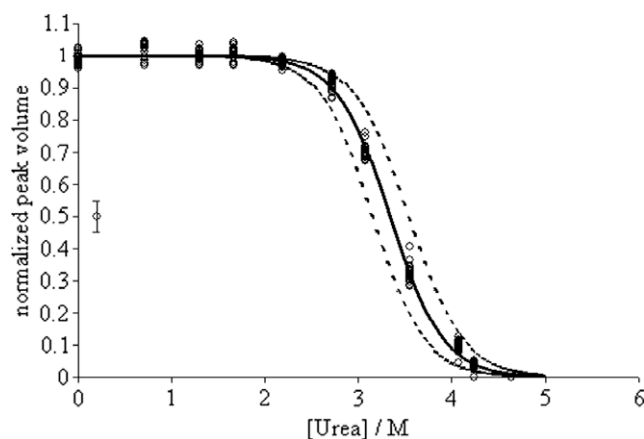


Fig. 5. Same data of Fig. 4 scaled with respect to the mean pre-denaturation baseline and normalized to a unity drop (circles). The continuous line represents the best fit according to the two-state unfolding model. Dotted lines represent the confidential limit (95%).

as described above, provides a reasonable estimate of the conformational stability of the protein [32]. Residues T110, I111 and R120 are characterized by the highest ΔG_{op} values (Table 1), which translate into an average value of $7.78 \pm 0.03 \text{ kcal mol}^{-1}$ (corrected for the proline *cis/trans* isomerisation equilibrium) [32], representing a reasonable estimate of conformational stability. This value is in good agreement with $\Delta G_{\text{U}}(\text{H}_2\text{O})$ derived from urea unfolding experiments ($7.20 \pm 0.25 \text{ kcal mol}^{-1}$), considering the reported slight bias towards higher values for HX stabilities with respect to unfolding energies measured by urea denaturation [33]. This difference is here ascribed to solvent isotopic effects as $\Delta G_{\text{U}}(\text{H}_2\text{O})$ can differ by $\pm 0\text{--}2 \text{ kcal mol}^{-1}$ in H_2O compared to D_2O [34].

Urea unfolding and singular behavior analysis

The unfolding of cL-BABP indicated that most analyzed protein residues, evenly distributed along the sequence and on the structural elements, exhibited a consistent behavior (Figs. 5 and 6),

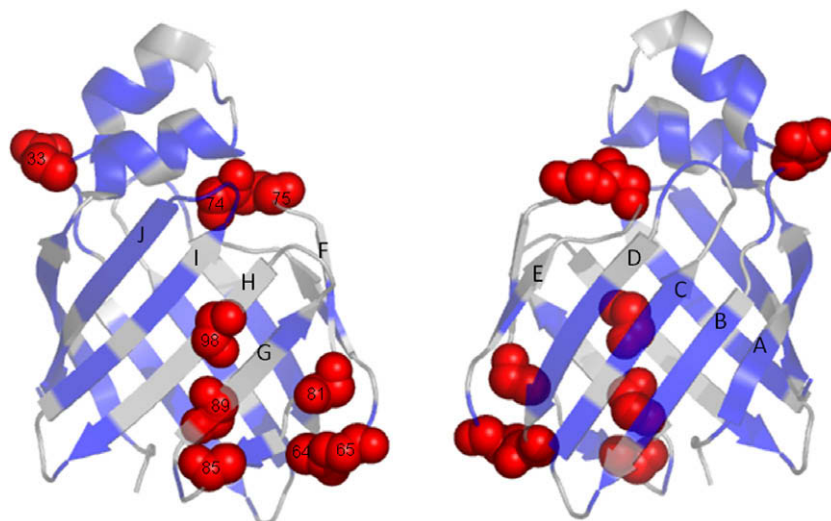


Fig. 6. Residues following a similar unfolding pattern, exhibiting a cooperative unfolding, characterized by $\Delta G_U(\text{H}_2\text{O}) = 7.20 \pm 0.25 \text{ kcal mol}^{-1}$, $m = 2.15 \pm 0.05 \text{ kcal mol}^{-1} \text{ M}^{-1}$, are mapped in blue onto the protein ribbon. Subset A, comprising residues exhibiting a singular behavior are indicated with red spheres labeled with their residue number. Secondary structures elements are labeled. (For interpretation of the references to colour in this figure legend, the reader is referred to the web version of this paper.)

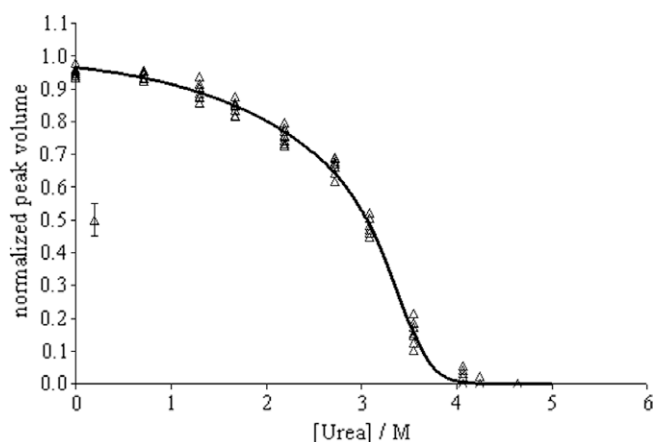


Fig. 7. Peak volumes of cL-BABP residues D33, L64, G65, D74, G75, T81, A85, L89, H98 measured from ^1H - ^{15}N HSQC experiments obtained at 500.13 MHz are reported as a function of urea concentration, scaled with respect to the mean pre-denaturation baseline (see Fig. 4) and normalized to a unity drop (triangles). The continuous line represents the best fit according to the aggregation model assuming $n = 2$.

defining a cooperative one-step unfolding transition. A minor group of residues, mainly located in G and H strands and in DE, EF, FG loops (subset A, residues colored in red in Fig. 6), did not fit in the general model, mainly because the corresponding NMR signals showed a steeper decrease at low urea concentrations (Fig. 7). It should be noted that the same protein regions also display other singular properties: they contain residues broadened beyond detection and residues exhibiting chemical exchange contributions, as derived from ^{15}N relaxation measurements [15] (Fig. 9). These data altogether point to the presence of a protein patch characterized by high plasticity and therefore exhibiting higher urea susceptibility. Indeed it was recently underlined that the most vulnerable sites to act as unfolding initiation sites would be those where the expenditure of energy would be the least, namely the regions which are conformationally more dynamic in the folded protein [35]. Thus residues of subset A, belonging to a protein region characterized by larger R_2 values, compared to the average, may indeed act as unfolding initiation sites.

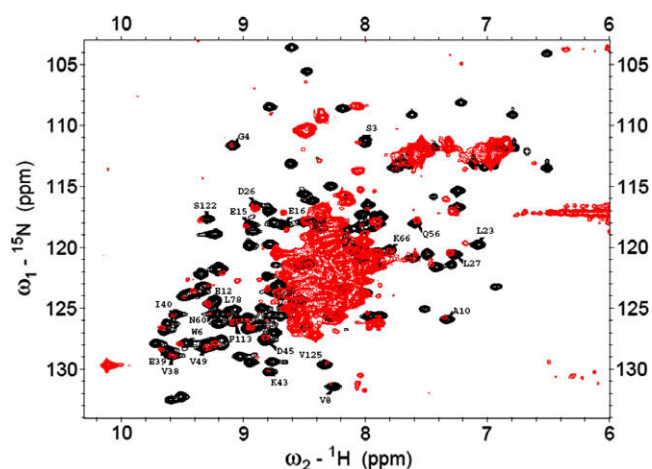


Fig. 8. Superposition of ^1H - ^{15}N HSQC experiments obtained for cL-BABP at 0 M (black) and 4.3 M urea (red). The assignment of most of the native peaks still observable at 4.3 M urea is reported. (For interpretation of the references to colour in this figure legend, the reader is referred to the web version of this paper.)

In light of the results reported for other proteins of the family [7,13,36], it was first hypothesized that the observed behavior could be explained with the formation of an intermediate state. All the available data, including those relative to subset A, were analyzed with a three-state model $N \rightleftharpoons I \rightleftharpoons U$. However, such an analysis led to inconsistent results, indicating a mid-point for the $N \rightleftharpoons I$ transition beyond 4.5 M urea, which is certainly a nonsense, as at this molar concentration the protein is substantially completely unfolded.

Other models were therefore considered in an attempt to explain this behavior and to relate it to protein functionality.

A naïve interpretation might start from the assumption that the baseline slope for residues of subset A is sharper than for the others: this would imply a larger scaling of the data to reproduce the expected sigmoidal decay. However, this analysis, invoking for few residues a pre-transition baseline closely merged to the transition region, would let the baseline artificially absorb all variability in the data and does not seem to represent any reasonable physical model.

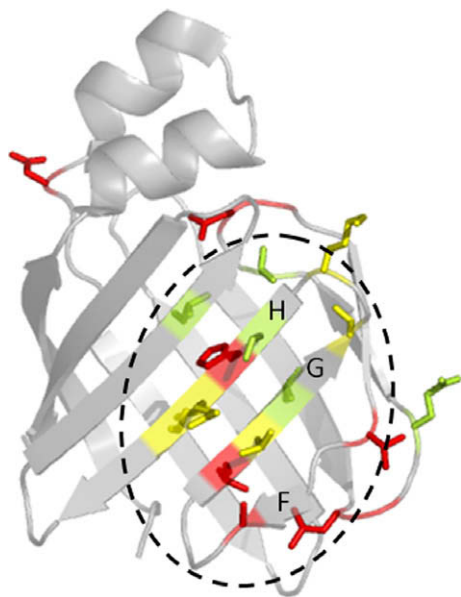
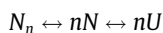


Fig. 9. Residues exhibiting a singular behavior to unfolding are depicted in red (D33, L64, G65, D74, G75, T81, A85, L89, H98), residues broadened beyond detection are depicted in yellow (V90, S93, K95, E99, Q100) and residues exhibiting an R_{ex} contribution, identified from high R_2/R_1 values, are depicted in lemon green (E67, T72, T81, L89, T91, S97 and I111). A dotted line indicates the overall protein patch exhibiting a singular behavior, centered around F–H strands, as indicated. (For interpretation of the references to colour in this figure legend, the reader is referred to the web version of this paper.)

An alternative hypothesis was then made that the singular behaviors could depend on intermolecular interactions. DSC measurements indicated the presence of aggregation upon temperature change and preliminary mutational studies suggested that aggregation phenomena may take place, involving the singular protein patch here identified. Indeed, when residue H98, whose protonation equilibria were hypothesized to modulate slow motions in the micro to milliseconds timescale [15], was mutated to a tyrosine, the soluble protein expression, usually highly efficient (180 mg/l), was very low (few mg/l) and the protein was found essentially in the inclusion bodies. The tiny amount of soluble protein was highly aggregated, as clearly shown by the broad line width observed in homonuclear 2D TOCSY spectra (Fig. S3). Several independent studies provided clear evidence that inclusion bodies, comparable to amorphous aggregates, result from specific intermolecular interactions between conformational ensembles [37]. Thus a single mutation highlighted the presence of an aggregation-prone fragment where unfolding may result in solvent exposure, which in turn leads to the formation of transient aggregates. A more complex explanation of the observed behavior was therefore attempted, based on the hypothesis, derived from the described observations, that a fraction of the protein molecules could form transient aggregates at the concentration level considered in this work. Addition of urea would reduce this fraction and affect mainly the region of the protein where the residues of subset A are located. The formation of transient aggregates with a role in folding has been already discussed by others [38].

A model describing this behavior can be represented by the following equilibrium:



where N_n would account for the transient protein aggregates (see Material and methods).

For $n \geq 2$ the results of the global fit seemed satisfactory (see Fig. 7). The $N \rightleftharpoons U$ transition parameters imposed were the same

as those obtained previously for the $N \rightleftharpoons U$ transition of the main set of residues (i.e. $\Delta G_U(H_2O) = 7.20 \text{ kcal mol}^{-1}$, $m = 2.15 \text{ kcal mol}^{-1} \text{ M}^{-1}$). The parameters relevant to the transition related to reversible formation of intermolecular interactions are not reported since their meaning depends on the stoichiometry of aggregation which is not known. It should be remarked that this intermolecular interactions should be transient in nature as, at neutral pH, a correlation time (τ_c) of $7.2 \pm 0.3 \text{ ns}$ was determined, on the basis of heteronuclear T_1/T_2 ratios measured at 600 MHz, in line with the presence of an essentially monomeric protein.

The above interpretations represent an attempt to explain the behavior of this flexible patch of the protein where plasticity and/or a spread distribution of native microstates could be associated with aggregation. It should be stressed that a spread distribution of native microstates is important for translocation across membranes. This process is particularly relevant for this protein family, as it was shown that some intracellular lipid binding proteins are involved in enhancing the transcriptional activities of nuclear receptors with which they share a common ligand. This task is accomplished, in the case of three FABPs (cellular retinoic acid binding protein, adipocyte FABP and keratinocyte FABP [39,40]), through nuclear translocation in reply to conformational changes triggered by fatty acid binding. The presence of nuclear localisation (NLS) and nuclear export (NES) consensus motifs was reported to regulate this process and it has been suggested that this mechanism could be a characteristic of the family [40]. These consensus signal motifs were identified at the level of tertiary structure in the region 29–33 (BABP numbering) for the NLS and at the level of residues 66, 86 and 91, for the NES. Interestingly, all these residues are among those displaying a singular behavior and we propose here that the high plasticity may compete with aggregation phenomena. In this line, a study on the nuclear translocation mechanism in cL-BABP and its correlation with flexible domains is in progress.

Comparison with intracellular lipid binding protein (iLBP) unfolding

The folding mechanism of few members of the iLBP family, namely cellular retinoic acid binding protein I (CRABP-I) [41–46], intestinal fatty acid binding protein (IFABP) and ileal lipid binding protein (iLBP) [7,9,12,14,47] were previously reported and are here compared with the unfolding data obtained for cL-BABP. In the initial work on CRABP-I, the role played by β -turns in folding was underlined. It was shown [48] that the 24-residue hairpin containing both DE and EF turns displayed native-like conformational propensity, even in the presence of 7 M urea, suggesting that this region of the protein may be structured in the unfolded state and may therefore be important in narrowing the conformational space at the earliest stages in folding. This result applied also to another member of the iLBP, intestinal fatty acid binding protein, where residual structure was found in regions spatially close to turn DE and turn EF [2,6]. Especially EF turn was proposed to influence early steps in folding and it was suggested to be formed in the rate-determining transition state ensemble. A cluster of hydrophobic residues near turn DE is conserved throughout the iLBP family and is shown to include residues in the different strands both in the front and in the back sheets.

Interestingly a completely different behavior was detected for cL-BABP where persistent structure was localized in the N-terminal sheet, at the level of β A–D strands while residues located in DE, EF loops and in β -F,G,H strands revealed a highly dynamic behavior (subset A), characterized by singular response to urea unfolding.

More specifically, rat intestinal FABP (IFABP) and ileal lipid binding protein (iLBP), are the two homologous proteins exhibiting the highest structural similarity with cL-BABP (RMSD < 1.9 Å) in spite of the low sequence identity (< 20%).

The value of $\Delta G_U(H_2O)$ reported for ILBP [7] and IFABP [36] are relatively lower than for cL-BABP ($\Delta G_U(H_2O)$ 4.39 ± 0.26 and 5.28 ± 0.56 kcal mol⁻¹, respectively). ILBP and IFABP were shown to behave as “*N* \rightleftharpoons *U* folders” by fluorescence and CD [7], with only the unfolded and native states being present at equilibrium at significant concentrations. However, refolding kinetics of ILBP revealed the presence of an “on path” intermediate with the structural properties of a molten globule, while for IFABP a burst phase and at least two intermediates were identified along the folding pathways [7,36]. Interestingly, all the NMR equilibrium experiments confirmed that these intermediates were NMR-undetectable, even though clear evidence of their existence was given, such as the loss of overall peak intensity around the denaturation midpoint, attributable to intermediates exchanging on a time scale undetectable by NMR [13,47]. The hypothesis was advanced that differences in the number of buried hydrophobic atoms in folding initiation sites are responsible for the different type of intermediates observed for homologous rat IFABP and ILBP [49]. To test this hypothesis, mutations were made that swapped a more hydrophobic residue for a more hydrophilic residue in the respective cores of the two proteins, and indeed it was concluded that altering the characteristics and/or size of residues within an initiating core of hydrophobic interactions is critical to the types of intermediates observed during protein folding [49].

In this line many differences can be envisaged for cL-BABP and IFABP. For instance, native-like structures, detected in IFABP at very high denaturant concentrations (5.5–6.5 M urea), involved residues in or near DE, IJ and helix I-helix II turns [6], i.e. in regions completely different from those identified at 4.3 M urea in cL-BABP, namely A–D strands. Another important difference is that the extended aromatic cluster of IFABP, comprising eight phenylalanines, is not conserved in cL-BABP, where only F2, F17, F47 and F62 are present and the remaining F55, F68, F93 and F128 are substituted with proline, alanine and two serines, respectively. More specifically in IFABP the distance between two phenylalanine residues (F68 and F93), located in the two different β -sheets enclosing the central cavity, was monitored through two-dimensional ¹⁹F–¹⁹F nuclear Overhauser effect and did not change up to 4 M urea [13]. None of these extended side chains is present in cL-BABP in the corresponding position.

In conclusion, in light of all the reported results, it is reasonable to hypothesize that the different plasticity of these proteins may dictate different paths along the folding landscape before reaching the same final structure.

Acknowledgments

This research was supported by FIRB 2003 (Project No. RBNE03PX83), and Cariverona Foundation. L.R. thanks CNR-RSTL 2007 (Code No. 779) for financial support. CIRMMMP (Consorzio Interuniversitario di Risonanze Magnetiche di Metalloproteine Paramagnetiche) is gratefully acknowledged. Authors are indebted to Lucia Zetta and Alberto Schiraldi for useful discussion and to Clelia Cogliati for help in sample preparation. Helena Kovacs is gratefully acknowledged for precious help in setting the SOFAST experiments.

Appendix A. Supplementary data

Supplementary data associated with this article can be found, in the online version, at doi:10.1016/j.abb.2008.10.017.

References

- [1] A. Chmurzynska, *Journal of Applied Genetics* 47 (2006) 39–48.
- [2] I.J. Ropson, C. Frieden, *Proceedings of the National Academy of Sciences of the United States of America* 89 (1992) 7222–7226.
- [3] I.J. Ropson, J.I. Gordon, C. Frieden, *Biochemistry* 29 (1990) 9591–9599.
- [4] K. Kim, C. Frieden, *Protein Science* 7 (1998) 1821–1828.
- [5] K. Kim, R. Ramanathan, C. Frieden, *Protein Science* 6 (1997) 364–372.
- [6] M.E. Hodsdon, C. Frieden, *Biochemistry* 40 (2001) 732–742.
- [7] P.M. Dalessio, I.J. Ropson, *Biochemistry* 39 (2000) 860–871.
- [8] K. Chattopadhyay, E.L. Elson, C. Frieden, *Proceedings of the National Academy of Sciences of the United States of America* 102 (2005) 2385–2389.
- [9] K. Chattopadhyay, C. Frieden, *Proteins* 63 (2006) 327–335.
- [10] K. Chattopadhyay, S. Saffarian, E.L. Elson, C. Frieden, *Proceedings of the National Academy of Sciences of the United States of America* 99 (2002) 14171–14176.
- [11] K. Chattopadhyay, S. Zhong, S.R. Yeh, D.L. Rousseau, C. Frieden, *Biochemistry* 41 (2002) 4040–4047.
- [12] H. Li, C. Frieden, *Biochemistry* 45 (2006) 6272–6278.
- [13] H. Li, C. Frieden, *The Journal of Biological Chemistry* 280 (2005) 38556–38561.
- [14] H. Li, C. Frieden, *Biochemistry* 44 (2005) 2369–2377.
- [15] L. Ragona, M. Catalano, M. Luppi, D. Cicero, T. Eliseo, J. Foote, F. Fogolari, L. Zetta, H. Molinari, *The Journal of Biological Chemistry* 281 (2006) 9697–9709.
- [16] M. Guariento, D. Raimondo, M. Assfalg, S. Zanzoni, P. Pesente, L. Ragona, A. Tramontano, H. Molinari, *Proteins* 70 (2008) 462–472.
- [17] N.S. Tan, N.S. Shaw, N. Vinckenbosch, P. Liu, R. Yasmin, B. Desvergne, W. Wahli, N. Noy, *Molecular and Cellular Biology* 22 (2002) 5114–5127.
- [18] M. Nakahara, N. Furuya, K. Takagaki, T. Sugaya, K. Hirota, A. Fukamizu, T. Kanda, H. Fujii, R. Sato, *The Journal of Biological Chemistry* 280 (2005) 42283–42289.
- [19] P. Schanda, E. Kupce, B. Brutscher, *Journal of Biomolecular NMR* 33 (2005) 199–211.
- [20] R. Keller (Ed.), *The Computer Aided Resonance Assignment Tutorial* 2004.
- [21] D.S. Wishart, C.G. Bigam, J. Yao, F. Abildgaard, H.J. Dyson, E. Oldfield, J.L. Markley, B.D. Sykes, *Journal of Biomolecular NMR* 6 (1995) 135–140.
- [22] J. Clarke, A.R. Fersht, *Folding & design* 1 (1996) 243–254.
- [23] W.H. Press, B.P. Fannery, S.A. Teukolsky, W.T. Vetterling (Eds.), *Numerical Recipes*, Cambridge University Press, 1989.
- [24] C.N. Pace, *Methods in Enzymology* 131 (1986) 266–280.
- [25] M. Gal, P. Schanda, B. Brutscher, L. Frydman, *Journal of the American Chemical Society* 129 (2007) 1372–1377.
- [26] Y. Bai, J.S. Milne, L. Mayne, S.W. Englander, *Proteins* 17 (1993) 75–86.
- [27] C.N. Pace, K.L. Shaw, *Proteins (Suppl. 4)* (2000) 1–7.
- [28] O.K. Ganesh, T.B. Green, A.S. Edison, S.J. Hagen, *Biochemistry* 45 (2006) 13585–13596.
- [29] S.J. Hagen, C.W. Carswell, E.M. Sjolander, *Journal of Molecular Biology* 305 (2001) 1161–1171.
- [30] M. Sadqi, D. Fushman, V. Munoz, *Nature* 442 (2006) 317–321.
- [31] S. Capaldi, M. Guariento, G. Saccomani, D. Fessas, M. Perduca, H.L. Monaco, *The Journal of Biological Chemistry* 282 (2007) 31008–31018.
- [32] B.M. Huyghues-Despointes, J.M. Scholtz, C.N. Pace, *Nature Structural Biology* 6 (1999) 910–912.
- [33] B.M. Huyghues-Despointes, C.N. Pace, S.W. Englander, J.M. Scholtz, *Methods in Molecular Biology (Clifton N.J.)* 168 (2001) 69–92.
- [34] G.I. Makhatadze, G.M. Clore, A.M. Gronenborn, *Nature Structural Biology* 2 (1995) 852–855.
- [35] A. Chatterjee, P.M. Krishna Mohan, A. Prabhu, A. Ghosh-Roy, R.V. Hosur, *Biochimie* 89 (2007) 117–134.
- [36] I.J. Ropson, J.A. Boyer, P.M. Dalessio, *Biochemistry* 45 (2006) 2608–2617.
- [37] Z. Ignatova, L.M. Gierasch, *Biochemistry* 44 (2005) 7266–7274.
- [38] M. Silow, M. Oliveberg, *Proceedings of the National Academy of Sciences of the United States of America* 94 (1997) 6084–6086.
- [39] R.J. Sessler, N. Noy, *Molecular Cell* 18 (2005) 343–353.
- [40] S.D. Ayers, K.L. Nedrow, R.E. Gillilan, N. Noy, *Biochemistry* 46 (2007) 6744–6752.
- [41] M. Sukumar, L.M. Gierasch, *Folding & design* 2 (1997) 211–222.
- [42] K.S. Rotondi, L.M. Gierasch, *Biochemistry* 42 (2003) 7976–7985.
- [43] P.L. Clark, Z.P. Liu, J. Rizo, L.M. Gierasch, *Nature Structural biology* 4 (1997) 883–886.
- [44] Z.P. Liu, J. Rizo, L.M. Gierasch, *Biochemistry* 33 (1994) 134–142.
- [45] P.L. Clark, B.F. Weston, L.M. Gierasch, *Folding & design* 3 (1998) 401–412.
- [46] P.L. Clark, Z.P. Liu, J. Zhang, L.M. Gierasch, *Protein Science* 5 (1996) 1108–1117.
- [47] H. Li, C. Frieden, *Proceedings of the National Academy of Sciences of the United States of America* 104 (2007) 11993–11998.
- [48] K.S. Rotondi, L.F. Rotondi, L.M. Gierasch, *Biophysical Chemistry* 100 (2003) 421–436.
- [49] P.M. Dalessio, J.A. Boyer, J.L. McGettigan, I.J. Ropson, *Biochemistry* 44 (2005) 3082–3090.

Supplementary Material

Table S1: Summary of k_{ex} derived from H/D SOFAST NMR experiments at $p^2\text{H}$ 6.5.

Residue	$k_{\text{ex}} \cdot p^2\text{H } 6.4^{\text{a}}$ (min^{-1})
W6	2.20E-03
V8	6.90E-03
Y9	3.00E-04
A10	8.70E-03
F17	3.20E-03
L18	1.40E-03
K19	4.10E-03
A20	1.04E-02
L21	4.50E-03
V38	3.00E-04
E39	8.00E-04
I40	1.40E-03
D46	2.90E-03
F47	2.40E-03
V49	1.40E-03
T50	6.60E-03
F62	1.55E-02
L78	6.40E-03
H83	-
L89	8.50E-03
T91	-
E101	4.70E-03

M107	1.10E-03
------	----------

V108	-
------	---

T110	6.49E-05
------	----------

I111	1.00E-04
------	----------

T112	7.00E-04
------	----------

F113	5.60E-03
------	----------

L118	2.00E-04
------	----------

I119	3.40E-03
------	----------

R120	3.00E-04
------	----------

S122	6.80E-03
------	----------

K123	1.10E-02
------	----------

V125	4.10E-03
------	----------

Figure legends

Figure S1. $\Delta G_{u,i}^{H2O}$ vs m_i plot singled out from the best fit of the sigmoidal decay of the NMR signal for most of the residues, using the simple one-step model (major group, full circles, no labelling for each amino acid was applied for visual reasons). We observe a straight line trend with a slope ($[U]_{d,i} = \Delta G_i^{H2O} / m_i$ ratio) of about 3.35 M.

The insert reports the histogram of the population of residues vs the χ^2 fit parameter (full black bars; the smaller the χ^2 the better the fitting). For a minor group of residues (white bars) a simple one-step model was not able to reproduce the early trend of the sigmoidal data decay, thus giving an unsatisfactory fit. The relevant parameters, however, are also reported in the $\Delta G_{u,i}^{H2O}$ vs m_i plot (empty squares, minor group). The ΔG_i^{H2O} and m_i values in this case are much lower and the $\Delta G_i^{H2O} / m_i$ ratio was dispersed around a mean value of about 2.6 M.

Figure S2. Micro-DSC traces (0.5 °C/min scanning rate) obtained from cL-BABP in acetate 50mM, NaCl 150 mM, pH 6 (curve a) and in Tris-HCl 20mM, NaCl 150 mM, pH 7.5 (curve b) buffer solutions. Curve c corresponds to the measurement of cL-BABP protein with the free cysteine methylated with CH₃I in Tris-HCl 20mM, NaCl 150 mM, pH 7.5 buffer solution.

Figure S3. Fingerprint region of the ¹H TOCSY spectra acquired at 500 MHz (isotropic mixing =80ms) for wt cL-BABP 0.5 mM in 30 mM phosphate buffer at pH 7 (left panel) and H98Y cL-BABP (soluble fraction) in the same solution conditions

Figure S1

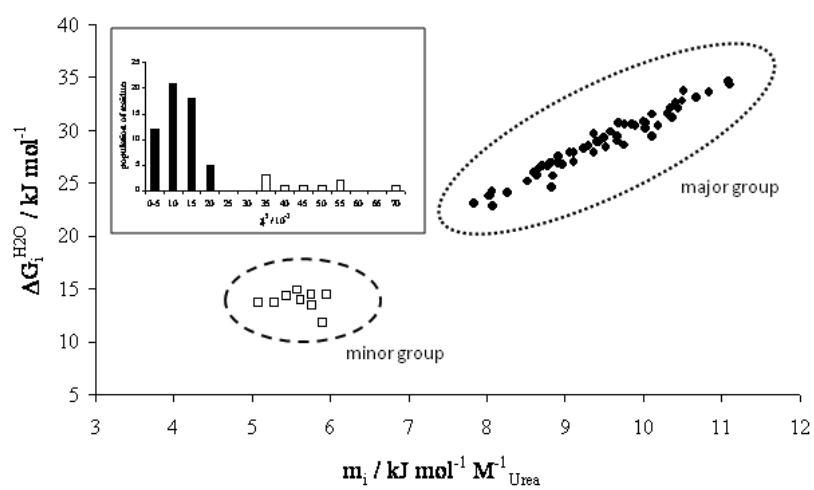


Figure S2

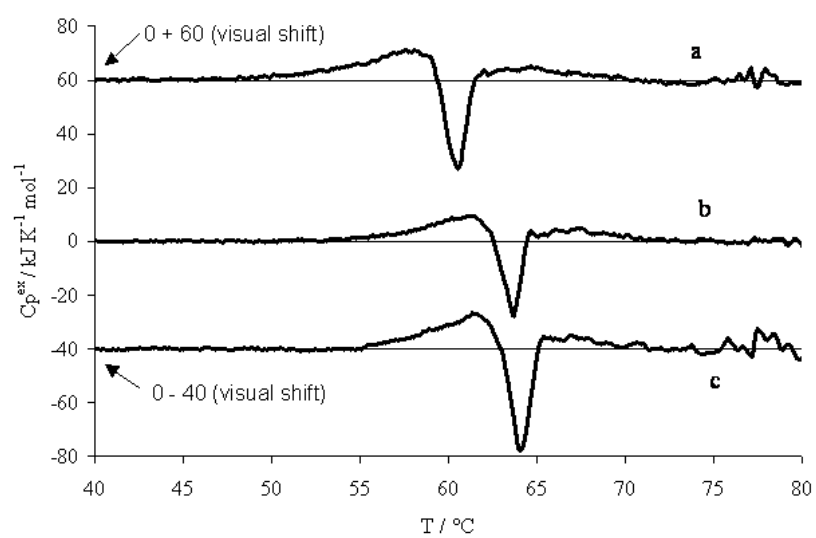
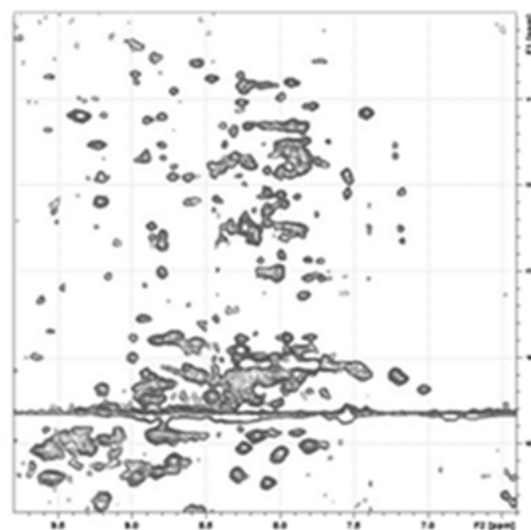
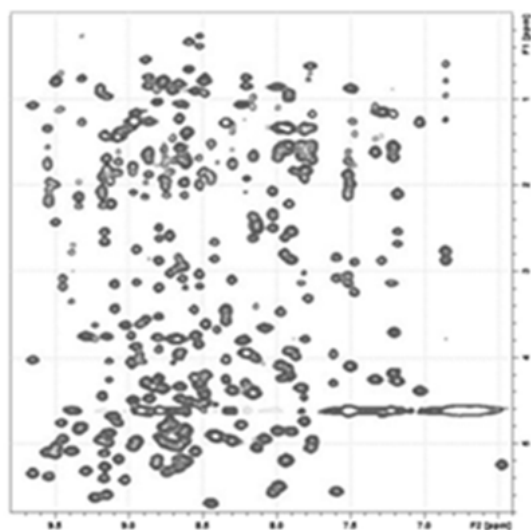


Figure S3



CHAPTER 5: Publication 2

Title: Towards the elucidation of molecular determinants of cooperativity in the liver bile acid binding protein.

Short presentation of publication 2:

Allosteric regulation is a common mechanism for the control of most metabolic and signal transduction pathways. An allosteric response is defined as the effect that binding of one ligand to a protein has on the affinity of the protein and/or catalysis for a second ligand. When more than one ligand is present and all the ligands are chemically identical is referred to as homotropic allostery or homotropic cooperativity. A well known cooperative response is represented by the mechanism that allows a protein to bind a ligand with increased affinity after a first binding event. A protein that has been suggested to undergo such an allosteric activation is the chicken liver bile acid-binding protein (cL-BABP). cL-BABP can bind two molecules of bile acids in two distinct binding sites (site 1 and site 2) and it has been proposed that the binding of the first bile acid molecule determines an allosteric activation for the second binding event, through the stabilisation of a protein open conformation. This mechanism implies the presence of a dynamic ensemble of conformations in the apo form. The ligand has thus the capability of selecting one existing conformation thus shifting the equilibrium towards the most stable holo form. It has been suggested that the protonation/deprotonation of the residue His 98 is the event triggering the conformational mobility of apo cL-BABP.

As reported in the following publication, to verify this hypothesis we introduced a mutation at the 98 position, substituting the histidine residue with a glutamine. The substitution was achieved at DNA level by site direct mutagenesis and the mutated protein was produced with high yield through classical recombinant protein expression methods.

First of all the structural and stability properties of the mutated protein, defined H98Q, was analyzed. The NH backbone signals of apo and holo (complexed with bile acids) forms at two pH (7.2 and 5.6) were obtained by recording and analyzing three-dimensional ^1H , ^{15}N -NOESY-HSQC and ^1H , ^{15}N -TOCSY-HSQC spectra. The analysis of the combined (proton and nitrogen) chemical shift changes of H98Q mutant with respect to the wild type (WT) protein evidenced that, in the apo form, the C-terminal half is the region mostly affected by the mutation while the N-terminal part resulted unchanged. In the holo form differences were located in residues near the substitution and, interestingly, in the 55-59 segment which is far from the mutation and close to the ligand in site 1. The structural integrity was confirmed by the retrieval of practically 100% of through-space backbone connectivities observed in the WT protein through NOESY crosspeaks. H98Q mutant stability was assessed by hydrogen exchange (HX) measurements using SOFAST NMR techniques which revealed a slightly reduced stability ($\Delta G_{\text{op}} = 6.60 \pm 0.05 \text{ kcal mol}^{-1}$) for the mutant with respect to the WT protein ($\Delta G_{\text{op}} = 7.78 \pm 0.03 \text{ kcal mol}^{-1}$).

To investigate the bile acid binding capacity of the mutant mass spectrometry measurements were carried out on both H98Q and WT proteins in the presence of different amounts of bile acid. The presence of a singly ligated protein was identified in the mutant that was almost absent in the case of WT protein. The binding site occupancy was further investigated through NMR spectroscopy by titrating both H98Q and WT proteins with ^{15}N -enriched ligands and by acquiring ^1H , ^{15}N -HSQC experiments. The main effect of the mutation on the binding properties was represented by a severe line broadening of the signal corresponding to the ligand bound to site 1, possibly indicating an increased exchange with the unbound ligand. NMR diffusion measurements, performed on WT and mutant proteins, identified an increased diffusion coefficient for the ligand in site 1, indicating a reduced average occupancy of the corresponding binding site.

The pKa for the two histidine residues (83 and 98) present in the WT protein was also estimated. A series of ^1H , ^{13}C HSQC experiments were acquired on a WT cL-BABP sample at different pHs and the chemical shift of the C ϵ 1-H ϵ 1 resonance was monitored. By fitting the chemical shift variations as a function of pH it was possible to calculate a pKa of 6.0 for His 83. For His 98 a pKa of 4.8-4.5 was determined in spite of an incomplete titration curve due to protein unfolding at pH lower than 4.5. The affinity of the WT and H98Q mutant for bile acids was also monitored as a function of pH. We recorded several ^1H , ^{15}N -HSQC spectra on a ^{15}N -labeled bile acid/unlabeled protein complex at different pH values and monitored the signals of the ligands. In both cases the signals of the bound ligands decreased with the lowering of pH and the mutant showed a more dramatic drop in intensity. This behaviour confirmed the role of His98 in stabilizing the complex at neutral pH.

The protein backbone dynamics of the H98Q mutant was also investigated by standard ^{15}N relaxation measurements. R_1 and R_2 relaxation rates were estimated and were compared with the data previously obtained for the WT protein. His 98 substitution caused a global quenching of the motions observed in D,E,G,H and I strands for the WT protein.

All the data collected suggested that the introduced mutation determines a disruption of the cooperative binding mechanism thus confirming the central role of His 98.

My contribution in this work was in the planning and performance of the experiments, in the data analysis and discussion and in the writing of the scientific paper.

Towards the elucidation of molecular determinants of cooperativity in the liver bile acid binding protein

Massimo Pedò,¹ Mariapina D'Onofrio,¹ Pasquale Ferranti,² Henriette Molinari,¹ and Michael Assfalg^{1*}

¹Dipartimento di Biotecnologie, Università di Verona, Strada le Grazie 15, Verona 37134, Italy

²Dipartimento di Scienza degli Alimenti, Università di Napoli Federico II, Parco Gussone, Portici 80055, Italy

ABSTRACT

Bile acid binding proteins (BABPs) are cytosolic lipid chaperones contributing to the maintenance of bile acid homeostasis and functional distribution within the cell. Liver BABPs act in parallel with ileal transporters to ensure vectorial transport of bile salts in hepatocytes and enterocytes, respectively. We describe the investigation of ligand binding to liver BAPP, an essential step in the understanding of intracellular bile salt transport. Binding site occupancies were monitored in NMR titration experiments using ¹⁵N-labelled ligand, while the relative populations of differently bound BAPP forms were assessed by mass spectrometry. This site-specific information allowed the determination of intrinsic thermodynamic parameters and the identification of an extremely high cooperativity between two binding sites. Protein-observed NMR experiments revealed a global structural rearrangement which suggests an allosteric mechanism at the basis of the observed cooperativity. The view of a molecular tool capable of buffering against significant concentrations of free bile salts in a large range of solution conditions emerges from the observed pH-dependence of binding. We set to determine the molecular determinants of cooperativity by analysing the binding properties of a protein containing a mutated internal histidine. Both mass spectrometry and NMR experiments are consistent with an overall decreased binding affinity of the mutant, while the measured diffusion coefficients of ligand species reveal that the affinity loss concerns essentially one of the two binding sites. We therefore identified a mutation able to disrupt energetic communication functional to efficient binding and conclude that the buried histidine establishes contacts that stabilize the ternary complex.

Proteins 2009; 77:718–731.
© 2009 Wiley-Liss, Inc.

Key words: NMR spectroscopy; fatty acid binding protein; bile acids; cooperativity; molecular recognition.

INTRODUCTION

In recent years, the discovery of a number of transporter proteins expressed in the liver and intestine, specifically involved in bile acid transport, has led to an improved understanding of bile acid homeostasis and the enterohepatic circulation. Bile acid transporters are now recognized to play central roles in driving bile flow, as well as in the adaptation to various pathological conditions, with complex regulation of activity and function in the nucleus, cytoplasm and membrane.^{1–4}

In a continuous effort to understand the mechanism underlying bile acid binding and release within the cytoplasm,^{5–8} we describe here the study of protein-ligand interactions involving liver bile acid binding proteins (L-BABPs), which act in parallel with the ileal transporters (ileal bile acid binding proteins, I-BABPs) to ensure vectorial transport of bile salts within hepatocytes and enterocytes, respectively.⁵

L-BABPs are small molecular mass proteins (14–15 kDa), belonging to the fatty acid binding protein family (FABP), exhibiting the typical and well-conserved fold of the family in which 10 strands of antiparallel β -sheets surround the hydrophobic ligand binding cavity and two short α -helices are located between the first and second strands.

I-BABP was shown to be able to bind bile salts with high cooperativity, although the structural basis of this feature was not clarified.^{9–11} Given the expected parallel function of L-BABP in hepatocytes, we reasoned that the identification of a similar cooperative binding mechanism would firmly set this parallelism, adding new insights into the mechanism of molecular recognition of this protein subfamily.

In a previous work, focussed on the structural and dynamic properties of chicken liver bile acid binding protein (cL-

Additional Supporting Information may be found in the online version of this article.

Grant sponsor: FIRB 2003; Grant number: RBNE03PX83; Grant sponsors: Fondazione Cariverona, Università degli Studi di Verona, CIRMMP.

*Correspondence to: Michael Assfalg, Dipartimento di Biotecnologie, Università degli Studi di Verona, Strada le Grazie, 15 Verona 37134, Italy. E-mail: michael.assfalg@univr.it

Received 9 April 2009; Revised 21 May 2009; Accepted 1 June 2009

Published online 9 June 2009 in Wiley InterScience (www.interscience.wiley.com).

DOI: 10.1002/prot.22496

BABP),^{6,8} we underlined the role of the mobility of the unbound protein, which resulted capable of sampling a variety of conformations including the one stabilized by the ligands.¹² Molecular dynamics simulations suggested that this “allosteric” binding behavior was related to the protonation state of a buried histidine (H98), involved (at least in the simulation) in triggering a conformational change, namely the opening/closure of the EF loop, at the protein open end.⁸ Specifically, the highly conserved H98 was indicated to play a central role in establishing a network of hydrogen bonds and salt bridges among buried residues, defining a sort of continuous polar “spine” connecting remote strands of the apo protein.⁶ In this light the comparison of the binding behavior of WT cL-BABP and its mutant devoid of H98 should provide the means to describe the elusive structural and dynamic basis of cooperativity in the ternary complex with two bile salt molecules. Indeed, understanding the molecular mechanism underlying cooperativity is an elaborate task involving detailed characterization of all ligated states. Precious insights would be gained by analyzing the structure and dynamics of singly bound states, in comparison with data obtained for the free and fully loaded forms. However, due to cooperativity, the half-saturated states are always too scarcely populated to allow direct experimental observation. One way to circumvent this problem is to reduce binding affinity of one of the sites by amino acid mutation. This strategy, here delineated and followed, has found extensive application for example, in NMR studies of EF-hand calcium binding proteins.^{13,14}

To analyse ligand binding and gain insight into the structural basis of cooperativity we have applied to both the WT protein and its mutant H98Q a variety of NMR methods which provide site-specific and residue-specific information, both at the structural and dynamic level. The data are complemented with a mass spectrometry analysis of the bound proteins.

MATERIALS AND METHODS

Sample preparation

The expression plasmid for H98Q cL-BABP was obtained from that of wild type (WT) cL-BABP using the Quickchange (Stratagene) mutagenesis kit. The proteins were expressed as previously reported.⁸ A number of different mutations at position 98 were essayed but they resulted in scarcely soluble proteins. All the protein preparations were checked by 1D ¹H NMR prior and after delipidation and reproducible spectra were always obtained for the apo and undelipidated forms. The molecular weight of the two proteins and the extent of labeling was verified by MALDI mass analysis.¹⁵ N Glycochexodeoxycholic acid was prepared as described.⁷

Mass spectrometry analysis

A Q-TOF Ultima mass spectrometer (Waters, UK) was used in positive ion mode in all the experiments. Ten microliter aliquots of reaction mixture were introduced through a Rheodyne external loop injector into the ion source at a flow rate set to 1 μ L/min with a Phoenix 20 CU HPLC pump driving a gas-tight syringe connected to the instrument capillary. To obtain maximum signals for the ions of the putative PL complex, a solution of 5% acetonitrile in 10 mM ammonium hydrogen carbonate, pH 7.0, was used. Different conditions of orifice voltage (20–80 V) and temperature (40–100°C) were employed. The declustering potential of 40 V and the source temperature of 60°C were the chosen parameters at which no fragmentation was observed and optimum sensibility for the adduct peaks was attained. Spraying was achieved, using nitrogen as the nebulizing gas, by charging the probe at 3.6 kV. Calibration of the mass scale was performed using the multiple charged ions of horse heart myoglobin from a separate sample introduction in positive ion mode. Full-scan mass spectra were generated in continuous data-acquisition mode. The spectra reported are an average of three scans from m/z 800 to m/z 2200 (scan rate = 10 s/scan). Quantitative analysis of components was performed by integration of the multiple-charge ions relative to each species.

NMR spectroscopy

NMR experiments were performed on a Bruker Avance 500 spectrometer equipped with a triple-resonance TXI probe and incorporating x, y, z-axes gradient coils, and on a Bruker Avance III 600 spectrometer equipped with a pulsed field gradient triple resonance TCI cryoprobe. Standard sequence schemes employing pulsed field gradients were employed to achieve suppression of the solvent signal and spectral artefacts. Direct and indirect dimensions were normally apodized by use of 90°-shifted squared sine-bell functions (for ¹³C- and ¹⁵N-edited dimensions) or Lorentzian-to-Gaussian functions (for the ¹H dimension), followed by zero filling and Fourier transform. The NMR data were processed with Topspin 2.1 (Bruker) and analyzed with the same software or with Sparky 3.11 (UCSF). Typical samples contained 0.5 mM protein dissolved in 30 mM phosphate (Na₂HPO₄/NaH₂PO₄), 90% H₂O-10% D₂O. Proton chemical shifts were referenced to external 3-(trimethylsilyl)-3,3,2,2-d₄-propionic acid while nitrogen chemical shifts were referenced indirectly as described.¹⁵ The measurement temperature was 298 K.

Resonance assignments

The complete resonances assignments for both apo- and holo WT cL-BABP were available.^{6,8} Amide resonances assignments for apo- and holo H98Q cL-

BABP at pH 7.2 and 5.6 were performed by following titration data (addition of ligand or change of pH) as well as by analysis of four pairs of TOCSY- ^1H , ^{15}N -HSQC and NOESY- ^1H , ^{15}N -HSQC spectra registered at 600 MHz. The NOESY mixing time was set to 120 ms, while a spin-lock time of 70 ms was used for TOCSY spectra.

Hydrogen exchange experiments

Hydrogen exchange (HX) experiments were performed at p^2H 7.4 and 6.5 ($\text{p}^2\text{H} = \text{pH}^{\text{read}} + 0.4$). The HX sample of ^{15}N H98Q cL-BABP was prepared by dissolving 5 mg of delipidated and lyophilized protein in 600 μL of 30 mM deuterated $\text{Na}_2\text{HPO}_4/\text{NaH}_2\text{PO}_4$ buffer. The final protein concentration was about 0.4 mM. The solution was centrifuged briefly at 4°C to remove insoluble protein and transferred to a 5 mm NMR tube. SOFAST-HSQC experiments¹⁶ were collected at 600 MHz, with 700 points in F_2 and 192 complex data points in F_1 , using a relaxation delay of 0.5 s with 2 scans and a duty cycle of 7.6%. The band-selective ^1H excitation (PC9) and refocusing (RSNOB) pulses were centered at 4.7 ppm with pulse lengths of 2.33 and 0.817 ms, respectively. ^{13}C decoupling was achieved through an adiabatic pulse (smoothed chirp) employing a 500 μs pulse, where the carbon hard pulse corresponded to 11.8 μs . An adiabatic pulse of 2 ms was employed for ^{15}N decoupling, where the nitrogen hard pulse was 45 μs . HX rate constants were determined by fitting cross-peak volumes to a first-order exponential decay:

$$I(t) = I(0) \exp(-k_{\text{ex}}t) \quad (1)$$

where $I(t)$ represents the volume of the cross-peaks at the time t , $I(0)$ the volume of the cross peak at $t = 0$; k_{ex} is the observed rate of hydrogen exchange and t is the time in minutes. Here, $t = 0$ corresponded to 11 minutes and k_{ex} greater than 10^{-1} min^{-1} could not be determined.

The free energy of opening, ΔG_{op} , was derived from the equation:

$$\Delta G_{\text{op}} = -RT \ln K_{\text{op}} \quad (2)$$

where $K_{\text{op}} = k_{\text{ex}}/k_{\text{ch}}$ and k_{ch} for model random peptides was derived as previously reported.¹⁷ Data were fitted with the program Sigmaplot (Jandel Scientific).

Protein-ligand titration experiments

Protein stock solution concentrations were determined by UV while the ligand stock solutions were determined by measuring dry weights using a microbalance. Random samples were prepared three times and the error on the measured points reflects the standard deviation from the mean. Typical ^1H - ^{15}N HSQC spectra were collected at 500 MHz with 40 and 12 ppm spectral widths and 256

and 1024 complex points along the $F_1(^{15}\text{N})$ and $F_2(^1\text{H})$ frequency dimensions, respectively. Thirty-two scans were collected for observation of ^{15}N -labelled protein samples, while 256 scans were used for observation of ^{15}N -labelled ligand. For the latter experiments only 64 data points in $F_1(^{15}\text{N})$ were collected.

The NMR isotherms for the binding of ^{15}N -GCDA to WT cL-BABP were analyzed simultaneously according to the site-specific model depicted in the inset of Figure 5. The volumes for the unbound, site 1, and site 2 resonances were estimated considering the following relationships:

$$P = C_P / [1 + (\kappa_1 + \kappa_2)L + c_{12}(\kappa_1 + \kappa_2)L^2]; \quad PL' = P \cdot \kappa_1 \cdot L; \\ PL'' = P \cdot \kappa_2 \cdot L; \quad PL_2 = P \cdot c_{12}\kappa_1\kappa_2 \cdot L^2 \quad (3)$$

where κ_1 and κ_2 are the intrinsic affinity constants and c_{12} the intrinsic cooperativity factor. P , L , PL' , PL'' , PL_2 are the equilibrium concentrations of unbound protein, unbound ligand, protein singly bound at site 1, protein singly bound at site 2, and double bound protein, respectively. C_P is the total protein concentration. L was obtained by solving analytically the cubic equation:

$$c_{12}\kappa_1\kappa_2 \cdot L^3 + [(2C_P - C_L)c_{12}\kappa_1\kappa_2 + \kappa_1 + \kappa_2] \cdot L^2 \\ + [(C_P - C_L)(\kappa_1 + \kappa_2) + 1] \cdot L - C_L = 0 \quad (4)$$

where C_L is the total ligand concentration. This analysis was performed with an in-house Matlab routine.

The Hill coefficient determined at half saturation was used as an indicator of macroscopic cooperativity and calculated as:

$$n_H = 2/[1 + (K_{d2}/K_{d1})^{0.5}] \quad (5)$$

where K_{d1} and K_{d2} are the stepwise dissociation constants, related to the intrinsic affinity constants as follows: $K_{d1} = 2/(\kappa_1 + \kappa_2)$, $K_{d2} = (c_{12}K_{d1}\kappa_1\kappa_2)^{-1}$.

pH titration experiments

A sample of 0.5 mM ^{15}N , ^{13}C -labelled WT cL-BABP was subjected to pH change and an echo-antiecho ^1H , ^{13}C -HSQC spectrum was acquired at 600 MHz at each of the following pH values: 4.1, 4.7, 5.0, 5.1, 5.5, 5.9, 6.2, 6.7, 7.1, 7.5, and 7.9. Thirty-two scans were acquired over a spectral window of 12 (F_2 , ^1H) \times 40 (F_1 , ^{13}C) ppm. The nitrogen carrier was placed at 125 ppm, and the INEPT delay was set to 1,725 ms. The pH change was obtained by dialysis of the protein sample against the already equilibrated buffer. At selected titration points, a long-range ^1H , ^{15}N -HSQC experiment was acquired, optimized for observation of histidine resonances with the following parameters: spectral window of 17 (F_2 , ^1H) \times 110 (F_1 , ^{15}N) ppm, 512 scans, nitrogen carrier centered at 205 ppm, INEPT delay 17 ms. Carbon and nitrogen decoupling was introduced. The long-range

^1H , ^{15}N -HSQC experiments were also recorded on a ^{15}N -labelled H98Q cL-BABP sample at different pH values. The pH-dependent proton and carbon chemical shift values obtained from the above ^1H , ^{13}C -HSQC spectra of the wild type protein were used to estimate the pK_a values of the corresponding histidine residues according to the following equation:

$$\delta_{\text{obs}} = \delta_{\text{d}} + (\delta_{\text{p}} - \delta_{\text{d}}) / [1 + 10^{(\text{pH} - \text{pK}_a)}] \quad (6)$$

where δ_{p} and δ_{d} are the chemical shifts of the protonated and the deprotonated state, respectively.

Two series of standard 1D and ^1H , ^{15}N -HSQC (only first increment) titrations were also performed on the following holo protein samples: (i) WT cL-BABP + ^{15}N -GCDA ($P:L = 1:2$) at pH 7.3, 5.9, 4.5, 4.3, 4.1, 3.4, 3.0, 7.1; (ii) H98Q cL-BABP + ^{15}N -GCDA ($P:L = 1:2$) at pH 6.9, 6.1, 5.1, 4.4, 4.2, 4.1, 3.3. The pH was changed by step-wise addition of small aliquots of hydrochloric acid to the samples in starting neutral conditions.

Relaxation rate analysis

^{15}N relaxation experiments were acquired at 600 MHz on ^{15}N -labelled 0.4 mM protein samples, both at pH 7.0 and 5.6. The pulse program for longitudinal relaxation measurements, run in interleaved fashion and including a water flip-back scheme, was provided by Daniel Cicero (University of Rome, Tor Vergata), while for transverse relaxation measurements a standard Bruker two-dimensional pulse sequence was used. Nine delays (0.01, 0.18, 0.36, 0.54, 0.72, 0.90, 1.08, 1.26, 1.44 s) were used for R_1 measurements, and nine delays (16.96, 33.92, 50.88, 67.84, 101.76, 135.68, 169.6, 220.48, and 237.44 ms) were used for R_2 measurements. The delay in the Carr-Purcell-Meiboom-Gill pulse train was set to 0.45 ms. The number of scans was set to 8 for R_1 and to 16 for R_2 , the relaxation delay between each scan was 3 s, and a matrix of 2048×128 data points was acquired. Cross-peaks were integrated with Topspin 2.1, and the volume sets were fitted to a monoexponential decaying function using two adjustable parameters within the program RELAXFIT.¹⁸ The errors were obtained using a standard Monte Carlo approach.

Diffusion experiments

^{15}N -edited diffusion experiments were performed on samples of WT cL-BABP: ^{15}N -GCDA and H98Q cL-BABP: ^{15}N -GCDA with protein:ligand ratios of 1:3, to determine the diffusion coefficients of protein-bound ligands compared to that of the free molecules. The pulse program was obtained by combining the standard HSQC pulse scheme with a PFG-STE (pulsed field gradient stimulated echo) module employing bipolar gradients¹⁹ (the pulse program was kindly provided by Daniel Cicero, University of Rome "Tor Vergata"). PFG-STE-HSQC

spectra were conducted with 4 ms of gradient pulse duration (δ) and 80 ms between the pulsed field gradients (Δ). The spectral widths and other acquisition parameters were the same as those used for standard HSQC experiments for ligand observation. However, 512 or 768 scans per t_1 -point and a recycling delay of 3 s were used. The gradient strength was calibrated from a diffusion experiment on 99.8% D_2O using the HDO apparent diffusion coefficient of $1.902 \times 10^{-9} \text{ m}^2\text{s}^{-1}$ at 25°C . The length of all delays and pulses was held constant while the gradient strength was varied from 2 to 95% of its maximum value (53.5 G/cm). The measured signal volumes as a function of the applied gradient were fitted to the following equation using a nonlinear least squares minimization:

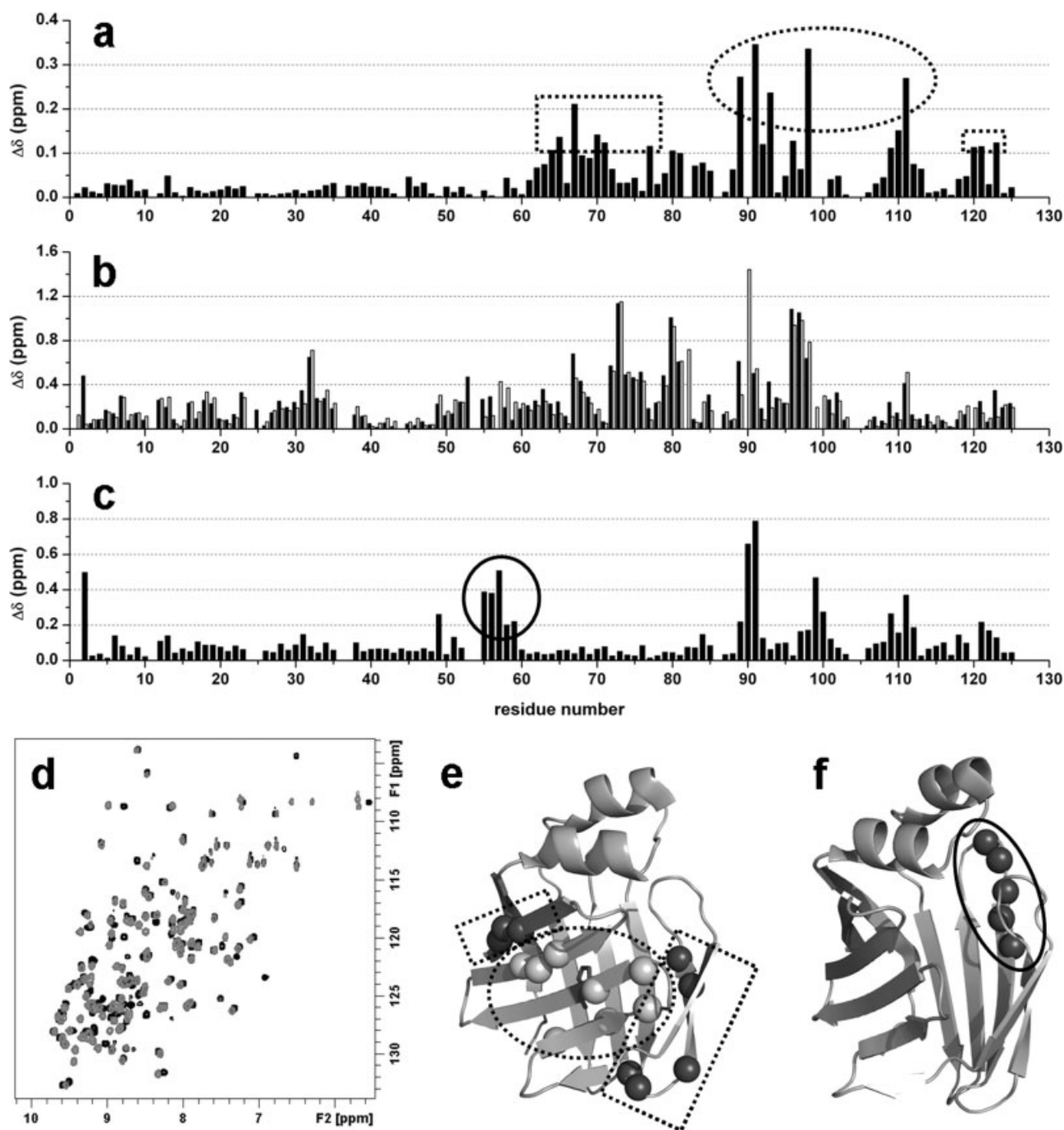
$$I = I(0) \exp[-D\gamma^2 G^2 \delta^2 (\Delta - \delta/3 - \tau/2)] \quad (7)$$

where D is the translational diffusion coefficient, γ is ^1H gyromagnetic ratio, G is the gradient strength, Δ and δ are defined as above and τ is the gradient pulse separation.

RESULTS

Structure and stability assessment of WT and H98Q cL-BABP

The assignments for free and bound WT cL-BABP were determined previously at pH 7.2 and 5.6, and the solution structures for apo and holo forms at pH 7.2 were reported.^{6,8,20} The choice of amino acid for mutation of the highly conserved H98 was based on the observation of the presence of glutamine in few sequences of proteins belonging to the BABP family. The H98Q mutant exhibited well-dispersed ^1H , ^{15}N -HSQC spectra [Fig. 1(d)] and the assignment of the amide resonances was obtained at pH 7.2 and 5.6 by use of ^1H , ^{15}N -NOESY-HSQC and ^1H , ^{15}N -TOCSY-HSQC spectra. The comparison of peak positions of the wild type and mutant proteins is reported in Figure 1 for both apo and holo proteins. It appears that while the N-terminal half of the apo protein is completely unaffected by the mutation, non-negligible perturbations are evidenced in the C-terminal half, with combined (proton and nitrogen) chemical shift changes up to 0.4 ppm. While the largest effects are, as expected, close to the mutation (residues in the range 89–111, belonging to strands G, H, I), the changes involve also distant residues located mainly in loop DE and strands E, F [Fig. 1(a,e)]. The distances of these residues from position 98 are larger than 6.5 Å, implying that these effects cannot be merely ascribed to ring current shifts²¹ but are indicative of an altered stereo-electronic environment. The NOESY patterns of H98Q cL-BABP residues which showed chemical shift variations compared to the wild type protein were analyzed. Practi-

**Figure 1**

Chemical shift differences between H98Q cL-BABP and WT cL-BABP. (a) Combined average chemical shift differences ($\Delta\delta = \{[(\Delta\delta_H)^2 + (\Delta\delta_N/5)^2]/2\}^{1/2}$) between the apo proteins; the protein regions displaying the largest differences are highlighted with a dotted circle (residues close to the mutation) and dotted rectangles (residues far from the mutation). (b) Chemical shift differences between holo and apo H98Q cL-BABP (grey bars) and holo and apo WT cL-BABP (black bars). (c) Chemical shift differences between holo H98Q cL-BABP and holo WT cL-BABP; the largest differences experienced by residues (56–59) distant from the mutation site are highlighted by a circle. (d) Superposition of ^1H , ^{15}N -HSQC spectra of apo WT (black) and H98Q cL-BABP (grey). (e) Mapping of the chemical shift differences of plot (a) on the three-dimensional structure of cL-BABP, the residues experiencing the largest perturbations are represented by spheres, histidine98 is represented in sticks. (f) Mapping of the chemical shift differences of plot (c) on the three-dimensional structure of cL-BABP, the residues experiencing the largest perturbations are represented by spheres.

cally 100% of through-space backbone connectivities observed in the wild type protein were retrieved in the spectrum of the mutant, thereby assessing the substantial invariance of the protein scaffold following the mutation.

As already noticed previously⁷, saturation of the binding sites in the WT protein by two bile salt molecules produces chemical shift perturbations that are spread over the whole protein and are as large as 1.4 ppm. The same comparison was here performed on the H98Q mutant by recording an HSQC spectrum on a $P:L = 1:4$ sample. The assignment of amide HN signals was made by use of $^1\text{H}, ^{15}\text{N}$ -NOESY-HSQC and $^1\text{H}, ^{15}\text{N}$ -TOCSY-HSQC spectra at pH 7.2 and 5.6. By plotting the chemical shift differences observed upon binding for both the WT protein and the mutant [Fig. 1(b)], it appears that the corresponding residues are affected by essentially the same amount of shift change. However, the direct chemical shift comparison [Fig. 1(c)] indicates that the shift change upon binding is not always in the same frequency direction. While most of the differences appear in the region of the mutation (residues 90, 91, 99, 100, 109, and 111) and may be accounted for by the loss of the aromatic ring current effect and by changes in local electronic distribution, interesting differences are observed for residues in the 55–59 segment [loop CD and strand D, Fig. 1(f)], a region that faces the carboxylate group of ligand bound to site 1 in the wild type protein.

It was not possible to assess the stability of the H98Q mutant with respect to the wild type protein by differential scanning calorimetry (DSC) because, for the latter, the curve was compromised by a strong exothermic signal related to transient aggregation, occurring at a temperature lower than the melting temperature.²² We therefore used NMR techniques to monitor protein stability. Hydrogen exchange (HX) measurements allow to simultaneously determine the solvent less accessible core of the protein, to infer dynamic information, and finally to derive thermodynamic parameters to be associated with the free energy of unfolding under bimolecular exchange (EX2) conditions.^{23,24} Hydrogen/deuterium exchange rates have been previously measured for WT cL-BABP, using SOFAST NMR techniques,^{25,26} and related to the free energy of opening ΔG_{op} ²² (see Experimental section). Similar exchange experiments were repeated here for the H98Q mutant and the EX2 condition verified. Peak volumes for selected residues are plotted as a function of time in Supporting Information Figure 1. It has been proposed that the average of the three largest ΔG_{op} values, derived from HX measurements, provides a reasonable estimate of the conformational stability of the protein.²⁷ In the WT protein, residues T110, I111, and R120 are characterized by the largest ΔG_{op} values, which translate into an average value of $7.78 \pm 0.03 \text{ kcal mol}^{-1}$ (corrected for the proline cis/trans isomerization equilibrium).²⁷ Following the same approach (Supporting Information Table 1) an average value of $6.60 \pm$

Table 1

Diffusion Coefficients Measured Simultaneously for the Signals Corresponding to Free and Bound Ligand in Samples Containing ^{15}N -GCDA/Unlabelled Protein in a Ratio of 3:1, 25°C

Protein	Diffusion coefficients ($\times 10^{-6} \text{ m}^2 \text{ s}^{-1}$)		
	Free	Site 1	Site 2
WT-cLBABP	3.60 ± 0.18	2.28 ± 0.05	1.60 ± 0.03
H98Q-cLBABP	3.90 ± 0.20	3.23 ± 0.18	1.60 ± 0.06

$0.05 \text{ kcal mol}^{-1}$ was obtained for the mutant, indicating a slightly reduced stability with respect to WT. Very similar figures were found for the two proteins, with the exception of residues located in β -strands I and J, showing lower ΔG_{op} values for the mutant. A further observation is that hydrogen bonds involving amides of V8 (V8-P36 strand A-B), A10 (A10-R121 strand A-J), L21 (L21-F17 helix I-helix I), and F113 (F113-V116 strand I-J) are more labile, if present, in the mutant protein, as their protection factors could not be evaluated because the corresponding peaks were present only in the first HSQC experiment. This provides an indication of a weakening of H-bond patterns at the level of the protein cavity entrance.

Binding stoichiometry and ligand environments

Mass spectrometry analysis carried out on WT and H98Q cL-BABP samples containing different amounts of glycochenodeoxycholic acid (GCDA) showed the presence of adducts with maximum two bound ligand molecules for both proteins. Interesting differences between the spectra of the WT and H98Q cL-BABP were however observed in terms of the relative populations of single- and double-bound proteins. An equimolar ligand/H98Q BABP mixture indicated the presence of adducts with one (PL1) and two (PL2) molecules of GCDA bound per protein monomer with a relative abundance of about 1:1 [Fig. 2, top]. The same experiment carried out with the wild type protein showed that the adduct with two GCDA molecules was quantitatively much more abundant than the adduct with a single GCDA molecule [PL2:PL1 ratio 4:1, Fig. 2, bottom]. When the ligand/H98Q BABP ratio was raised to 5:1, the adduct with two GCDA molecules became dominant, as in the wild type protein. These data indicate that the population of singly ligated species is negligible for WT cL-BABP and that the introduced mutation was indeed effective in increasing the population of the half-saturated state.

The $^1\text{H}, ^{15}\text{N}$ -HSQC spectra of WT and H98Q cL-BABP in the presence of excess ^{15}N -labelled GCDA show two distinct resonances corresponding to bound ligands, in addition to the signal of the free bile salt [Fig. 3(a) and references^{6,7}]. The peak positions are similar in the two

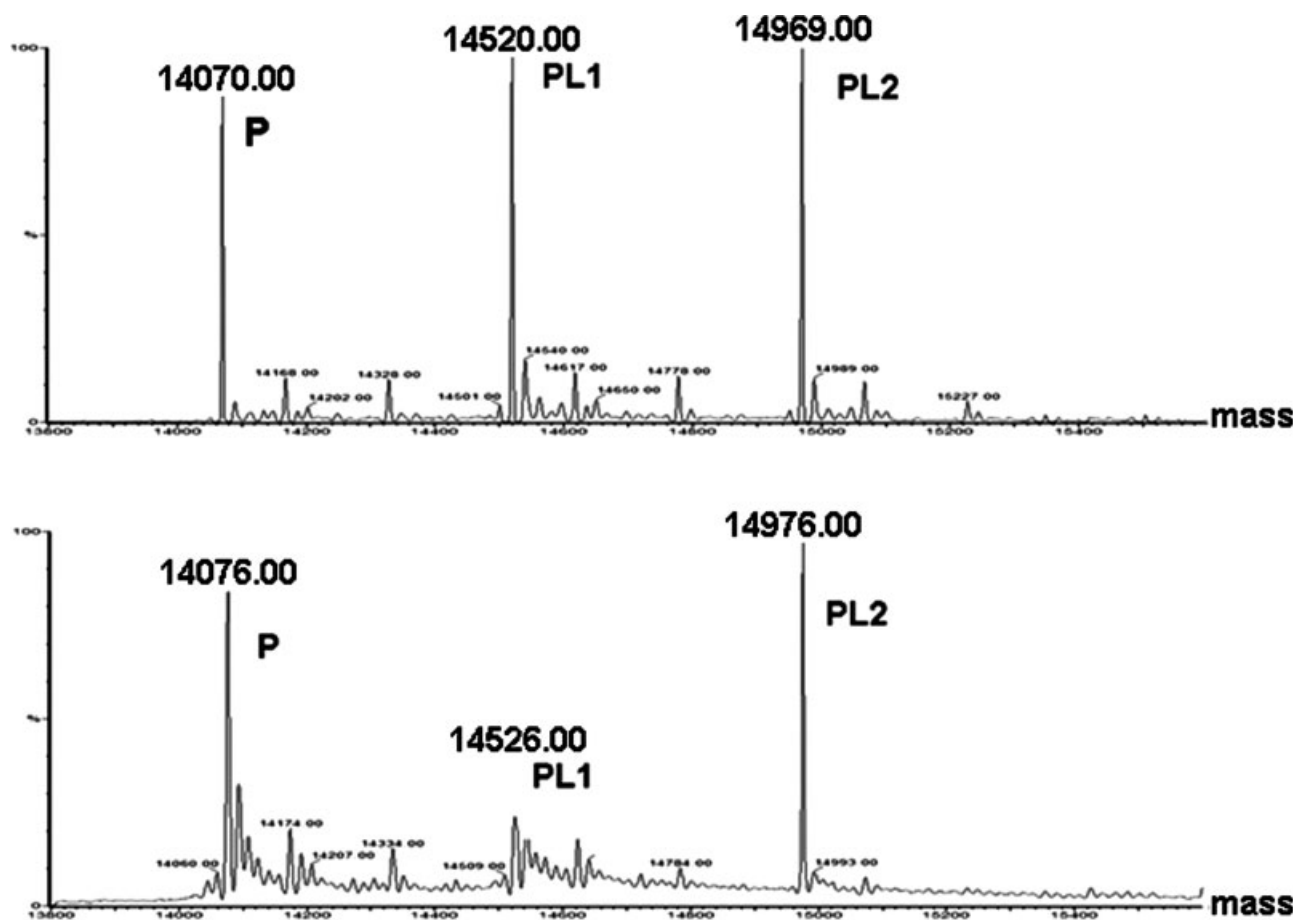


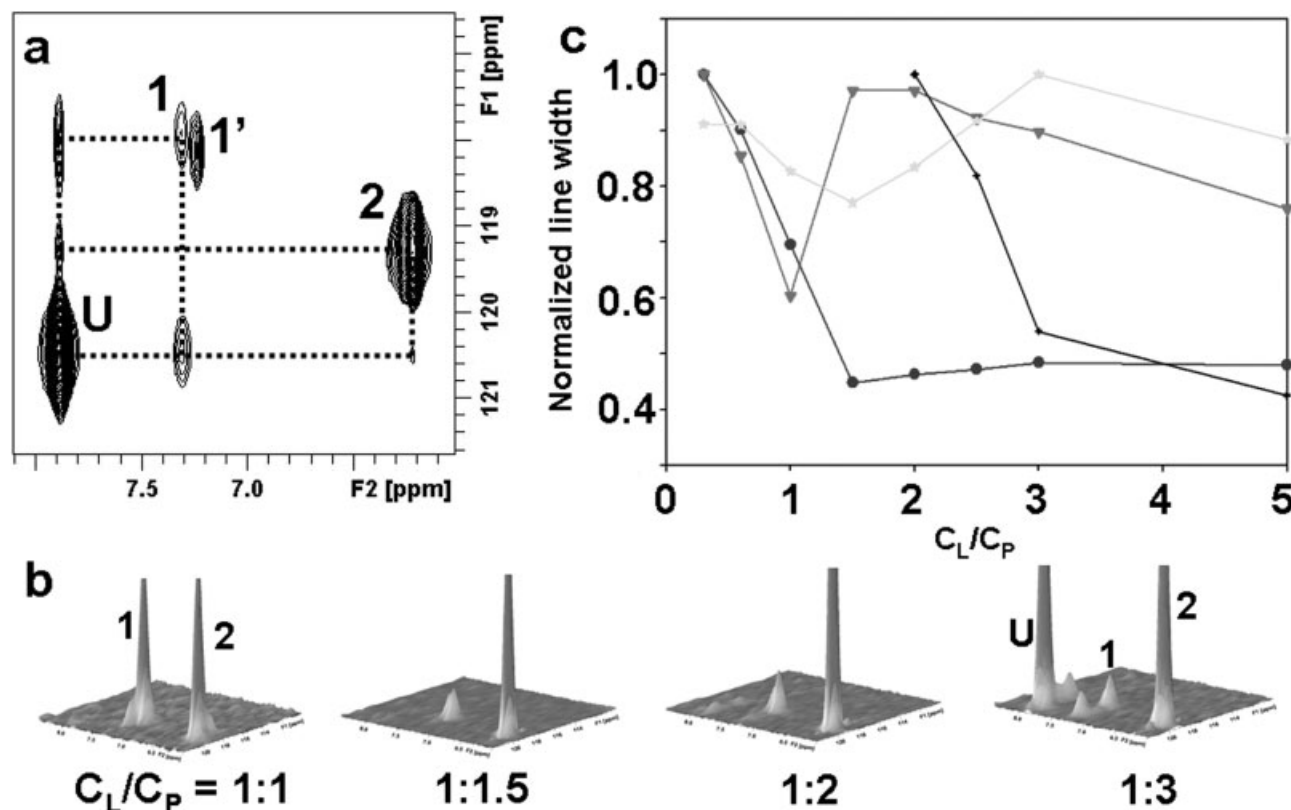
Figure 2

Binding stoichiometry of cL-BABP with GCDA. Transformed ESI spectrum of the equimolar GCDA/H98Q-BABP (top) and equimolar GCDA/WT-BABP (bottom) mixtures. The spectra reveal the presence of complexes with one (PL1) and two (PL2) molecules of GCDA bound per BABP monomer (P). The relative abundance of the various complex forms, measured on the basis of the respective multicharged ion intensities, is different in the two spectra.

proteins, however in the mutant the peak corresponding to site 1 is shifted towards lower field in both the proton and nitrogen dimensions by 0.08 and 0.30 ppm, respectively, while the peak corresponding to site 2 appears 0.12 ppm high-field in the nitrogen dimension and 0.62 down-field in the proton axis. No major rearrangements of the ligand binding sites were therefore inferred, and the two signals were attributed to the ligand bound to sites 1 (δH 7.3 ppm, δN 118.1 ppm) and 2 (δH 6.2 ppm, δN 119.3 ppm) as in the WT protein adduct. A third resonance almost overlapping the signal of site 1 was always present during the titration of the mutant and was attributed to an alternative conformation, denoted as site 1' (δH 7.2 ppm, δN 118.1 ppm). Site 1' was identified in the WT protein as well, but only at a later stage of the titration (protein:ligand ratio 1:2).

Binding site occupancy and energetics

The determination of the relative populations of the differently bound protein species was assayed by titration analyses with ^{15}N -enriched ligands on both the WT and mutant proteins. Titration of WT cL-BABP with ^{15}N -GCDA shows ligand signals with constant line widths and chemical shift position indicating slow chemical exchange processes on the NMR time scale. Thus, the peak intensities are proportional to the binding site occupancies [Fig. 4] and can be used in principle to derive the intrinsic thermodynamic parameters.^{9–11} The experimental data were here analyzed considering a general 1:2 binding model. The intensity ratio between the two signals corresponding to the bound molecules is close to one at any protein:ligand ratio, indicating that the two binding sites are always similarly occupied.

**Figure 3**

Titration of ^{15}N -GCDA into unlabelled H98Q cL-BABP. (a) Contour plot of the ^1H , ^{15}N -HSQC spectrum of a sample containing a protein:ligand ratio of 1:2; U: unbound ligand, 1: ligand bound to site1, 1': ligand bound to site1', 2: ligand bound to site 2. (b) three-dimensional plots of a series of ^1H , ^{15}N -HSQC spectra registered at the indicated protein:ligand ratios; same labelling as in (a); C_L : total ligand concentration, C_P : total protein concentration. (c) Plot of normalized line widths of signals from spectra as in (b); triangles: ligand bound to site 1, asterisks: site 1', circles: site 2, diamonds: unbound ligand.

Titration curve simulations indicated that two different models were consistent with the data: two identical independent high affinity sites, and two highly cooperative binding sites. Because of the absence of changes in spectral positions of peaks corresponding to bound ligands over a range of relative protein/ligand concentrations, the signals of singly bound protein forms could not be resolved by NMR and were considered degenerate with those of the double-bound protein. On the other hand, the reported mass spectrometry observation of scarcely populated singly bound WT protein adducts over a range of protein/ligand molar ratios allows the unambiguous selection of the most appropriate model involving two cooperatively linked binding sites. The estimated intrinsic (site-specific) dissociation constant values are $1.0 \pm 0.3 \times 10^{-3} \text{ M}$ for both sites, with a microscopic cooperativity factor of $9.8 \pm 0.4 \times 10^3$ (extremely high cooperativity), translating into a macroscopic positive cooperativity described by a Hill coefficient at half saturation of 1.98.

In the case of the H98Q mutant, a ^{15}N -labelled GCDA was titrated into unlabelled protein covering protein:

ligand ratios of 1:0.3 to 1:5. The chemical shift positions of free (U) and bound ligands (1, 1' and 2) remained unchanged during the titration. A substantial line broadening was instead observed for the signals of ligands bound to the mutant [Fig. 3(b,c)] and was ascribed to chemical exchange occurring on a faster regime. The observed line width variations of all the ligand resonances on changing protein:ligand ($P:L$) ratio prevent a straightforward determination of binding site occupancies from signal intensities and, consequently, the derivation of the thermodynamic parameters. However, the following observations can be made: (i) up to a $P:L$ ratio of 1:1 the signal line widths of sites 1 and 2 decrease of the same amount and exchange peaks between them are visible, while site 1' exhibits only a minor line width variation; (ii) at $P:L$ ratios $>1:1.5$ the resonance line width of site 2 is substantially unchanged, while a faster chemical exchange is present between site 1 and another unknown position (possibly the free ligand) broadening the signal of site 1 and decreasing its intensity; (iii) at $P:L >1:2$ the resonance of the free ligand becomes detectable, in agree-

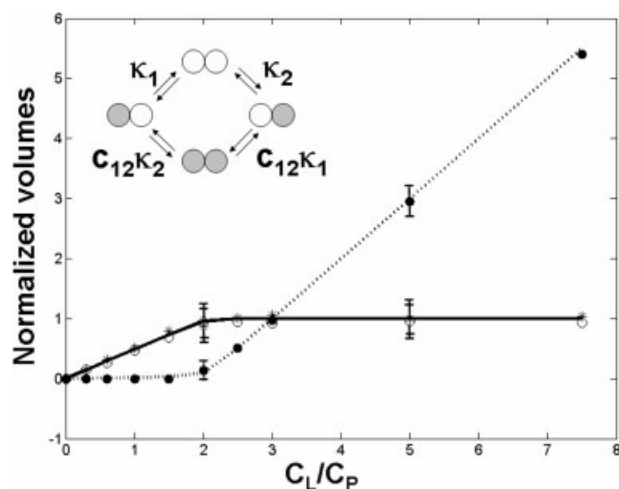


Figure 4

NMR observed titration of ^{15}N -GCDA into unlabelled WT cL-BABP. Ligand peak volumes are plotted as a function of ligand:protein ratio; empty circles: ligand bound to site 1, asterisks: site 2, filled circles: unbound ligand; C_L : total ligand concentration, C_P : total protein concentration; dotted and continuous lines represent the simulation of experimental data according to a two-site binding model as depicted in the inset, the dotted line corresponding to the equilibrium concentration of unbound ligand $[L]$, the continuous lines to the concentrations of the single and double bound protein $[\text{PL}'] + [\text{PL}'L'']$ and $[\text{PL}'] + [\text{PL}'L'']$ (overlapped curves); κ_1 and κ_2 : intrinsic affinity constants relative to the binding to sites 1 and 2, c_{12} : intrinsic cooperativity factor. Error bars are reported for measurements performed at protein:ligand ratios of 1:2 and 1:5, estimated on the basis of experiments performed in triplicate.

ment with the determined 1:2 stoichiometry, and exchange with site 1 is observed.

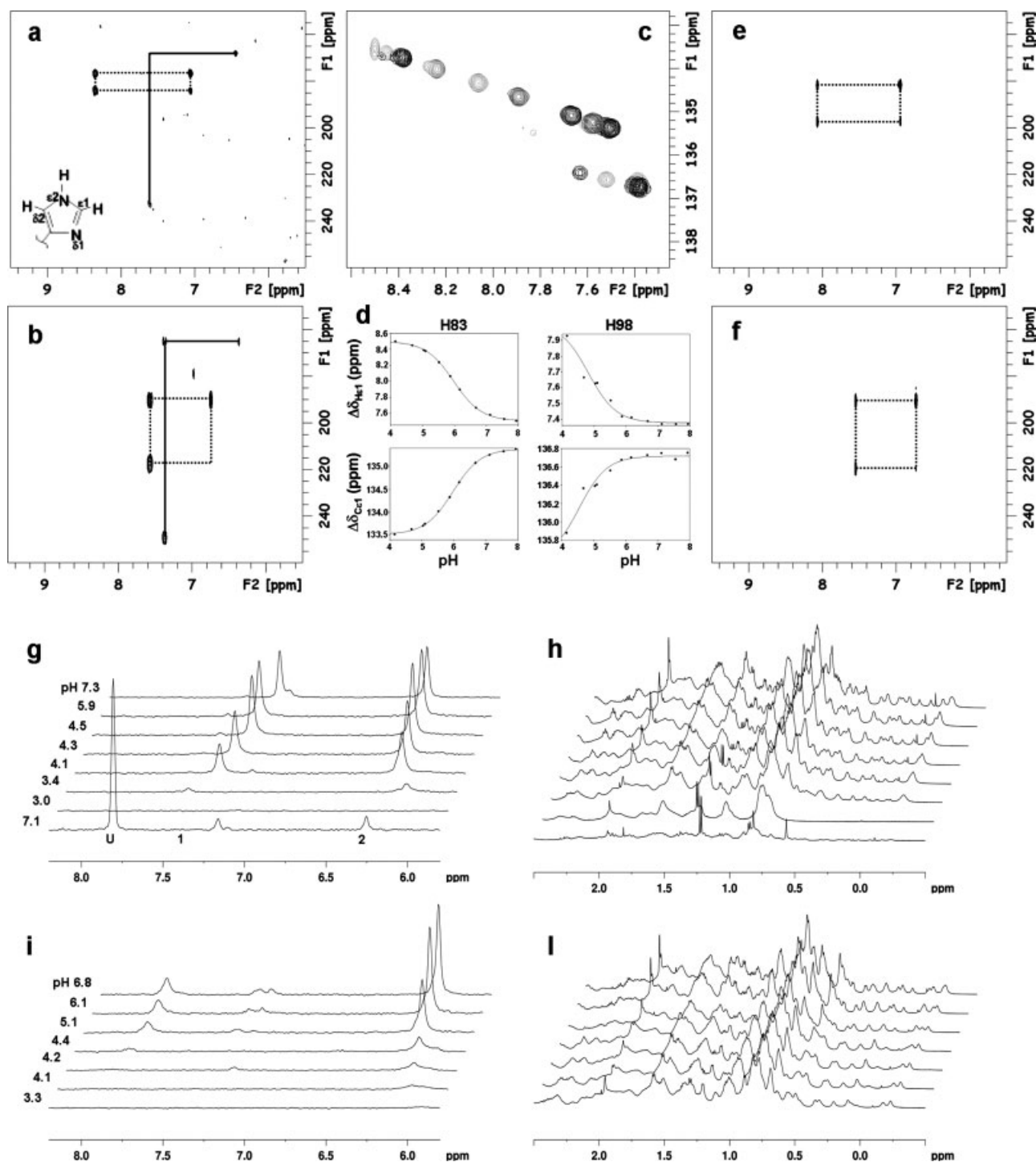
Because of the impossibility of evaluating binding site occupancies directly from peak intensities, site-specific diffusion measurements were essayed. In a previous work we used a PFG-STE pulse sequence concatenated with a one-dimensional version of an F_1 -edited, F_3 -filtered three-dimensional HMQC-NOESY⁶ for measuring diffusion of two chenodeoxycholic acid (CDA) molecules bound to a double labelled WT cL-BABP. The calculated diffusion coefficients (performed on a 1:3 WT protein:ligand adduct) indicated that one molecule is essentially tumbling with the protein, while the other one is slightly exchanging with the free species.²⁸ Here we used an alternative pulse program where the PFG-STE sequence was combined with an HSQC element.²⁹ By observing the signal decays of ^{15}N -labelled ligand in 1:3 WT-cL-BABP/ligand and 1:3 H98Q-cL-BABP/ligand complexes (see Supporting Information Figure 2), the diffusion coefficients of unbound and bound ligands were obtained (Table I). While no change is observed for site 2 in the mutant compared to the wild type protein, a substantial increase of the diffusion coefficient is measured for the ligand in site 1, translating in a reduced average occupancy of this binding site.

pH-dependent equilibria for apo proteins

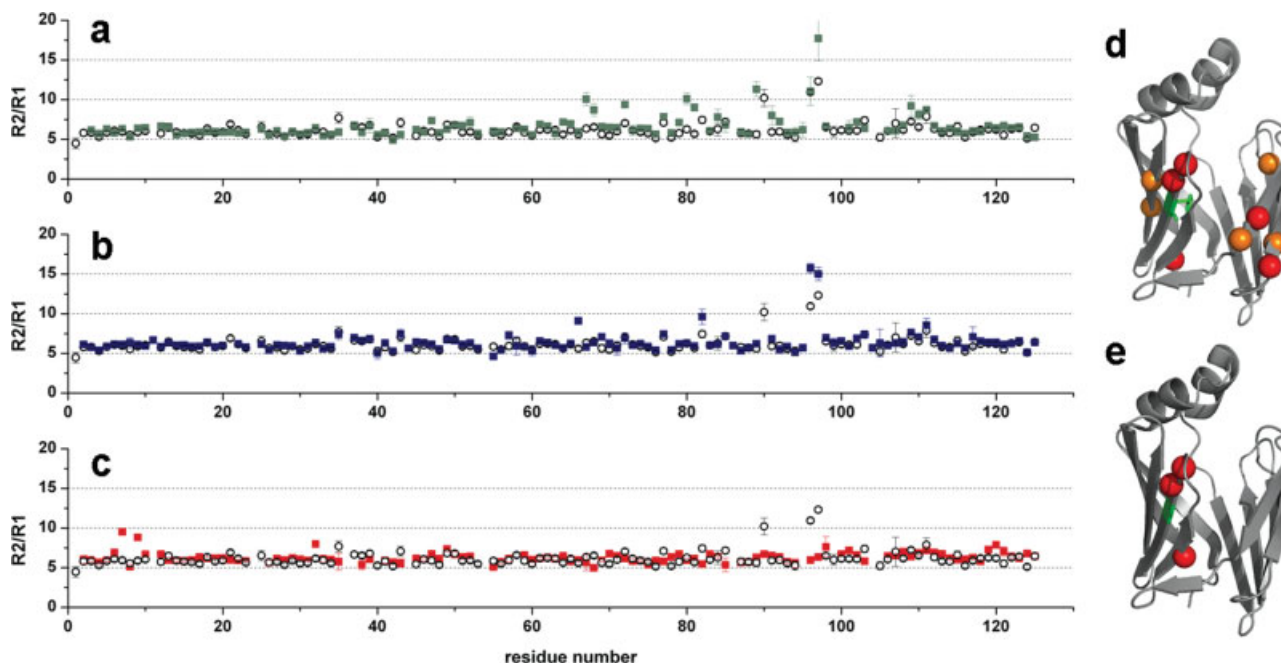
A key feature of the histidine amino acid residue is its imidazole side chain, which can exist in a variety of different protonation states, namely a pair of neutral tautomer forms and a fully protonated charged conformer. A dynamic equilibrium among these various states is often present, as histidines generally have pK_a values close to physiological pH, allowing them to act as switches in proteins in a variety of different contexts. These equilibria can be investigated via long-range (^2J , ^3J -optimized) ^1H , ^{15}N -HSQC spectra linking the protons H δ_2 and H ϵ_1 with nitrogens via two-bond (H δ_2 -N ϵ_2 , H ϵ_1 -N ϵ_2 , and H ϵ_1 -N δ_1) and three-bond (H δ_2 -N δ_1) scalar couplings.³⁰ The characteristic shifts expected for different tautomers and/or protonated forms can be followed as a function of pH.³⁰ Figures 5(a,b) report the long-range HSQC spectra obtained for apo WT cL-BABP at pH 5.1 and 7.1. At neutral pH, H98 exhibits the typical pattern of the N ϵ tautomer. H83, instead, displays a pattern consistent with an equilibrium between N ϵ and N δ tautomers. The downfield shift observed for H ϵ_1 and H δ_2 protons suggests that a minor amount of protonated form is present. At pH 5.1 H83 is essentially protonated while at the same pH the resonance pattern displayed by H98 is only slightly changed with respect to neutral conditions. These data reflect a different pK_a for the two histidines. To determine the acid dissociation constants for the histidines, the signals of C ϵ_1 -H ϵ_1 in ^1H , ^{13}C -HSQC spectra were monitored as a function of pH [Fig. 5(c)]. By plotting chemical shift values against pH, the curves obtained for H83 appear monophasic [Fig. 5(d)], and allow the calculation of a pK_a equal to 6.0. The titration curve for H98 is not complete because of the occurrence of protein unfolding at pH lower than 4.5, but a pK_a of 4.8–4.5 could be determined. Long-range ^1H , ^{15}N -HSQC spectra were also recorded on the H98Q mutant at different pH values [Fig. 5(e,f)], indicating that the pK_a of H83 remains unaltered following mutation of H98.

pH-dependent affinity

Ligand binding was previously shown to stabilize the neutral state of H98 in WT cL-BABP, displaying a pH-independent chemical shift down to pH 4. Given that clear spectral changes should be seen at a pH of $\text{pK}_a \pm 1$, the pK_a of H98 in the holo protein is expected to be lower than 3. Examples are reported in the literature where histidine residues span more than 8 log units of pK_a values from <2.8 to >10.4 .³¹ By thermodynamic linkage, the decrease in the pK_a value of H98 upon ligand binding necessitates that the contribution by H98 toward substrate binding will be favoured with an increase in pH. We therefore investigated the behavior of the protein adduct upon changing pH, and compared the data with those obtained for the adduct of the mutant.


Figure 5

Histidine protonation equilibria and pH-dependent ligand binding. (a, b, e, f) ^1H , ^{15}N long-range HSQC spectra of (a) WT cL-BABP at pH 5.1, (b) WT cL-BABP at pH 7.1, (e) H98Q cL-BABP at pH 5.6, (f) H98Q cL-BABP at pH 7.0. The signals of H98 (where present) are connected by continuous lines, those of H83 by dotted lines. (c) Superimposed portion of ^1H , ^{13}C -HSQC spectra, optimized for observation of histidines, showing the signals of H98 (lower set) and H83 (upper set) of WT cL-BABP on decreasing pH at the following values (high- to lowfield): 7.9, 7.5, 7.1, 6.7, 6.2, 5.9, 5.5, 5.0, 4.7, 4.1. (d) Titration data obtained by plotting the shift position of the signals from spectra in (c) versus pH; the curves represent the least-squares fit of the experimental data. (g) one-dimensional projections of ^1H , ^{15}N -HSQC spectra recorded on a ^{15}N -GCDA / WT cL-BABP sample ($P:L=1:2$) at the indicated pH values. (h) High-field region of 1D spectra recorded in the same conditions as in (g). (i) 1D projections of ^1H , ^{15}N -HSQC spectra recorded on a ^{15}N -GCDA/H98Q cL-BABP sample ($P:L=1:2$) at the indicated pH values. (j) High-field region of 1d spectra recorded in the same conditions as in (i).

**Figure 6**

Heteronuclear relaxation analysis upon mutation, pH change, and binding. ^{15}N amide transverse to longitudinal relaxation rates ratios (R_2/R_1) are plotted as a function of residue number for (a) dark green squares: WT cL-BABP pH 7.0, empty circles: H98Q cL-BABP; (b) blue squares: H98Q cL-BABP pH 5.6, empty circles: H98Q cL-BABP pH 7.0. (c) red squares: holo H98Q cL-BABP pH 7.0, empty circles: apo H98Q cL-BABP pH 7.0. (d) Mapping of the largest R_2/R_1 values (red>orange) observed for WT cL-BABP at pH 7.0. (e) Mapping of the largest R_2/R_1 values (red>orange) observed for H98Q cL-BABP at pH 7.0.

Figure 5(g,h) shows that the intensities of resonances from ligands bound to WT decrease on lowering the pH. At pH 3.4 they start to disappear, but this event is not accompanied by an appearance of the signal from the free ligand. Thus, the observed behavior can be explained either by an increased dissociation kinetics (lower affinity at low pH), or by the onset of less specific interactions which again imply lower binding affinity. As far as the results on the mutant are concerned, the intensity decrease of bound ligands upon lowering pH is more dramatic [Fig. 5(i,j)]. These data confirm the critical role of H98 in stabilizing the complex in neutral, but even more in acidic solution, in agreement with a functional need of the system to buffer free bile salts in a variety of solution conditions.

Mapping the binding sites and the protein conformational changes

In addition to the analysis of the difference in protein resonances between unbound and fully bound forms, complete HSQC titrations were performed by adding unlabelled ligand to ^{15}N -labelled proteins, to monitor the changes along the whole binding coordinate. In the case of WT, ligand binding determines chemical exchange which is slow in the NMR time scale for the majority of

analysed residues (≈ 60 residues). A generalized faster exchange has been detected for the H98Q mutant. Signals of the great majority of the analysed residues appeared in the slow-to-intermediate exchange regime. These differences imply larger k_{off} and dissociation constant for the mutant protein.

Protein dynamics

Protein backbone dynamics was investigated by standard amide ^{15}N relaxation measurements. In a previous study, a thorough model-free analysis was performed on the wild type protein.⁸ A number of residues were shown to be characterized by enhanced mobility in the μs -ms time scale, reflected also in larger ratios of transverse to longitudinal heteronuclear relaxation rates (R_2/R_1). These residues are located in strands G, H, I, that include H98, as well as in strands D, E at the opposite side of the β -clam. Substitution of histidine with glutamine attenuates protein mobility in the GHI region, and quenches motions in the DE patch [Fig. 6(a,d,e)]. It was noted earlier that the dynamics of WT cL-BABP is enhanced on lowering the pH. This same behavior was here assessed for the mutant [Fig. 6(b)]. Finally, the addition of bile acids was shown to quench the mobility of the carboxy-terminus in the WT protein⁸ and in the H98Q mutant

[Fig. 6(c)]. Interestingly, ligand binding caused very localized but significant dynamic changes in the N-terminal part of H98Q cLBABP which are likely to be ascribed to conformational exchange. The latter effect could not be assessed for the WT protein due to spectral overlap of the corresponding residues 7 and 9, however some useful observation can be derived both from the structural comparison between apo and holo WT protein and from the results of molecular dynamics simulations.³² It appears that an hydrogen bond network involving residue 9 is lost upon binding and that strand A (N-terminus) is the only strand not making contact with the ligand. Thus, as a result of binding, the residual flexibility is expected to be conveyed to the N-terminus.

DISCUSSION

cL-BABP shows the ability to bind two bile salt molecules with high affinity, as deduced from the combined analysis of titration experiments, diffusion data, and mass spectrometry. We were able to perform a consistent analysis of the WT protein, estimating a microscopic cooperativity factor of 10^4 (corresponding to a free energy change of -23 kJ mol^{-1} at $T = 298 \text{ K}$), to be compared with the value of 10^3 (-17 kJ mol^{-1} at $T = 283 \text{ K}$) reported for the structurally similar human ileal BABP.^{9,11} These figures place both these members of the BABP protein family among the examples of proteins displaying very high positive cooperativity.

A variety of experiments performed on the H98Q mutant are consistent with significant decreased affinity. Indeed, all monitored NMR signals of both the ligands and the protein appeared in slow exchange regime in the case of WT cL-BABP, while several line broadening effects were observed in the case of the mutant, implying faster processes and reduced affinities. Particularly dramatic is the increase in line width of the signal corresponding to ligand bound to site 1 of H98Q cL-BABP when the total ligand concentration exceeds the total protein concentration. Because of the attendant invariance of the signal line width for ligand in site 2, it can be excluded that the observed broadening is due to exchange between the two sites, while on the basis of the diffusion measurements it is clear that chemical exchange with the unbound form is significant. An additional contribution to the increase of line width of the ligand in site 1 could be due to the exchange among multiple conformations implying a loss of specificity and, possibly, of affinity. On the other hand, the measured diffusion coefficients suggest that the affinity of GCDA to site 2 has remained practically unchanged. Finally, mass spectrometry data show that in equimolar protein-ligand mixtures the singly bound species is remarkably more populated in the mutant with respect to the WT protein, indicating that the second binding event becomes less favourable. It

can therefore be concluded that cooperativity of binding is dramatically reduced in the investigated mutant protein.

Several experiments have been performed to study the structural basis for the binding cooperativity in the WT protein and the change in affinity of the H98Q mutant. The titration data reporting the changes in protein signals at increasing ligand concentration show peak perturbations for a large number of residues. A cluster analysis was performed to attempt a rationalization of the observed changes. To this purpose we monitored chemical shift variations, although presented with a slow or slow-to-intermediate exchange regime, analyzing the changes of the most prevalent species (apo or holo). This approach did not aim at interpreting the titration curves in terms of binding models, but was focused on the identification of behavioral classes of residues. This analysis, when performed on the WT protein, did not produce clusters attributable to specific regions of the protein, thus confirming that binding is accompanied by a global structural rearrangement which suggests an allosteric mechanism at the basis of the observed cooperativity. On the contrary, the same analysis revealed for H98Q cL-BABP the existence of two populated clusters (Fig. 7), mainly composed of residues which could be localized close to site 1 (cluster 1) or to site 2 (cluster 2) (based on the analysis of the WT structure). Cluster 2 appears to be well defined, and the majority of residues belonging to it exhibit a very similar chemical shift transition upon binding. On the contrary, cluster 1 is more heterogeneous, and only one-third of its components is represented by residues making contact with site 1. This unbiased statistical analysis roughly distinguishes between the two sites and reflects some properties of site 1 detected by the NMR experiments. Indeed, site 1 has been reported to exhibit conformational multiplicity and exchange contributions with other species. The clusterization observed for the mutant protein is another proof that modification of H98 promotes the disruption of the concerted structural rearrangement of residues observed for the wild type protein, consistent with a cooperative binding model.

Although the binding sites appear essentially preserved after modification of the buried histidine, the structural perturbations observed at the level of strands D, E and loop DE contribute to an altered binding capability. Because this region is distant from the position of H98, its involvement in establishing cooperativity implies the existence of a communication network. In this line, the comparison of the dynamic properties is informative. ¹⁵N relaxation data of WT and H98Q cL-BABP show a substantial attenuation of motions upon mutation at the level of loops DE, EF and strands E and F, that is the regions where common coordinated motions were previously detected in WT protein by essential dynamics.³² These results confirm the role of H98 in triggering a con-

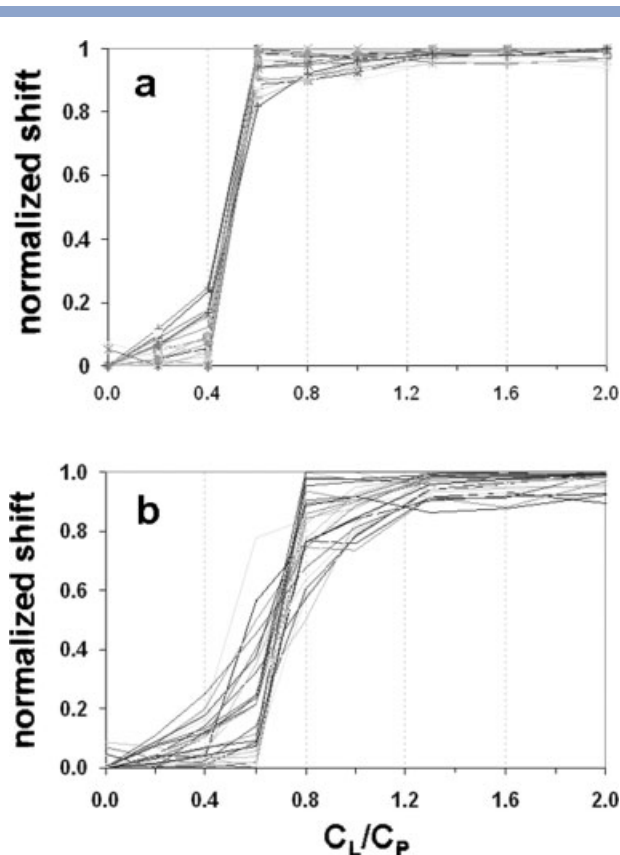


Figure 7

Cluster analysis for the binding of GCDA to ^{15}N -labelled H98Q cL-BABP. Proton amide normalized chemical shift changes are plotted as a function of the total ligand: protein ratio (C_L/C_P). (a) residues in cluster 1 are: 13, 23, 31, 51, 58, 61, 62, 64, 69, 70, 72, 75, 78, 79, 85, 89, 92, 95, 96, 97, 99, 101, 111, 119, 123. (b) residues in cluster 2 are: 2, 5, 7, 8, 9, 10, 12, 14, 16, 17, 18, 19, 20, 27, 28, 29, 30, 32, 34, 38, 42, 50, 52, 60, 63, 71, 74, 102, 103, 112, 121, 124, 125.

formational variability functional to allosteric binding, as previously suggested.

Because histidine protonation equilibria are often at the basis of functional changes in proteins,³⁰ we have also considered the effect of protonation in modulating backbone dynamics. On the basis of the measured acid dissociation constants of the two histidines ($\text{p}K_a^{\text{H}83} = 6.0$, $\text{p}K_a^{\text{H}98} \sim 4.5$), previously observed changes in cL-BABP mobility between pH 7 and 5⁸ were related to protonation equilibria affecting either residues. The observation of an analogous trend in mobility for the H98Q mutant excludes an additional role of H98 in terms of determining pH-dependent protein dynamics.

In summary, the observed decrease in binding efficiency of cL-BABP after removal of the internal histidine is accompanied by localized structural perturbations and quenching of motions at distant sites, proving that a concerted conformational variability in remote regions of the unbound protein favours cooperative binding. The

decreased affinity for GCDA at site 1 could be related to the essential role of the histidine in providing an ideal hydrophobic environment for the binding of two ligands, with a pronounced stabilising effect of the bound species at lower pH. It appears that the deprotonated neutral form of H98 is necessary, together with the nearby I111, to establish an hydrophobic bridge linking the two bile salt molecules and favouring the binding cooperativity.

The experimental evidence of highly cooperative systems both in the hepatocytes and enterocytes supports the validity of a parallel mode of action of the two proteins in the different tissues, in the frame of the enterohepatic circulation.

CONCLUSION

NMR spectroscopy provides a tool to investigate binding energetics and mechanisms in a site-specific manner, as evidenced by the described ligand-based and protein-based experiments. The investigated system appears non-trivial due to the proximity of two strongly coupled binding sites in the same protein cavity. However, through the comparative analysis of structural and dynamic features of the wild type protein and the H98Q mutant, a buried histidine residue was shown to be a determinant of efficient multisite binding of bile acids, allowing the protein to sample a wider conformational space and stabilising the bound state. The presented data led to the possibility of selecting an appropriate binding model for L-BABP, for which the thermodynamic parameters could be evaluated. Further investigations on amino acid side-chain dynamics will address additional details of this cooperative binding mechanism. The knowledge of the molecular basis of interaction is of value not only for the understanding of physiological mechanisms but is also crucial for an appropriate design of drug delivery in the hepatobiliary system.

ACKNOWLEDGEMENT

The authors are indebted to Serena Zanzoni for help in sample preparation.

REFERENCES

1. Mi LZ, Devarakonda S, Harp JM, Han Q, Pellicciari R, Willson TM, Khorasanizadeh S, Rastinejad F. Structural basis for bile acid binding and activation of the nuclear receptor FXR. *Molecular cell* 2003;11:1093–1100.
2. Alrefai WA, Gill RK. Bile acid transporters: structure, function, regulation and pathophysiological implications. *Pharmaceutical Res* 2007;24:1803–1823.
3. Chawla A, Saez E, Evans RM. Don't know much bile-ology. *Cell* 2000;103:1–4.
4. Thomas C, Pellicciari R, Pruzanski M, Auwerx J, Schoonjans K. Targeting bile-acid signalling for metabolic diseases. *Nature Rev* 2008;7:678–693.
5. Guariento M, Raimondo D, Assfalg M, Zanzoni S, Pesente P, Ragona L, Tramontano A, Molinari H. Identification and functional

- characterization of the bile acid transport proteins in non-mammalian ileum and mammalian liver. *Proteins* 2008;70:462–472.
6. Eliseo T, Ragona L, Catalano M, Assfalg M, Paci M, Zetta L, Molinari H, Cicero DO. Structural and dynamic determinants of ligand binding in the ternary complex of chicken liver bile acid binding protein with two bile salts revealed by NMR. *Biochemistry* 2007;46:12557–12567.
 7. Tomaselli S, Ragona L, Zetta L, Assfalg M, Ferranti P, Longhi R, Bonvin AM, Molinari H. NMR-based modeling and binding studies of a ternary complex between chicken liver bile acid binding protein and bile acids. *Proteins* 2007;69:177–191.
 8. Ragona L, Catalano M, Luppi M, Cicero D, Eliseo T, Foote J, Fogolari F, Zetta L, Molinari H. NMR dynamic studies suggest that allosteric activation regulates ligand binding in chicken liver bile acid-binding protein. *J Biol Chem* 2006;281:9697–9709.
 9. Tochtrop GP, Richter K, Tang C, Toner JJ, Covey DF, Cistola DP. Energetics by NMR: site-specific binding in a positively cooperative system. *Proc Natl Acad Sci USA* 2002;99:1847–1852.
 10. Toke O, Monsey JD, DeKoster GT, Tochtrop GP, Tang C, Cistola DP. Determinants of cooperativity and site selectivity in human ileal bile acid binding protein. *Biochemistry* 2006;45:727–737.
 11. Toke O, Monsey JD, Cistola DP. Kinetic mechanism of ligand binding in human ileal bile acid binding protein as determined by stopped-flow fluorescence analysis. *Biochemistry* 2007;46:5427–5436.
 12. Chaudhury S, Gray JJ. Conformer selection and induced fit in flexible backbone protein-protein docking using computational and NMR ensembles. *J Mol Biol* 2008;381:1068–1087.
 13. Hu H, Sheehan JH, Chazin WJ. The mode of action of centrins. Binding of Ca^{2+} and a peptide fragment of Kar1p to the C-terminal domain. *J Biol Chem* 2004;279:50895–50903.
 14. Linse S, Chazin WJ. Quantitative measurements of the cooperativity in an EF-hand protein with sequential calcium binding. *Protein Sci* 1995;4:1038–1044.
 15. Wang Y, Wishart DS. A simple method to adjust inconsistently referenced ^{13}C and ^{15}N chemical shift assignments of proteins. *J Biomol NMR* 2005;31:143–148.
 16. Schanda P, Kupce E, Brutscher B. SOFAST-HMQC experiments for recording two-dimensional heteronuclear correlation spectra of proteins within a few seconds. *J Biomol NMR* 2005;33:199–211.
 17. Clarke J, Fersht AR. An evaluation of the use of hydrogen exchange at equilibrium to probe intermediates on the protein folding pathway. *Fold Des* 1996;1:243–254.
 18. Fushman D, Cahill S, Cowburn D. The main-chain dynamics of the dynamin pleckstrin homology (PH) domain in solution: analysis of ^{15}N relaxation with monomer/dimer equilibration. *J Mol Biol* 1997;266:173–194.
 19. Li Y, Kim S, Brodsky B, Baum J. Identification of partially disordered peptide intermediates through residue-specific NMR diffusion measurements. *J Am Chem Soc* 2005;127:10490–10491.
 20. Vasile F, Ragona L, Catalano M, Zetta L, Perduca M, Monaco H, Molinari H. Solution structure of chicken liver basic fatty acid binding protein. *J Biomol NMR* 2003;25:157–160.
 21. Gomes JA, Mallion RB. Aromaticity and ring currents. *Chem Rev* 2001;101:1349–1383.
 22. D'Onofrio M, Ragona L, Fessas D, Signorelli M, Ugolini R, Pedo M, Assfalg M, Molinari H. NMR unfolding studies on a liver bile acid binding protein reveal a global two-state unfolding and localized singular behaviors. *Archives Biochem Biophys* 2009;481:21–29.
 23. Bruix M, Ribo M, Benito A, Laurents DV, Rico M, Vilanova M. Destabilizing mutations alter the hydrogen exchange mechanism in ribonuclease A. *Biophys J* 2008;94:2297–2305.
 24. Huang JR, Craggs TD, Christodoulou J, Jackson SE. Stable intermediate states and high energy barriers in the unfolding of GFP. *J Mol Biol* 2007;370:356–371.
 25. Schanda P, Forge V, Brutscher B. Protein folding and unfolding studied at atomic resolution by fast two-dimensional NMR spectroscopy. *Proc Natl Acad Sci USA* 2007;104:11257–11262.
 26. Gal M, Schanda P, Brutscher B, Frydman L. UltraSOFAST HMQC NMR and the repetitive acquisition of 2D protein spectra at Hz rates. *J Am Chem Soc* 2007;129:1372–1377.
 27. Huyghues-Despointes BM, Pace CN, Englander SW, Scholtz JM. Measuring the conformational stability of a protein by hydrogen exchange. *Methods Mol Biol* 2001;168:69–92.
 28. Hsu EW, Mori S. Analytical expressions for the NMR apparent diffusion coefficients in an anisotropic system and a simplified method for determining fiber orientation. *Magn Reson Med* 1995;34:194–200.
 29. Buevich AV, Baum J. Residue-specific real-time NMR diffusion experiments define the association states of proteins during folding. *J Am Chem Soc* 2002;124:7156–7162.
 30. Hass MA, Hansen DF, Christensen HE, Led JJ, Kay LE. Characterization of conformational exchange of a histidine side chain: protonation, rotamerization, and tautomerization of His61 in plastocyanin from *Anabaena variabilis*. *J Am Chem Soc* 2008;130:8460–8470.
 31. Schubert M, Poon DK, Wicki J, Tarling CA, Kwan EM, Nielsen JE, Withers SG, McIntosh LP. Probing electrostatic interactions along the reaction pathway of a glycoside hydrolase: histidine characterization by NMR spectroscopy. *Biochemistry* 2007;46(25):7383–7395.
 32. Eberini I, Rocco AG, Ientile AR, Baptista AM, Gianazza E, Tomaselli S, Molinari H, Ragona L. Conformational and dynamics changes induced by bile acids binding to chicken liver bile acid binding protein. *Proteins* 2008;71:1889–1898.

Supplementary material

Table 1. (Summary of k_{ex} , protection factors (P) and ΔG_{op} derived from hydrogen exchange SOFAST NMR experiments at p ³ H 7.4. Hydrogen bonds involving the listed residues are reported).							
Residue	k_{ex} (min ⁻¹)		P (x10 ³)		ΔG_{op} (kcal mol ⁻¹)		H-bond (NH-OC)
	WT	H98Q	WT	H98Q	WT	H98Q	
W6	8.70±0.2E-03	11.6±0.3E-03	29.06	21.796	6.09	5.92	W6-V38
Q7	5.50±0.2E-03	9.2±0.9E-03	66.45	39.723	6.58	6.27	Q7-K123
V8	5.98±0.29E-02		2.16		4.55		V8-P36
Y9	1.36±0.12E-03		117.3		6.91		Y9-R121
A10	4.89±0.21E-02		9.41		5.42		A10-R121
F17	2.93±0.06E-02	5.2±0.2E-02	5.8	3.267	5.13	4.79	F17-N13
L18	2.23±0.04E-02	3.98±0.1E-02	5.64	3.161	5.11	4.77	L18-Y14
K19	5.20±0.16E-02		4.43		4.97		K19-E15
A20	1.17±0.15E-01		4.61		4.99		A20-E16
L21	6.01±0.19E-02		1.81		4.44		L21-F17
V38	2.80±0.1E-03	4.6±0.2E-03	17.13	10.427	5.77	5.48	V38-W6
E39	6.70±0.2E-03	9.4±0.1E-03	13.89	9.899	5.65	5.45	E39-T50
I40	7.60±0.2E-03	9.0±0.1E-03	7.16	6.050	5.26	5.16	I40-G4
Q41		19.3±0.6E-03		14.364		5.67	Q41-N48
D46	2.11±0.4E-02	2.45±0.4E-02	6.53	5.624	5.20	5.11	D46-K43
F47	1.77±0.3E-02	1.64±0.4E-02	8.94	9.648	5.39	5.43	F47-F62
V48	8.4±0.2E-03	10.1±0.1E-03	11.21	9.326	5.52	5.41	V48-Q41
V49	8.60±0.2E-03	13.4±0.3E-03	6.86	4.403	5.23	4.97	V49-N60
T50	3.63±0.18E-02		7.03		5.24		T50-E39
F62	4.65±0.37E-02	5.15±0.3E-02	10.21	9.220	5.47	5.41	F62-F47
L64	3.43±0.2E-02	3.61±0.2E-02	5.03	4.776	5.05	5.02	L64-D45
I70		25.9±0.5E-03		1.958		4.49	I70-L78
L78	2.71±0.06E-02	3.83±0.1E-02	5.29	3.744	5.08	4.87	L78-I70
V82		6.8±0.3E-02		1.898		4.47	
H83	2.35±0.09E-02		30.69		6.12		H83-V90
L89	2.43±0.1E-02	3.36±0.1E-02	5.9	4.268	5.14	4.95	L89-Q100
T91	3.73±0.29E-02	3.92±0.3E-02	6.84	6.506	5.23	5.20	T91-H98
E99		3.28±0.2E-02		6.207		5.17	E99-T110
Q100		3.32±0.2E-02		10.112		5.46	Q100-L89
E101	1.45±0.04E-02		13.79		5.65		E101-V108
M107	4.60±0.2E-03		62.27		6.54		M107-S122
V108	1.10±0.1E-03	4.4±0.3E-03	95.39	23.848	6.79	5.97	V108-E101
E109		45±0.5E-04		20.679		5.88	E109-R120
T110	3.48±0.1E-04	43±0.3E-04	721.37	58.381	7.99	6.50	T110-E99
I111	1.53±0.1E-04	20±0.3E-04	790.94	60.507	8.08	6.52	I111-L118
T112	7.50±0.2E-03	25.1±0.3E-03	27.64	8.259	6.06	5.34	T112-S97
F113	1.88±0.06E-02		20.08		5.87		F113-V116
L118	6.03±1.92E-04	41±0.4E-04	285.93	42.053	7.44	6.30	L118-I111
I119	4.33±0.17E-02	5.31±0.2E-02	1.09	0.887	4.14	4.02	I119-E12
R120	5.19±0.25E-04	62±0.5E-04	555.11	46.468	7.83	6.36	R120-E109
S122	1.69±0.04E-02	2.23±0.1E-02	94.39	71.537	6.78	6.62	S122-M107
K123	2.75±0.14E-02	2.33±0.1E-02	27.11	31.991	6.0	6.14	K123-Q7
V125	3.17±0.06E-02	3.01±0.5E-02	0.07	0.071	2.49	2.52	V125-T5

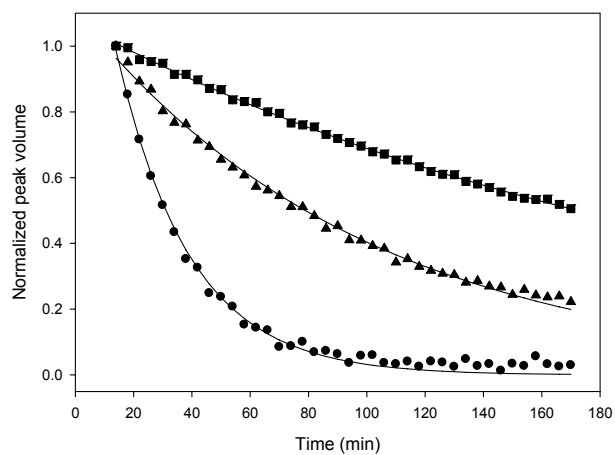


Figure 1. Representative examples of exchange kinetics of three residues (L18 ●, V48 ▲, V108 ■) obtained from the analysis of HX from SOFAST NMR experiments performed at 600 MHz on a 0.4 mM sample of H98Q cL-BABP, p²H 7.5.

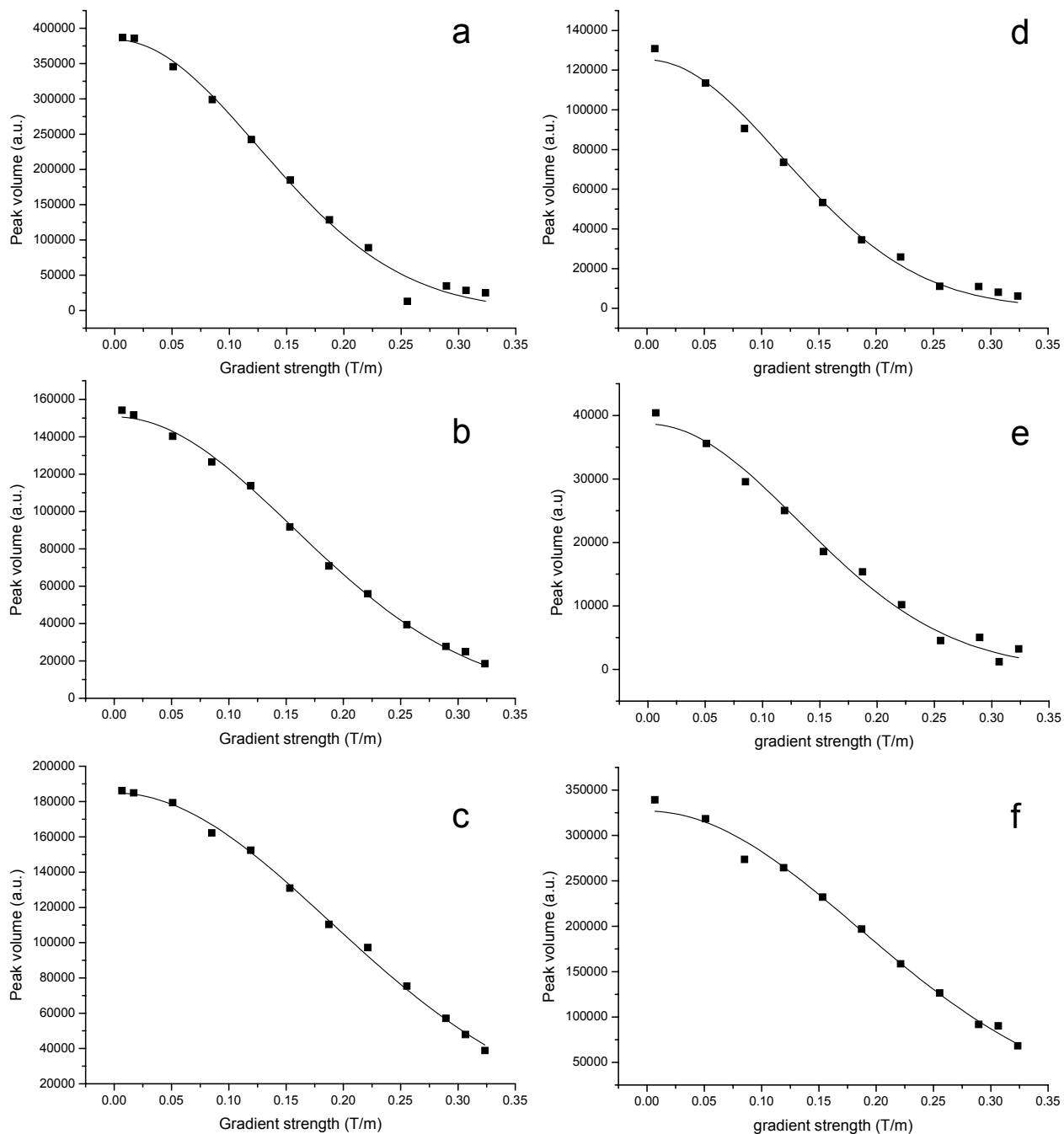


Figure 2. Diffusion measurements. The data represent the gradient-dependent intensity decay of the STE-HSQC signals measured on a sample containing (a-c) WT cL-BABP:¹⁵N-GCDA 1:3 and (d-f) H98Q cL-BABP:¹⁵N-GCDA 1:3. a, d) Unbound ligand. b, e) Ligand bound to site 1. c, f) Ligand bound to site 2. The pH was 7.2 and the measurement temperature 298 K. The abscissae represent the effective applied gradient strength, corrected for the sine shape.

CHAPTER 6: Publication 3

Title: Disulfide bridge regulates ligand-binding site selectivity in liver bile acid-binding proteins.

Short presentation of publication 3:

Bile acids are very important molecules that helps the absorption of lipids and lipid soluble vitamins in the intestine. Bile acids recently received a renewed attention due to the discovery of their signaling functions and their implication in pathogenesis of several diseases. Bile acids undergo a recirculation pathway between the liver and the intestine through a mechanism called “enterohepatic circulation”. In this process the proteins belonging to the bile acid-binding protein (BABP) family play a crucial role as cytosolic transporters of bile acids.

The bile acid binding ability of some mammalian and non mammalian BABPs have been investigated and three-dimensional structures of apo and holo forms have been resolved both by NMR spectroscopy and X-Ray diffraction techniques. The structural basis of ligand binding have also been studied and, from literature data, it appears that the most characterized BABP is that from human intestine (hI-BABP). Non mammalian liver BABPs present binding similarities with intestinal BABPs. In most non mammalian liver BABPs a disulfide bridge is present between the two cysteines at positions 80 and 91, while two forms are known for chicken liver BABP (cL-BABP), in which a threonine or a cysteine can be present at position 91. The threonine carrying form, devoid of disulphide bridge, has been extensively characterized both from the structural and the binding points of view.

The form carrying cysteine, named cL-BABP/S-S, was produced and further characterised through NMR spectroscopy and mass spectrometry experiments with the aim of establishing the determinants of binding stoichiometry and site preferences. The results are reported in the following paper.

The DNA sequence coding for cL-BABP/S-S protein was obtained by site specific mutagenesis, starting from a plasmid coding for the protein carrying the threonine at position 91. The protein of interest was expressed and purified by classical recombinant protein production protocols.

The obtained protein was analyzed for the bile acid binding capacity by titrating an unlabeled protein sample with ^{15}N -labeled bile acids and recording ^1H , ^{15}N -HSQC experiments. Glycochenodeoxycholic (GCDA) and glycocholic (GCA) acids were used and for both of them a stoichiometry of 2 bile acid molecules for protein molecule was identified. The binding stoichiometry was also confirmed by mass spectrometry measurements on a sample of cL-BABP/S-S complexed with GCDA. In the GCDA titration experiments other signals were visible at frequencies close to both site 1 or site 2 resonances, while in the case of GCA only an additional peak was visible close to site 1 chemical shift. All these additional resonances were attributed to heterogeneous binding of the ligands.

Exchange phenomena were also analyzed employing a sample containing a protein:bile acid ratio 1:3. The self-diffusion coefficients (D) were measured by NMR spectroscopy and by comparison with the data previously obtained for cL-BABP, it was clear that relevant exchange processes were present

between ligand at site 1 and the free ligand, while ligand at site 2 was not apparently involved in exchange processes.

Bile acid binding preferences were also analyzed through NMR competition experiments. Previous observation indicated that the cL-BABP does not exhibit site selectivity for GCDA and GCA. Competition experiments were performed by adding the unlabeled competitor to a sample of unlabeled cL-BABP/S-S protein complexed with ^{15}N -labeled ligand. These experiments pointed out the preference of GCDA for site 1 and GCA for site 2 and evidenced a higher affinity of cL-BABP/S-S for GCDA.

The effect of the introduction of a disulfide bridge was also investigated by evaluating the NH chemical shift changes in cL-BABP/S-S with respect to cL-BABP. The complete resonance assignment of cL-BABP/S-S was achieved by recording and analyzing three-dimensional ^1H , ^{15}N -TOCSY- and ^1H , ^{15}N -NOESY-HSQC experiments. The residues mostly affected by the introduction of the disulfide bridge are located in the E, F, G and H strands, close to the substitution site. The complete assignment of the complex cL-BABP/S-S -CDA was also obtained. The comparison of the chemical shift changes between apo and holo cL-BABP/S-S indicated that the regions mostly affected by ligand binding are located in the C-terminal part of the protein.

Measurement of ^{15}N T_1 and T_2 relaxation was performed with the aim to evaluate the impact of the disulfide bridge on the protein mobility. The main effect on the apo protein was the propagation of motions to the N-terminal region. In cL-BABP/S-S-CDA complex the motions were not quenched by ligand binding, as observed for the protein devoid of disulphide bridge. Enhance motions were instead observed.

The study of cL-BABP/S-S and cL-BABP variants indicated that the presence of the disulfide bridge introduced site selectivity which was mostly affected by protein mobility rather than by steric factors.

My contribution in this work was the development of a protocol for the expression and the purification of the ^{13}C ^{15}N doubly labeled cL-BABP/S-S protein and the production of the protein samples employed in the study.

Disulfide bridge regulates ligand-binding site selectivity in liver bile acid-binding proteins

Clelia Cogliati¹, Simona Tomaselli¹, Michael Assfalg², Massimo Pedò², Pasquale Ferranti³, Lucia Zetta¹, Henriette Molinari² and Laura Ragona¹

¹ Laboratorio NMR, Istituto per lo Studio delle Macromolecole, CNR, Milan, Italy

² Dipartimento di Biotecnologie, Università di Verona Strada le Grazie, Verona, Italy

³ Dipartimento di Scienza degli Alimenti, Università di Napoli Federico II, Portici, Italy

Keywords

backbone dynamics; disulfide bridge; intracellular lipid-binding protein; molecular recognition; NMR

Correspondence

L. Ragona, Lab. NMR, Istituto per lo Studio delle Macromolecole, CNR, Via Bassini, 15, 20133, Milano, Italy

Fax: +39 02 23699620

Tel: +39 02 23699619

E-mail: laura.ragona@ismac.cnr.it

H. Molinari, Dipartimento di Biotecnologie, Università degli Studi di Verona, Strada le Grazie, 15, 37134 Verona, Italy

Fax: +39 0458027929

Tel: +39 0458027901

E-mail: henriette.molinari@univr.it

(Received 3 July 2009, revised 17 August 2009, accepted 18 August 2009)

doi:10.1111/j.1742-4658.2009.07309.x

Bile acid-binding proteins (BABPs) are cytosolic lipid chaperones that play central roles in driving bile flow, as well as in the adaptation to various pathological conditions, contributing to the maintenance of bile acid homeostasis and functional distribution within the cell. Understanding the mode of binding of bile acids with their cytoplasmic transporters is a key issue in providing a model for the mechanism of their transfer from the cytoplasm to the nucleus, for delivery to nuclear receptors. A number of factors have been shown to modulate bile salt selectivity, stoichiometry, and affinity of binding to BABPs, e.g. chemistry of the ligand, protein plasticity and, possibly, the formation of disulfide bridges. Here, the effects of the presence of a naturally occurring disulfide bridge on liver BABP ligand-binding properties and backbone dynamics have been investigated by NMR. Interestingly, the disulfide bridge does not modify the protein-binding stoichiometry, but has a key role in modulating recognition at both sites, inducing site selectivity for glycocholic and glycochenodeoxycholic acid. Protein conformational changes following the introduction of a disulfide bridge are small and located around the inner binding site, whereas significant changes in backbone motions are observed for several residues distributed over the entire protein, both in the apo form and in the holo form. Site selectivity appears, therefore, to be dependent on protein mobility rather than being governed by steric factors. The detected properties further establish a parallelism with the behaviour of human ileal BABP, substantiating the proposal that BABPs have parallel functions in hepatocytes and enterocytes.

Introduction

Bile acids (BAs) are vital components of many biological processes and play an important role in the pathogenesis of numerous common diseases [1], but the specific mechanisms coupling intracellular BAs to biological targets are not well understood. BAs circulate between the liver and intestine through a mechanism

known as 'enterohepatic circulation', which is a tightly regulated process, particularly by BAs themselves. BA-binding proteins (BABPs), belonging to the intracellular lipid-binding protein (iLBP) family, play a vital role in the enterohepatic circulation as cytoplasmic transporters of BAs. Understanding the mechanism

Abbreviations

BA, bile acid; BABP, bile acid-binding protein; CA, cholate; CDA, chenodeoxycholate; CSP, chemical shift perturbation; GCA, glycocholic acid; GCDA, glycochenodeoxycholic acid; I-BABP, human ileal bile acid-binding protein; iLBP, intracellular lipid-binding protein; L-BABP, chicken liver bile acid-binding protein.

nism regulating these interactions is a key step in providing a model for the transfer of BAs from the cytoplasm to the nucleus for delivery to nuclear receptors, and can be used to inspire the design of therapeutic agents for the treatment of metabolic disorders, such as obesity, type 2 diabetes, hyperlipidaemia, and atherosclerosis [1–3].

BABPs are characterized by a conserved β -barrel structure, formed by two orthogonal β -sheets, and a helix–loop–helix motif defining, with flexible loops, the so-called protein open end, delimiting the entrance to the barrel cavity. BABPs from various organisms have been shown to bind bile salts with differences in ligand selectivity, binding affinity, stoichiometry, and binding mechanism. The two most extensively characterized BABPs, namely human ileal BABP (I-BABP) and chicken liver BABP (L-BABP), share the common property of binding two bile salt molecules with weak intrinsic affinities and strong positive cooperativity [4–6]. I-BABP, unlike L-BABP, displays remarkable site selectivity for the two main glycoconjugated BAs, glycocholic acid (GCA) and glycochenodeoxycholic (GCDA). A number of factors have been shown to modulate ligand binding, e.g. the chemistry of the ligand and the nature of the protein residues [7,8]. A prominent role for protein plasticity was suggested for L-BABP, where binding was found to be regulated by a dynamic process and accompanied by a global conformational rearrangement [9]. Essential dynamics analysis of the molecular dynamics trajectories obtained for L-BABP indicated that the portal area is the region mostly affected by complex formation, and that the major concerted motions involve the structural elements of the open end, which are dynamically coupled in different ways, whether in the presence or in the absence of the ligands [10]. Another source of ligand-binding variability may be introduced by the presence of disulfide bridges. Indeed, several cases have been reported in the literature for members of the iLBP family where the introduction/removal of a disulfide bridge was responsible for changes in ligand-binding stoichiometry and affinities. The removal of a disulfide bond in rat lipocalin-type prostaglandin D synthase slightly increased the binding affinity for biological ligands, by leading to a less compact barrel pocket and allowing a higher number of residues to contribute to ligand binding [11]. In the cellular retinoic acid-binding protein I, the introduction of a disulfide bond abolished the structural mobility of the portal region, thus leading to irreversible retinoic acid binding [12].

Most liver BABPs belonging to nonmammalian species have a disulfide bond involving the conserved

Cys80 and the cysteine at position 91. For L-BABP, two forms are known, in which residue 91 can be either a threonine or a cysteine, although all the studies presented up to now have dealt with the form devoid of the disulfide bridge [5,9,13–15]. The presence of a disulfide bridge in the protein scaffold of the homologous liver zebrafish BABP (69.8% identity, calculated with CLUSTALW) was correlated with the binding stoichiometry [16], which varied from one ligand molecule, with a disulfide bridge, to two ligand molecules, with the Cys80–Cys91 disulfide bridge removed. On this basis, in a continuous effort to establish the determinants of binding stoichiometry and site selectivity in this protein family, the T91C L-BABP protein, with a Cys80–Cys91 disulfide bridge, has been studied by different NMR and MS approaches.

The role of the disulfide bridge in ligand binding and the backbone dynamics of L-BABP has been investigated here by combining different labelling strategies for both ligand and protein with appropriate NMR experiments. The results clearly show that, although the binding stoichiometry is conserved, site selectivity for GCDA and GCA, which is not observable in the absence of the disulfide bridge, is now present. Changes in motion propagation within the β -barrel, induced by the disulfide bridge, have been mapped onto the BABP apo structure, and the effects of the binding of the two most abundant glycoconjugated bile salts on the backbone conformation and dynamics have been clearly assessed.

Results

Effect of disulfide bridge on binding properties

Binding site occupancies

$^1\text{H}/^{15}\text{N}$ -HSQC spectra were collected on isotopically enriched physiological glycine conjugates, GCA and GCDA (differing only in the presence of a hydroxyl group at position 12; Fig. S1), complexed with unlabelled T91C L-BABP at different protein/ligand ratios (1 : 0.3, 1 : 0.6, 1 : 1, 1 : 1.5, 1 : 2, 1 : 2.5, and 1 : 3), in order to monitor the number and occupancy of individual binding sites. The spectra obtained for ^{15}N -GCDA revealed the presence of two main resonances, corresponding to ^{15}N -GCDA bound to two distinct binding sites, denoted site 1 (7.17, 117.3 p.p.m.) and site 2 (6.0, 117.5 p.p.m.), whose chemical shifts did not change during the titration, suggesting the presence of a slow exchange regime (Fig. 1A). A few other cross-peaks with chemical shifts very close to those of peak 1 and peak 2 were visible, and were ascribed to heterogeneous binding at site 1

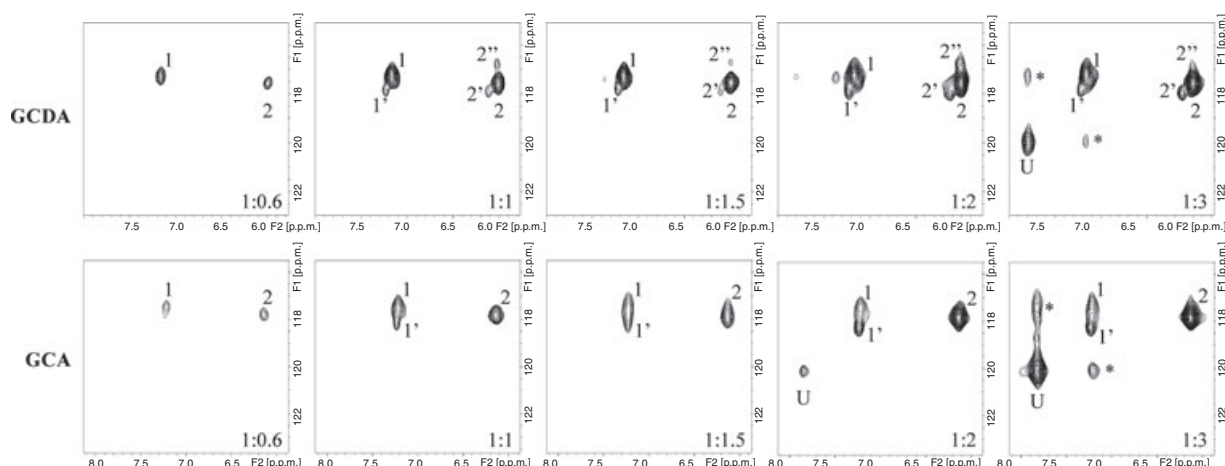


Fig. 1. $[^{15}\text{N}]\text{GCDA}$ and $[^{15}\text{N}]\text{GCA}$ in complex with T91C L-BABP. 2D $^1\text{H}/^{15}\text{N}$ -HSQC spectra at different protein/ligand ratios (1 : 0.6, 1 : 1, 1 : 1.5, 1 : 2, and 1 : 3) were recorded at 298 K at 500 MHz. The resonances corresponding to the unbound ligand and to binding sites 1 and 2 are indicated as U, 1, and 2, respectively. The satellite peaks of site 1 and site 2 are also marked. Asterisks indicate exchange peaks.

and site 2. The unbound resonance (7.8, 119.8 p.p.m.) was visible at protein/ligand ratios higher than 1 : 2, together with exchange peaks between the unbound and site 1 cross-peaks. During the titration, site 1 and site 2 ^1H linewidths were substantially unchanged (Fig. 2A). The quantitative volume analysis of these predominant forms indicated that binding site occupancies reached a plateau value, for both sites, at a protein/ligand ratio of 1 : 2 (Fig. 2B). The NMR data thus indicate that, even in the presence of the disulfide bridge, L-BABP maintains the ability to bind two GCDA molecules, at variance with the homologous zebrafish protein. This result was corroborated by MS analysis of T91C L-BABP in complex with GCDA, indicating the presence of the doubly ligated form in solution at a protein/ligand ratio 1 : 2 (data not shown). Similar NMR results were obtained for GCA, and $^1\text{H}/^{15}\text{N}$ -HSQC NMR titration experiments, performed on the unlabelled T91C L-BABP with increasing amounts of $[^{15}\text{N}]\text{GCA}$, indicated the presence of the three cross-peaks named site 1 (7.2, 117.5 p.p.m.), site 1' (7.2 and 118.0 p.p.m.), and site 2 (6.122, 117.81 p.p.m.) (Fig. 1B). The cross-peak annotated as site 1' was probably due to the presence of slightly different populations of GCA at this site. The resonance corresponding to the unbound ligand became visible at a protein/ligand ratio of 1 : 2 (7.8 and 120.1 p.p.m.) and exhibited exchange cross-peaks with site 1. The chemical shifts of GCA resonances did not change during the titration, whereas for some of them a variation in linewidth was observed (Fig. 2C), suggesting the presence of a slow to intermediate exchange regime. Site 2 and free GCA resonances exhibited a linewidth

decrease upon an increase in protein/ligand ratio. This behaviour is consistent with exchange with free ligand being abolished as saturation is approached [17]. The changes in linewidths did not allow a quantitative determination of site 2 occupancy. The site 1 linewidth (~ 33 Hz), which was broader than that of site 1' (~ 22 Hz), is attributable to exchange with free ligand, as supported by the observation of exchange peaks for site 1 and unbound GCA. Both site 1 and site 1' linewidths did not decrease as saturation was approached, thus confirming the presence of conformational heterogeneities of the bound states at superficial sites.

Detection of ligand exchange phenomena

The temperature dependence of GCDA and GCA resonances was investigated in the range 280–305 K on samples with a protein/ligand ratio of 1 : 3 (Fig. 3A,B). In both cases, a slow exchange regime on the NMR chemical shift time scale was observed for site 2, which exhibited, upon temperature increase, decreased linewidths, reflecting the shorter protein correlation time at higher temperatures. In contrast, site 1 and the unbound resonances exhibited line broadening upon temperature increase, further confirming the involvement of ligand bound to site 1 in exchange phenomena with the free ligand. Interestingly, at all the investigated temperatures, the resonance of the unbound GCA showed a similar linewidth but a higher intensity with respect to GCDA, reflecting a minor overall affinity of GCA for T91C L-BABP.

One alternative way of detecting exchange phenomena between the different species in solution is through

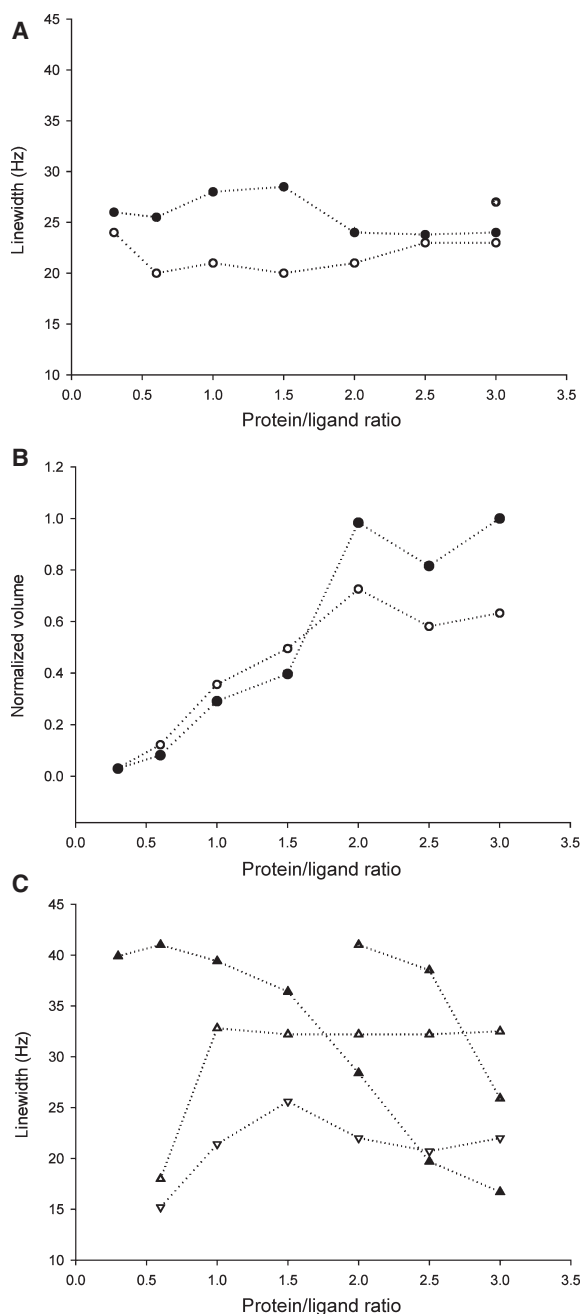


Fig. 2. Analysis of GCA and GCDA resonances at different protein/ligand ratios. Plots of linewidths (A) and volume (B) of amide proton resonances of GCDA as a function of protein/ligand ratio: site 1 (empty circle); site 2 (filled black circle); unbound ligand (filled grey symbols). Plots of linewidths (C) of amide proton resonances of GCA as a function of protein/ligand ratio: site 1 (empty triangle up); site 1' (empty triangle down); site 2 (filled black triangle); unbound ligand (filled grey triangle).

the measurement of the self-diffusion coefficient (*D*) [18]. It is expected that a ligand molecule that is in exchange with the free form will show a *D*-value that

is a linear combination of those of the free ligand and the protein. Diffusion experiments were performed on wild-type and T91C L-BABP complexed with GCDA and GCA at protein/ligand ratios of 1 : 3, and the diffusion coefficients, calculated from the analysis of signal decay as a function of the applied gradient, are reported in Table 1. From comparison of these values with those previously obtained for the free ligand ($3.97 \times 10^{-6} \text{ cm}^2 \text{ s}^{-1}$) and the protein ($1.04 \times 10^{-6} \text{ cm}^2 \text{ s}^{-1}$) [14], it is possible to conclude that exchange processes between bound and free forms are relevant for site 1 and negligible for site 2 for both wild-type and T91C L-BABP. However, in the presence of the Cys80–Cys91 disulfide bridge, the diffusion values of ligand bound to site 1 were lower than those of the wild-type protein, suggesting a higher affinity for both ligands at site 1 of T91C L-BABP.

Site selectivity

Previous observations indicated that wild-type L-BABP did not show any site selectivity for GCDA and GCA, and revealed a higher affinity for GCDA at both sites. T91C L-BABP site selectivity for the two bile salts was investigated in competition experiments, in which unlabelled GCDA was added to a solution containing a T91C L-BABP/[^{15}N]GCA molar ratio of 1 : 2. One-dimensional first increments of the 2D $^1\text{H}/^{15}\text{N}$ correlation spectra for the sample containing an equimolar mixture of [^{15}N]GCA and unlabelled GCDA (Fig. 4A) showed that the peak corresponding to site 2 was sharpened but its intensity was marginally affected by GCDA addition. In contrast, [^{15}N]GCA bound to site 1 was completely displaced by the unlabelled GCDA as its resonance disappeared. This behaviour clearly indicates that the presence of a disulfide bridge had introduced site selectivity. Such an effect was confirmed by the complementary competition experiment, in which the unlabelled GCA was added to a solution containing a T91C L-BABP/[^{15}N]GCDA molar ratio of 1 : 2 (Fig. 4B). In agreement with the selectivity of GCA for site 2, complete disappearance of the resonance of [^{15}N]GCDA at site 2 was expected. However, only a 60% reduction of this resonance intensity was observed, which can be explained by a general overall higher affinity of T91C L-BABP for GCDA than for GCA. Interestingly, the presence of GCA at site 2 favoured one secondary form at a superficial site, characterized by chemical shifts close to the site 1' resonance, previously observed in $^1\text{H}/^{15}\text{N}$ -HSQC spectra of the T91C L-BABP–GCDA complex (Fig. 1). The change of the population at site 1 in

Fig. 3. Stacked plot showing the temperature dependence of the BA amide ^1H resonances in the temperature range 280–305 K. One-dimensional first increment of the 2D $^1\text{H}/^{15}\text{N}$ -HSQC spectra collected on T91C L-BABP- ^{15}N GCDA (left) and T91C L-BABP- ^{15}N GCA (right) complexes, using a protein/ligand molar ratio of 1 : 3.

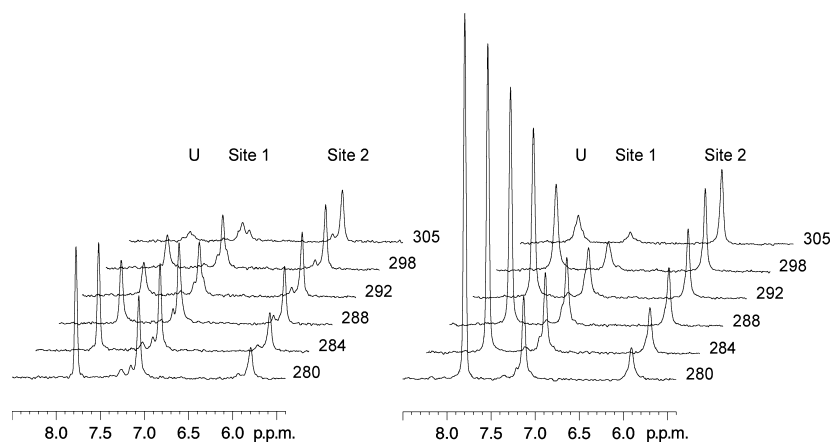


Table 1. Diffusion coefficients of bile salt species. *D*-values measured for free CDA and holo L-BABP are $3.97 \times 10^{-6} \text{ cm}^2 \text{ s}^{-1}$ and $1.04 \times 10^{-6} \text{ cm}^2 \text{ s}^{-1}$, respectively [14]. Errors in *D*-values were estimated to be of the order of $10^{-8} \text{ cm}^2 \text{ s}^{-1}$ from the fitting procedure.

	Site 1 ($\times 10^{-6} \text{ cm}^2 \text{ s}^{-1}$)	Site 2 ($\times 10^{-6} \text{ cm}^2 \text{ s}^{-1}$)
Wild type-GCDA	2.28	1.57
Wild type-GCA	2.43	1.28
T91C-GCDA	2.02	1.45
T91C-GCA	2.20	1.48

favour of a new site 1' suggests that site 1' is the preferred orientation of GCDA at the superficial site when GCA is bound to site 2.

The different intensities exhibited by the unbound GCA and GCDA reflect, as previously observed for wild-type protein, both the lower affinity of the protein for GCA and the onset of different equilibria between monomeric and micellar bile salts. Indeed, the critical micellar concentration of GCDA (2.4 mM) is significantly lower than that of GCA (10 mM) [19], and the broader linewidth of the resonance of unbound GCDA (22 Hz) with respect to that of unbound GCA (15 Hz) can be explained by the equilibrium between free monomeric and micellar GCDA. The comparison of ^1H -spectra of the two protein samples at a protein/GCDA/GCA ratio of 1 : 2 : 2 indicated that the final holo state is independent of the order of addition of the bile salts and supports the results of competition data.

In summary, competition experiments pointed to a site preference of GCDA for site 1 and of GCA for site 2 in T91C L-BABP, together with a higher affinity of the protein for GCDA.

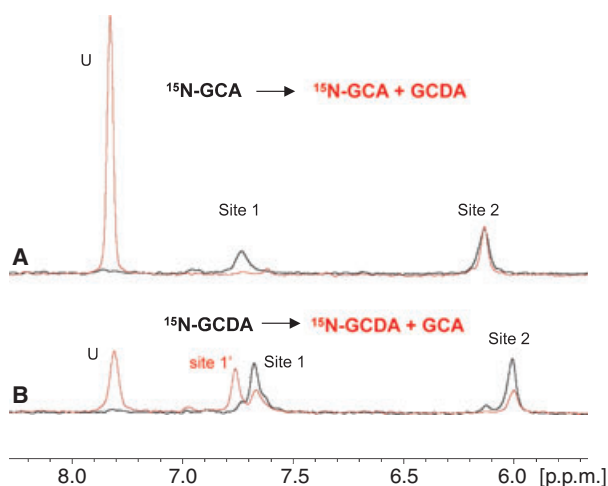


Fig. 4. Bile salt site selectivity experiments. One-dimensional first increment of the 2D $^1\text{H}/^{15}\text{N}$ -HSQC spectra collected on: ^{15}N GCDA in a 1 : 2 T91C L-BABP/GCA molar ratio [(A), black line]; ^{15}N GCDA in the presence of equimolar amounts of unlabelled GCDA (T91C L-BABP/GCA/GCDA molar ratio of 1 : 2 : 2) [(A), red line]; ^{15}N GCDA in a 1 : 2 T91C L-BABP/GCDA molar ratio [(B), black line]; ^{15}N GCDA in the presence of equimolar amounts of unlabelled GCA (T91C L-BABP/GCA/GCDA molar ratio of 1 : 2 : 2) [(B), red line]. The resonances corresponding to the unbound ligand are indicated as U.

Conformational changes induced by disulfide bridge in the apo and holo forms of T91C L-BABP

The effect of the disulfide bond introduction on the structure of the apo protein was investigated by monitoring the $^1\text{H}/^{15}\text{N}$ chemical shifts changes observed in T91C L-BABP with respect to the wild type. The resonance assignment of signals from backbone and side chains atoms of the apo form of T91C L-BABP was performed using standard 3D heteronuclear triple resonance NMR experiments, as described in Experimental

procedures, together with a combination of 2D and 3D TOCSY and NOESY HSQC spectra recorded at pH 5.6 and pH 7.2. The observed shift of some cross-peaks, induced by acidic pH, allowed the assignment of resonances that substantially overlapped at neutral pH. Backbone amide resonance assignment was complete at 93%, and resonances of residues Thr72, Met73, Lys77, Leu78, Asn86, Leu89, Lys95 and Phe96 could not be unequivocally assigned, owing to signal overlap and/or broadening.

The secondary structure of T91C L-BABP is substantially unchanged with respect to the wild-type protein. In particular, the secondary structural elements, as derived with TALOS [20], include 10 antiparallel β -strands and two α -helices in the following regions: 5–8 (strand A), 14–18 (helix 1), 25–29 (helix 2), 34–43 (strand B), 46–53 (strand C), 56–60 (strand D), 66–71 (strand E), 76–85 (strand F), 88–92 (strand G), 96–103 (strand H), 105–113 (strand I), and 116–124 (strand J).

The analysis of chemical shift perturbation (CSP) induced by the introduction of a disulfide bridge showed that the most significant changes occurred at the level of strand E (Ala68, Asp69, and Ile71), strand F (Lys79, Cys80, Thr81, and Leu84), strand G (Ser93), and strand H (His98) (Fig. 5). All of the mentioned residues are in close proximity to the disulfide bridge connecting strand F and strand G, except for Ile71, which is, however, contiguous with the 68–69 region affected by the mutation.

The T91C L-BABP–chenodeoxycholate (CDA) complex was characterized by NMR, and the assignment of backbone amide resonances, performed on a protein/ligand sample of molar ratio 1 : 3, was complete at 95% (missing assignments for Ala1, Gln7, Ile37, Asn86, Gln100, and Asn105). Resonance assignments of apo and holo forms of the protein have been reported in BiomagResBank (accession numbers 16310 and 16309 for the apo and holo proteins, respectively).

Comparison of the chemical shifts of the apo and holo forms of T91C L-BABP indicated that the regions mostly affected by binding are mainly located at the C-terminal part of the protein, at the level of Lys76,

Thr81, and Val82 (strand F), Val90, Lys92, and Ser93 (strand G), Glu94 (loop GH), Phe96, Ser97, and His98 (strand H), together with a few residues in the N-terminal region, namely Arg32 (helix II), Thr57 (strand D), and Glu67 (strand E) (Fig. 6). Comparison of CSP induced by complex formation in wild-type and T91C L-BABP (Fig. 6A) indicated that the same protein regions are affected by ligand binding, confirming a conserved binding mode. A few differences were, however, observed for some residues gathered around the ligand bound at site 2 (Fig. 6B), closer to the disulfide bridge.

Backbone dynamics of apo and holo forms of T91C L-BABP

Backbone dynamics were investigated for the apo and holo forms of T91C L-BABP to assess the relevance of backbone motions to ligand-binding properties.

^{15}N T_1 and T_2 relaxation values were calculated for the apo form of T91C L-BABP, and several residues, namely Arg32, Lys52, Phe62, Thr71, Asp74, Cys91, Lys92, Glu94, Ser97, His98, Gln100, Gly104, Glu109, Ile111 and Gly115, showed high T_1/T_2 ratios, indicative of conformational exchange processes on the microsecond and millisecond time scales (Fig. 8). Interestingly, the introduction of the new disulfide bond, connecting strand F and strand G, did not reduce conformational motions, which, on the contrary, were extended to the N-terminal regions of the protein, as a result of changes in motion propagation, within the β -barrel (Fig. S2).

The relaxation experiments were also performed on a holo T91C L-BABP–CDA sample at a protein/ligand ratio of 1 : 3. In these conditions, the protein is substantially saturated and a negligible population of the free protein is present, as derived from the analysis of titration experiments performed on the ^{15}N -labelled protein (data not shown). As a consequence, the detected exchange contribution can be related to protein conformational motions rather than to free-bound exchange. Analysis of T_1/T_2 ratios

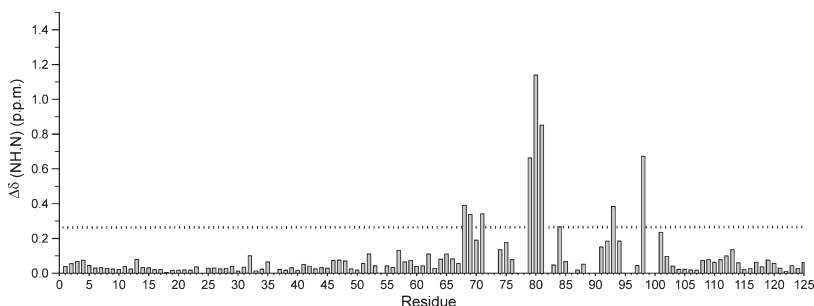


Fig. 5. CSP upon disulfide bridge introduction. Chemical shift differences between apo T91C L-BABP and wild-type (WT) L-BABP, at pH 7 and 298 K, calculated as $\Delta\delta(\text{HN},\text{N}) = [(\Delta\delta\text{HN}(\text{T91C} - \text{WT})^2 + \Delta\delta\text{N}(\text{T91C} - \text{WT})^2/25)/2]^{1/2}$ are plotted versus residue number. The dotted line corresponds to the mean value plus one standard deviation.

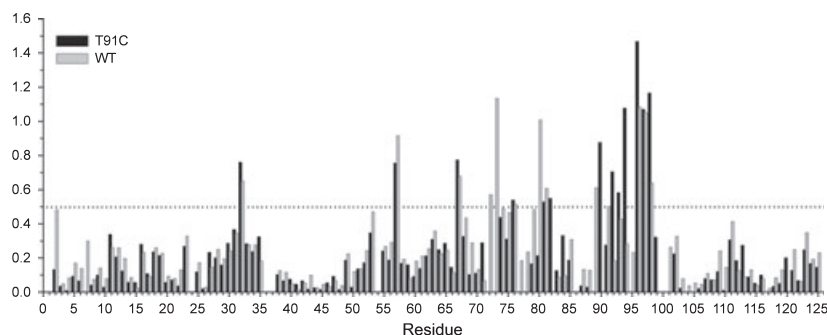
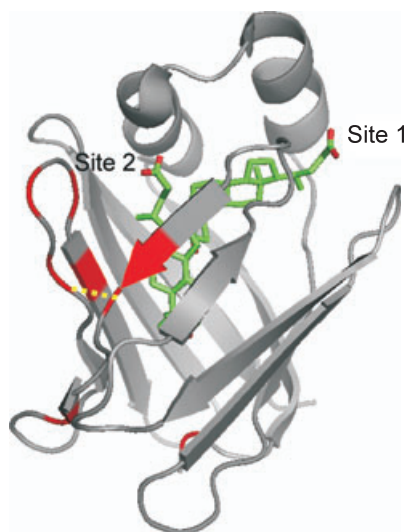


Fig. 6. Chemical shift changes upon CDA binding at pH 7 and 298 K. (A) Chemical shift differences between apo and holo resonances for T91C (black) and wild-type (WT) (grey) L-BABP, calculated as $\Delta\delta(\text{HN},\text{N}) = [(\Delta\delta\text{HN}(\text{T91C} - \text{WT})^2 + \Delta\delta\text{N}(\text{T91C} - \text{WT})^2/25)/2]^{1/2}$, are plotted versus residue number. The dotted line corresponds to the mean value plus one standard deviation of T91C L-BABP CSP. (B) Residues showing the major differences upon introduction of a disulfide bridge (Phe2, Lys79, Cys80, Leu84, Lys92, Glu94, Phe96, and His98) are coloured in red on the ribbon representation of L-BABP. The two ligands are coloured in green, and the position of the disulfide bridge is in yellow.



showed that slow motions were not quenched upon ligand binding. Indeed, high T_1/T_2 ratios were observed for Tyr9 and Gln11 (strand A), Arg32 (helix II), Val90, Lys92, and Glu94 (strand G), Phe96 and Ser97 (strand H), Phe113 (strand I), and Arg120 and Val125 (strand J) (Fig. 7). We can conclude that, at variance with what was observed for wild-type protein (Fig. S2), T91C L-BABP complexation with CDA

enhanced backbone motions that were already present in the apo protein, except for residues belonging to strand C and strand D and to loop EF and loop IJ.

In view of the physiological relevance of bile salt conjugation, which prevents passive diffusion of bile salts across cell membranes, the NMR analysis was extended to glycoconjugates, namely GCDA and GCA. Both homotypic complexes (T91C

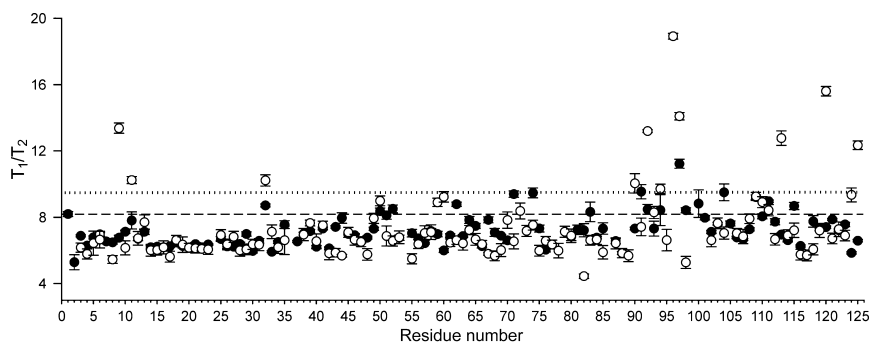


Fig. 7. Comparison of T_1/T_2 ratios for apo and holo T91C L-BABP. ^{15}N amide T_1/T_2 values as a function of residue number measured at 298 K. Filled black circles: apo T91C L-BABP. Empty circles: T91C L-BABP/CDA at a molar ratio of 1 : 3. Dashed and dotted lines correspond to the mean value plus one standard deviation of apo and holo T91C L-BABP, respectively. Error bars are shown.

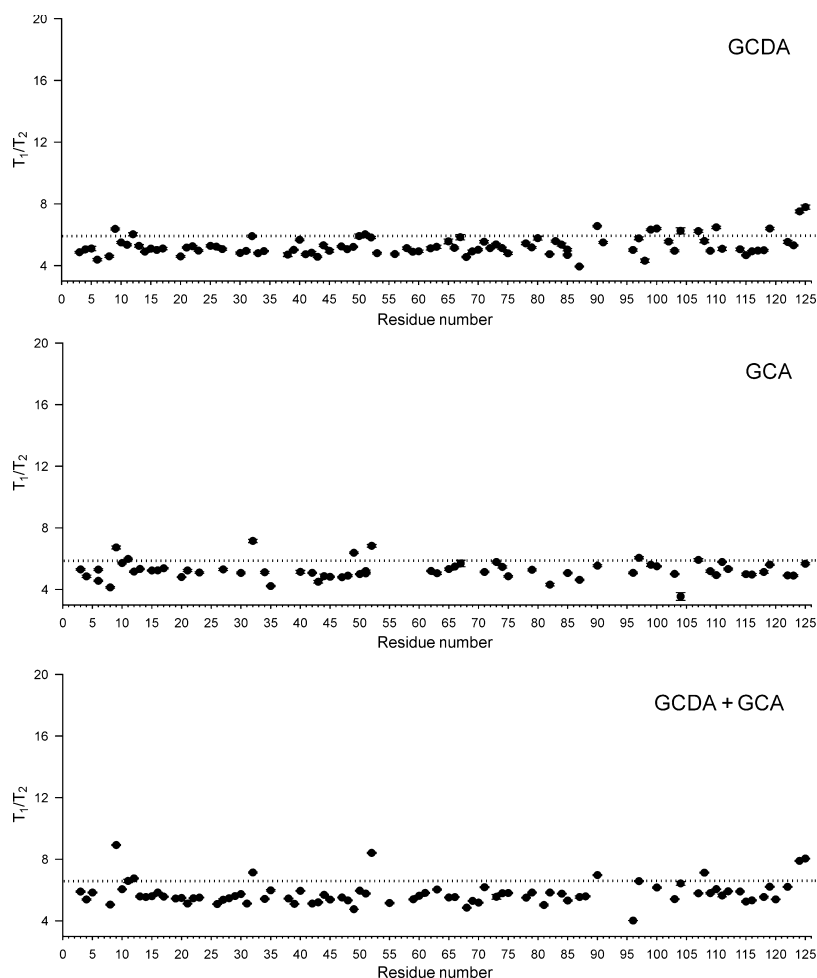


Fig. 8. T_1/T_2 ratios for T91C L-BABP complexed with the different glycerol derivatives. $[^{15}\text{N}]$ amide T_1/T_2 values as a function of residue number measured at 298 K. Upper panel: T91C L-BABP/GCDA at a molar ratio of 1 : 3. Middle panel: T91C L-BABP/GCA at a molar ratio of 1 : 3. Lower panel: T91C L-BABP/GCDA/GCA at a molar ratio of 1 : 1.5 : 1.5. Dotted lines correspond to the mean value plus one standard deviation of the data. Error bars are shown.

L-BABP/GCDA molar ratio of 1 : 3 and T91C L-BABP/GCA molar ratio of 1 : 3) and the heterotypic complex (T91C L-BABP/GCDA/GCA molar ratio of 1 : 1.5 : 1.5) were characterized according to their relaxation properties. Interestingly, substantial quenching of the motions was observed in the presence of all the glycine derivatives, independent of the hydroxylation pattern (Fig. 8). A few residues at the C-terminal end showed T_1/T_2 ratios higher than one standard deviation for the T91C L-BABP–GCDA complex, whereas the same behavior was observed for residues at the N-terminal end for the T91C L-BABP–GCA complex.

Discussion

Several examples have been reported in the literature, for members of the lipocalin family, where the introduction/removal of a disulfide bridge was responsible for changes in ligand-binding stoichiometry and affini-

ties [11,12,21]. In intracellular proteins, disulfide bonds are generally transiently formed, owing to the reductive nature of the cellular environment. It has been shown that transient disulfide bonds are generally not essential for structural integrity, but can contribute to protein function. Reversible disulfide bridge formation within intracellular proteins can give rise to local and/or global conformational changes that may lead to distinct binding and functional properties [22,23]. In line with this, we have shown here that the presence of a disulfide bridge, while maintaining the same binding stoichiometry, induces changes in binding ability, site selectivity and dynamic properties of L-BABP. Thus, the study of a recombinant protein with a stable disulfide bridge helps in clarifying the role of transient intracellular disulfide bonds.

Both NMR analysis and MS data confirmed the ability of T91C L-BABP to bind two GCDA or GCA molecules, indicating that both protein forms are competent for efficient BA binding and transport within

the cell. These results differ from the recently reported data for the homologous liver zebrafish protein, where the introduction of a disulfide bridge resulted, intriguingly, in a singly ligated protein, with the cholate occupying the more superficial binding site [16].

Exchange peaks observed in $^1\text{H}/^{15}\text{N}$ -HSQC spectra of holo proteins, together with diffusion experiments, showed that exchange processes between bound and free forms are relevant for site 1 and negligible for site 2, independently of the presence of a disulfide bridge (Table 1). The introduction of a disulfide bridge induced significant changes in the GCA exchange regime for ligand bound to site 1, whose resonance was observable at all the investigated protein/ligand ratios, at variance with the wild-type protein [5]. In line with this observation is the trend of diffusion coefficients measured for GCA bound to T91C and wild-type L-BABP, pointing to a higher affinity of this ligand for T91C L-BABP site 1 (Table 1).

The most relevant feature emerging from the analysis presented here is the ability of the disulfide bridge to modulate recognition at both sites. Indeed, no site selectivity was previously observed for wild-type L-BABP [5], whereas it is now clear that when T91C L-BABP is incubated with only GCDA or GCA, both binding sites are occupied, but when the two bile salts, differing only in hydroxylation at position 12, are present, GCDA preferentially binds to site 1 and GCA to site 2. Site selectivity is, however, observed only when both GCDA and GCA are present, suggesting that it does not derive from steric exclusion of one bile salt from a specific site.

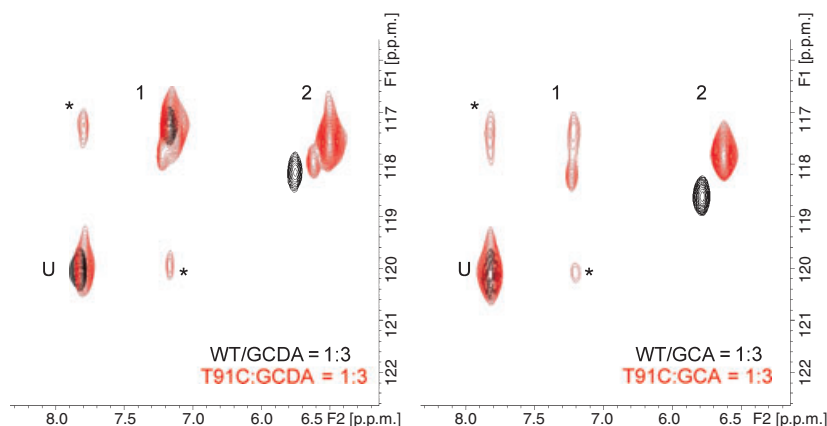
Protein observation was required in order to investigate the structural basis of these varied ligand-binding properties. Both CSP (Fig. 5) and TALOS analysis on the apo protein indicated that no significant change in 3D structure occurred. The comparison of CSP for

the holo forms of T91C L-BABP and the wild type (Fig. 6A) indicated that the same protein regions are generally involved in ligand binding, even if all the residues showing significantly different CSP values in the two proteins were gathered around the ligand bound at site 2 (Fig. 6B). This result is in perfect agreement with data derived from ligand observation (Fig. 9), revealing significant changes in the chemical shifts of site 2 resonance for both bile salts. This behaviour is ascribed to local changes in the chemical environment due to the introduction of a disulfide bridge, which involves two residues that are in contact with the ligand bound to the 'internal' binding site in the holo wild-type structure (Protein Data Bank ID: 2JN3 [14]).

Protein dynamics is largely influenced by the presence/absence of the disulfide bridge. Indeed, the presence of the disulfide bridge favoured the propagation of slow motions from the C-terminal region of the molecule to the N-terminal β -sheet in the apo protein, and enhanced backbone motions in the T91C L-BABP–CDA complex, at variance with the behaviour of the wild-type protein, where the binding of this ligand was accompanied by substantial quenching of motions (Fig. S2).

Molecular dynamics simulation studies revealed differently coupled correlated motions for some iLBPs, depending on the presence and the type of ligand [10,24]. These data prompted us to evaluate the effect of BA glycosylation and hydroxylation pattern on protein conformational motions. Interestingly, all glycodeivative mixtures were efficient in reducing backbone dynamics (Fig. 8), possibly as a consequence of the onset of more favourable interactions between the glycine moiety and the protein portal region. Indeed, comparison of the CSP in the presence of CDA or GCDA (Fig. 10) suggested that the most affected

Fig. 9. Comparison of ^{15}N GCDA and ^{15}N GCA in complex with wild-type (WT) L-BABP (black) and T91C L-BABP (red). Superposition of 2D $^1\text{H}/^{15}\text{N}$ -HSQC spectra of ^{15}N GCDA (left panel) and ^{15}N GCA (right panel) at a 1 : 3 protein/ligand molar ratio. The resonances corresponding to the unbound ligand and to binding sites 1 and 2 are indicated as U, 1, and 2, respectively. Exchange peaks between site 1 and unbound resonance are labelled with asterisks.



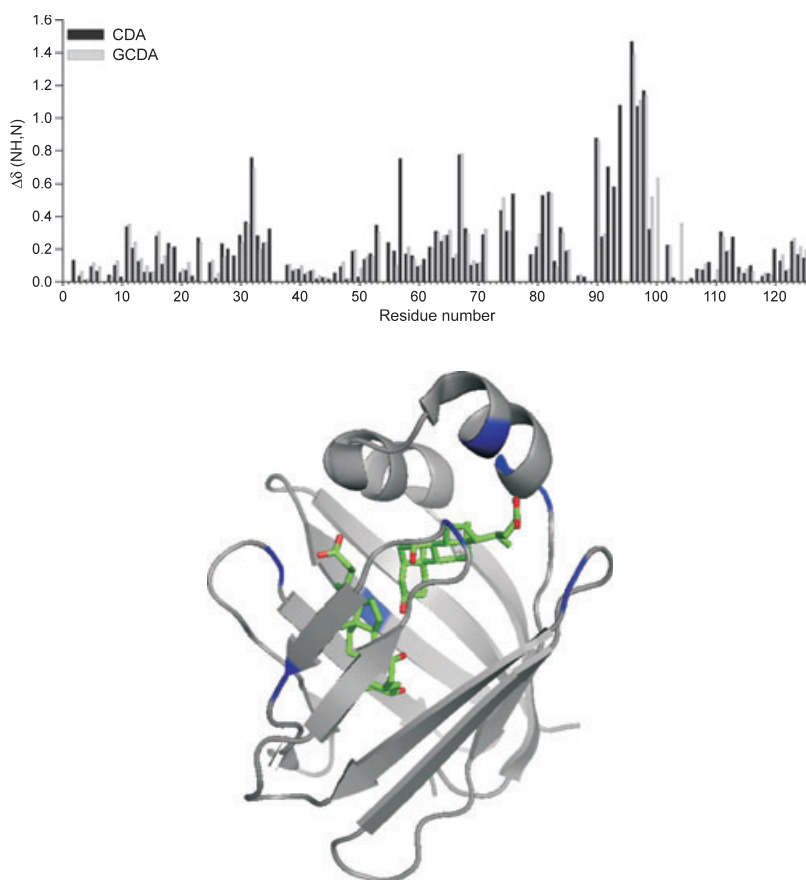


Fig. 10. Chemical shift changes upon CDA or GCDA binding to T91C L-BABP at pH 7 and 298 K. (A) Chemical shift differences between apo and holo resonances for CDA (black) and GCDA (grey) complexes, calculated as $\Delta\delta(\text{HN},\text{N}) = [(\Delta\delta\text{HN}(\text{T91C} - \text{WT})^2 + \Delta\delta\text{N}(\text{T91C} - \text{WT})^2/25)/2]^{1/2}$, are plotted versus residue number. The dotted line corresponds to the mean value plus one standard deviation of T91C L-BABP CSP. (B) Residues showing the major differences in the two complexes (Tyr9, Leu27, Gln42, Val49, Thr50, Thr59, Asp74, Cys80, Lys86, and Arg124) are coloured in blue on a ribbon representation of L-BABP.

residues are located at the level of the portal area, as expected in response to the protrusion of the glycine moieties, and at the level of strand F, strand H and strand I, in close contact with the ligand bound at site 2. Specifically, the chemical shift variation observed at the portal area for Arg32 and Asp33 suggests a different positioning of helix II in the two complexes. Arg32, characterized by high T_1/T_2 values in the apo protein and in all of the investigated holo proteins, thus plays a key role in regulating the positioning of the helix-loop-helix motif with respect to the β -barrel in order to accommodate the different BAs.

Analysis of relaxation data obtained for the glycode-derivatives showed that GCDA was able to quench motions affecting the protein open end (helical and loop EF regions), whereas the bound GCA mostly influenced the C-terminal region of the protein, in agreement with the site selectivity observed for the two ligands. Interestingly, the heterotypic complex, in which the proper ligand is expected to be located at the corresponding binding site, still presented a few residues with high T_1/T_2 values, especially at the N-terminal end. The competition data (Fig. 4) indi-

cated that GCDA preferentially populates site 1' when GCA is bound to site 2, and this different orientation at the superficial site may induce different motional properties at the N-terminal end. The detected site preferences and changes in chemical shifts in heterotypic complexes further establish a parallelism with the behaviour observed for I-BABP and its mutants [8], thus substantiating the previous proposal that BABPs exert a parallel function in hepatocytes and enterocytes [4,25].

In conclusion, it is shown here that the introduction of a disulfide bond makes the protein competent for site selectivity. NMR data indicated that protein conformational changes induced by the disulfide bond are small and gathered around the inner binding site, whereas significant changes in backbone motions are observed for several residues distributed over the entire protein. Site selectivity appears, therefore, to be governed by protein mobility, rather than by steric factors related to the hydroxylation pattern of the ligand, in agreement with what has been observed for other BABPs [4,25]. These results once more underline the tight connection between ligand-binding phenomena

and protein mobility in this protein family, and set the basis for further NMR kinetic studies based on line-shape analysis and relaxation dispersion measurements.

Experimental procedures

Protein expression and purification

The expression plasmid for T91C L-BABP was obtained from that of wild-type L-BABP using the Quickchange (Stratagene, La Jolla, CA, USA) mutagenesis kit. The presence of the desired mutation was confirmed by plasmid sequencing. Recombinant T91C L-BABP was expressed from *Escherichia coli* and purified to homogeneity as previously described [9]. Delipidated T91C L-BABP was obtained in a yield of 95 mg·L⁻¹ of rich medium. ¹³C/¹⁵N labelling was achieved using M9 minimal media containing 1 g·L⁻¹ ¹⁵NH₄Cl and 4 g·L⁻¹ ¹³C-enriched glucose, following protocols reported in the literature [26]. ¹⁵N-labelled and ¹³C/¹⁵N-labelled T91C L-BABP were both obtained with a 75 mg·L⁻¹ yield from minimal media. Protein concentrations for sample preparation were determined spectrophotometrically. The presence of the disulfide bridge was confirmed by MS.

NMR sample preparation

NMR studies on the apo protein were performed on 0.5 mM ¹⁵N/¹³C-labelled samples of T91C L-BABP dissolved in 30 mM potassium phosphate buffer in 95% H₂O/5% D₂O. The pH of the solutions was 5.6 or 7.2. Unenriched BAs and [24-¹³C]glycocholate were purchased from Sigma (St Louis, MO, USA). [¹⁵N]Glycine conjugates of CDA and CA were prepared as previously reported [5]. The titration of the unlabelled T91C L-BABP with increasing amounts of [¹⁵N]GCDA or [¹⁵N]GCA was performed at seven protein/ligand ratios (1 : 0.3, 1 : 0.6, 1 : 1, 1 : 1.5, 1 : 2, 1 : 2.5, and 1 : 3), and the preparation of the holo protein samples was performed following a procedure previously described [5]. Protein–ligand complexes were analysed at pH 7.2 on 0.5 mM T91C L-BABP samples; each protein/ligand molar ratio sample was prepared and analysed twice, in order to minimize errors.

NMR data collection and analysis

NMR spectra were acquired at 298 K on Bruker DMX 500 and Avance III 600 spectrometers equipped with a 5 mm TCI cryoprobe and a Z-field gradient. Data were processed with NMRPIPE [27] and visualized with NMRVIEW [28]. For the assignment of apo and holo protein resonances, the following experiments, ¹H/¹⁵N-TOCSY HSQC and NOESY HSQC, together with HNCACB, CBCA(CO)NH and

HN(CA)CO were performed. The secondary structure elements were derived with the software TALOS [20] from chemical shift data of HN, N, HA, CA and CB nuclei.

¹⁵N-relaxation experiments

¹⁵N-relaxation experiments for apo and holo (T91C L-BABP/CDA molar ratio of 1 : 3) samples were acquired at 600 MHz at pH 7.2. A dataset of 14 variable delays (2.5, 20, 60, 100, 150, 200, 300, 400, 600, 800, 1000, 1500, 1700 and 2500 ms) was used for T₁ measurements, and a dataset of nine variable delays (16.96, 33.92, 50.88, 67.84, 101.76, 135.68, 169.6, 220.48 and 237.44 ms) was used for T₂ measurements. For T91C L-BABP in complex with CDA, a dataset of nine variable delays (0.01, 180, 360, 540, 720, 900, 300, 1080, 1260 and 1440 ms) was used for T₁ measurements, and a dataset of seven variable delays (16.96, 33.92, 50.88, 67.84, 101.76, 220.48 and 237.44 ms) was used for T₂ measurements. T₁ and T₂ values were determined for 112 nonoverlapping cross-peaks. For the holo T91C L-BABP in complex with GCDA, GCA, or both (T91C L-BABP/GCDA/GCA molar ratio of 1 : 1.5 : 1.5), 10 variable delays (10, 60, 180, 300, 450, 600, 740, 900, 1100, 1200 and 1400 ms) were used for T₁ measurements, and 10 variable delays (16.98, 33.16, 49.74, 66.32, 82.9, 99.48, 132.64, 149.22, 198.96 and 232.12 ms) were used for T₂ measurements, recorded at 500 MHz. T₁ and T₂ relaxation values were estimated for 92, 65 and 86 residues for the complex with GCDA, the complex with GCA, and the heterotypic complex, respectively.

Titration experiments

¹H/¹⁵N-HSQC spectra of unlabelled protein complexed with labelled ligands were acquired with a ¹H spectral width of 6510 Hz and 1024 points, zero-filled to a total of 2048 points. Relaxation delays of 1.7 s were employed. In the ¹⁵N dimension, 256 increments were collected, with a sweep width of 2032 Hz, zero-filled to a total of 1024 points.

The linewidth dependence of ligand ¹H resonances as a function of protein/ligand ratio was followed through the first increment of 2D ¹H/¹⁵N-HSQC spectra recorded under identical conditions (8000 points on a sweep width of 6510 Hz).

The temperature dependence of the BA amide ¹H resonances for a protein/ligand molar ratio of 1 : 3 was followed through the first increment of a 2D ¹H/¹⁵N-HSQC spectrum collected with 8000 points on a sweep width of 6510 Hz in the temperature range 280–305 K.

Diffusion experiments

¹⁵N-edited diffusion experiments were performed on samples of wild-type and T91C L-BABP in complex with

[¹⁵N]GCDA and [¹⁵N]GCA at protein/ligand ratios of 1 : 3, in order to determine the diffusion coefficients of protein-bound ligands as compared with those of the free molecules. The pulse program was obtained by combining the standard HSQC pulse scheme with a pulsed-field gradient stimulated echo module employing bipolar gradients under the same conditions previously reported [6]. The measured signal volumes as a function of the applied gradient were fitted to the following equation, using a nonlinear least squares minimization:

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

where D is the translational diffusion coefficient, γ is the ¹H gyromagnetic ratio, G is the gradient strength, Δ and δ are as defined above, and τ is the gradient pulse separation.

Acknowledgements

S. Zanzoni and M. Guariento are gratefully acknowledged for help in protein expression and purification. We are grateful to R. Longhi for providing the BA [¹⁵N]glycine conjugates. This research was supported by FIRB 2003 (Project No. RBNE03PX83), Cariverona Foundation. The University of Verona is acknowledged for financial support in the acquisition of the NMR Bruker Avance 600 MHz spectrometer equipped with a cryoprobe. L. Ragona thanks CNR-RSTL 2007 (Code No. 779) for financial support. C. Cogliati was supported by a grant 'Sovvenzione Globale INGENIO' from 'Fondo Sociale Europeo, Ministero del Lavoro e della Previdenza Sociale and Regione Lombardia'. CIRMMP (Consorzio Interuniversitario di Risonanze Magnetiche di Metalloproteine Paramagnetiche) is gratefully acknowledged.

References

- Houten SM, Watanabe M & Auwerx J (2006) Endocrine functions of bile acids. *EMBO J* **25**, 1419–1425.
- Chawla A, Saez E & Evans RM (2000) Don't know much bile-ology. *Cell* **103**, 1–4.
- Thomas C, Pellicciari R, Pruzanski M, Auwerx J & Schoonjans K (2008) Targeting bile-acid signalling for metabolic diseases. *Nat Rev Drug Discov* **7**, 678–693.
- Tochtrop GP, DeKoster GT, Covey DF & Cistola DP (2004) A single hydroxyl group governs ligand site selectivity in human ileal bile acid binding protein. *J Am Chem Soc* **126**, 11024–11029.
- Tomaselli S, Ragona L, Zetta L, Assfalg M, Ferranti P, Longhi R, Bonvin AM & Molinari H (2007) NMR-based modeling and binding studies of a ternary complex between chicken liver bile acid binding protein and bile acids. *Proteins* **69**, 177–191.
- Pedò M, D'Onofrio M, Ferranti P, Molinari H & Assfalg M (2009) Towards the elucidation of molecular determinants of cooperativity in the liver bile acid binding protein. *Proteins Struct Funct Bioinformatics* doi:10.1002/prot.22496.
- Tochtrop GP, Bruns JL, Tang C, Covey DF & Cistola DP (2003) Steroid ring hydroxylation patterns govern cooperativity in human bile acid binding protein. *Biochemistry* **42**, 11561–11567.
- Toke O, Monsey JD, DeKoster GT, Tochtrop GP, Tang C & Cistola DP (2006) Determinants of cooperativity and site selectivity in human ileal bile acid binding protein. *Biochemistry* **45**, 727–737.
- Ragona L, Catalano M, Luppi M, Cicero D, Eliseo T, Foote J, Fogolari F, Zetta L & Molinari H (2006) NMR dynamic studies suggest that allosteric activation regulates ligand binding in chicken liver bile acid-binding protein. *J Biol Chem* **281**, 9697–9709.
- Eberini I, Guerini Rocco A, Ientile AR, Baptista AM, Gianazza E, Tomaselli S, Molinari H & Ragona L (2008) Conformational and dynamics changes induced by bile acids binding to chicken liver bile acid binding protein. *Proteins* **71**, 1889–1898.
- Liu J, Chenyun G, Yihe Y & Donghai L (2008) Effects of removing a conserved disulfide bond on the biological characteristics of rat lipocalin-type prostaglandin D synthase. *Biochimie* **90**, 1637–1646.
- Sjoelund V & Kaltashov IA (2007) Transporter-to-trap conversion: a disulfide bond formation in cellular retinoic acid binding protein I mutant triggered by retinoic acid binding irreversibly locks the ligand inside the protein. *Biochemistry* **46**, 13382–13390.
- Vasile F, Ragona L, Catalano M, Zetta L, Perduca M, Monaco H & Molinari H (2003) Solution structure of chicken liver basic fatty acid binding protein. *J Biomol NMR* **25**, 157–160.
- Eliseo T, Ragona L, Catalano M, Assfalg M, Paci M, Zetta L, Molinari H & Cicero DO (2007) Structural and dynamic determinants of ligand binding in the ternary complex of chicken liver bile acid binding protein with two bile salts revealed by NMR. *Biochemistry* **46**, 12557–12567.
- Nichesola D, Perduca M, Capaldi S, Carrizo ME, Righetti PG & Monaco HL (2004) Crystal structure of chicken liver basic fatty acid-binding protein complexed with cholic acid. *Biochemistry* **43**, 14072–14079.
- Capaldi S, Guariento M, Saccomani G, Fessas D, Perduca M & Monaco HL (2007) A single amino acid mutation in zebrafish (*Danio rerio*) liver bile acid-binding protein can change the stoichiometry of ligand binding. *J Biol Chem* **282**, 31008–31018.
- Reibarkh M, Malia TJ & Wagner G (2006) NMR distinction of single- and multiple-mode binding of small-molecule protein ligands. *J Am Chem Soc* **128**, 2160–2161.

- 18 Hsu EW & Mori S (1995) Analytical expressions for the NMR apparent diffusion coefficients in an anisotropic system and a simplified method for determining fiber orientation. *Magn Reson Med* **34**, 194–200.
- 19 Nakashima T (2002) Potentiometric study on critical micellization concentrations (CMC) of sodium salts of bile acids and their amino acid derivatives. *Colloids Surf B Biointerfaces* **24**, 103–110.
- 20 Cornilescu G, Delaglio F & Bax A (1999) Protein backbone angle restraints from searching a database for chemical shift and sequence homology. *J Biomol NMR* **13**, 289–302.
- 21 Capaldi S, Perduca M, Faggion B, Carrizo ME, Tava A, Ragona L & Monaco HL (2007) Crystal structure of the anticarcinogenic Bowman–Birk inhibitor from snail medic (*Medicago scutellata*) seeds complexed with bovine trypsin. *J Struct Biol* **158**, 71–79.
- 22 Piotukh K, Kosslick D, Zimmermann J, Krause E & Freund C (2007) Reversible disulfide bond formation of intracellular proteins probed by NMR spectroscopy. *Free Radic Biol Med* **43**, 1263–1270.
- 23 Thangudu RR, Manoharan M, Srinivasan N, Cadet F, Sowdhamini R & Offmann B (2008) Analysis on conservation of disulphide bonds and their structural features in homologous protein domain families. *BMC Struct Biol* **8**, 55.
- 24 Woolf TB, Grossfield A & Tychko M (2000) Differences between apo and three holo forms of the intestinal fatty acid binding protein seen by molecular dynamics computer calculations. *Biophys J* **78**, 608–625.
- 25 Guariento M, Raimondo D, Assfalg M, Zanzoni S, Pesente P, Ragona L, Tramontano A & Molinari H (2008) Identification and functional characterization of the bile acid transport proteins in non-mammalian ileum and mammalian liver. *Proteins* **70**, 462–472.
- 26 Marley J, Lu M & Bracken C (2001) A method for efficient isotopic labeling of recombinant proteins. *J Biomol NMR* **20**, 71–75.
- 27 Delaglio F, Grzesiek S, Vuister GW, Zhu G, Pfeifer J & Bax A (1995) NMRPipe: a multidimensional spectral processing system based on UNIX pipes. *J Biomol NMR* **6**, 277–293.
- 28 Johnson B.A. (2004) Using NMRView to visualize and analyze the NMR spectra of macromolecules. *Methods Mol Biol* **278**, 313–352.

Supporting information

The following supplementary material is available:

Fig. S1. Chenodeoxycholic acid structure.

Fig. S2. Comparison of T_1/T_2 ratio for apo and holo T91C and wild-type proteins.

This supplementary material can be found in the online version of this article.

Please note: As a service to our authors and readers, this journal provides supporting information supplied by the authors. Such materials are peer-reviewed and may be re-organized for online delivery, but are not copy-edited or typeset. Technical support issues arising from supporting information (other than missing files) should be addressed to the authors.

Supplementary Material

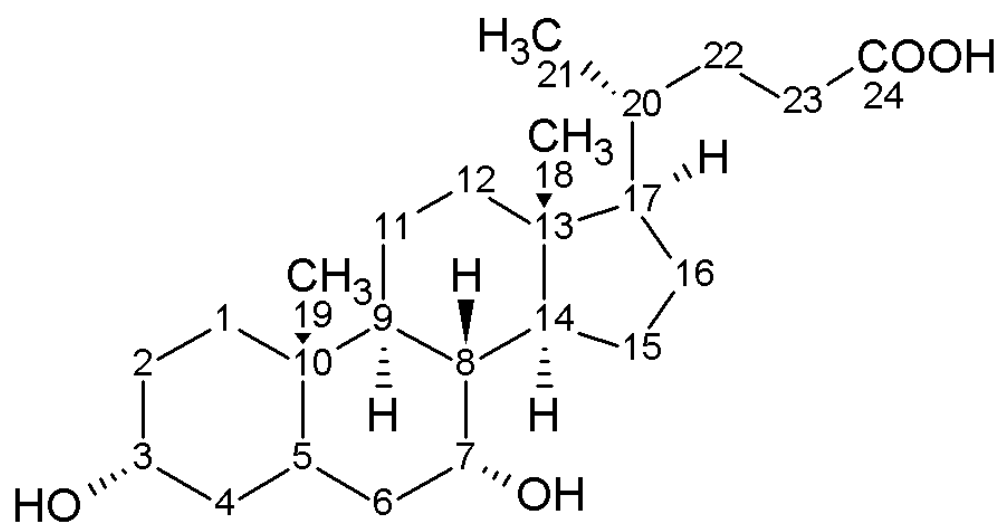


Figure S1 Chenodeoxycholic acid structure. Carbon atom numbering is reported. Cholic acid has a hydroxyl group at position 12.

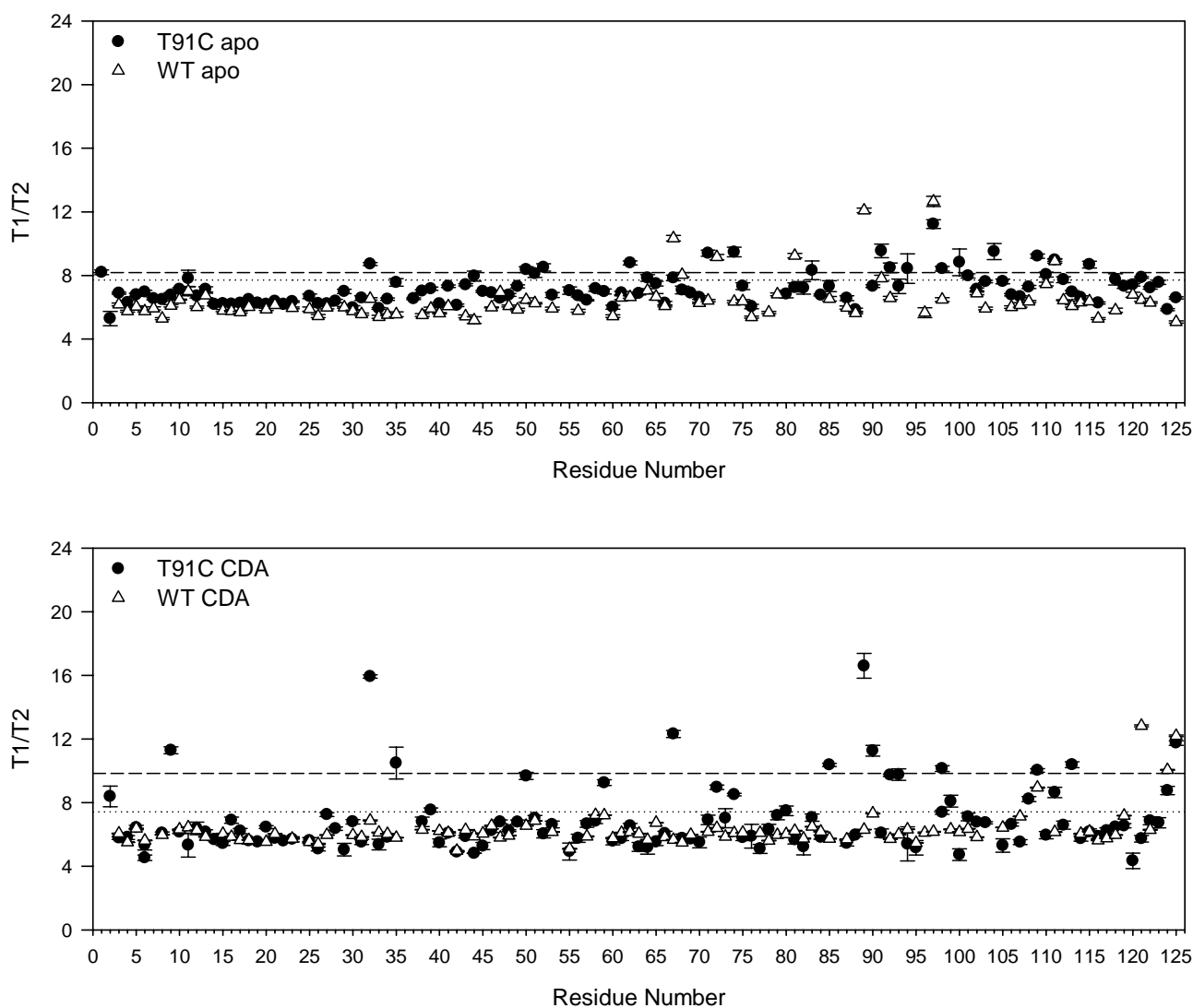


Figure S2 Comparison of T_1/T_2 ratio for apo and holo T91C and WT L-BABPs. Upper panel: ^{15}N amide T_1/T_2 values as a function of residue number measured at 298 K for apo T91C (black filled circles) and apo WT (empty triangles) at 600 MHz. Lower panel: ^{15}N amide T_1/T_2 values protein:CDA 1:3 complexes for T91C (black filled circles) and WT (empty triangles) proteins. Dashed and dotted lines correspond to the mean value plus one standard deviation T91C and WT proteins, respectively. Error bars are shown.

CHAPTER 7: Publication 4

Title: NMR studies reveal the role of biomembranes in modulating ligand binding and release by intracellular Bile Acid-Binding Proteins.

Short presentation of publication 4:

The peripheral interaction of proteins with biological membranes has been reported since a long time and the functional relevance of this interaction has been highlighted for several proteins. Bile acid-binding protein from chicken liver (cL-BABP) was shown to interact with membrane mimetic vesicles, namely anionic phospholipids vesicles, mainly via electrostatic interactions. It was shown that, upon membrane binding, cL-BABP undergoes a partial loss of structure.

We performed several NMR experiments on cL-BABP/membrane system, in the presence of the physiological ligand, i.e., glycochenodeoxycholic acid (GCDA), with the aim of clarifying the functional relevance of the membrane binding properties of cL-BABP. The results are reported in the following publication.

The phospholipid of choice was the 1,2-dimyristoyl-sn-glycero-3-phospho-(1'-rac-glycerol) (DMPG) which is a negatively charged phospholipid commonly used to mimic anionic membrane in *in vitro* studies.

As the first step we verified the interaction of cL-BABP with anionic lipid vesicles recording ^1H , ^{15}N -TROSY experiments on a sample of ^{15}N cL-BABP in the presence of DMPG vesicles, employing a protein:DMPG ratio of 1:100. The obtained spectrum displayed broadened resonances with a limited chemical shift dispersion, consistent with the suggestion that, upon binding to vesicles, the protein undergoes partial unfolding. The binding of the protein to the vesicles was also confirmed by filtration experiments. In addition the electrostatic nature of this interaction was verified by recovering the typical pattern of resonances for a correctly folded ^{15}N cL-BABP upon increasing the ionic strength in the sample.

A sample of ^{15}N cL-BABP in the presence of DMPG vesicles was then titrated with a GCDA concentrated solution, to investigate whether the protein maintains its ligand binding capability in the membrane bound state. During the titration, the appearance of the typical resonances of the GCDA bound protein were observed, indicating that cL-BABP, though associated with the vesicles, is competent for ligand binding and that the binding event reduces the strength of the interaction between the protein and the vesicles.

To evaluate the reversibility of the protein/vesicles interaction in the presence of GCDA, a titration experiment was further performed, starting from holo cL-BABP. A concentrated DMPG vesicles solution was titrated into a 1:1 ^{15}N cL-BABP:GCDA sample and, for each titration point, a ^1H , ^{15}N -TROSY experiment was recorded. The resonances of the holo protein were decreasing in intensity during the titration while broad signals, typical of the vesicles associated protein, appeared. This result indicates that holo cL-BABP can interact with DMPG vesicles.

The interaction of holo cL-BABP with the vesicles was also monitored from the ligand point of view by repeating the DMPG titration experiment starting from a 1:1 cL-BABP: ^{15}N GCDA sample. The stepwise addition of DMPG vesicles induced the release of the GCDA from the protein as evident from the progressive intensity reduction of the bound ligand signals and the concomitant increasing of the free ligand resonance. By monitoring the signal intensity variation during the titration it was highlighted that binding site 1 was the preferential position from which the ligand was released. This preference was also confirmed through 2D ^{15}N -edited NOESY experiments in which exchange phenomena between ligand in binding site 1 and the free ligand were identified but not between ligand in binding site 2 and free ligand.

The intensity reduction trend of the holo ^{15}N cL-BABP signals during the titration with DMPG vesicles was monitored to locate, on the protein structure, the regions involved in membrane interaction. A small group of resonances showed a distinct behaviour characterized by a marked reduction in intensity already at an early stage of the titration, possibly indicating a direct interaction with the lipid vesicles. We mapped on the protein structure the residues whose intensity was reduced at least of 50% already at the protein:DMPG:GCDA ratio of 1:20:1. All the selected residues, namely S3, G4, Q41, K43, D45, K103 and G104 are close to each other and identify a positive patch on the protein surface, possibly participating to the direct interaction with the anionic vesicles.

Starting from the collected data it was possible to define an equilibrium involving cL-BABP, GCDA and DMPG vesicles. We described two states, one in which apo cL-BABP is associated with the vesicles and another in which the GCDA binding induce the detachment of the protein from the membrane. The switch between the two states is therefore driven by the presence of the ligand.

Finally we proposed a mechanism for the intracellular transport of bile acids mediated by cL-BABP based on the described equilibrium. In the proposed model, an hepatic polarized cell is thought to be spanned by an intracellular bile acid gradient with an higher concentration close to the basolateral membrane (site for the uptake of bile acids from the blood circulation) and a lower concentration near the apical membrane (site from which bile acids are secreted into the bile canaliculus). The ongoing bile acids gradient induce the loading of cL-BABP with bile acids and the consequent detachment from the membrane at the basolateral membrane, while at the apical membrane it determines the association of cL-BABP with the membrane with the release of the carried bile acids.

My contribution in this work was in the planning and performance of the experiments, in the data analysis and discussion and in the writing of the scientific paper.

NMR Studies Reveal the Role of Biomembranes in Modulating Ligand Binding and Release by Intracellular Bile Acid Binding Proteins

Massimo Pedò¹, Frank Löhr², Mariapina D'Onofrio¹, Michael Assfalg^{1*}, Volker Dötsch² and Henriette Molinari¹

¹Department of Biotechnology,
University of Verona,
Strada Le Grazie 15,
37134 Verona, Italy

²Goethe University,
Max-von-Laue-Strs. 9,
61438 Frankfurt, Germany

Received 20 July 2009;
received in revised form
5 October 2009;
accepted 8 October 2009
Available online
22 October 2009

Bile acid molecules are transferred vectorially between basolateral and apical membranes of hepatocytes and enterocytes in the context of the enterohepatic circulation, a process regulating whole body lipid homeostasis. This work addresses the role of the cytosolic lipid binding proteins in the intracellular transfer of bile acids between different membrane compartments. We present nuclear magnetic resonance (NMR) data describing the ternary system composed of the bile acid binding protein, bile acids, and membrane mimetic systems, such as anionic liposomes. This work provides evidence that the investigated liver bile acid binding protein undergoes association with the anionic membrane and binding-induced partial unfolding. The addition of the physiological ligand to the protein-liposome mixture is capable of modulating this interaction, shifting the equilibrium towards the free folded holo protein. An ensemble of NMR titration experiments, based on nitrogen-15 protein and ligand observation, confirm that the membrane and the ligand establish competing binding equilibria, modulating the cytoplasmic permeability of bile acids. These results support a mechanism of ligand binding and release controlled by the onset of a bile salt concentration gradient within the polarized cell. The location of a specific protein region interacting with liposomes is highlighted.

© 2009 Elsevier Ltd. All rights reserved.

Edited by J. Bowie

Keywords: liver bile acid binding protein; NMR; liposomes; protein-membrane interaction; lipid transport

Introduction

Bile acids are physiological detergents that generate bile flow and facilitate intestinal absorption and transport of lipids, nutrients and vitamins. They are also signaling molecules and inflammatory agents that rapidly activate nuclear receptors and cell signaling pathways that regulate lipid, glucose and energy metabolism.¹ The enterohepatic circulation

of bile acids, fundamentally composed of secretion from the liver and absorption from the intestine, exerts important physiological functions in feedback inhibition of bile acid synthesis and in control of whole body lipid homeostasis.²

Detailed knowledge is available about the transporter proteins at the membranes of hepatocytes and enterocytes involved in vectorial bile acid movement under both normal physiological and pathological conditions.^{3,4,5} In contrast, the process of intracellular transfer of bile acids between the basolateral and the apical membrane is largely unexplored.

In general, cytoplasmic transport of small molecules occurs mostly by diffusion, which reflects random motions of molecules along a concentration gradient.⁶ In the case of lipids, diffusion of the free molecules is limited by their poor solubility in water. Soluble cytosolic proteins, such as the intracellular lipid binding proteins (iLBP), provide an aqueous

*Corresponding author. E-mail address:
michael.assfalg@univr.it.

Abbreviations used: iLBP, intracellular lipid binding proteins; BABP, bile acid binding protein; L-BABP, chicken liver BABP; DMPG, dimyristoyl phosphatidylglycerol; TROSY, transverse relaxation optimized spectroscopy; GCDA, glycochenodeoxycholic acid; FABP, fatty acid binding protein; DMPG, 1,2-dimyristoyl-sn-glycero-3-phospho-(1'-rac-glycerol).

carrier system to catalyze lipid transport through intracellular water layers. The efficiency of ligand uptake, transport, and release by iLBP determines the permeability of the cytoplasm. If the permeability of the cytoplasm is much less than that of the membrane, then a large concentration gradient will form at steady state. The effective diffusion of the protein-bound ligand is modulated by a variety of cell-dependent factors, such as tortuosity, viscosity, and molecular crowding. Particularly, cytoplasmic membrane components may cause obstructions along the diffusional path. Protein-membrane interactions are thus bound to affect intracellular lipid diffusion.

Within the framework of a broader project aimed at elucidating the mechanism by which bile acid binding proteins (BABP), belonging to the iLBP family, participate in the vectorial transport of bile acids, we set up an experimental approach to model this mechanism, exploiting the potentiality offered by NMR to address the systems under study in near-physiological environment.^{7,8,9,10}

The BABP from chicken liver (L-BABP) is one of the most extensively characterized BABPs in terms of structure,^{11,12,13} dynamics,¹⁴ and ligand-binding,^{12,15} and has proved an excellent system to investigate BABP-membrane interactions.^{16,17,18} On the basis of Fourier transform infrared spectroscopy, it was proposed that protein-membrane binding is accompanied by protein conformational changes taking place at the membrane-solvent interface.¹⁹ Further studies addressed the simulation of L-BABP-membrane interactions via 30 ns molecular dynamics calculations, and derived a model of interaction.¹⁶

We present a thorough NMR investigation on the role of membranes in bile acid binding to BABPs and propose a mechanism, based on interaction studies performed in the presence of dimyristoyl phosphatidylglycerol (DMPG) liposomes, consistent with reported physiological data. Given the role of L-BABP in bile acid distribution within the cell, the proposed description of chemical equilibria involving three molecular components (protein, liposome, and ligand) constitutes a step forward in the understanding of the mechanism of bile acid binding and release by L-BABP. The approach described here provides the possibility of monitoring both protein and ligand signals in a site-specific manner, and is of general applicability in the study of protein/ligand/membrane interactions.

Results

Protein-membrane association and ligand-induced dissociation

We have investigated an intracellular lipid binding protein in the presence of ligands and of membrane mimetic systems, with the aim of describing the protein ligand binding properties in a near-physio-

logical environment. We have used the liver bile acid binding protein (L-BABP) from chicken, a protein that can be expressed in *Escherichia coli* cells in high yield and has been extensively characterized in terms of bile salt binding *in vitro*.^{12,14,20}

Membrane mimetic systems were chosen to study the role of biomembranes in modulating bile acid binding to BABPs. Bilayer and multilayer vesicles formed by synthetic phospholipids are popular membrane models used in biophysical studies.¹⁰ We used the negatively charged DMPG liposomes as a model of anionic cell membrane components. All the experiments were done at 303 K, i.e. under conditions in which DMPG has been reported to be in a well defined liquid-crystalline phase.²¹

A [¹H,¹⁵N]-transverse relaxation optimized spectroscopy (TROSY) experiment was performed on a sample of apo-L-BABP:DMPG at a molar ratio of 1:100. The spectrum shown in Fig. 1a reveals that the resonances expected for the folded protein are absent, while the observable signals appear severely broadened and exhibit limited chemical shift dispersion particularly in the proton dimension. Interestingly, one well resolved resonance at 10.5 ppm (¹H) and 130 ppm (¹⁵N) corresponds to the random coil values of the aromatic nitrogen of the Trp-6 sidechain. This resonance is indeed shifted to 9.63 ppm and 126.3 ppm in the folded apo protein (Fig. 1b). The characteristic position of this Trp sidechain resonance has already been exploited to follow urea-induced unfolding of L-BABP.²² A fingerprint region obtained of unfolded (in the presence of 5 M urea) apo-L-BABP is reported in Fig. 1c for comparison. The above observations, together with the notion that most protein residues of the folded protein are part of well defined secondary structure elements, favor the view of the formation of membrane-bound partially folded species, rather than large supramolecular structures where only the flexible loops are visible. This view is consistent with previously reported infrared spectroscopy data describing the coupling between L-BABP unfolding and the phase state of certain phospholipid vesicles.¹⁸

To provide a further confirmation of the interaction between L-BABP and liposomes in our preparations, we carried out a filtration assay based on an established protocol.¹⁹ The upper chamber of a concentration device was loaded with apo-L-BABP and an apo-L-BABP/DMPG mixture. When the pure L-BABP sample was used, it was possible to recover the protein quantitatively in the flow-through, indicating that L-BABP can cross the membrane (50 kDa cutoff). In contrast, no protein could be recovered in the flow-through in the presence of liposomes. We verified that, upon salt addition, the [¹H,¹⁵N]-TROSY spectrum of the folded monomeric apo-L-BABP was reproduced (Supplementary Data Fig. S1).

The main focus of the present study was to assess the ability of the membrane-associated protein to take up bile salts: to achieve this goal, the

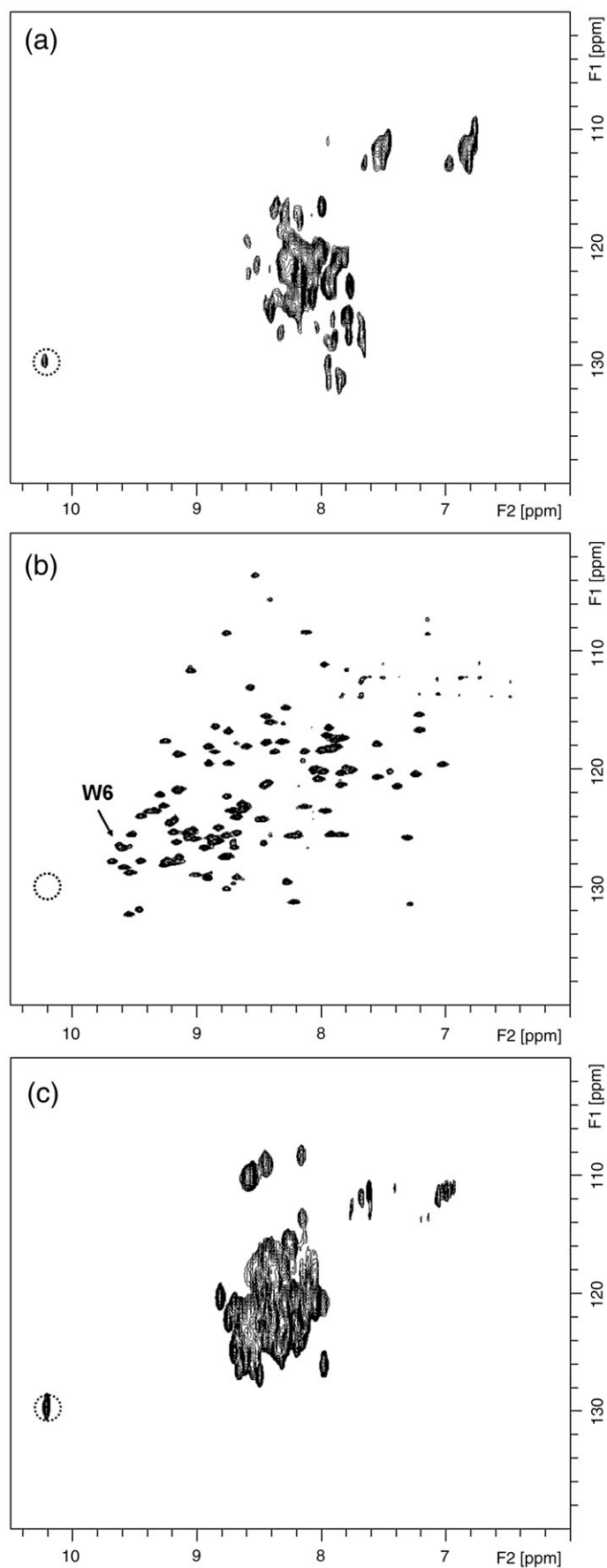


Fig. 1. $[\text{}^1\text{H}, \text{}^{15}\text{N}]$ -TROSY spectra of (a) apo-L-BABP:DMPG 1:100 and (b) purified apo-L-BABP. (c) The $[\text{}^1\text{H}, \text{}^{15}\text{N}]$ -TROSY spectrum of apo-L-BABP in the presence of 5 M urea is reported for comparison. The position of the sidechain signal of Trp6 in the unfolded protein is evidenced by a dotted circle, while the corresponding native resonance is labeled W6. Experiments were done at 600 MHz and 30 °C.

physiological ligand glycochenodeoxycholic acid (GCDA) was titrated to the protein/liposome mixture. L-BABP was shown earlier to bind two bile salt molecules with high affinity within its internal cavity.^{15,20} The protein-ligand binding equilibrium has been thoroughly studied by our group,^{12,14,15,20} using a variety of NMR experiments to observe both the ligand and the protein. The widespread protein backbone HN chemical shift perturbations observed on ligand binding suggested an allosteric behavior. Experiments performed using ¹⁵N-labeled GCDA and unlabeled protein revealed the possibility of obtaining site-specific binding information, as the ligand produced distinct signals for the free form and for two bound states. Following ligand resonances in protein-ligand titration experiments, it was possible to ascertain positive binding cooperativity between the two sites. The results obtained in the presence of DMPG are compared to the data obtained in the absence of phospholipids.

A titration of the protein:DMPG 1:100 mixture was done with increasing amounts of GCDA up to a protein:ligand ratio of 1:4. Upon GCDA addition the typical signals of holo-L-BABP become observable (Fig. 2a and b). The signals of the folded protein become more intense at higher ligand/protein ratios, relative to those of the unfolded apo protein, as revealed by selected 1D cross-sections obtained at P/DMPG/GCDA ratios of 1:100:2 and 1:100:4 (Fig. 2c). Thus, the presence of bile salt produces a shift of the equilibrium towards the ligand-bound folded protein at the expense of the population of the membrane-bound unfolded apo-BABP. The observation of holo protein signals in the presence of the physiological ligand GCDA confirms the association of the apo protein with the membrane and provides additional evidence of the membrane-induced loss of structure of L-BABP.

Protein-ligand binding and membrane-induced ligand release

To test the reversibility of the L-BABP/ligand/liposome interaction, two additional titrations of holo-L-BABP (¹⁵N L-BABP:unlabeled(u)-GCDA and uL-BABP:¹⁵NGCDA of equimolar concentrations) were done with increasing amounts of DMPG, from P:DMPG:L 1:1:1 to P:DMPG:L 1:100:1. Both the ¹⁵N-enriched protein and the ¹⁵N-enriched ligand were monitored in separate NMR experiments under identical conditions. The choice of using an equimolar complex of L-BABP-GCDA was dictated by the need for the ligand-observed NMR experiment to start the DMPG titration under solution conditions where no free ligand resonance was present, in order to clearly identify the appearance of the released ligand and to monitor the increase in its NMR signal intensity.

The titration of the L-BABP:¹⁵NGCDA 1:1 complex with DMPG was followed by monitoring the changes in the signals of the enriched ligand. As already mentioned, in a membrane-free solution

two distinct resonances are observed for the bound ¹⁵NGCDA in [¹H,¹⁵N]-heteronuclear single quantum coherence (HSQC) spectra, corresponding to the ligand bound to the more internal site 2 (δ 6.3 ppm and 118.3 ppm) and to the more superficial site 1 (δ 7.3 ppm and 117.4 ppm). An ensemble of rotating frame Overhauser effect spectroscopy (ROESY) and diffusion experiments indicated that only site 1 ligand exhibited exchange with the unbound ligand (located at δ 7.9 ppm and 120.2 ppm), while some slow exchange among the two bound forms was detected.¹² Selected spectra of the ligand along the titration are reported in Fig. 3a–d, indicating that, upon increasing the concentration of DMPG, the resonance of the free GCDA appears and increases in intensity, while the signal intensities of the bound species decrease. Inspection of Fig. 3e, which reports a plot of ligand peak intensities as a function of added DMPG, points to a differential behavior of the two ligands bound to sites 1 and 2. Indeed, the signal of site 1 ligand shows a steady decrease in intensity, while the resonance of site 2 ligand shows a maximum intensity at P:DMPG:L = 1:5:1 and then decays similarly to site 1 ligand. The curves in Fig. 3e could be interpreted as the ligand exiting site 1 of doubly ligated protein molecules and saturating site 2. After saturation of site 2, at P:DMPG:L = 1:5:1, the increased concentration of liposome induces ligand release also from this binding site, as confirmed by the appearance of the free ligand signal starting from P:DMPG:L = 1:10:1. These results identify binding site 1 as the preferred path of ligand release.

This notion is confirmed by the 2D ¹⁵N-edited NOESY experiment shown in Fig. 4. A mixing time of 20 ms allowed the observation of strong exchange signals between site 1 and free ligand but not between unbound ligand and site 2. Crosspeaks between site 1 and site 2 are also absent even in the presence of larger amounts of DMPG and with longer mixing times, suggesting that the exchange, if present, occurs at a different rate.

The same titration experiment was then done with ¹⁵N-enriched protein and unlabeled ligand. Upon addition of DMPG it was observed that the ¹⁵N amide resonances of the protein were shifting towards the unfolded form and the intensity decrease of the folded resonances along the titration could be monitored. The spectrum obtained at a P:DMPG:GCDA 1:100:1 ratio (mol/mol/mol) is reported in Fig. 5a and selected 1D cross-sections showing the stepwise decrease of the protein signal intensity upon DMPG addition are reported in Fig. 5b and c. Thus, the ligand-protein-membrane equilibria are shown to be reversible.

We attempted to determine the identity of residues involved in the interaction with the liposomes. The resonances corresponding to the membrane-bound unfolded protein could not be assigned but it was possible to analyze the signals of the folded holo protein upon the addition of DMPG. While the majority of residues displayed a very similar behavior in terms of peak intensity

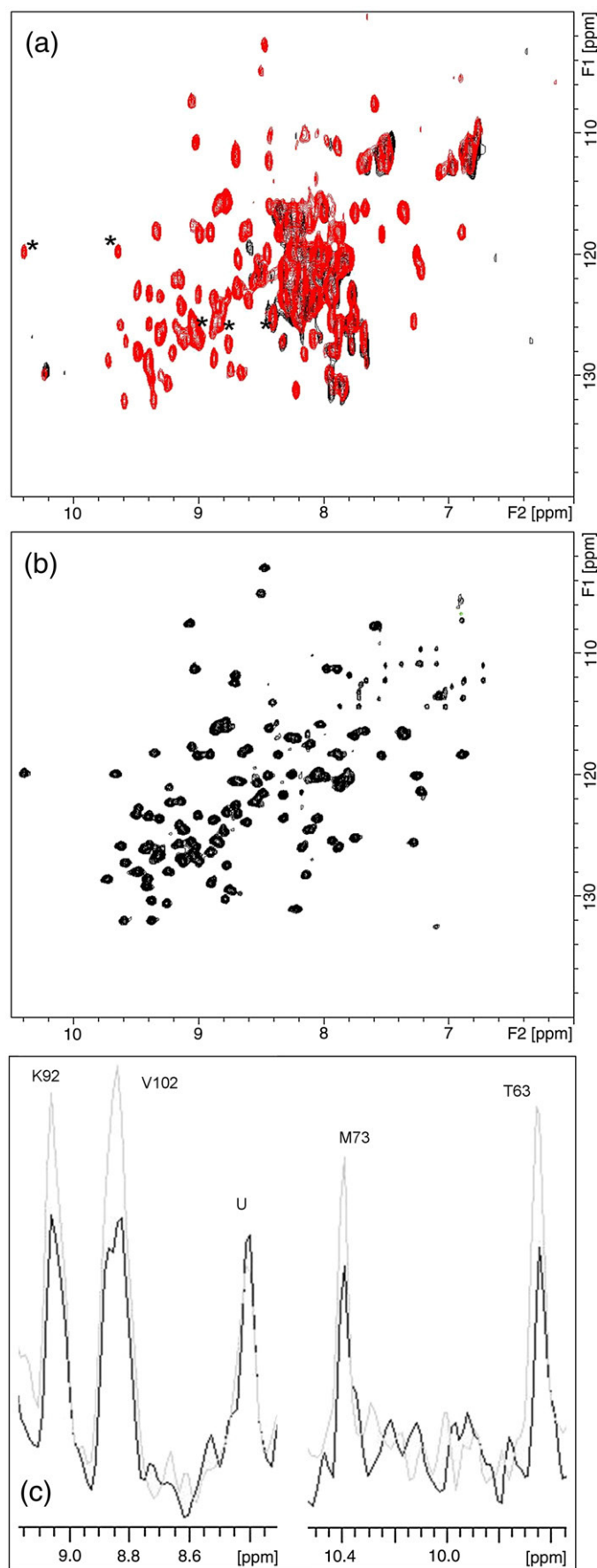


Fig. 2. (a) $[\text{}^1\text{H}, \text{}^{15}\text{N}]$ -TROSY spectrum of ^{15}N -L-BABP:DMGP:GCDA 1:100:2 (red contours) superimposed with the spectrum of ^{15}N -L-BABP:DMGP 1:100 (black contours) for comparison. Asterisks denote peaks for which ^1H cross-sections are reported in (c). (b) The $[\text{}^1\text{H}, \text{}^{15}\text{N}]$ -TROSY spectrum of ^{15}N -L-BABP:GCDA 1:2. (c) 1D cross-sections of selected resonances at P: L ratio 1:2 (black) and 1:4 (grey). Assignments are reported, the symbol U indicates an unfolded resonance.

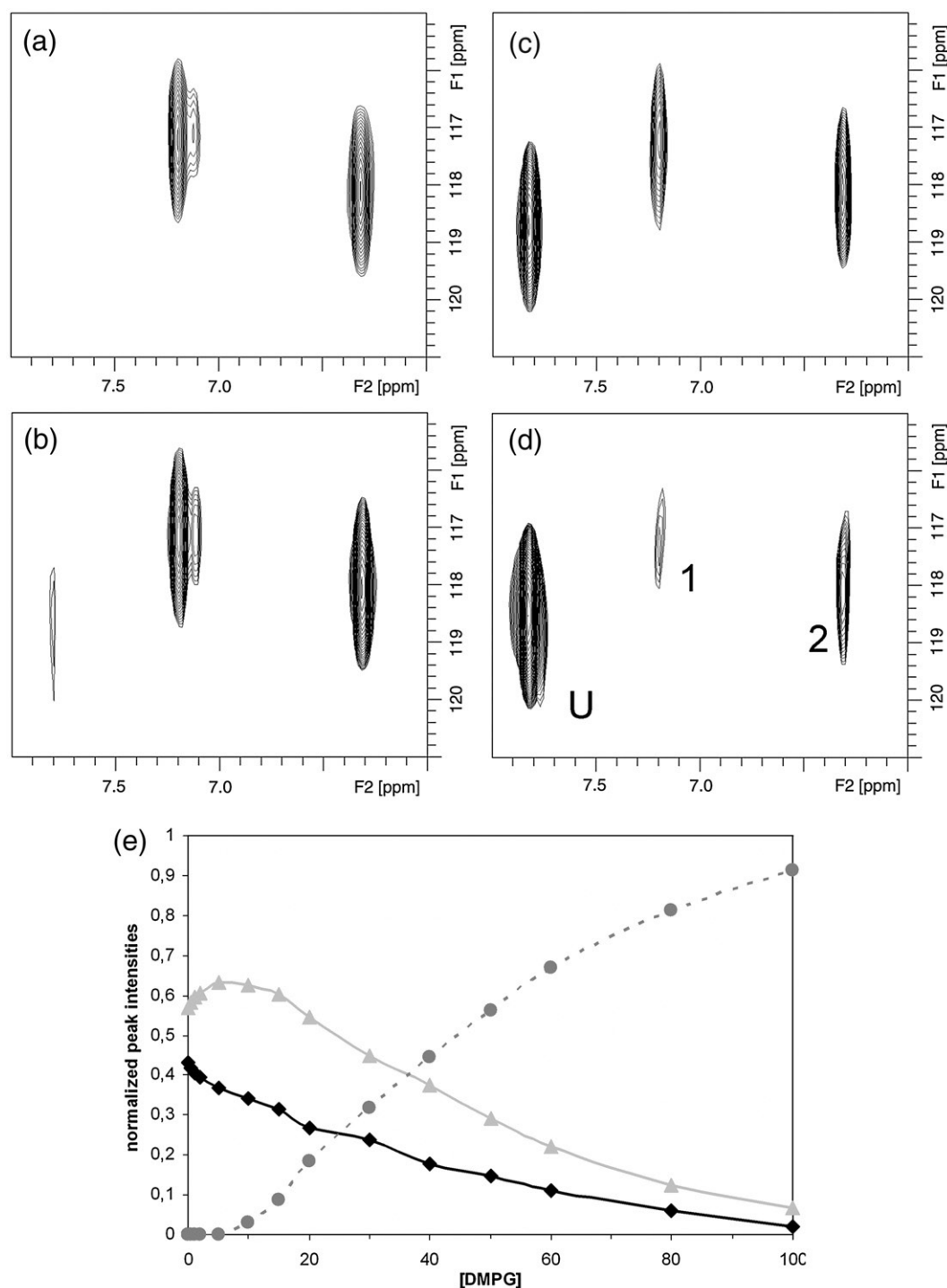


Fig. 3. $[^1\text{H}, ^{15}\text{N}]$ -TROSY spectra of (a) L-BABP:DMPG: ^{15}N -GCDA 1:0:1, (b) 1:10:1, (c) 1:50:1, and (d) 1:100:1. (e) Plot of the normalized ^{15}N -GCDA peak intensity ((♦) ligand in site 1, (▲) ligand in site 2, and (●) unbound ligand) as a function of DMPG concentration. Peaks denoted as U, 1 and 2 represent the unbound ligand, and ligand bound to site 1 and to site 2, respectively.

decay, a number of signals showed a different trend (Fig. 6a). The general decay is due to (i) sample dilution, (ii) increased viscosity of the sample on addition of phospholipids, and (iii) onset of the protein-membrane association equilibrium, occurring in a slow-to-intermediate exchange regime, as evidenced by the observed line broadening. The

exchange regime between the holo protein and the membrane-bound apo protein implies that the observed signal at the position of the holo protein resonance is partly mixed with the corresponding resonance of the membrane-bound apo protein. The different intensity decay then reflects the interaction of residues in the latter protein form with the

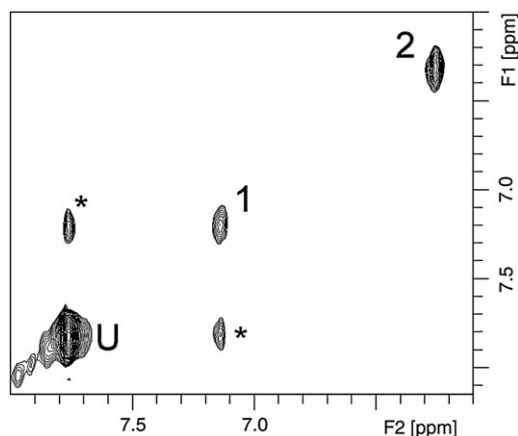


Fig. 4. The 2D ^{15}N -edited NOESY of L-BABP:DMPG: $^{15}\text{NGCDA}$ 1:50:1 acquired with a mixing time of 20 ms at 600 MHz. Peaks denoted U, 1, and 2 represent the unbound $^{15}\text{NGCDA}$, and $^{15}\text{NGCDA}$ bound to site 1 and to site 2, respectively. The asterisks denote exchange peaks.

liposomes. The resonances exhibiting 50% intensity reduction already at a P: DMPG:L ratio of 1:20:1, namely those of S3, G4, Q41, Q42, K43, D45, K103, and G104, correspond to residues located at the protein anti-portal region. We infer that these residues identify a protein patch interacting with the negatively charged membrane (Fig. 6b). This finding is in agreement with previously reported data of a 30 ns molecular dynamics simulation of L-BABP in the presence of dilauroylphosphoserine (DLPS),¹⁶ showing that Q41, Q42, K43 and K66 were consistently in contact with the membrane in the final stages of the simulations. In the computational study, however, no indication of protein unfolding was observed.

Previous urea unfolding studies on apo-L-BABP showed that the unfolding profiles of the majority of protein residues were consistent with a two-state unfolding process.²² However, few amino acids displayed a singular behavior, not directly amenable to the presence of a folding intermediate, which was

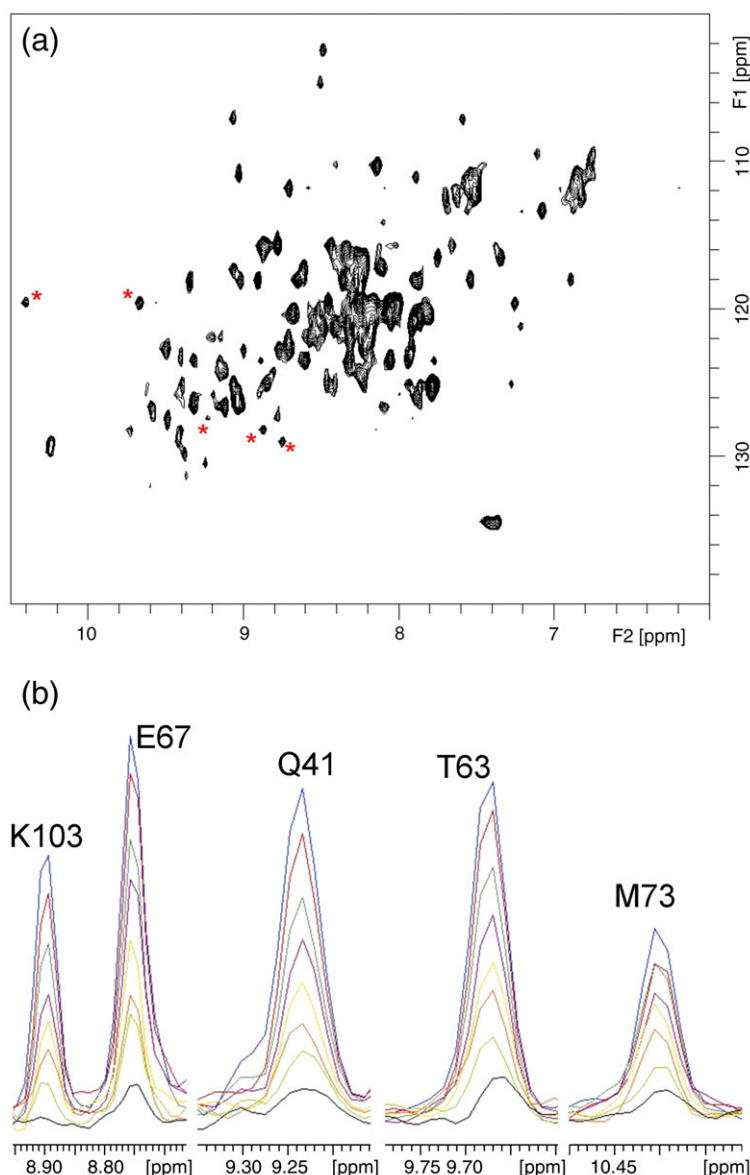


Fig. 5. (a) The $[\text{}^1\text{H}, \text{}^{15}\text{N}]$ -TROSY spectrum of ^{15}N -L-BABP:DMPG:GCDA 1:100:1 (asterisks denote peaks whose ^1H cross-sections are reported in (b)). (b) The 1D cross-sections of selected resonances from spectra recorded on a 1:1 ^{15}N -L-BABP:GCDA complex in the presence of incremental amounts of DMPG, at protein:DMPG ratios in the range of 1:10 (blue) to 1:100 (black).

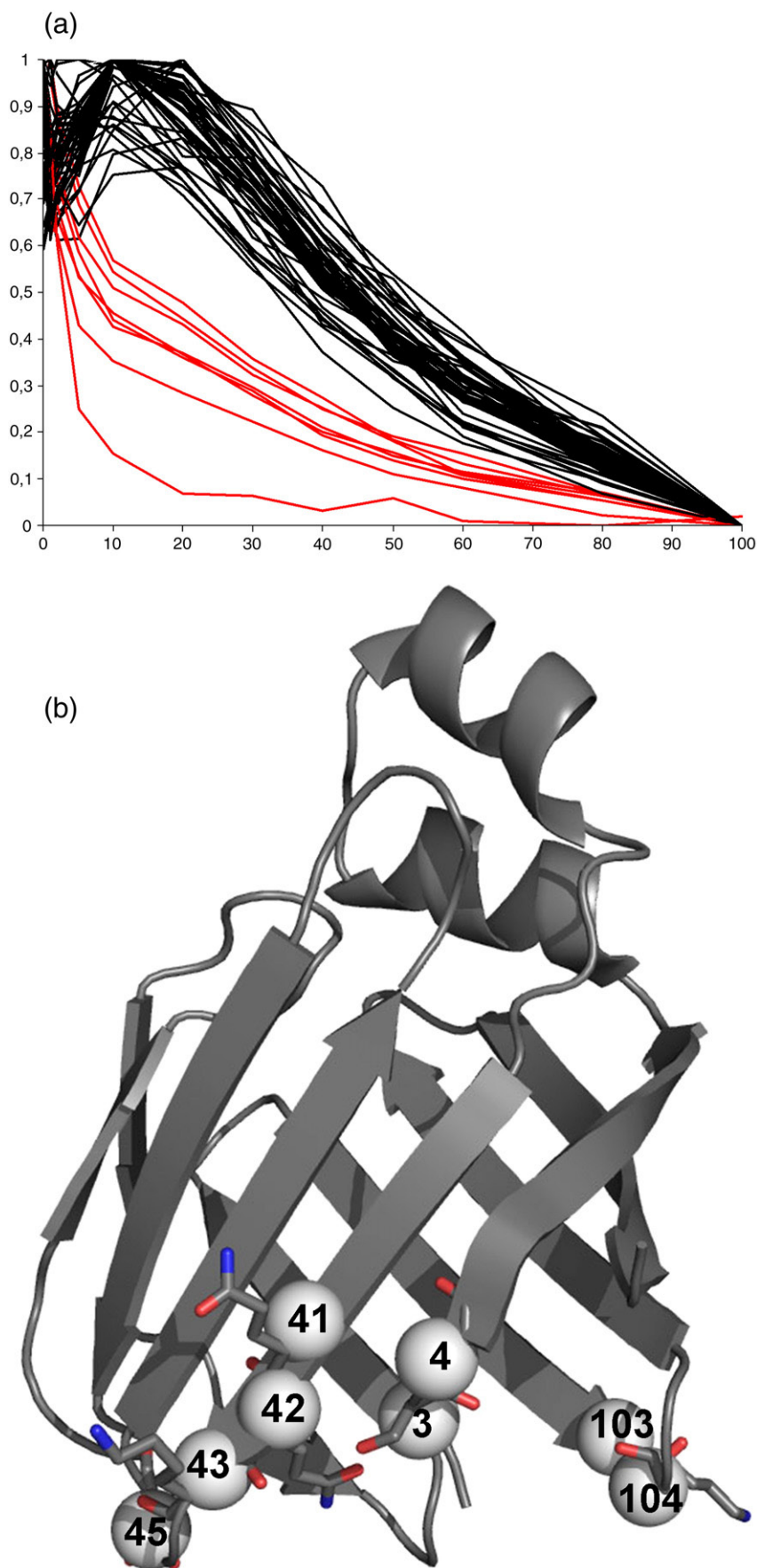


Fig. 6. Protein residues interacting with liposomes. Plot of protein amide normalized peak intensities *versus* the concentration of DMPG: two groups are identified on the basis of the peak intensity change upon the addition of DMPG. (a) Red curves correspond to residues exhibiting 50% intensity decrease at L-BABP:DMPG:GCDA of 1:20:1, black curves represent the behavior of the remaining residues. (b) Residues corresponding to the red curves in (a) are drawn as sticks and their backbone N_H atoms are shown as spheres on the protein structure. The residue number is indicated.

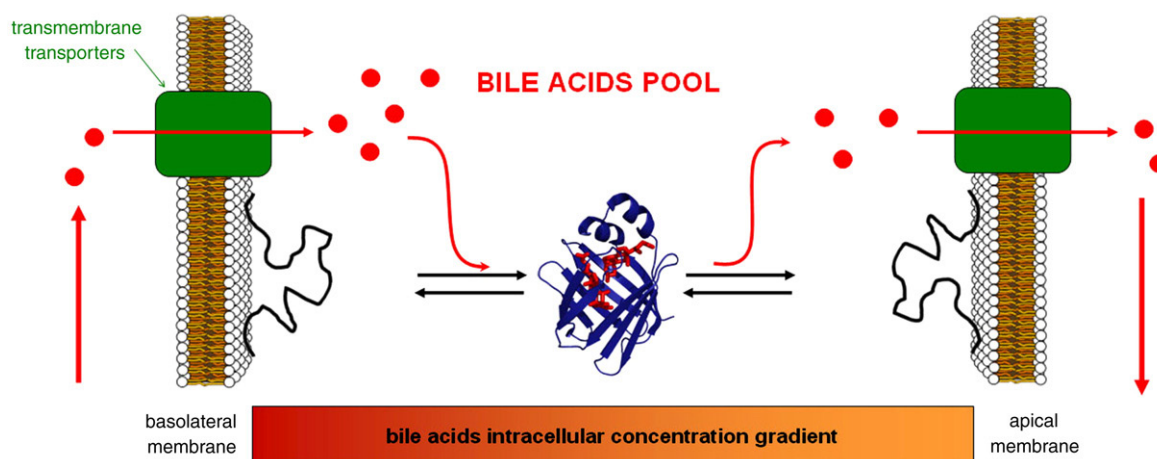
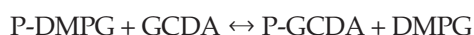
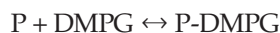


Fig. 7. Cartoon depicting the proposed mechanism for bile acid binding and release, dependent on the onset of an intracellular GCDA gradient in cells exhibiting vectorial bile acid transport. The PDB code of the folded L-BABP:GCDA complex is 2JN3.

interpreted as due to an interplay between dynamic features and formation of transient aggregates. The identity of these residues is different from that of the residues outlined here that show stronger intensity decrease on liposome addition. This notion rules out the possibility that the observed intensity trend reflects some greater tendency to lose structure or participation in protein-protein interactions.

Taken together, these data are consistent with a global picture in which L-BABP (P) is in equilibrium between an apo membrane-bound unfolded state (P-DMPG) and a holo free-in-solution state (P-GCDA):



Phospholipid vesicles and bile salt ligands are thus exclusive competitors for binding to the protein.²³

Discussion

L-BABP/membrane/lipid interactions are largely unexplored. There are a number of reports on the interaction of fatty acid binding proteins (FABPs) with phospholipid membranes (micelles or liposomes).^{23,24,25} While the intracellular transport of bile acids and fatty acids follows different routes, the putative cytosolic carriers of both types of ligands belong to the same protein family and display strong structural similarity. Studies on FABPs show that, despite a strong similarity in the overall fold of the proteins in the family, the different distribution of amino acid sidechains determines distinct membrane binding properties. Membrane-active FABPs interact directly with membranes during fatty acid uptake and release, while membrane-inactive FABPs exchange fatty acids with the aqueous solution.

Here, we show that L-BABP is able to interact with anionic phospholipid membranes in its apo form, while it appears to have little or no affinity for phospholipids after binding bile salts. Thus, the membrane and the ligand establish competing binding equilibria that are part of the factors that modulate the cytoplasmic permeability of bile acids. The ensemble of NMR titration experiments on the present system allows us to propose a model (Fig. 7) describing a GCDA gradient formed in the cytosol. GCDA accumulates in proximity of the basolateral membrane, where GCDA is internalized through the action of membrane transporters, while its concentration decreases in proximity of the apical membrane, where an efficient bile acids efflux is provided. The GCDA gradient in hepatocytes would shift the equilibrium toward the holo free-in-solution state in the proximity of the basolateral membrane, enabling simultaneous detachment from the membrane and bile acids loading of the protein, and toward the apo membrane-bound state in the proximity of the apical membrane allowing the release of bile acids. The GCDA gradient could thus be at the basis of the binding and release mechanism, ensuring bile acid vectorial transport between membranes. The membrane-active behavior of L-BABP would allow the cell to best control the efficiency of diffusional mobility.

The present model of a GCDA gradient formation is consistent with the physiological data of the vectorial transport taking place in polar hepatocyte cells, a mechanism contributing to ensure efficient enterohepatic circulation.²⁶ Because of a complex cellular architecture, hepatocyte polarity is not easy to study and is difficult to maintain *in vitro*. In this line, our study provides a description that complements the reported data and represents an important result to be tested on further BABPs.

It was shown earlier that both ileal and liver BABPs are able to bind two bile salt molecules in a highly cooperative manner.^{27,28} This observation

might suggest that these BABPs act as intracellular lipid chaperones and as cytosolic buffer agents against toxic levels of free bile salts. In the presence of such a strong affinity for the ligand, there must be an efficient mechanism of bile salt release. The possibility that the acidic conditions found in the proximity of cellular membranes could trigger conformational changes that are responsible of a decreased affinity for bile salts has been explored.²⁰ The NMR data showed no apparent major structural rearrangement of the holo protein or a clear evidence of ligand release on lowering the pH, so it is still possible that some unfolded protein conformation with little or no affinity for the ligands is stabilized in acidic conditions, favoring dissociation of the complex, as proposed.¹⁶ Whether the specific pH conditions at the water-membrane interface are determinant in ligand release, our data point to the existence of a prevalent unfolded conformation when the protein binds to the membrane. In our model, the latter structure becomes energetically favored when the concentration of free ligand is lower, e.g. in proximity of the bile salt metabolic sink, thus determining a decrease in the population of ligand-bound conformation. Unfolding effects simply due to macromolecular crowding can be ruled out because we have verified that high molecular mass polyethylene glycol and analogous compounds do not induce protein unfolding.

A few models of FABP-membrane interaction have been proposed. For example, the interaction of liver FABP with anionic phospholipids at low ionic strength was shown to involve both electrostatic and non-polar interactions.^{29,30} A phospholipid molecule is then suggested to interact with the central cavity of the protein, replacing the fatty acid ligand. The interaction with the membrane appears to be mediated by residues located at the portal region of the protein, with the N-terminal region undergoing a conformational change. However, clear evidence of a structural rearrangement is missing, and could be transient in nature. In the case of L-BABP, the present results point to a binding epitope located in the anti-portal region where basic and polar residues are likely to establish ionic interactions with negatively charged phospholipids. These data provide experimental validation of an earlier computational study.¹⁶ Furthermore, the spectral fingerprint of the membrane-bound protein describes a molecular species with poor tertiary fold.

In terms of membrane interaction, the ensemble of data bears some correlation with the well-characterized cytochrome *c*,³¹ suggesting that L-BABP may behave essentially as a peripheral membrane protein in its unliganded state. The same protein in its ligated state acquires a full cytosolic nature. It has to be mentioned that ileal bile acid binding proteins from other species have been identified as peripheral membrane-associated,³² and functionally associated with proteins involved in the uptake of bile acids by the intestinal brush border membrane.³³ These data, together with our results, suggest that the common capacity of bile acid binding protein to associate with

membranes can have a functional relevance in the mechanism of cytosolic bile acids transport in the liver and in the intestine. Further investigation of related proteins and membranes of different compositions in the presence of putative ligands or xenobiotic drugs are needed for a more detailed characterization of this intriguing lipid-carrier system.

Materials and Methods

Sample production and purification

L-BABP was expressed and purified as described.¹⁴ For the production of ¹⁵N-labeled L-BABP, host cells were grown in M9 minimal medium using ¹⁵NH₄Cl as sole nitrogen source. All the protein preparations were checked by 1D ¹H NMR before and after delipidation and reproducible spectra were always obtained for the apo and undelipidated forms. The molecular mass of the proteins and the extent of labeling were verified by matrix-assisted laser desorption/ionization (MALDI) mass analysis. ¹⁵N-enriched glycochenodeoxycholic acid was prepared as described.¹⁹

Liposomes preparation

Lyophilized lipid sodium salts were used as starting material. The lipid used was 1,2-dimyristoyl-sn-glycero-3-phospho-(1'-rac-glycerol) (DMPG) obtained from Avanti Polar Lipids. The required amount of DMPG was hydrated in distilled water containing 10% (v/v) ²H₂O. The lipid suspension was then heated to 37 °C and processed by several sonication and stirring steps until all aggregates were dissolved. Liposomes were prepared by extrusion through a polycarbonate filter (pore diameter 100 nm) using an extruder device from Avanti Polar Lipids.

Sample preparation

Typical protein samples contained 0.2 mM protein dissolved in 30 mM phosphate (Na₂HPO₄/NaH₂PO₄), 90% (v/v) H₂O, 10% (v/v) ²H₂O, pH 7.0. Samples used to monitor the L-BABP/liposomes interaction were prepared by resuspending a lyophilized L-BABP aliquot directly in a liposome solution of known concentration. The L-BABP:DMPG ratio was adjusted to the desired value by addition of a concentrated liposome solution. A 120 mM DMPG solution was used for the titration experiments starting from a 0.23 mM L-BABP sample.

Filtration assay

Samples of pure L-BABP and L-BABP/DMPG liposomes mixture (1:100 protein/lipid ratio (mol/mol)) were loaded into the upper chamber of Amicon® Ultra 50K devices (Millipore). The filtration process was carried out with centrifugation at 500 g at 303 K until 50% of the initial volume was eluted. The concentration of protein was quantified by measuring the absorbance at 280 nm.

NMR spectroscopy

Measurements were recorded on Bruker Avance spectrometers operating at 600.13 MHz, 700.13 MHz and

800.13 MHz all equipped with a cryoprobe. All [^1H , ^{15}N]-TROSY,³⁴ [^1H , ^{15}N]-HSQC, and 2D ^{15}N -edited nuclear Overhauser effect spectroscopy (NOESY) experiments (mixing times of 20 ms and 100 ms) were done using standard Bruker library sequences. Experiments acquired in the presence of DMPG liposomes, were recorded at 303 K where the bilayer is in a liquid-crystalline phase.¹⁷ Standard sequence schemes with pulsed field gradients were used to achieve suppression of the solvent signal and spectral artifacts. Direct and indirect dimensions were normally apodized by use of 90°-shifted squared sine-bell functions or Lorentzian-to-Gaussian functions, followed by zero filling and Fourier transform. The NMR data were processed with Topspin 2.1 (Bruker) and analyzed with the same software or with SPARKY 3.11 (T.D. Goddard and D.G. Kneller, SPARKY3, University of Carolina, San Francisco).

Data bank codes

Assignments for apo- and holo-L-BABP are available from the Biological Magnetic Resonance Data Bank (BMRB codes 6642 and 15084).

Acknowledgements

EU-NMR Infrastructures (contract # RII3-026145) are acknowledged for providing access to NMR spectrometers. This research was supported by FIRB 2003 (Project no. RBNE03PX83) and the Cariverona Foundation. The University of Verona is acknowledged for financial support in the acquisition of the Bruker NMR Avance 600 MHz spectrometer equipped with a cryoprobe. Sina Reckel is acknowledged for help in the preparation of liposomes.

Supplementary Data

Supplementary data associated with this article can be found, in the online version, at [doi:10.1016/j.jmb.2009.10.014](https://doi.org/10.1016/j.jmb.2009.10.014)

References

- Alrefai, W. A. & Gill, R. K. (2007). Bile acid transporters: structure, function, regulation and pathophysiological implications. *Pharm. Res.* **24**, 1803–1823.
- Chiang, J. Y. (2009). Bile acids: regulation of synthesis. *J. Lipid Res.* **50**, 1955–1966.
- Pellicciari, R., Camaioni, E., Costantino, G., Formentini, L., Sabbatini, P., Venturoni, F. *et al.* (2008). On the way to selective PARP-2 inhibitors. Design, synthesis, and preliminary evaluation of a series of isoquinolinone derivatives. *ChemMedChem*, **3**, 914–923.
- Kosters, A. & Karpen, S. J. (2008). Bile acid transporters in health and disease. *Xenobiotica*, **38**, 1043–1071.
- Kullak-Ublick, G. A., Stieger, B. & Meier, P. J. (2004). Enterohepatic bile salt transporters in normal physiology and liver disease. *Gastroenterology*, **126**, 322–342.
- Weisiger, R. A. (1996). When is a carrier not a membrane carrier? The cytoplasmic transport of amphipathic molecules. *Hepatology*, **24**, 1288–1295.
- Pielak, G. J., Li, C., Miklos, A. C., Schlesinger, A. P., Slade, K. M., Wang, G. F. & Zigoneanu, I. G. (2009). Protein nuclear magnetic resonance under physiological conditions. *Biochemistry*, **48**, 226–234.
- Selenko, P. & Wagner, G. (2007). Looking into live cells with in-cell NMR spectroscopy. *J. Struct. Biol.* **158**, 244–253.
- Serber, Z., Selenko, P., Hansel, R., Reckel, S., Lohr, F., Ferrell, J. E., Jr *et al.* (2006). Investigating macromolecules inside cultured and injected cells by in-cell NMR spectroscopy. *Nat. Protoc.* **1**, 2701–2709.
- Xu, Y., Yushmanov, V. E. & Tang, P. (2002). NMR studies of drug interaction with membranes and membrane-associated proteins. *Biosci. Rep.* **22**, 175–196.
- Vasile, F., Ragona, L., Catalano, M., Zetta, L., Perduca, M., Monaco, H. & Molinari, H. (2003). Solution structure of chicken liver basic fatty acid binding protein. *J. Biomol. NMR*, **25**, 157–160.
- Eliseo, T., Ragona, L., Catalano, M., Assfalg, M., Paci, M., Zetta, L. *et al.* (2007). Structural and dynamic determinants of ligand binding in the ternary complex of chicken liver bile acid binding protein with two bile salts revealed by NMR. *Biochemistry*, **46**, 12557–12567.
- Nichesola, D., Perduca, M., Capaldi, S., Carrizo, M. E., Righetti, P. G. & Monaco, H. L. (2004). Crystal structure of chicken liver basic fatty acid-binding protein complexed with cholic acid. *Biochemistry*, **43**, 14072–14079.
- Ragona, L., Catalano, M., Luppi, M., Cicero, D., Eliseo, T., Foote, J. *et al.* (2006). NMR dynamic studies suggest that allosteric activation regulates ligand binding in chicken liver bile acid-binding protein. *J. Biol. Chem.* **281**, 9697–9709.
- Tomaselli, S., Ragona, L., Zetta, L., Assfalg, M., Ferranti, P., Longhi, R. *et al.* (2007). NMR-based modeling and binding studies of a ternary complex between chicken liver bile acid binding protein and bile acids. *Proteins: Struct. Funct. Genet.* **69**, 177–191.
- Villarreal, M. A., Perduca, M., Monaco, H. L. & Montich, G. G. (2008). Binding and interactions of L-BABP to lipid membranes studied by molecular dynamic simulations. *Biochim. Biophys. Acta*, **1778**, 1390–1397.
- Galassi, V., Nolan, V., Villarreal, M. A., Perduca, M., Monaco, H. L. & Montich, G. G. (2009). Kinetics of lipid-membrane binding and conformational change of L-BABP. *Biochem. Biophys. Res. Commun.* **382**, 771–775.
- Decca, M. B., Perduca, M., Monaco, H. L. & Montich, G. G. (2007). Conformational changes of chicken liver bile acid-binding protein bound to anionic lipid membrane are coupled to the lipid phase transitions. *Biochim. Biophys. Acta*, **1768**, 1583–1591.
- Nolan, V., Perduca, M., Monaco, H. L., Maggio, B. & Montich, G. G. (2003). Interactions of chicken liver basic fatty acid-binding protein with lipid membranes. *Biochim. Biophys. Acta*, **1611**, 98–106.
- Pedò, M., D'Onofrio, M., Ferranti, P., Molinari, H. & Assfalg, M. (2009). Towards the elucidation of molecular determinants of cooperativity in the liver bile acid binding protein. *Proteins: Struct. Funct. Genet.* **77**, 718–731.
- Schneider, M. F., Marsh, D., Jahn, W., Kloesgen, B. & Heimburg, T. (1999). Network formation of lipid membranes: triggering structural transitions by chain melting. *Proc. Natl Acad. Sci. USA*, **96**, 14312–14317.
- D'Onofrio, M., Ragona, L., Fessas, D., Signorelli, M., Ugolini, R., Pedo, M. *et al.* (2009). NMR unfolding studies on a liver bile acid binding protein reveal a global two-state unfolding and localized singular behaviors. *Arch. Biochem. Biophys.* **481**, 21–29.

23. Storch, J. & Bass, N. M. (1990). Transfer of fluorescent fatty acids from liver and heart fatty acid-binding proteins to model membranes. *J. Biol. Chem.* **265**, 7827–7831.
24. Kim, H. K. & Storch, J. (1992). Free fatty acid transfer from rat liver fatty acid-binding protein to phospholipid vesicles. Effect of ligand and solution properties. *J. Biol. Chem.* **267**, 77–82.
25. Kim, H. K. & Storch, J. (1992). Mechanism of free fatty acid transfer from rat heart fatty acid-binding protein to phospholipid membranes. Evidence for a collisional process. *J. Biol. Chem.* **267**, 20051–20056.
26. Cassio, D., Macias, R. I., Grosse, B., Marin, J. J. & Monte, M. J. (2007). Expression, localization, and inducibility by bile acids of hepatobiliary transporters in the new polarized rat hepatic cell lines, Can 3-1 and Can 10. *Cell Tissue Res.* **330**, 447–460.
27. Tomaselli, S., Zanzoni, S., Ragona, L., Gianolio, E., Aime, S., Assfalg, M. & Molinari, H. (2008). Solution structure of the supramolecular adduct between a liver cytosolic bile acid binding protein and a bile acid-based gadolinium(III)-chelate, a potential hepatospecific magnetic resonance imaging contrast agent. *J. Med. Chem.* **51**, 6782–6792.
28. Tochtrop, G. P., Richter, K., Tang, C., Toner, J. J., Covey, D. F. & Cistola, D. P. (2002). Energetics by NMR: site-specific binding in a positively cooperative system. *Proc. Natl Acad. Sci. USA*, **99**, 1847–1852.
29. Davies, J. K., Thumser, A. E. & Wilton, D. C. (1999). Binding of recombinant rat liver fatty acid-binding protein to small anionic phospholipid vesicles results in ligand release: a model for interfacial binding and fatty acid targeting. *Biochemistry*, **38**, 16932–16940.
30. Hagan, R. M., Worner-Gibbs, J. & Wilton, D. C. (2008). The interaction of liver fatty-acid-binding protein (FABP) with anionic phospholipid vesicles: is there extended phospholipid anchorage under these conditions? *Biochem. J.* **410**, 123–129.
31. Tuominen, E. K., Wallace, C. J. & Kinnunen, P. K. (2002). Phospholipid-cytochrome c interaction: evidence for the extended lipid anchorage. *J. Biol. Chem.* **277**, 8822–8826.
32. Kramer, W., Girbig, F., Gutjahr, U., Kowalewski, S., Jouvenal, K., Muller, G. *et al.* (1993). Intestinal bile acid absorption. Na(+)-dependent bile acid transport activity in rabbit small intestine correlates with the coexpression of an integral 93-kDa and a peripheral 14-kDa bile acid-binding membrane protein along the duodenum-ileum axis. *J. Biol. Chem.* **268**, 18035–18046.
33. Nakahara, M., Furuya, N., Takagaki, K., Sugaya, T., Hirota, K., Fukamizu, A. *et al.* (2005). Ileal bile acid-binding protein, functionally associated with the farnesoid X receptor or the ileal bile acid transporter, regulates bile acid activity in the small intestine. *J. Biol. Chem.* **280**, 42283–42289.
34. Pervushin, K., Riek, R., Wider, G. & Wuthrich, K. (1997). Attenuated T2 relaxation by mutual cancellation of dipole-dipole coupling and chemical shift anisotropy indicates an avenue to NMR structures of very large biological macromolecules in solution. *Proc. Natl Acad. Sci. USA*, **94**, 12366–12371.

SUPPLEMENTARY MATERIAL

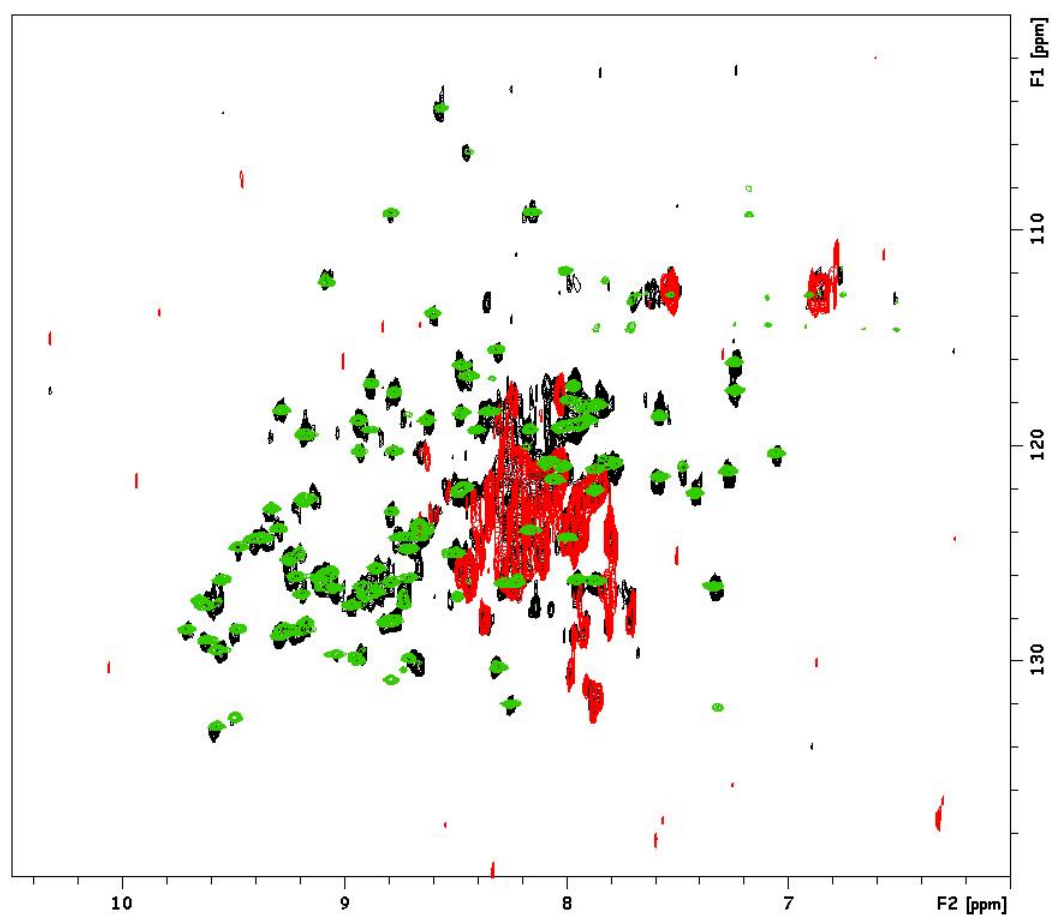


Figure S1. Superimposed $[^1\text{H},^{15}\text{N}]$ -TROSY spectra of: L-BABP:DMPG 1:100 (red), the same sample after addition of 0.1M NaCl (black), and apo L-BABP (green). Experiments were performed on 0.3mM protein in 20mM phosphate buffer, pH 7.0, at 600 MHz and 30°C.

CHAPTER 8: Conclusions

In recent years bile acids have gained renewed interest because of their involvement in several important biological mechanisms. Bile acids are well known detergent molecules acting, in the intestinal tract, as fat and vitamins solubilising agents. However bile acids present another, even most intriguing property, that is the ability to act as signalling molecules. The involvement in significant processes makes bile acids and related metabolism a relevant topic to investigate.

The transport mechanism of bile acids in the body have been characterized and several transmembrane transporters have been identified. The intracellular transport has been also investigated, however our knowledge about protein mediated intracellular transport is still low.

The study presented here is aimed at the NMR characterization of the bile acid intracellular transporter present in the chicken liver, namely cL-BABP. This protein can be considered as a model for the class of intracellular bile acid transporters belonging to the intracellular lipid binding protein (iLBP) family.

The structural stability and the unfolding mechanism was initially studied through hydrogen-deuterium (HX) exchange experiments and urea unfolding experiments in equilibrium conditions. The most stable protein region was identified in the A-C, I and J strands and a ΔG_{op} of 7.78 ± 0.03 kcal mol⁻¹ was estimated. During the urea induced protein unfolding experiments some signals displayed a sharp decrease in intensity already at low urea concentration. The corresponding residues are located in the G and H strands and in DE, EF and FG loops, which are regions displaying a high mobility in the folded structure. These residues were suggested to represent an unfolding initiation site during the urea induced protein denaturation process. The protein unfolding behaviour was different from other members of the iLBP family so that it appears that closely related proteins present different unfolding pathways and stability although sharing a high three-dimensional structure similarity.

The project continued with the analysis of the cooperative binding mechanism that was previously proposed to govern the ligand binding process. We produced and analyzed a mutated cL-BABP protein carrying a glutamine (Q) residue at 98 position, substituting the original histidine (H). The mutant H98Q, was thoroughly analyzed, in its apo and holo form, through several NMR spectroscopy experiments and by mean of mass spectrometry (MS). This investigation indicated a disruption of the cooperativity, resulting in the reduction of the affinity of the ligand for binding site 1. The decreased affinity for the ligand bound in site 1 was further confirmed through diffusion measurements: an increased diffusion coefficient (D), approaching the value of the unbound ligand, was indeed detected for the ligand bound to site 1. In addition MS analysis revealed the presence of a mass entity correspondent to a singly ligated protein, again indicating a disruption of the cooperativity. The structural alterations deriving from the substitution of His 98 were mainly located in the C-terminal half of the protein although long distance effect were identified for the holo protein in the 55-59 segment which is close to the ligand bound to site 1. The mutation also affected the protein mobility as a quenching of the slow motions observed for the WT protein in D, E, G, H and I strands was observed.

The analysis of the H98Q mutant pointed out the involvement of His 98 in long distance interactions and in the determination of the mobility observed in the WT protein. Finally the alteration of the two site ligand binding confirmed the central role of His 98 in the cooperative bile acid binding mechanism. Successively we studied a naturally occurring variant of cL-BABP carrying a cysteine, instead of a threonine, in position 91. The form carrying cysteine, named cL-BABP/S-S, presents a disulfide bridge between the cysteines in position 80 and 91. The effect of the presence of the disulfide bridge on the protein structure and binding was investigated, resulting in the introduction of a site-selective binding mechanism in the presence of a mixture of GCDA and GCA.

The final part of the project was focused on the study of the interaction of cL-BABP with anionic vesicles in the presence of bile acids. cL-BABP was shown to associate, via electrostatic interaction, with 1,2-dimyristoyl-sn-glycero-3-phospho-(1'-rac-glycerol) (DMPG) liposomes and to undergo partial unfolding. The ^1H - ^{15}N TROSY spectrum of the liposome associated ^{15}N cL-BABP was indeed characterized by broadened signals with limited frequency dispersion. The interaction with the membrane mimetics was shown to be weakened in the presence of the physiological ligand. Indeed, upon GCDA addition, the spectrum corresponding to a correctly folded holo protein was obtained. Holo cL-BABP was also shown to interact with DMPG liposomes suggesting the presence of an equilibrium that involves the membrane associated apo cL-BABP and the free in solution holo cL-BABP. Holo cL-BABP was shown to release again its ligand upon interaction with DMPG liposomes and the site 1 was identified as the preferential position for ligand release. The residues involved in the liposome interaction were also identified and, mapped on the protein structure, they localized in a positively charged patch on the protein surface, in the antiportal region. The characterization of the interaction of cL-BABP with anionic liposomes led us to propose a model for the protein mediated intracellular transport of bile acids in the hepatocytes. In this model bile acids transport mediated by cL-BABP is driven by the intracellular bile acids concentration gradient. The high bile acids concentration at the basolateral hepatocyte pole shifts the equilibrium toward the free in solution holo cL-BABP form allowing the detachment of the protein from the membrane and the correspondent loading of bile acids. On the contrary at the apical pole the low bile acids concentration shifts the equilibrium toward the membrane associated apo cL-BABP determining the release of the carried bile acids.

As the fluidity and charge of natural membranes were shown to represent an important factor in biological processes, the aim of the future work will be the investigation of the role played by different membrane mimetic vesicles on the described intracellular transport mechanism.

Bibliography

1. Capaldi, S. et al. The X-ray structure of zebrafish (*Danio rerio*) ileal bile acid-binding protein reveals the presence of binding sites on the surface of the protein molecule. *J Mol Biol* 385, 99-116 (2009).
2. Vaezeslami, S., Mathes, E., Vasileiou, C., Borhan, B. & Geiger, J.H. The structure of Apo-wild-type cellular retinoic acid binding protein II at 1.4 Å and its relationship to ligand binding and nuclear translocation. *J Mol Biol* 363, 687-701 (2006).
3. Thomas, C., Pellicciari, R., Pruzanski, M., Auwerx, J. & Schoonjans, K. Targeting bile-acid signalling for metabolic diseases. *Nat Rev Drug Discov* 7, 678-93 (2008).
4. Hofmann, A.F. Chemistry and enterohepatic circulation of bile acids. *Hepatology* 4, 4S-14S (1984).
5. Monte, M.J., Marin, J.J., Antelo, A. & Vazquez-Tato, J. Bile acids: chemistry, physiology, and pathophysiology. *World J Gastroenterol* 15, 804-16 (2009).
6. Furuhashi, M. & Hotamisligil, G.S. Fatty acid-binding proteins: role in metabolic diseases and potential as drug targets. *Nat Rev Drug Discov* 7, 489-503 (2008).
7. Toke, O. et al. Determinants of cooperativity and site selectivity in human ileal bile acid binding protein. *Biochemistry* 45, 727-37 (2006).
8. Thompson, J.D., Higgins, D.G. & Gibson, T.J. CLUSTAL W: improving the sensitivity of progressive multiple sequence alignment through sequence weighting, position-specific gap penalties and weight matrix choice. *Nucleic Acids Res* 22, 4673-80 (1994).
9. Eliseo, T. et al. Structural and dynamic determinants of ligand binding in the ternary complex of chicken liver bile acid binding protein with two bile salts revealed by NMR. *Biochemistry* 46, 12557-67 (2007).
10. Hofmann, A.F. & Hagey, L.R. Bile acids: chemistry, pathochemistry, biology, pathobiology, and therapeutics. *Cell Mol Life Sci* 65, 2461-83 (2008).
11. Coello, A., Meijide, F., Nunez, E.R. & Tato, J.V. Aggregation behavior of bile salts in aqueous solution. *J Pharm Sci* 85, 9-15 (1996).
12. Russell, D.W. The enzymes, regulation, and genetics of bile acid synthesis. *Annu Rev Biochem* 72, 137-74 (2003).
13. Hylemon, P.B. & Harder, J. Biotransformation of monoterpenes, bile acids, and other isoprenoids in anaerobic ecosystems. *FEMS Microbiol Rev* 22, 475-88 (1998).
14. Hofmann, A.F. et al. A proposed nomenclature for bile acids. *J Lipid Res* 33, 599-604 (1992).

15. Chiang, J.Y. Regulation of bile acid synthesis: pathways, nuclear receptors, and mechanisms. *J Hepatol* 40, 539-51 (2004).
16. Hofmann, A.F. The continuing importance of bile acids in liver and intestinal disease. *Arch Intern Med* 159, 2647-58 (1999).
17. Chiang, J.Y. Bile acid regulation of gene expression: roles of nuclear hormone receptors. *Endocr Rev* 23, 443-63 (2002).
18. Perissi, V. & Rosenfeld, M.G. Controlling nuclear receptors: the circular logic of cofactor cycles. *Nat Rev Mol Cell Biol* 6, 542-54 (2005).
19. Song, C., Hiipakka, R.A. & Liao, S. Selective activation of liver X receptor alpha by 6alpha-hydroxy bile acids and analogs. *Steroids* 65, 423-7 (2000).
20. Francis, G.A., Fayard, E., Picard, F. & Auwerx, J. Nuclear receptors and the control of metabolism. *Annu Rev Physiol* 65, 261-311 (2003).
21. Chiang, J.Y., Kimmel, R., Weinberger, C. & Stroup, D. Farnesoid X receptor responds to bile acids and represses cholesterol 7alpha-hydroxylase gene (CYP7A1) transcription. *J Biol Chem* 275, 10918-24 (2000).
22. Kullak-Ublick, G.A., Stieger, B. & Meier, P.J. Enterohepatic bile salt transporters in normal physiology and liver disease. *Gastroenterology* 126, 322-42 (2004).
23. Horie, T., Mizuma, T., Kasai, S. & Awazu, S. Conformational change in plasma albumin due to interaction with isolated rat hepatocyte. *Am J Physiol* 254, G465-70 (1988).
24. Trauner, M. & Boyer, J.L. Bile salt transporters: molecular characterization, function, and regulation. *Physiol Rev* 83, 633-71 (2003).
25. Hagenbuch, B. & Dawson, P. The sodium bile salt cotransport family SLC10. *Pflugers Arch* 447, 566-70 (2004).
26. Zhu, Q.S. et al. Inhibition of human m-epoxide hydrolase gene expression in a case of hypercholanemia. *Biochim Biophys Acta* 1638, 208-16 (2003).
27. Hagenbuch, B. & Meier, P.J. The superfamily of organic anion transporting polypeptides. *Biochim Biophys Acta* 1609, 1-18 (2003).
28. Hagenbuch, B. & Meier, P.J. Organic anion transporting polypeptides of the OATP/ SLC21 family: phylogenetic classification as OATP/ SLCO superfamily, new nomenclature and molecular/functional properties. *Pflugers Arch* 447, 653-65 (2004).
29. Meier, P.J. & Stieger, B. Bile salt transporters. *Annu Rev Physiol* 64, 635-61 (2002).
30. Gerke, P.M. & Vore, M. Regulation of expression of the multidrug resistance-associated protein 2 (MRP2) and its role in drug disposition. *J Pharmacol Exp Ther* 302, 407-15 (2002).

-
31. Bahar, R.J. & Stolz, A. Bile acid transport. *Gastroenterol Clin North Am* 28, 27-58 (1999).
 32. Weinman, S.A., Carruth, M.W. & Dawson, P.A. Bile acid uptake via the human apical sodium-bile acid cotransporter is electrogenic. *J Biol Chem* 273, 34691-5 (1998).
 33. Amelsberg, A., Jochims, C., Richter, C.P., Nitsche, R. & Folsch, U.R. Evidence for an anion exchange mechanism for uptake of conjugated bile acid from the rat jejunum. *Am J Physiol* 276, G737-42 (1999).
 34. Monaco, H.L. The liver bile acid-binding proteins. *Biopolymers* (2009).
 35. Guariento, M. et al. Identification and functional characterization of the bile acid transport proteins in non-mammalian ileum and mammalian liver. *Proteins* 70, 462-72 (2008).
 36. Nakahara, M. et al. Ileal bile acid-binding protein, functionally associated with the farnesoid X receptor or the ileal bile acid transporter, regulates bile acid activity in the small intestine. *J Biol Chem* 280, 42283-9 (2005).
 37. Glatz, J.F. & van der Vusse, G.J. Cellular fatty acid-binding proteins: their function and physiological significance. *Prog Lipid Res* 35, 243-82 (1996).
 38. Flower, D.R. The lipocalin protein family: structure and function. *Biochem J* 318 (Pt 1), 1-14 (1996).
 39. Marcelino, A.M., Smock, R.G. & Gierasch, L.M. Evolutionary coupling of structural and functional sequence information in the intracellular lipid-binding protein family. *Proteins* 63, 373-84 (2006).
 40. Ockner, R.K., Manning, J.A., Poppenhausen, R.B. & Ho, W.K. A binding protein for fatty acids in cytosol of intestinal mucosa, liver, myocardium, and other tissues. *Science* 177, 56-8 (1972).
 41. Scapin, G. et al. Chicken liver basic fatty acid-binding protein (pI = 9.0). Purification, crystallization and preliminary X-ray data. *FEBS Lett* 240, 196-200 (1988).
 42. Gantz, I. et al. Gastrotropin: not an enterooxyntin but a member of a family of cytoplasmic hydrophobic ligand binding proteins. *J Biol Chem* 264, 20248-54 (1989).
 43. Vodenlich, A.D., Jr. et al. Identification of the 14 kDa bile acid transport protein of rat ileal cytosol as gastrotropin. *Biochem Biophys Res Commun* 177, 1147-54 (1991).
 44. Nicesola, D. et al. Crystal structure of chicken liver basic fatty acid-binding protein complexed with cholic acid. *Biochemistry* 43, 14072-9 (2004).
 45. Beringhelli, T. et al. Interaction of chicken liver basic fatty acid-binding protein with fatty acids: a ¹³C NMR and fluorescence study. *Biochemistry* 40, 12604-11 (2001).
 46. Capaldi, S. et al. Crystal structure of axolotl (*Ambystoma mexicanum*) liver bile acid-binding protein bound to cholic and oleic acid. *Proteins* 64, 79-88 (2006).
-

47. Ragona, L. et al. NMR dynamic studies suggest that allosteric activation regulates ligand binding in chicken liver bile acid-binding protein. *J Biol Chem* 281, 9697-709 (2006).
48. Likic, V.A., Juranic, N., Macura, S. & Prendergast, F.G. A "structural" water molecule in the family of fatty acid binding proteins. *Protein Sci* 9, 497-504 (2000).
49. Likic, V.A. & Prendergast, F.G. Dynamics of internal water in fatty acid binding protein: computer simulations and comparison with experiments. *Proteins* 43, 65-72 (2001).
50. Ricchiuto, P. et al. Structural and dynamic roles of permanent water molecules in ligand molecular recognition by chicken liver bile acid binding protein. *J Mol Recognit* 21, 348-54 (2008).
51. Nolan, V., Perduca, M., Monaco, H.L., Maggio, B. & Montich, G.G. Interactions of chicken liver basic fatty acid-binding protein with lipid membranes. *Biochim Biophys Acta* 1611, 98-106 (2003).
52. Villarreal, M.A., Perduca, M., Monaco, H.L. & Montich, G.G. Binding and interactions of L-BABP to lipid membranes studied by molecular dynamic simulations. *Biochim Biophys Acta* 1778, 1390-7 (2008).
53. Tochtrop, G.P., Bruns, J.L., Tang, C., Covey, D.F. & Cistola, D.P. Steroid ring hydroxylation patterns govern cooperativity in human bile acid binding protein. *Biochemistry* 42, 11561-7 (2003).
54. Tochtrop, G.P. et al. Energetics by NMR: site-specific binding in a positively cooperative system. *Proc Natl Acad Sci U S A* 99, 1847-52 (2002).
55. Hofmann, A.F. & Small, D.M. Detergent properties of bile salts: correlation with physiological function. *Annu Rev Med* 18, 333-76 (1967).
56. Knisely, A.S. et al. Hepatocellular carcinoma in ten children under five years of age with bile salt export pump deficiency. *Hepatology* 44, 478-86 (2006).
57. Parks, D.J. et al. Bile acids: natural ligands for an orphan nuclear receptor. *Science* 284, 1365-8 (1999).
58. De Fabiani, E. et al. Coordinated control of cholesterol catabolism to bile acids and of gluconeogenesis via a novel mechanism of transcription regulation linked to the fasted-to-fed cycle. *J Biol Chem* 278, 39124-32 (2003).
59. Ma, K., Saha, P.K., Chan, L. & Moore, D.D. Farnesoid X receptor is essential for normal glucose homeostasis. *J Clin Invest* 116, 1102-9 (2006).
60. Stayrook, K.R. et al. Regulation of carbohydrate metabolism by the farnesoid X receptor. *Endocrinology* 146, 984-91 (2005).

-
61. Zhang, Y. et al. Activation of the nuclear receptor FXR improves hyperglycemia and hyperlipidemia in diabetic mice. *Proc Natl Acad Sci U S A* 103, 1006-11 (2006).
 62. Haussinger, D., Hallbrucker, C., Saha, N., Lang, F. & Gerok, W. Cell volume and bile acid excretion. *Biochem J* 288 (Pt 2), 681-9 (1992).
 63. Baquet, A., Hue, L., Meijer, A.J., van Woerkom, G.M. & Plomp, P.J. Swelling of rat hepatocytes stimulates glycogen synthesis. *J Biol Chem* 265, 955-9 (1990).
 64. Saha, N., Stoll, B., Lang, F. & Haussinger, D. Effect of anisotonic cell-volume modulation on glutathione-S-conjugate release, t-butylhydroperoxide metabolism and the pentose-phosphate shunt in perfused rat liver. *Eur J Biochem* 209, 437-44 (1992).
 65. Haussinger, D. The role of cellular hydration in the regulation of cell function. *Biochem J* 313 (Pt 3), 697-710 (1996).
 66. Fromm, H., Eschler, A., Tollner, D., Canzler, H. & Schmidt, F.W. [In vivo dissolving of gall-stones: the effect of chenodeoxycholic acid. (author's transl)]. *Dtsch Med Wochenschr* 100, 1619-24 (1975).
 67. Crouse, J.R., 3rd. Hypertriglyceridemia: a contraindication to the use of bile acid binding resins. *Am J Med* 83, 243-8 (1987).
 68. Kast, H.R. et al. Farnesoid X-activated receptor induces apolipoprotein C-II transcription: a molecular mechanism linking plasma triglyceride levels to bile acids. *Mol Endocrinol* 15, 1720-8 (2001).
 69. Hirokane, H., Nakahara, M., Tachibana, S., Shimizu, M. & Sato, R. Bile acid reduces the secretion of very low density lipoprotein by repressing microsomal triglyceride transfer protein gene expression mediated by hepatocyte nuclear factor-4. *J Biol Chem* 279, 45685-92 (2004).
 70. Evans, R.M., Barish, G.D. & Wang, Y.X. PPARs and the complex journey to obesity. *Nat Med* 10, 355-61 (2004).
 71. Pineda Torra, I. et al. Bile acids induce the expression of the human peroxisome proliferator-activated receptor alpha gene via activation of the farnesoid X receptor. *Mol Endocrinol* 17, 259-72 (2003).
 72. Savkur, R.S., Bramlett, K.S., Michael, L.F. & Burris, T.P. Regulation of pyruvate dehydrogenase kinase expression by the farnesoid X receptor. *Biochem Biophys Res Commun* 329, 391-6 (2005).
 73. Kawamata, Y. et al. A G protein-coupled receptor responsive to bile acids. *J Biol Chem* 278, 9435-40 (2003).
-

-
74. Watanabe, M. et al. Bile acids induce energy expenditure by promoting intracellular thyroid hormone activation. *Nature* 439, 484-9 (2006).
 75. Pain, J.A. et al. Prevention of postoperative renal dysfunction in patients with obstructive jaundice: a multicentre study of bile salts and lactulose. *Br J Surg* 78, 467-9 (1991).
 76. Calmus, Y. et al. Differential effects of chenodeoxycholic and ursodeoxycholic acids on interleukin 1, interleukin 6 and tumor necrosis factor-alpha production by monocytes. *Hepatology* 16, 719-23 (1992).
 77. Inagaki, T. et al. Regulation of antibacterial defense in the small intestine by the nuclear bile acid receptor. *Proc Natl Acad Sci U S A* 103, 3920-5 (2006).
 78. Sewnath, M.E., van der Poll, T., van Noorden, C.J., ten Kate, F.J. & Gouma, D.J. Cholestatic interleukin-6-deficient mice succumb to endotoxin-induced liver injury and pulmonary inflammation. *Am J Respir Crit Care Med* 169, 413-20 (2004).
 79. Huang, W. et al. Nuclear receptor-dependent bile acid signaling is required for normal liver regeneration. *Science* 312, 233-6 (2006).
 80. Alpini, G. et al. Bile acid feeding induces cholangiocyte proliferation and secretion: evidence for bile acid-regulated ductal secretion. *Gastroenterology* 116, 179-86 (1999).
 81. Billington, D., Evans, C.E., Godfrey, P.P. & Coleman, R. Effects of bile salts on the plasma membranes of isolated rat hepatocytes. *Biochem J* 188, 321-7 (1980).
 82. Sokol, R.J. et al. Role of oxidant stress in the permeability transition induced in rat hepatic mitochondria by hydrophobic bile acids. *Pediatr Res* 49, 519-31 (2001).
 83. Sokol, R.J. et al. Vitamin E reduces oxidant injury to mitochondria and the hepatotoxicity of taurochenodeoxycholic acid in the rat. *Gastroenterology* 114, 164-74 (1998).
 84. Krahenbuhl, S., Talos, C., Fischer, S. & Reichen, J. Toxicity of bile acids on the electron transport chain of isolated rat liver mitochondria. *Hepatology* 19, 471-9 (1994).
 85. Sokol, R.J. et al. Human hepatic mitochondria generate reactive oxygen species and undergo the permeability transition in response to hydrophobic bile acids. *J Pediatr Gastroenterol Nutr* 41, 235-43 (2005).
 86. Sokol, R.J., Winklhofer-Roob, B.M., Devereaux, M.W. & McKim, J.M., Jr. Generation of hydroperoxides in isolated rat hepatocytes and hepatic mitochondria exposed to hydrophobic bile acids. *Gastroenterology* 109, 1249-56 (1995).
 87. Lemasters, J.J. et al. The mitochondrial permeability transition in cell death: a common mechanism in necrosis, apoptosis and autophagy. *Biochim Biophys Acta* 1366, 177-96 (1998).
-

-
88. Spivey, J.R., Bronk, S.F. & Gores, G.J. Glycochenodeoxycholate-induced lethal hepatocellular injury in rat hepatocytes. Role of ATP depletion and cytosolic free calcium. *J Clin Invest* 92, 17-24 (1993).
 89. Rodrigues, C.M., Fan, G., Wong, P.Y., Kren, B.T. & Steer, C.J. Ursodeoxycholic acid may inhibit deoxycholic acid-induced apoptosis by modulating mitochondrial transmembrane potential and reactive oxygen species production. *Mol Med* 4, 165-78 (1998).
 90. Faubion, W.A. et al. Toxic bile salts induce rodent hepatocyte apoptosis via direct activation of Fas. *J Clin Invest* 103, 137-45 (1999).
 91. Yerushalmi, B., Dahl, R., Devereaux, M.W., Gumprich, E. & Sokol, R.J. Bile acid-induced rat hepatocyte apoptosis is inhibited by antioxidants and blockers of the mitochondrial permeability transition. *Hepatology* 33, 616-26 (2001).
 92. Iizaka, T., Tsuji, M., Oyamada, H., Morio, Y. & Oguchi, K. Interaction between caspase-8 activation and endoplasmic reticulum stress in glycochenodeoxycholic acid-induced apoptotic HepG2 cells. *Toxicology* 241, 146-56 (2007).
 93. Attili, A.F., Angelico, M., Cantafora, A., Alvaro, D. & Capocaccia, L. Bile acid-induced liver toxicity: relation to the hydrophobic-hydrophilic balance of bile acids. *Med Hypotheses* 19, 57-69 (1986).
 94. Bernstein, H., Bernstein, C., Payne, C.M., Dvorakova, K. & Garewal, H. Bile acids as carcinogens in human gastrointestinal cancers. *Mutat Res* 589, 47-65 (2005).
 95. Payne, C.M. et al. The stress-response proteins poly(ADP-ribose) polymerase and NF-kappaB protect against bile salt-induced apoptosis. *Cell Death Differ* 5, 623-36 (1998).
 96. Crowley-Weber, C.L. et al. Development and molecular characterization of HCT-116 cell lines resistant to the tumor promoter and multiple stress-inducer, deoxycholate. *Carcinogenesis* 23, 2063-80 (2002).
 97. Hofmann, A.F. Cholestatic liver disease: pathophysiology and therapeutic options. *Liver* 22 Suppl 2, 14-9 (2002).
 98. Bomzon, A., Holt, S. & Moore, K. Bile acids, oxidative stress, and renal function in biliary obstruction. *Semin Nephrol* 17, 549-62 (1997).
 99. Goddard, T.D. & Kneller, D.G. (University of California, San Francisco).
 100. Gordon, S.R. *Fundamentals of Protein NMR Spectroscopy* (Springer, Dordrecht, 2006).
 101. Hore, P.J. *Nuclear Magnetic Resonance* (Oxford University Press, Oxford, 1995).
 102. Wüthrich, K. *NMR of Proteins and Nucleic Acids* (John Wiley & Sons, New York, 1986).
 103. Wishart, D.S., Sykes, B.D. & Richards, F.M. Relationship between nuclear magnetic resonance chemical shift and protein secondary structure. *J Mol Biol* 222, 311-33 (1991).
-

104. Lipari, G. & Szabo, A. Nuclear magnetic resonance relaxation in nucleic acid fragments: models for internal motion. *Biochemistry* **20**, 6250-6 (1981).

FIGURE LIST

Figure list of chapter 1:

Figure 1.1 Physical-chemical characteristics of bile acids.	13
Figure 1.2 Bile acids biosynthesis pathways.	15
Figure 1.3 Enterohepatic circulation of bile acids.	20
Figure 1.4 Three-dimensional structure of a typical intracellular lipid-binding protein.	23
Figure 1.5 Sequence alignment of some known L-BABPs sequences.	25
Figure 1.6 Three-dimensional structure of cL-BABP in complex with two CDCA molecules.	26
Figure 1.7 Homology model of human I-BABP with GCA and GCDA docked in the binding pocket.	27
Figure 1.8 Three-dimensional structure of ileal BABP from zebrafish complexed with cholic acid.	28

Figure list of chapter 2:

Figure 2.1 Schematic representation of pET24d(+) plasmid.	35
Figure 2.2 Cloning/expression region of pET24d(+) plasmid.	35

Figure list of chapter 3:

Figure 3.1 Orientation of nuclear magnetic dipoles in the absence and presence of a magnetic field.	48
Figure 3.2 Energy separation between the ground and excited states as a function of the magnetic field strength.	48
Figure 3.3 Precessional motion, effect of a RF pulse on the magnetization vector and data sampling during acquisition of FID.	50
Figure 3.4 Nuclear spin coupling in a ^{13}C -H group.	53
Figure 3.5 Spectral density function as a function of the correlation time .	54
Figure 3.6 Spin-lattice relaxation process.	56
Figure 3.7 Spin-spin relaxation process.	57

Figure 3.8 Energy levels for a couple of $\frac{1}{2}$ spins.	57
Figure 3.9 Origin of the NOE in a homonuclear $\frac{1}{2}$ spin system.	59
Figure 3.10 The one-pulse NMR experiment.	61
Figure 3.11 Two-dimensional and three-dimensional NMR experiments.	62
Figure 3.12 Steps of a general 2D NMR experiment.	63
Figure 3.13 Pulse sequence diagram for COSY, DQF-COSY and TOCSY experiments.	63
Figure 3.14 Pulse sequence diagram for ^1H - ^{15}N experiment.	64
Figure 3.15 Example of an ^1H - ^{15}N HSQC spectrum.	65
Figure 3.16 Pulse sequence diagram for NOESY experiment.	65
Figure 3.17 Pulse sequence diagram for T_1 inversion recovery experiment.	66
Figure 3.18 DPMG basic pulse sequence.	66
Figure 3.19 Basis of TROSY NMR experiment.	68

TABLE LIST

Tables of chapter 1:

Table 1.5 Family of fatty acid-binding proteins.	24
---	----

Tables of chapter 2:

Table 2.1 Primer sequence used in mutagenesis procedures.	34
Table 2.2 Site-directed mutagenesis volumes and stock concentrations.	39
Table 2.3 Site-directed mutagenesis cycling parameters.	40

Tables of chapter 3:

Table 3.1 Properties of several NMR active nuclei.	48
Table 3.2 Random coil ^1H chemical shifts for the 20 amino acid residues.	52
Table 3.3 Homonuclear and heteronuclear coupling constants.	53

Appendices

Appendix 1

The sequences of the proteins employed in this study are reported below. Mutated residues are marked in red.

chicken Liver Bile-Acid Binding Protein (cL-BABP)

```

      10      20      30      40      50
1  A F S G T W Q V Y A Q E N Y E E F L K A L A L P E D L I K M A R D I K P I V E I Q Q K G D D F V V T S K T P
      70      80      90     100     110
61 S F T L G K E A D I T T M D G K K L K C T V H L A N G K L V T K S E K F S H E Q E V K G N E M V E T I T F G
      . . . . |
121 R S K R V

```

chicken Liver Bile-Acid Binding Protein H98Q mutant (cL-BABP H98Q)

```

      10      20      30      40      50
1  A F S G T W Q V Y A Q E N Y E E F L K A L A L P E D L I K M A R D I K P I V E I Q Q K G D D F V V T S K T P
      70      80      90     100     110
61 S F T L G K E A D I T T M D G K K L K C T V H L A N G K L V T K S E K F S Q E Q E V K G N E M V E T I T F G
      . . . . |
121 R S K R V

```

chicken Liver Bile-Acid Binding Protein disulfide bridge carrying form (cL-BABP/S-S)

```

      10      20      30      40      50
1  A F S G T W Q V Y A Q E N Y E E F L K A L A L P E D L I K M A R D I K P I V E I Q Q K G D D F V V T S K T P
      70      80      90     100     110
61 S F T L G K E A D I T T M D G K K L K C T V H L A N G K L V C K S E K F S H E Q E V K G N E M V E T I T F G
      . . . . |
121 R S K R V

```

Appendix 2

The list of the resonances for the backbone NH groups assigned for the study of cL-BABP H98Q are reported below.

Assignments of NH groups of apo cL-BABP H98Q.

residue	pH 7.2		pH 5.6	
	HN	N	HN	N
ALA1	125,059	8,712	125,089	8,682
PHE2	112,952	8,419	113,109	8,391
SER3	111,384	8,019	111,212	7,923
GLY4	111,816	9,084	111,75	9,07
THR5	117,858	8,354	117,937	8,348
TRP6	127,816	9,509	127,872	9,506
GLN7	123,87	9,407	123,941	9,41
VAL8	131,526	8,337	131,554	8,341
TYR9	124,222	9,522	124,232	9,52
ALA10	126,094	7,363	126,126	7,363
GLN11			120,234	8,888
GLU12	124,835	9,284	124,823	9,28
ASN13	118,697	9,229	118,719	9,225
TYR14	118,308	7,914	118,374	7,918
GLU15	118,346	8,966	118,357	8,969
GLU16	118,435	8,7	118,434	8,693
PHE17	123,551	8,203	123,548	8,205
LEU18	118,025	8,502	118,029	8,497
LYS19	117,747	7,913	117,768	7,914
ALA20	125,707	7,955	125,742	7,955
LEU21	116,982	7,283	116,993	7,282
ALA22	118,611	7,985	118,637	7,987
LEU23	119,945	7,118	119,961	7,113
PRO24				
GLU25	122,469	8,813	122,535	8,813
ASP26	116,684	8,906	116,67	8,909
LEU27	120,74	7,284	120,723	7,285
ILE28	121,67	7,895	121,717	7,896
LYS29	117,702	7,871	117,701	7,885
MET30	115,648	7,258	115,657	7,256
ALA31	121,065	8,048	121,162	8,046
ARG32	116,701	7,983	116,726	7,983
ASP33	117,404	8,007	117,386	8,013
ILE34	120,901	7,592	120,943	7,591
LYS35	129,268	8,762	129,295	8,763
PRO36				
ILE37	118,751	8,037	118,797	8,051
VAL38	129,025	9,565	129,033	9,574

residue	pH 7.2		pH 5.6	
	HN	N	HN	N
GLU39	128,443	9,604	128,5	9,618
ILE40	125,842	9,571	125,797	9,568
GLN41	127,913	9,247	128,158	9,26
GLN42	125,781	8,249	125,625	8,262
LYS43	130,353	8,812	130,376	8,81
GLY44			117,58	8,93
ASP45	127,475	8,831	127,723	8,858
ASP46	120,623	8,039	120,702	8,056
PHE47	122,243	9,214	122,223	9,21
VAL48	121,51	8,482	121,558	8,484
VAL49	128,375	9,333	128,434	9,339
THR50	127,857	9,706	127,915	9,709
SER51	122,424	9,346	122,466	9,346
LYS52	126,908	9,621	127,009	9,629
THR53	116,328	8,461	116,355	8,462
PRO54				
ARG55	111,727	7,996	111,744	7,993
GLN56	118,111	7,599	118,124	7,597
THR57	123,75	8,845	123,546	8,836
VAL58	127,891	8,862	127,423	8,824
THR59	125,97	8,957	126,029	8,966
ASN60	125,623	9,242	125,674	9,235
SER61	116,825	8,803	116,889	8,813
PHE62	118,213	8,145	118,156	8,154
THR63	119,329	9,359	119,343	9,385
LEU64	125,988	9,166	126,062	9,181
GLY65	108,414	8,996	108,204	8,899
LYS66	120,614	7,809	120,734	7,815
GLU67	126,418	8,706	126,678	8,735
ALA68	128,438	9,268	128,51	9,234
ASP69	120,942	8,482	121,148	8,498
ILE70	125,501	8,939	125,611	8,937
THR71	123,383	9,304	123,414	9,307
THR72	119,427	8,95	119,439	8,948
MET73	118,331	8,763	118,316	8,753
ASP74	115,053	8,25	115,092	8,261
GLY75	108,475	8,152	108,499	8,154
LYS76	121,684	7,423	121,678	7,424

residue	pH 7.2		pH 5.6	
	HN	N	HN	N
LYS77	119,075	8,04	119,103	8,032
LEU78	125,817	9,114	125,787	9,109
LYS79	124,095	8,507	124,055	8,512
CYS80	123,642	9,231	123,605	9,233
THR81	113,626	8,485	113,61	8,473
VAL82	127,848	8,977	128,264	9,093
HIS83	123,812	8,782	122,652	8,796
LEU84	123,476	8,825	124,049	8,901
ALA85	129,318	9,127	129,164	9,125
ASN86			124,103	9,567
GLY87	103,684	8,619	103,601	8,613
LYS88	120,258	7,902	120,347	7,924
LEU89	124,585	8,53	124,688	8,54
VAL90	129,462	9,43	129,106	9,328
THR91	125,469	9,199	126,214	9,252
LYS92	126,047	8,852	125,82	8,762
SER93	118,755	8,591	119,317	8,657
GLU94	121,628	8,747	121,758	8,76
LYS95	115,185	8,592	115,136	8,571
PHE96	119,736	7,451	119,861	7,46
SER97	119,172	8,144	119,361	8,142
HIS98	123,497	8,616	123,684	8,658
GLU99	125,053	8,82	125,312	8,846
GLN100	122,27	8,886	122,146	8,855
GLU101	123,608	8,73	123,688	8,734

residue	pH 7.2		pH 5.6	
	HN	N	HN	N
VAL102	126,785	8,742	126,999	8,725
LYS103	129,536	8,95	129,805	8,975
GLY104			116,724	9,078
ASN105			125,277	8,901
GLU106	120,239	8,097	120,214	8,11
MET107	123,543	8,781	123,602	8,789
VAL108	126,954	8,918	126,989	8,924
GLU109	128,705	9,196	128,727	9,191
THR110	120,449	8,979	120,376	8,981
ILE111	130,481	9,452	130,451	9,456
THR112	123,279	9,222	123,209	9,223
PHE113	126,524	9,069	126,528	9,069
GLY114	115,869	8,514	115,847	8,509
GLY115	105,78	8,489	105,861	8,484
VAL116	123,871	8,045	123,861	8,041
THR117	123,494	8,705	123,433	8,7
LEU118	132,521	9,56	132,525	9,561
ILE119	125,422	7,911	125,47	7,912
ARG120	126,802	9,363	126,79	9,358
ARG121	125,694	8,85	125,982	8,855
SER122	117,876	9,267	117,88	9,269
LYS123	119,297	8,501	119,27	8,502
ARG124	125,871	8,31	125,774	8,305
VAL125	129,786	8,311	129,57	8,301

Assignments of NH groups of holo cL-BABP H98Q complexed with GCDA (protein/GCDA ratio of 1:4).

residue	pH 7.2		pH 5.6	
	HN	N	HN	N
ALA1	124,581	8,668	124,442	8,565
PHE2	112,814	8,475	112,761	8,38
SER3	111,289	7,989	110,951	7,811
GLY4	111,636	9,065	111,416	8,969
THR5	117,822	8,221	117,67	8,145
TRP6	127,414	9,502	127,205	9,423
GLN7	122,311	9,28	122,13	9,207
VAL8	131,363	8,245	131,106	8,17
TYR9	123,532	9,495	123,305	9,409
ALA10	125,759	7,317	125,516	7,24
GLN11				
GLU12	123,325	9,166	123,064	9,087
ASN13	117,138	9,091	116,841	9,011
TYR14	118,324	7,937	118,148	7,866
GLU15	118,625	8,934	118,436	8,861
GLU16	118,814	9,117	118,571	9,033
PHE17	124,217	8,471	124,003	8,401
LEU18	116,705	8,209	116,504	8,128
LYS19	116,558	7,702	116,381	7,627
ALA20	126,154	7,93	125,918	7,855
LEU21	116,847	7,358	116,674	7,28
ALA22	118,978	7,919	118,759	7,846
LEU23	118,55	6,952	118,361	6,875
PRO24				
GLU25				
ASP26	116,599	8,904	116,398	8,83
LEU27	121,841	7,206	121,65	7,131
ILE28	120,693	7,854	120,257	7,774
LYS29	116,981	7,799	116,779	7,728
MET30	117,115	7,402	116,932	7,327
ALA31	120,293	8,381	120,098	8,302
ARG32	121,667	7,732	121,529	7,652
ASP33	116,19	8,302	116,008	8,238
ILE34	120,125	7,212	119,935	7,133
LYS35	130,923	8,977	130,687	8,909
PRO36				
ILE37				
VAL38	128,435	9,4	128,43	9,336
GLU39	128,245	9,527	128,092	9,451
ILE40	125,978	9,618	125,68	9,535
GLN41	128,003	9,251	127,874	9,184
GLN42	126,158	8,192	125,698	8,125
LYS43	130,401	8,794	130,19	8,716
GLY44	117,597	8,899	117,28	8,821

residue	pH 7.2		pH 5.6	
	HN	N	HN	N
ASP45	127,583	8,821	127,397	8,76
ASP46	120,417	8,001	120,463	7,94
PHE47	122,443	9,243	122,222	9,163
VAL48	121,767	8,503	121,598	8,432
VAL49	127,365	9,05	127,338	8,991
THR50	129,202	9,812	129,066	9,731
SER51	122,263	9,056	122,083	8,979
LYS52	125,996	9,434	125,808	9,361
THR53				
PRO54				
ARG55	111,336	7,962	111,18	7,877
GLN56	118,589	7,515	118,4	7,441
THR57	121,357	8,6	121,214	8,528
VAL58	125,567	8,783	125,267	8,716
THR59	126,834	9,344	126,626	9,272
ASN60	124,626	9,091	124,31	9,009
SER61	115,827	8,828	115,714	8,756
PHE62	116,904	8,229	116,667	8,157
THR63	120,163	9,763	119,926	9,702
LEU64	127,071	9,177	126,849	9,12
GLY65	107,589	9,067	107,05	8,914
LYS66	121,144	7,881	121,068	7,813
GLU67	129,87	8,826	129,728	8,774
ALA68	130,922	9,252	130,684	9,154
ASP69	120,074	8,307	119,977	8,249
ILE70	126,891	9,081	126,99	9,005
THR71	123,809	9,326	123,595	9,244
THR72	116,12	8,775	115,974	8,707
MET73	119,974	10,435	119,754	10,354
ASP74	112,103	8,689	111,92	8,614
GLY75	107,863	7,617	107,674	7,543
LYS76	120,411	8,034	120,129	7,945
LYS77	119,796	8,081	119,57	8,009
LEU78	124,278	9,195	124,049	9,119
LYS79	121,715	8,386	121,557	8,317
CYS80	118,126	8,652	117,867	8,59
THR81	118,102	8,675	117,922	8,622
VAL82	132,556	9,509	132,099	9,439
HIS83	123,897	8,775	122,741	8,73
LEU84	122,254	8,702	122,463	8,654
ALA85	130,171	9,39	129,942	9,313
ASN86			124,128	9,468
GLY87	103,051	8,537	102,909	8,464
LYS88	120,367	7,852	120,564	7,804

residue	pH 7.2		pH 5.6	
	HN	N	HN	N
LEU89	122,986	8,355	122,792	8,275
VAL90	119,74	9,042	119,362	8,888
THR91	123,098	8,696	123,531	8,652
LYS92	126,14	9,04	125,833	8,951
SER93	118,934	8,395	118,957	8,331
GLU94	120,048	8,689	119,594	8,647
LYS95	114,069	8,467	113,97	8,393
PHE96	113,641	7,091	113,51	7,031
SER97	112,607	8,548	112,522	8,461
HIS98	128,017	9,386	128,078	9,339
GLU99	124,57	8,651	124,435	8,584
GLN100	124,333	9,151	123,97	9,035
GLU101	124,712	8,858	124,509	8,781
VAL102	125,183	8,776	125,202	8,668
LYS103	129,21	8,925	129,257	8,863
GLY104			116,481	8,992
ASN105			125,551	8,928
GLU106	120,161	8,075	119,97	8,009
MET107	123,811	8,882	123,58	8,818

residue	pH 7.2		pH 5.6	
	HN	N	HN	N
VAL108	126,835	8,978	126,611	8,9
GLU109	129,321	9,136	129,188	9,058
THR110	121,2	9,031	120,921	8,952
ILE111	127,357	9,255	126,913	9,157
THR112	123,966	9,359	123,845	9,291
PHE113	126,862	9,12	126,728	9,05
GLY114	116,4	8,539	116,249	8,478
GLY115	105,142	8,556	104,974	8,48
VAL116	123,713	8,081	123,478	8,004
THR117	123,59	8,774	123,521	8,707
LEU118	131,993	9,459	131,683	9,371
ILE119	124,804	7,742	124,557	7,662
ARG120	125,723	9,501	125,447	9,423
ARG121	126,497	9,084	126,457	9,014
SER122	118,01	9,2	117,772	9,124
LYS123	120,223	8,624	119,997	8,549
ARG124	125,141	8,139	124,785	8,054
VAL125	128,912	8,218	128,506	8,131