doi:10.1016/S0022-2836(03)00968-9 J. Mol. Biol. (2003) 332, 965–977

JMB

Available online at www.sciencedirect.com

SCIENCE DIRECT®

ELSEVIER

Water Dynamics in the Large Cavity of Three Lipid-binding Proteins Monitored by ^{17}O Magnetic Relaxation Dispersion

Kristofer Modig[1], Martin Rademacher[2], Christian Lücke[2] and Bertil Halle[1]*

[1]*Department of Biophysical Chemistry, Lund University P.O. Box 124, SE-22100 Lund Sweden*

[2]*Institut für Biophysikalische Chemie, J. W. Goethe-Universität, DE-60439 Frankfurt am Main, Germany*

**Corresponding author*

Intracellular lipid-binding proteins contain a large binding cavity filled with water molecules. The role played by these water molecules in ligand binding is not well understood, but their energetic and dynamic properties must be important for protein function. Here, we use the magnetic relaxation dispersion (MRD) of the water ^{17}O resonance to investigate the water molecules in the binding cavity of three different lipid-binding proteins: heart fatty acid-binding protein (H-FABP), ileal lipid-binding protein (I-LBP) and intestinal fatty acid-binding protein (I-FABP). Whereas about half of the crystallographically visible water molecules appear to be expelled by the ligand, we find that ligand binding actually increases the number of water molecules within the cavity. At 300 K, the water molecules in the cavity exchange positions on a time-scale of about 1 ns and exchange with external water on longer time-scales (0.01–1 μs). Exchange of water molecules among hydration sites within the cavity should be strongly coupled to ligand motion. Whereas a recent MD simulation indicates that the structure of the cavity water resembles a bulk water droplet, the present MRD results show that its dynamics is more than two orders of magnitude slower than in the bulk. These findings may have significant implications for the strength, specificity and kinetics of lipid binding.

Keywords: fatty acid-binding proteins; hydration; dynamic cluster model; water residence time; water exchange

Introduction

Lipid-binding proteins (LBPs) are found in many tissues and in intracellular as well as extracellular locations.[1] Here, we are concerned with cytoplasmic intracellular LBPs from three different subfamilies:[2–4] porcine ileal lipid-binding protein (I-LBP), rat intestinal fatty acid-binding protein (I-FABP) and bovine heart fatty acid-binding protein (H-FABP). The cytoplasmic LBPs occur abundantly; for example, H-FABP makes up 3% of the cytoplasmic protein content in rat heart.[5] The LBPs are thought to be responsible for uptake and translocation of long-chain lipids, but additional functional roles have been proposed. LBPs have thus been implicated in modulation of cell growth and differentiation and in lipid metabolism.[1,5,6] Cytoplasmic LBPs typically have fatty acid dissociation constants in the nanomolar to micromolar range and the binding affinity is inversely related to the aqueous solubility of the fatty acid.[5–8]

Despite a sequence identity of only 30% among the LBPs investigated here, they have essentially the same secondary and tertiary structure as all other cytoplasmic LBPs. In particular, they all contain a large internal lipid-binding cavity, enclosed

Present address: C. Lücke, Max-Planck-Forschungsstelle für Enzymologie der Proteinfaltung Abteilung NMR-Spektroskopie, Weinbergweg 22, DE-06120 Halle, Germany.

Abbreviations used: DCM, dynamic cluster model; H-FABP, heart fatty acid-binding protein; I-FABP, intestinal fatty acid-binding protein; I-LBP, ileal lipid-binding protein; LBP, lipid-binding protein; MD, molecular dynamics; MRD, magnetic relaxation dispersion; NMR, nuclear magnetic resonance; NOE, nuclear Overhauser effect; SCM, static cluster model.

E-mail address of the corresponding author: bertil.halle@bpc.lu.se

by ten anti-parallel β-strands. The crystal structures of I-FABP and H-FABP reveal a large number of ordered water molecules inside this cavity.[9–13] Other functionally significant structural features are the cavity "portal", lined by two short α-helices, and a hydrogen-bond defect, known as the "gap", between β-strands D and E.[2] The portal and the gap have been postulated as pathways for ligand and water exchange with the external solution.[2,14] In the case of water, this has been confirmed by molecular dynamics (MD) simulations.[15–19] In the region of the turn linking strands D and E, a hydrophobic pocket is present in all cytoplasmic LBPs. Despite the hydrophobicity of the nearby side-chains, this pocket is fully occupied by a highly ordered water molecule,[2,19–22] labelled W135 in the apo I-FABP structure (PDB code 1IFC). Upon fatty acid binding, several ordered water molecules appear to be expelled from the cavity. For I-FABP, the apo cavity is modeled with 20 ordered water molecules, whereas only eight are seen in the holo cavity.[9,12,14] Because of the close similarity between the apo and holo structures (1.24 Å RMSD for I-FABP[12]), buried water molecules located outside the binding cavity are unaffected by fatty acid binding. At present, no X-ray structure has been reported for apo H-FABP, but 14 ordered water molecules are modeled in the holo cavity.[10,13] The NMR structures of H-FABP and I-LBP have virtually the same overall fold as the crystal structures,[23–27] suggesting that the effects of ligand binding on cavity water are similar in solution and crystal.

The structural and dynamic properties of the cavity and its water molecules govern the primary functional event in the LBPs: the binding of lipids. Water molecules solvate the charged head group and non-polar alkyl chain of the bound fatty acid and provide a flexible environment that reduces the entropic penalty for ligand binding.[19,28] In addition, the dynamics of the cavity water control the kinetics of ligand entry and exit as well as structural fluctuations of the cavity itself. A previous water ^{2}H and ^{17}O magnetic relaxation dispersion (MRD) study of I-FABP revealed water dynamics on three time-scales.[21] First, water molecules in contact with the external protein surface are dynamically retarded by a factor 4–5 relative to bulk water, as for other globular proteins.[29,30] Second, a few highly ordered water molecules occupy hydration sites with long ($\gg 10$ ns) residence times. At least some of these sites appear to be located outside the binding cavity. Third, the 10–20 cavity water molecules give rise to a relaxation dispersion with a correlation time of about 1 ns. Surprisingly, the amplitude of this contribution is not reduced upon ligand binding, although the ligand is seen to displace ordered water molecules in the crystal structures. The nanosecond correlation time of the cavity water molecules allows two alternative interpretations: either the water molecules remain at definitive sites within the cavity and collectively exchange with bulk water as the cavity opens up, or they exchange positions within the cavity before exchanging with external solvent. In the first case, the nanosecond correlation time is the mean lifetime of the cavity in the closed state; in the second case, it is related to the mean water residence time in the individual intra-cavity hydration sites.

Following the I-FABP MRD study, several simulation studies have examined cavity hydration in I-FABP.[17–19] Among these, the simulations by Likić & Prendergast[18] and by Bakowies & van Gunsteren[19] are most relevant to our MRD work. These two MD studies both support the MRD results[21] in that only a few water molecules remain in the same hydration site for the duration (1–5 ns) of the simulation trajectory. Moreover, the MD studies report exchange among intra-cavity hydration sites on a time-scale of about 1 ns[19] or less.[18] However, there is substantial disagreement between the MD studies on the time-scale for exchange of cavity water with bulk water. Therefore, the MD studies do not resolve the interpretational ambiguity in the previous MRD study.[21]

We have now performed ^{17}O MRD measurements on two more LBPs: porcine I-LBP and bovine H-FABP, in both apo and holo forms. The aim was to better characterize the water molecules within the cavity and the effect of ligand binding. In particular, we wanted to know whether the counterintuitive increase in the MRD-derived cavity hydration upon ligand binding to I-FABP is a general feature of the LBP family. The new MRD results motivated us to develop a more general relaxation model that explicitly accounts for exchange among hydration sites within the binding cavity. In addition, we present a detailed analysis of cavities in the available X-ray and NMR structures, focusing on the effect of ligand binding.

For all the three investigated LBPs, we find that the water molecules remain inside the cavity for a much longer period (0.01–1 μs at 300 K) than the mean residence time (about 1 ns) in a given intra-cavity hydration site. Nevertheless, water dynamics within the binding cavity is two orders of magnitude slower than in bulk water. Surprisingly, a reduction in the number of cavity water molecules on ligand binding is not observed for any of the LBPs. On the contrary, the MRD data demonstrate that the holo proteins contain more cavity water molecules than the apo proteins. Although this water influx is not seen directly in the crystal structures, it is consistent with the large increase in cavity volume induced by ligand binding.

Results

We have measured the water ^{17}O longitudinal relaxation rate as a function of magnetic field strength in solutions of H-FABP and I-LBP in the apo form and in complex with ligand. In the case of H-FABP, the ligand was a mixture of fatty acids

and in the case of I-LBP is was chenodeoxycholate. Relaxation data were recorded at 27 °C (apo and holo forms) and at 4 °C (apo form only). We shall compare these results with published MRD results on I-FABP.[21] The water ^{17}O relaxation dispersion profile exclusively monitors the dynamics of water molecules in association with the protein.[31–33] In the case of LBPs, the frequency-dependent ^{17}O relaxation rate $R_1(\omega_0)$ is governed by a bi-Lorentzian spectral density $J(\omega_0)$ according to:[21,32,33]

$$R_1(\omega_0) = R_{bulk} + 0.2J(\omega_0) + 0.8J(2\omega_0) \quad (1)$$

$$J(\omega_0) = \alpha + \frac{\beta\tau_\beta}{1+(\omega_0\tau_\beta)^2} + \frac{\gamma\tau_\gamma}{1+(\omega_0\tau_\gamma)^2} \quad (2)$$

In equations (1) and (2), $\omega_0 = 2\pi\nu_0$ is the resonance frequency in angular frequency units. The two dispersive terms in equation (2), hereinafter referred to as the β and γ dispersions, arise from water molecules with residence times much longer than the rotational correlation time τ_R of the protein (β term) or with residence times shorter than the rotational correlation time (γ term). The parameters $\beta(\gamma)$ and $\tau_\beta(\tau_\gamma)$ are related to the number $N_\beta(N_\gamma)$ of water molecules and their mean residence time $\tau_{W\beta}(\tau_{W\gamma})$ through:

$$\beta = \frac{\omega_Q^2 N_\beta S_\beta^2}{N_T} \quad (3)$$

and

$$\frac{1}{\tau_\beta} = \frac{1}{\tau_{W\beta}} + \frac{1}{\tau_R} \quad (4)$$

Analogous relations hold for the γ term. In equation (3), S_β is the root-mean-square orientational order parameter for the N_β water molecules, $\omega_Q = 7.61 \times 10^6$ rad s^{-1} is the quadrupole coupling frequency for the ^{17}O nucleus in a protein-bound water molecule,[32] and N_T is the total number of water molecules per protein molecule, which is simply a measure of the protein concentration. Finally, the high frequency offset ("plateau") parameter α in equation (2) equals $R_{bulk}N_S\rho/N_T$, where N_S is the number of water molecules in contact with the protein surface and the dynamic retardation factor $\rho = \tau_S/\tau_{bulk} - 1$ is a relative measure of the water rotational correlation time at the protein surface and in bulk water. The number N_S may be estimated by dividing the solvent-accessible surface area of the protein with the effective area occupied by a water molecule, usually taken as 15 Å^2.[32] For typical globular proteins, $\rho = 4–5$.[29,30]

The relaxation theory outlined above is valid in the fast-exchange regime on the relaxation timescale, defined by the relation $\tau_W \ll (\omega_Q^2 S^2 \tau_R)^{-1}$.[32,33] For a highly ordered water molecule ($S = 0.9$) in a protein with $\tau_R = 7$ ns, this means that $\tau_W \ll 3$ μs. According to equation (4), a β dispersion with $\tau_\beta = \tau_R$ can only provide lower and upper bounds on the residence time $\tau_{W\beta}$: $\tau_R \ll \tau_{W\beta} \ll (\omega_Q^2 S^2 \tau_R)^{-1}$. For the γ term, on the other hand, the residence time $\tau_{W\gamma}$ is obtained directly from τ_γ using the analogue of equation (4).

In the previous I-FABP study,[21] bi-Lorentzian spectral densities were required to fit the MRD data. This is also the case for the present data. Single-Lorentzian fits yield anomalously large $N_S\rho$ values and anomalously short correlation times, just as for the I-FABP data.[21] These anomalies can be traced to the nanosecond dispersion step (γ dispersion), which, if not allowed for in the model, leads us to overestimate α and underestimate τ_β. This picture is reinforced by MRD measurements on the apo samples at 4 °C, where the γ dispersion can be more completely characterized. The *F*-test, based on the change in the χ^2 merit function when a second dispersion step is included,[34] yields probabilities of 0.83 and 0.92 for a second Lorentzian in the I-LBP and H-FABP data, respectively, without any parameter constraints. At 27 °C, the apo H-FABP sample yields a longer τ_β than expected ($\tau_\beta > \tau_R$) and the *F*-test probability for a second Lorentzian is 0.74.

Table 1 summarizes the properties of the four investigated samples. The ^{17}O MRD profiles for H-FABP and I-LBP at 27 °C and 4 °C are shown in Figures 1 and 2, respectively. The data have been subjected to bi-Lorentzian fits according to equations (1) and (2), with τ_β and $N_S\rho$ fixed, and the resulting parameter values are given in Table 2. As for I-FABP,[21] $N_S\rho$ was estimated with $N_S = 460$ (see Materials and Methods) and $\rho = 4.5$ at 27 °C,[29,30] while τ_β was estimated as described in Materials and Methods. For comparison, Table 2 includes parameter values obtained in the previous I-FABP study.[21]

The three most striking results in Table 2 are: (1) $N_\beta S_\beta^2$ is significantly larger for apo H-FABP than for the other two apo proteins; (2) $N_\gamma S_\gamma^2$ is smallest

Table 1. Protein characteristics

Sample	C_P[a] (mM)	N_T (10^3)	pH	Lipidation[b] (%)	τ_R (ns)[c]	τ_β (ns)[d]
H-FABP	0.59	93.0	7.4	19	7.3	8
H-FABP	1.16	46.9	7.5	83	7.3	6.4
I-LBP	0.79	69.2	8.0	7	–	6.4
I-LBP	0.50	110	8.0	100	7.3	6.4

[a] Protein concentration determined by complete amino acid analysis with a standard deviation of 1%.
[b] In the text, samples with high and low degree of lipidation are referred to as apo and holo, respectively.
[c] Rotational correlation time determined by ^{15}N relaxation. See the text for references and details. The τ_R values are scaled to 0% ^2H_2O. Values at 4 °C may be obtained by scaling with the factor $[\eta(277\text{ K})/\eta(300\text{ K})] \times (300/277) = 1.99$.
[d] The τ_β values were obtained as described in Materials and Methods. The corresponding τ_β values at 4 °C are obtained by scaling as above. In contrast to the previous measurements on I-FABP that were performed in 50% ^2H_2O,[21] we use 0% ^2H_2O in this study.

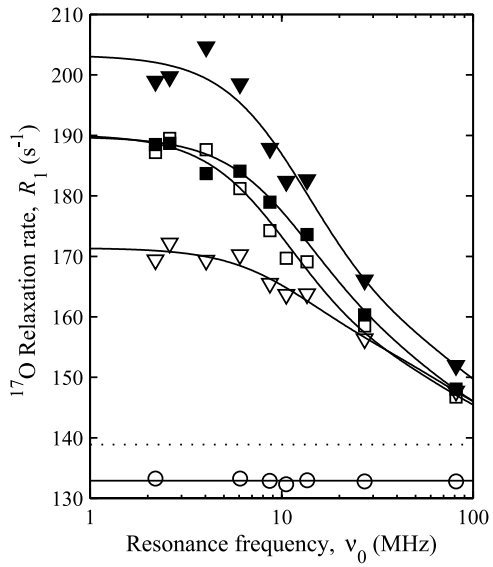


Figure 1. Water ^{17}O relaxation dispersion profiles from aqueous solutions of apo H-FABP (□), holo H-FABP (■), apo I-LBP (▽) and holo I-LBP (▼) at 27°C. The data have been normalized to the concentration of holo H-FABP (1.16 mM). The frequency-independent relaxation rate of bulk water (○) and the sum $R_{bulk} + \alpha$ (dotted line) are also shown. The estimated error bars are 0.5%, or approximately equal to the size of the data symbols. The curves are bi-Lorentzian fits with parameters given in Table 2.

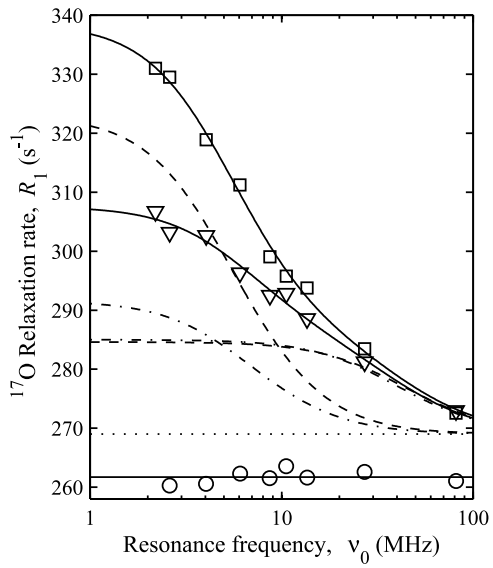


Figure 2. Water ^{17}O relaxation dispersion profiles from aqueous solutions of apo H-FABP (□) and apo I-LBP (▽) at 4 °C. The data have been normalized to the concentration of apo H-FABP (0.59 mM). The frequency-independent relaxation rate of bulk water (○) and the sum $R_{bulk} + \alpha$ (dotted line) are also shown. The estimated error bars are 0.5%, or approximately equal to the size of the data symbols. The continuous curves are bi-Lorentzian fits with parameters given in Table 2. The individual β (low frequency) and γ (high frequency) dispersion components are also shown for H-FABP (broken curves) and I-LBP (dash-dotted curves).

for I-FABP; and (3) the increase in $N_\beta S_\beta^2$ upon ligand binding is much larger for I-LBP than for the other proteins. Interestingly, there is no significant reduction in $N_\gamma S_\gamma^2$ for any of the three proteins upon ligand binding, as would be expected if the ligand expels water from the cavity without enlarging it. For I-LBP, we even find a significant increase in $N_\gamma S_\gamma^2$ upon ligand binding.

Because three of the samples are neither pure apo nor pure holo form (see Table 1), we plot $N_\beta S_\beta^2$ as a function of lipidation in Figure 3. This plot clearly brings out the differences between the three proteins. Whereas apo I-FABP and apo I-LBP have similar $N_\beta S_\beta^2$ values, $N_\beta S_\beta^2$ for apo H-FABP is at least one unit larger. For I-LBP, $N_\beta S_\beta^2$ increases strongly upon ligand binding, whereas only small changes are seen for I-FABP and H-FABP. Hereinafter, when referring to apo and holo $N_\beta S_\beta^2$ values, we use values extrapolated to 0% and 100% (see Figure 3).

Discussion

Modeling the γ dispersion

To quantitatively interpret the MRD data, we need to identify the origin of the γ dispersion. A prominent γ dispersion has been observed only for proteins with a large internal cavity: I-FABP,[21] H-FABP and I-LBP (this work) as well as β-lactoglobulin[35] and β-trypsin.[29] It is clear, therefore, that the γ dispersion from the LBPs reflects the dynamics of the water cluster in the large binding cavity. We shall consider two models that differ in the internal mobility of this water cluster. In the static cluster model (SCM), each water molecule resides in a given intra-cavity hydration site until

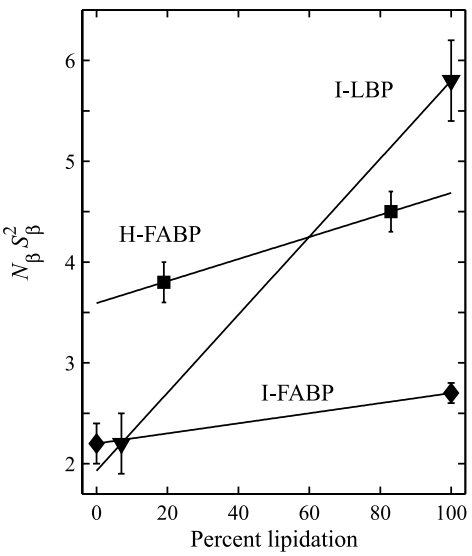


Figure 3. $N_\beta S_\beta^2$ as a function of lipidation at 27 °C for H-FABP (■), I-LBP (▼) and I-FABP (◆). The $N_\beta S_\beta^2$ values extrapolated to 0% (100%) lipidation are 3.6±0.3 (4.7±0.5) for H-FABP and 1.9±0.3 (5.8±0.4) for I-LBP.

Table 2. Result of bi-Lorentzian fits to ^{17}O MRD profiles at 4 °C and 27 °C

Protein	Lipidation (%)	T (°C)	$N_\beta S_\beta^2$	$N_\gamma S_\gamma^2$	τ_γ (ns)	$10^{-3}\, N_S \rho$ [a]	$N_I S_I^2 + N_C S_C^2$ [b]	max A_C^2 [c]
H-FABP	19	4	5.5 ± 0.3	12 ± 2	2.0 ± 0.5	[2.6]	17.5 ± 2.0	0.27 ± 0.04
I-LBP	7	4	2.8 ± 0.3	12 ± 2	2.2 ± 0.4	[2.6]	14.8 ± 2.0	0.13 ± 0.07
I-FABP[d]	0	4	2.5 ± 0.2	7.6 ± 0.2	3.1 ± 0.2	[2.6]	10.1 ± 0.3	0.16 ± 0.02
H-FABP	19	27	3.8 ± 0.2	11 ± 2	1.0 ± 0.3	[2.1]	14.6 ± 2.0	0.20 ± 0.03
H-FABP	83	27	4.5 ± 0.2	11 ± 1	1.1 ± 0.2	[2.1]	15.7 ± 1.1	0.24 ± 0.02
I-LBP	7	27	2.2 ± 0.3	12 ± 2	1.0 ± 0.2	[2.1]	13.9 ± 2.0	0.09 ± 0.02
I-LBP	100	27	5.8 ± 0.4	18 ± 3	0.8 ± 0.2	[2.1]	23.8 ± 3.0	0.21 ± 0.03
I-FABP[d]	0	27	2.2 ± 0.2	6.9 ± 0.3	1.1 ± 0.1	[2.1]	9.1 ± 0.4	0.15 ± 0.02
I-FABP[d]	100	27	2.7 ± 0.1	6.7 ± 0.8	1.1 ± 0.1	[2.1]	9.4 ± 0.8	0.20 ± 0.02

The errors correspond to one standard deviation, as estimated by 1000-point Monte Carlo simulations.[34] $N_S\rho$ values were fixed at the values given within brackets. The τ_β values were fixed at the values given in Table 1, and scaled to 4 °C according to $\tau_\beta \propto \eta(T)/T$.

[a] Based on $N_S = 460$ and $\rho = 4.5$ at 27 °C.

[b] Derived from equations (5a) and (5b), which is valid for the DCM.

[c] Upper bound on A_C^2, corresponding to $N_I S_I^2 = 1$.

[d] The I-FABP parameters were taken from Tables 2 and 3 of Wiesner *et al*.[21] and represent the mean of two samples at pH 7.4 (with palmitate or oleate as the ligand for holo I-FABP). The quoted uncertainty is the standard deviation of the mean.

the cavity transiently opens up sufficiently to allow all cavity water molecules to exchange with the external solvent. The correlation time τ_γ is then the mean water residence time in the cavity or, equivalently, the mean lifetime of the cavity in the "closed" state. In the dynamic cluster model (DCM), we allow for water interchange among the hydration sites within the cavity. The correlation time τ_γ is then essentially the mean water residence time in the individual hydration sites (see Supplementary Material). In general, the internal motion within the water cluster will not average out the anisotropic nuclear quadrupole coupling completely. The residual quadrupole coupling that remains after intra-cavity motions is averaged to zero when the water molecules escape from the cavity. If this exchange with external solvent is slower than protein tumbling, then the cavity water molecules will also contribute to the β dispersion, along with the conventional (singly buried) internal water molecules.

A brief derivation of the spectral density function $J(\omega_0)$ for the DCM, which includes the SCM as a special case, is given in the Supplementary Material. After several simplifying (but reasonable) assumptions, we find that the DCM spectral density can be expressed on the bi-Lorentzian form of equation (2) with the phenomenological amplitude parameters $N_\beta S_\beta^2$ and $N_\gamma S_\gamma^2$ (deduced from β and γ) related to the model parameters as:

$$N_\beta S_\beta^2 = N_I S_I^2 + N_C S_C^2 A_C^2 \quad (5a)$$

$$N_\gamma S_\gamma^2 = N_C S_C^2 (1 - A_C^2) \quad (5b)$$

Here, N_C is the number of water molecules in the binding cavity and N_I is the number of long-lived internal water molecules buried elsewhere in the protein. Furthermore, S_C is the root-mean-square order parameter for the hydration sites in the binding cavity, A_C is an order parameter that characterizes the anisotropy of the orientational distribution of these hydration sites, and S_I is the root-mean-square order parameter for the N_I other internal water molecules. In the limit $A_C = 0$, equations (5a) and (5b) correspond either to the SCM or to the DCM with an isotropic distribution of hydration sites. However, the physical significance of τ_γ is different in the two cases (see Supplementary Material).

A mobile water cluster

On the basis of the MRD data alone, it is not possible to discriminate between the two microdynamical models (SCM and DCM) for the water cluster in the binding cavity. However, a detailed analysis of the available protein structures, in particular the structural changes accompanying ligand binding, shows that only the DCM is consistent with both MRD and structural data. This conclusion is further supported by ^{13}C NMR relaxation[28] and MD simulation[19] studies of I-FABP.

We focus on the most striking MRD result: the threefold increase in $N_\beta S_\beta^2$ (from 1.9 to 5.8) on binding of bile acid in the I-LBP cavity (see Figure 3). For this observation to be consistent with the SCM, at least four new, highly ordered and long-lived, hydration sites must be created upon ligand binding. For H-FABP and I-FABP, at least one such ligand-induced site is required. Because water molecules with residence times longer than 10 ns are invariably located in cavities,[29,30] we searched for ligand-induced small cavities in the crystal and NMR structures of I-FABP, H-FABP and I-LBP (see Supplementary Material). In the 1.4 Å resolution crystal structure of human H-FABP in complex with oleate (PDB code 1HMS[13]), four water molecules are located in small cavities created by the ligand. In the 1.5 Å resolution crystal structure of rat I-FABP in complex with myristate (1ICM[12]), the only ligand-induced cavity appears to be empty. For I-LBP, the available structures were determined by NMR and, therefore, do not furnish information about cavity hydration. Although binding of

glycocholate to porcine I-LBP (1EIO[25]) creates a similar number of cavities as found in holo H-FABP, the predominantly non-polar character of these small cavities indicates a low water occupancy.

In summary, the lack of correlation between the ligand-induced increase in $N_\beta S_\beta^2$ and the structurally inferred number of ligand-trapped water molecules rules out the SCM. A further argument against the SCM comes from ^{13}C NMR relaxation[28] and MD simulation[19] studies, indicating that bound palmitate has considerable motional freedom within the binding cavity of I-FABP. Because this motion occurs on the nanosecond time-scale ($< \tau_R$), the ligand is not likely to trap water molecules long enough ($\gg \tau_R$) for them to contribute to the β dispersion.

In the DCM, an increase in $N_\beta S_\beta^2$ is not necessarily linked to an increase in the number of water molecules with long ($\gg \tau_R$) residence times in particular hydration sites. As seen from the second term in equation (5a), $N_\beta S_\beta^2$ can also increase if any or all of the parameters N_C, S_C and A_C increase upon ligand binding. As shown in the Supplementary Material, the N_C water molecules must then reside for a long time in the cavity ($\tau_C \gg \tau_R$), but may have short site residence times ($\tau_S \ll \tau_R$). The DCM scenario may seem unlikely, because the crystal structures show that the ligand expels a substantial fraction of the crystallographically identified water molecules from the binding cavity. However, the total number N_C of cavity water molecules might increase if the cavity is enlarged sufficiently by ligand binding. Furthermore, many of these water molecules may not be sufficiently positionally ordered to be X-ray visible.

To test the hypothesis that N_C increases upon ligand binding, we calculated the volume of the binding cavity in the apo and holo (after removal of the ligand coordinates) forms of I-FABP and I-LBP (see Supplementary Material). (The structure of apo H-FABP has not been reported.) Although cavity calculations have been reported for several LBPs,[2,36–38] a systematic comparison of the apo and holo forms has apparently not been presented. It is evident from Figure 4 that ligand binding greatly increases the cavity size. For I-FABP, the cavity volume increases from 520 Å^3 to 930 Å^3 on binding of myristate. The latter number was calculated for the holo protein after removal of the ligand, which has a van der Waals (vdW) volume of 220 Å^3. For I-LBP, the cavity volume increases from 400 Å^3 to 1010 Å^3 on binding of glycocholate, with a vdW volume of 460 Å^3. (The calculated ligand-induced increase in cavity volume for I-LBP may be affected by the slightly different refinement protocols used in the apo[24] and holo[25] structures.) These large cavity expansions might seem to be incompatible with the general observation that ligand binding to LBPs hardly affects the protein structure.[2] For example, the apo-holo RMSD for the backbone atoms of I-FABP is merely 0.4 Å.[9] However, when the cavity volume increases by 500 Å^3, the total protein volume (including the cavity) increases by less than 3% and the protein radius by less than 1%, or about 0.14 Å. This, together with the simple geometric fact that few atoms in a globular protein are far from the surface, explains the small apo–holo RMSD.

For both proteins, the cavity volume available to water molecules is larger in the holo form than in the apo form. Water molecules located within or at the surface of proteins occupy less space than in bulk water.[39] If cavity water molecules occupy 20–30 Å^3 per molecule on average, the ligand-induced increase in free cavity volume in I-FABP and I-LBP corresponds to five to ten additional water molecules. Evidence has been presented for the cooperative binding of two molecules of glycocholate to human I-LBP;[40] however, it is not clear whether the second binding site is located inside the cavity.[27,41] In summary, the available structural information indicates that N_C increases upon ligand binding. Together with concomitant increases in the order parameters A_C and S_C (see below), this can explain the observed increase in $N_\beta S_\beta^2$ upon ligand binding.

Several MD simulations of solvated LBPs have been carried out and three of these examined in detail, the properties of the water molecules in the binding cavity of I-FABP at ~300 K.[17–19] These simulations reveal a dynamic cavity whose water content undergoes large fluctuations on the nanosecond time-scale. With one or two exceptions, all cavity water molecules leave their hydration sites within 2 ns, but most of them appear to remain in the cavity much longer.[18,19] While this finding supports the DCM, the simulations are inconclusive in other respects. For example, Bakowies & van Gunsteren[19] find 25 internal water molecules in apo I-FABP that do not exchange with external solvent during their 5 ns trajectory, whereas Likić & Prendergast infer that 20 out of the 22 crystallographically identified cavity water molecules have cavity residence times (τ_C) in the range 0.6–2.0 ns,[18] which would support the SCM. This order-of-magnitude discrepancy between two state-of-the-art simulations that started from the same initial coordinates (the apo I-FABP crystal structure 1IFC with 22 internal water molecules) demonstrates that current MD simulations only provide a rough guide to water dynamics in the real protein. Simulation trajectories of 1–5 ns are clearly too short for characterizing dynamic processes on the nanosecond time-scale with statistical accuracy. There is also considerable ambiguity in the definition of water residence times[42] and in the classification of water molecules as internal or external.[18,43] Finally, the (different) semi-empirical force fields used in these simulations may not capture the subtle balance of interactions that control the gating of cavity water exchange through the small orifice in the gap region.[18,19] Moreover, the water potentials are designed to reproduce bulk

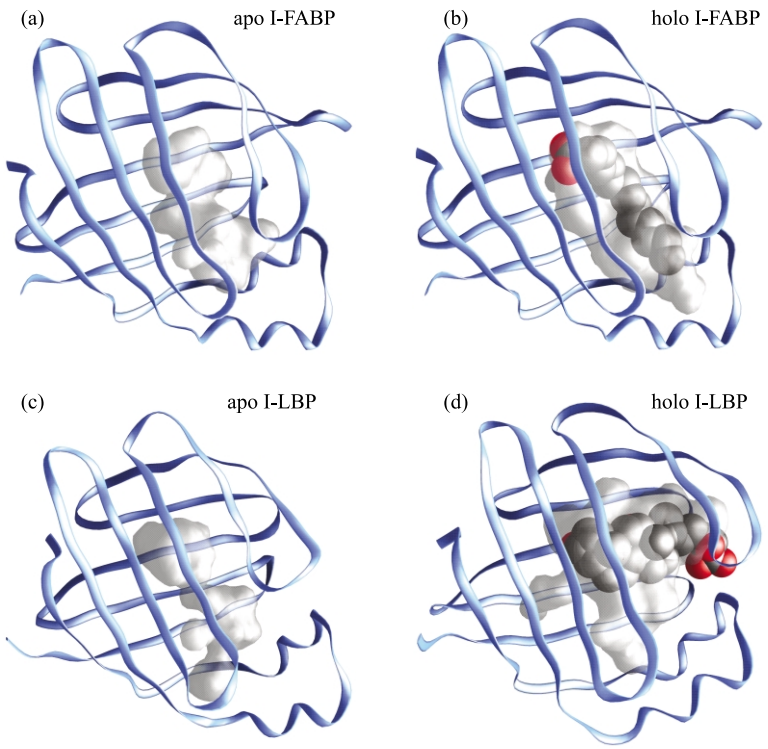


Figure 4. Binding cavities detected by the program GRASP with 1.2 Å probe and AMBER all atom vdW radii. The panels show (a) apo I-FABP (PDB code 1IFC), (b) holo (myristate) I-FABP (1ICM), (c) apo I-LBP (1EAL, model 1), and (d) holo I-LBP (1EIO, model 1). The cavities and ligands are colored gray with red oxygen atoms.

water properties and may therefore not be adequate for describing internal water.

Quantitative interpretation of the MRD data

Having argued that cavity water dynamics in the LBPs should conform more closely to the DCM than to the SCM, we now proceed to the quantitative analysis of the MRD data in terms of the DCM. Because τ_β matches the rotational correlation time τ_R of the protein (see Materials and Methods), we conclude that the protein contains $N_I + N_C$ long-lived water molecules, i.e. water molecules that on average reside within the protein much longer than τ_R. Of these water molecules, N_C reside in the binding cavity with short site residence times ($\tau_S < \tau_R$) and N_I reside in smaller cavities (within or outside the binding cavity) with long site residence times ($\tau_I \gg \tau_R$).

The numbers N_I and N_C cannot be determined from the MRD data, but their sum, weighted by the mean-square local order parameter, can be obtained since, according to equations (5a) and (5b), $N_I S_I^2 + N_C S_C^2 = N_\beta S_\beta^2 + N_\gamma S_\gamma^2$. Because $S_I^2, S_C^2 \leq 1$, this quantity (see Table 2) furnishes a lower bound on the total number, $N_I + N_C$, of long-lived water molecules.

Among the apo proteins, independent information about $N_I + N_C$ is available only for I-FABP, where crystallography and MD simulations both suggest a value in the range 20–25. Because the binding cavity is 23% smaller in apo I-LBP than in apo I-FABP, we expect that $N_I + N_C$ is in the range 15–20 for this protein. With these estimates of $N_I + N_C$, our results for $N_I S_I^2 + N_C S_C^2$ (see Table 2) imply that the internal water molecules are more highly ordered in apo I-LBP (and, presumably, also in apo H-FABP) than in apo I-FABP. The mean-square order parameter averaged over all long-lived water molecules, $(N_I S_I^2 + N_C S_C^2)/(N_I + N_C)$, can thus be estimated to $0.4(\pm 0.1)$ for I-FABP and $0.8(\pm 0.2)$ for apo I-LBP. This difference is consistent with the larger number of small cavities in I-LBP and H-FABP as compared to I-FABP (see Supplementary Material).

Upon ligand binding, the quantity $N_I S_I^2 + N_C S_C^2$ does not change significantly for I-FABP and H-FABP, but increases from $14(\pm 2)$ to $24(\pm 3)$ for I-LBP. As discussed above, the water-accessible cavity volume (after subtraction of the ligand volume) is larger in the holo form than in the apo form for both I-LBP and I-FABP (and, presumably, also for H-FABP). However, because the increase of water-accessible cavity volume is similar for I-FABP and I-LBP, the much larger increase in $N_I S_I^2 + N_C S_C^2$ for I-LBP cannot be explained by a much larger increase in N_C. (Note that the chenodeoxycholate ligand used for the MRD experiments is about 100 Å^3 smaller than the glycocholate ligand used for the NMR structure determination.) Instead, we believe that the exceptional increase in $N_I S_I^2 + N_C S_C^2$ for I-LBP results from the combined effect of increased N_I and S_C^2, due to more extensive water–ligand interactions (see Supplementary Material).

The cavity order parameter A_C^2, related to the orientational distribution of hydration sites within the cavity (see Supplementary Material), cannot be separately determined from the MRD data. However, an upper bound on A_C^2 can be established.

According to equations (5a) and (5b), $A_C^2 = (N_\beta S_\beta^2 - N_I S_I^2)/(N_\beta S_\beta^2 + N_\gamma S_\gamma^2 - N_I S_I^2)$, showing that A_C^2 decreases monotonically with increasing $N_I S_I^2$. The single water molecule (denoted W135 in the apo I-FABP crystal structure 1IFC[9]) buried in a small cavity formed by the loop between β-strands D and E is conserved throughout the LBP family and has been shown by NMR[20,22] and MD simulations[19,22] to be long-lived and highly ordered in all crystal structures. Accordingly, it should contribute about one unit to $N_\beta S_\beta^2$ for all proteins investigated here. We thus compute the upper bound on A_C^2 with $N_I S_I^2 = 1$. As seen from Table 2, this calculation yields A_C^2 bounds in the range 0.1–0.3 for all proteins. Such small values indicate that most water molecules explore a substantial fraction of the hydration sites before escaping from the cavity. For all LBPs, ligand binding increases the maximum A_C^2 value (see Table 2). This is as expected if the ligand makes the water-occupied part of the cavity less spherical (i.e. more elongated) and/or partitions the cavity into several sub-cavities, thereby restricting the number of inter-cavity hydration sites sampled by a given water molecule on time-scales shorter than τ_R.

In a recent NMR study, the titration behavior of histidine residues in the binding cavity of H-FABP (His93) and I-LBP (His99) was investigated.[44] For apo and holo H-FABP, His93 remained in the uncharged state down to the lowest investigated pH of 4.5 (where protein aggregation sets in). This was also the case for His99 in the holo I-LBP, whereas a rather low pK_a of 5.27 was determined for apo I-LBP. The observed destabilization of the charged form of these His residues indicates that the intra-cavity environment is significantly less polarizable than the external protein surface. This reduced polarity is partly a consequence of orientational constraints on intra-cavity water, as reflected in the order parameter S_C. The observability of the labile His proton resonances at near-neutral pH implies that these protons exchange slowly ($\gg$ ms) with water protons. On the other hand, intra-cavity water molecules contribute to the ^{17}O relaxation only if they exchange with external water on a time-scale of ca 1 μs or shorter. Therefore, water exchange out of the cavity is apparently not rate-limiting for His proton exchange with bulk water.

In the previous MRD study of I-FABP,[21] $N_\beta S_\beta^2$ was interpreted in terms of the SCM, i.e. without the second term in equation (5a). At least three (apo) or four (holo) long-lived water molecules were then required to account for the β dispersion. On the basis of the solvent-accessible area, thermal *B*-factor, and hydrogen-bonding status, parameters that govern the site residence time,[29] plausible candidates for these long-lived hydration sites were selected from the available crystal structures.[21] While there is general agreement that W135 (1IFC nomenclature, see above) is among the long-lived water molecules,[18–22] the other candidates are more uncertain. Both of the recent I-FABP simulations[18,19] tend to support the original interpretation of the MRD data, but, as discussed above, these simulations are really too short to reliably determine residence times in the nanosecond range. Bakowies & van Gunsteren[19] find quantitative agreement for both apo and holo I-FABP between the number of long-lived water molecules deduced from MRD and the number of water molecules with computed *B*-factors larger than 40 Å^2 and contact times (or site residence times) of about 1 ns (considerably shorter than τ_R).[19] We would thus not expect any of these water molecules to have long ($\gg \tau_R$) site residence times. For I-FABP, therefore, the accumulated evidence favors the DCM scenario advanced here, where the β dispersion is due to a single water molecule (W135) with long site residence time ($N_I S_I^2 \approx 1$) and a larger number of mobile and less ordered ($N_C S_C^2 A_C^2 \approx 2$) cavity water molecules with site residence times of about 1 ns and much longer residence time in the cavity.

In the case of H-FABP, a crystal structure is available only for the holo form.[13] A detailed hydration analysis of this structure in terms of accessibility, *B*-factor and hydrogen bonding shows that holo H-FABP contains several potentially long-lived hydration sites (see Supplementary Material). However, these sites are all created or affected by the bound fatty acid, which is not likely to remain in the same location for periods $\gg \tau_R$.[19,28] We therefore believe that the larger $N_\beta S_\beta^2$ for H-FABP as compared to I-FABP (see Table 2) is at least partly due to a larger contribution from labile cavity water molecules. Because the holo cavities in I-FABP and H-FABP differ by only 15 Å^3 in volume (see Supplementary Material), the difference must be attributed to a higher order parameter product $S_C^2 A_C^2$ in H-FABP (see equation (5a)).

The hydration of bovine H-FABP in apo and holo form has previously been investigated *via* the intermolecular nuclear Overhauser effect (NOE) to peptide NH protons.[20] This study reported a large number of negative NOEs (23 in apo, 21 in holo), which, according to the conventional interpretation,[45,46] indicates water residence times longer than 0.3 ns. Three of these NOEs can be attributed to the long-lived W166 (the equivalent to W135 in apo I-FABP) outside the binding cavity, while only two NOEs (with Arg106 and Ile62) can be matched with crystallographically identified water molecules in the binding cavity (W167 and W183) (see Supplementary Material). The remaining NOEs can probably be explained by a combination of long-range NOEs to mobile cavity water molecules and external water (B.H. & K.M., unpublished results), proton-exchange relayed magnetization transfer[46] and other artefacts.[47] Because the NOE data yield a product of factors, involving the number of contributing water molecules, their distances from the NH proton, and one or more order parameters and correlation times, their interpretation is highly model-dependent.

The lack of crystal structures makes it difficult

to assign potentially long-lived water molecules in I-LBP. A cavity analysis on the NMR structure of holo I-LBP[25] indicates few if any potentially long-lived hydration sites in the binding cavity (see Supplementary Material). This agrees with the finding of similar small $N_\beta S_\beta^2$ values for apo I-LBP and apo I-FABP (see Table 2) and suggests that the large increase in $N_\beta S_\beta^2$ upon ligand binding to I-LBP is mainly caused by increase in the number (N_C) and order ($S_C^2 A_C^2$) of mobile water molecules in the binding cavity.

For all the proteins investigated here at 300 K, τ_γ is close to 1 ns (see Table 2). According to equation (S19) in Supplementary Material, with $N_C \gg 1$ and $\tau_R \gg \tau_\gamma$ (see Table 1), we can identify τ_γ with the average residence time τ_S of water molecules in individual hydration sites within the cavity. The site residence time τ_S is closely related to the contact times computed from the most recent MD simulation of I-FABP[19] and reported to be in the range 0.8–4.0 ns (mean 1.3 ns for apo and 1.5 ns for holo) for most of the water molecules that remained inside the cavity during the entire 5 ns MD trajectory. (A few cavity water molecules had shorter contact times.) The close agreement between the average contact time and the MRD-derived site residence time τ_S provides strong support for our interpretation of the γ dispersion in terms of intra-cavity water exchange. Moreover, the MD simulation indicates that ligand binding has little effect on the water exchange dynamics within the cavity,[19] as also suggested by the invariance of τ_γ. Our results for H-FABP and I-LBP (see Table 2) indicate that the 1 ns exchange dynamics of cavity water molecules is a general feature of the LBPs. MRD and MD simulation also provide mutually consistent lower bounds on the characteristic time τ_C for water escape from the binding cavity: $\tau_C \gg \tau_R = 7$ ns from our MRD results and $\tau_C \gg 5$ ns from the MD simulation.[19] The finding that $\tau_S \ll \tau_C$ implies that water molecules explore a large part of the cavity before escaping into the external solvent, as also suggested by the small value of the cavity order parameter A_C^2 (see above).

Although the DCM highlights the site-to-site mobility of the water molecules in the large binding cavity, the translational dynamics of these water molecules is nevertheless strongly retarded compared to bulk water. The 1 ns site residence time may be contrasted with the 6 ps required to translate 2.8 Å (the nearest-neighbor separation) in bulk water at 300 K. This 170-fold dynamic retardation cannot be attributed to an anomalously high viscosity of the cavity water, because momentum transport in water is mediated mainly by intermolecular vibrations rather than by molecular diffusion. Yet, it appears likely that the spatial and hydrogen-bonding constraints responsible for the retardation of water dynamics will likewise retard ligand displacement within the cavity as well as fluctuations in cavity shape. Fluctuations in cavity size, on the other hand, should be more strongly coupled to water exchange between the cavity and the external solvent, which occurs mainly on a longer time-scale ($\tau_C \gg 7$ ns).

The displacement of positionally ordered cavity water by the ligand, as inferred from crystal structures, prompted the suggestion that I-FABP (and other LBPs) acts like a "molecular water pump",[14] where the expulsion of ordered cavity water molecules provides an entropic driving force for ligand binding. This view contrasts with our conclusion, supported by MRD data as well as cavity calculations, that ligand binding is actually accompanied by a net influx of water into the cavity. The ligand-induced cavity enlargement and water influx is likely to play a significant role in the subtle binding thermodynamics of the LBPs.[48]

Materials and Methods

Preparation of MRD samples

Recombinant bovine H-FABP and porcine I-LBP were expressed, purified and delipidated as described.[23,24,49] Dilute I-LBP and H-FABP solutions (0.07–0.2 mM) were concentrated with Centriprep 3 and Centricon 10 concentrators (Amicon). The solution volume was then doubled by addition of water enriched to 19% ^{17}O (Isotec), but without deuterium. (The water used in the previous MRD study[21] of I-FABP contained 52% ^2H.) The MRD samples contained 50 mM potassium phosphate buffer and 0.025% (w/v) sodium azide and the pH was pH 7.95 for I-LBP and 7.45 for H-FABP.

The holo H-FABP sample contained a mixture of endogenous lipids with approximate composition: oleic acid (25%), palmitic acid (38%), palmitoleic acid (25%) and 6% each of stearic acid and delta-C19:0 fatty acid. The overall degree of lipidation was 83%. The apo I-LBP sample had a lipidation degree of 7%, mainly consisting of palmitic acid with traces of stearic acid, oleic acid and palmitoleic acid. After the MRD measurements, this sample was diluted sevenfold with 100 mM phosphate buffer containing 0.08% azide. Bile acid (chenodeoxycholate) was then added in tenfold excess, whereupon the sample was reconcentrated (by repeated Centricon 10 centrifugation) and again diluted to double volume with ^{17}O-enriched water. The holo I-LBP sample thus prepared contained 50 mM phosphate, 0.04% azide and the pH was pH 8.0. The apo H-FABP preparation was delipidated on a Lipidex 5000 column (Canberra-Packard). The delipidated sample was concentrated and mixed with H_2 ^{17}O to give a MRD sample with 50 mM phosphate at pH 7.4. The degree of lipidation was determined to 19% by gas chromatography. Some protein that aggregated during the concentration and buffer exchange process was removed by centrifugation.

The protein concentration C_P in all MRD samples was determined by complete amino acid analysis. The number N_T of water molecules per protein molecule was obtained from the relation $N_T = [1/(6.022 \times 10^{-7} \times C_P/\text{mM}) - V_P/\text{Å}^3]/(V_w/\text{Å}^3)$, with $V_w = M_w/(\rho_w N_A) = 30$ Å^3 the volume per water molecule and the solvent-excluded protein volume, V_P, calculated with the program GRASP.[50] For calculation of the number N_S of water molecules in contact with the protein surface, the accessible surface areas of the two proteins were determined with GRASP for the mean of the NMR solution structures of I-LBP (PDB code 1EAL,[25] five structures) and H-FABP

(1BWY,[23] two structures randomly picked out of 25). The result is 6980 Å^2 for I-LBP and 6900 Å^2 for I-LBP. The sample properties are collected in Table 1.

Magnetic relaxation measurements

Using four different NMR spectrometers, we measured the ^{17}O longitudinal relaxation rate $R_1 = 1/T_1$ at nine magnetic field strengths. The measurements were performed with conventional Varian 600 Unity Plus, Bruker Avance DMX 100 and DMX 200 spectrometers and with an iron-core electromagnet (Drusch EAR-35N) equipped with field-variable lock and flux stabilizer and interfaced to a Bruker MSL 100 console. The ^{17}O resonance frequencies were 2.19, 2.60, 4.04, 6.06, 8.69, 10.56, 13.57, 27.13, 49.04 and 81.44 MHz. The sample temperature was 4 °C or 27 °C, maintained to within 0.1 deg. C by a thermostated airflow and recorded with a copper-constantan thermocouple referenced to an ice-bath. This procedure normally gives a reproducibility of 0.5% in R_1. An additional temperature control was provided by measurements of the frequency-independent ^{17}O relaxation rate of a reference sample, containing pure water of the same isotopic composition as the protein sample. The longitudinal relaxation rate was measured by the inversion recovery method, using a 16-step phase cycle and a sufficient number of transients to obtain a signal-to-noise ratio of more than 100. Each T_1 measurement used 20 delay times in random order. The standard deviation in the fitted T_1 value was generally better than 0.5%. Considering the reproducibility and fitting error, we estimate the precision of the individual R_1 values to be 0.5%, but somewhat inferior at lower fields. The bulk water R_1 was 132.9(±0.3) s^{-1} at 27 °C and 262(±1) s^{-1} at 4 °C.

In the apo H-FABP sample, a small amount of protein precipitated during the course of the MRD experiments. After recording the full MRD profile at 27 °C, we removed the precipitate by centrifugation and then repeated the measurements at 0.38 T, 0.45 T and 1.83 T. The results were, within the estimated standard deviation, identical to the original measurements. The presented R_1 values are averages of the two measurements.

The MRD profiles were analyzed with an in-house Matlab implementation of the Levenberg–Marquardt non-linear χ^2 minimization algorithm.[34] To estimate the uncertainty in the fitted parameters, we performed fits on a Monte-Carlo generated ensemble of 1000 data sets, subject to random Gaussian noise with 0.5% standard deviation. The quoted uncertainties always correspond to a confidence level of 68.3% (one standard deviation).

Rotational correlation time

Since the MRD profiles consist of two overlapping dispersion steps, it is not possible to determine both correlation times with high accuracy. To improve the accuracy of the derived parameters, we performed the bi-Lorentzian fits with the correlation time τ_β fixed at a value determined from independent data.

For I-FABP, τ_β was previously determined (from a Lorentzian fit to the low-frequency part of the MRD profile) to be 7.1 ns in 52% ^2H_2O at 27 °C,[21] which scales to 6.4 ns in pure H_2O. Similar low-frequency fits to the present (less extensive) MRD data yield τ_β consistent with this value (except for apo H-FABP; see below). This agreement is expected because τ_β should be close to the rotational correlation time τ_R, which is virtually the same for the three LBPs. The latter point was confirmed by hydrodynamic calculations with the program HYDRONMR.[51]

The hydrodynamic calculations were performed with the method of true extrapolation to zero minibead radius, with five different values of the minibead radius ranging from 1.2 Å to 2 Å. The effective atom radius was set to 2.2 Å, which yields $\tau_R = 6.4$ ns for holo I-FABP (PDB code 1ICM). In addition to holo I-FABP, hydrodynamic calculations were done on the following proteins: apo I-FABP (1IFC), holo H-FABP (1HMS, 1HMR and 1HMT), apo I-LBP (1EAL, five NMR models) and holo I-LBP (1EIO, five NMR models). For all these proteins, $\tau_R = 6.2(\pm 0.2)$ ns, except for holo I-LBP, where $\tau_R = 5.7$ ns. The ratio between the largest and smallest principal component of the rotational diffusion tensor is 1.3 for all the proteins. This slight rotational anisotropy does not affect the MRD profile significantly.

From ^{15}N relaxation, $\tau_R = 7.3$ ns (scaled to H_2O, 27 °C) for holo H-FABP (pH 5.8) and holo I-LBP (pH 5.0)[52] and for the apo form of human H-FABP (pH 7.5).[53] For water residence times in the MRD window, $\tau_R \ll \tau_W \ll (\omega_Q^2 S^2 \tau_R)^{-1}$ (see Results), the ^{17}O correlation time is expected to be about 10% shorter than τ_R.[33] The MRD-derived value $\tau_\beta = 6.4$ ns is therefore consistent with the ^{15}N results for τ_R. Accordingly, we use $\tau_\beta = 6.4$ ns as a fixed parameter in the bi-Lorentzian fits.

For apo H-FABP, the MRD data indicate a slightly, but significantly, larger value. We attribute this finding to some oligomerization in this sample, where a small amount of protein was observed to precipitate during the MRD experiments (see above). Oligomerization is not expected to affect the other model parameters significantly. Based on the low-frequency fits, we adopt the value $\tau_\beta = 8$ ns for apo H-FABP.

Cavity calculations

Cavities were analyzed in the following protein structures: apo I-LBP (PDB code 1EAL[24]), glycocholate I-LBP complex (1EIO[25]), apo I-FABP (1IFC[9]), myristate I-FABP complex (1ICM[12]), and oleate H-FABP complex (1HMS[13]). The last structure refers to human H-FABP, whereas the MRD experiments were done on bovine H-FABP. With a sequence identity of 89% (most of the 14 substitutions conserve the polar/non-polar residue character) and the same mode of fatty acid binding, the binding cavities of the human and bovine H-FABP orthologs should be very similar.

To detect small cavities created by the ligand, the holo forms of I-LBP, H-FABP and I-FABP were analyzed with the program VOIDOO[36] and united-atom vdW radii from the original AMBER force field.[54] The cavities were refined in up to 13 steps with an initial grid size of 0.3 Å and with a grid shrinkage factor of 0.9 per step. Reported cavity volumes refer to the probe-occupied volume, calculated by inserting a probe sphere of radius 1.2 Å at all possible grid points. This algorithm gives the volume not occupied by protein atoms, minus the interstitial volume inaccessible to the probe. The probe-occupied volume is the complement to the molecular surface, as defined by Connolly.[55] Other authors[38] have analyzed LBP cavities in terms of the volume accessible to the center of the probe sphere. The probe-occupied volume

calculated here is more appropriate for assessing the hydration capacity of a cavity.

For visualization, the program FLOOD[36] was used to fill the cavities with 0.3 Å radius dummy atoms placed on a 0.35 Å grid. The dummy atoms were then surfaced with GRASP.[50] Each cavity analysis was performed ten times by translating the grid 0.1 Å, 0.2 Å or 0.3 Å in the *X*, *Y*, and *Z* directions. For the NMR structure 1EIO, each of the five models was analyzed with 0.15 Å translations in each direction, yielding a total of 20 calculations. VOIDOO was also used to assess cavity polarity by listing all atoms that line the cavity.

The large ligand-binding cavity was analyzed with GRASP as well as VOIDOO and with different sets of vdW radii, always with a probe radius of 1.2 Å. GRASP was run with the vdW parameters from Connolly's program[55] (after addition of hydrogen atoms) and from the original united-atom AMBER[54] and CHARMM[56] force fields. VOIDOO was run with AMBER and CHARMM parameters only. To prevent "leakage" of probe spheres from the cavity, blocking oxygen atoms were inserted at appropriate positions (see Supplementary Material). Reported volumes refer to GRASP/AMBER calculations. Within the computational accuracy, VOIDOO/AMBER volumes were identical.

Acknowledgements

We thank Vladimir Denisov and Hans Lilja for spectrometer assistance. This work was supported by the Crafoord Foundation and the Swedish Research Council (to B.H.). M.R. and C.L. thank Heinz Rüterjans (University of Frankfurt) for access to various laboratory equipment.

References

1. Glatz, J. F. C. & van der Vusse, G. J. (1996). Cellular fatty acid-binding proteins: their function and physiological significance. *Prog. Lipid Res.* **36**, 243–282.
2. Banaszak, L., Winter, N., Xu, Z., Bernlohr, D. A., Cowan, S. & Jones, A. T. (1994). Lipid-binding proteins: a family of fatty acid and retinoid transport proteins. *Advan. Protein Chem.* **45**, 89–151.
3. Hohoff, C. & Spener, F. (1998). Fatty acid binding proteins and mammary-derived growth inhibitor. *Fett. Lipid.* **100**, 252–263.
4. Hanhoff, T., Lücke, C. & Spener, F. (2002). Insights into binding of fatty acids by fatty acid binding proteins. *Mol. Cell. Biochem.* **239**, 45–54.
5. Vork, M. M., Glatz, J. F. C., Surtel, D. A. M., Knubben, H. J. M. & van der Vusse, G. J. (1991). A sandwich enzyme linked immuno-sorbent assay for the determination of rat heart fatty acid-binding protein using the streptavin-biotin system. Application to tissue and effluent samples from normoxic rat heart perfusion. *Biochim. Biophys. Acta*, **1075**, 199–205.
6. Zimmerman, A. W. & Veerkamp, J. H. (2002). New insights into the structure and function of fatty acid-binding proteins. *Cell. Mol. Life Sci.* **59**, 1096–1116.
7. Richieri, G. V., Ogata, R. T. & Kleinfeld, A. M. (1994). Equilibrium constants for the binding of fatty acids with fatty acid-binding proteins from adipocyte, intestine, heart, and liver measured with the fluorescent probe ADIFAB. *J. Biol. Chem.* **269**, 23918–23930.
8. Zimmerman, A. W., van Moerkerk, H. T. B. & Veerkamp, J. H. (2001). Ligand specificity and conformational stability of human fatty acid-binding proteins. *Int. J. Biochem. Cell Biol.* **33**, 865–876.
9. Scapin, G., Gordon, J. I. & Sacchettini, J. C. (1992). Refinement of the structure of recombinant rat intestinal fatty acid-binding apoprotein at 1.2-Å resolution. *J. Biol. Chem.* **267**, 4253–4269.
10. Zanotti, G., Scapin, G., Spadon, P., Veerkamp, J. H. & Sacchettini, J. C. (1992). Three-dimensional structure of recombinant human muscle fatty acid-binding protein. *J. Biol. Chem.* **267**, 18541–18550.
11. Sacchettini, J. C., Scapin, G., Gopaul, D. & Gordon, J. I. (1992). Refinement of the structure of *Escherichia coli*-derived rat intestinal fatty acid binding protein with bound oleate to 1.75-Å resolution. *J. Biol. Chem.* **267**, 23534–23545.
12. Eads, J., Sacchettini, J. C., Kromminga, A. & Gordon, J. I. (1993). *Escherichia coli*-derived rat intestinal fatty acid-binding protein with bound myristate at 1.5 Å resolution and I-FABP(Arg106 → Gln) with bound oleate at 1.74 Å resolution. *J. Biol. Chem.* **268**, 26375–26385.
13. Young, A. C. M., Scapin, G., Kromminga, A., Patel, S. B., Veerkamp, J. H. & Sacchettini, J. C. (1994). Structural studies on human muscle fatty acid binding protein at 1.4 Å resolution: binding interactions with three C18 fatty acids. *Structure*, **2**, 523–534.
14. Sacchettini, J. C. & Gordon, J. I. (1993). Rat intestinal fatty acid binding protein. *J. Biol. Chem.* **268**, 18399–18402.
15. Woolf, T. B. (1998). Simulations of fatty acid-binding proteins suggest sites important for function. I. Stearic acid. *Biophys. J.* **74**, 681–693.
16. Woolf, T. B. & Tychko, M. (1998). Simulations of fatty acid-binding proteins. II. Sites for discrimination of monounsaturated ligands. *Biophys. J.* **74**, 694–707.
17. Woolf, T. B., 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.
18. Likic, V. A. & Prendergast, F. G. (2001). Dynamics of internal waters in fatty acid binding protein: computer simulations and comparison with experiments. *Proteins: Struct. Funct. Genet.* **43**, 65–72.
19. Bakowies, D. & van Gunsteren, W. F. (2002). Simulations of apo and holo-fatty acid binding protein: structure and dynamics of protein, ligand and internal water. *J. Mol. Biol.* **315**, 713–736.
20. Mesgarzadeh, A., Pfeiffer, S., Engelke, J., Lassen, D. & Rüterjans, H. (1998). Bound water in apo and holo bovine heart fatty-acid-binding protein determined by heteronuclear NMR spectroscopy. *Eur. J. Biochem.* **251**, 781–786.
21. Wiesner, S., Kurian, E., Prendergast, F. G. & Halle, B. (1999). Water molecules in the binding cavity of intestinal fatty acid binding protein: dynamic characterization by water ^{17}O and ^{2}H magnetic relaxation dispersion. *J. Mol. Biol.* **286**, 233–246.
22. Likic, V. A., Juranic, N., Macura, S. & Prendergast, F. G. (2000). A structural water molecule in the

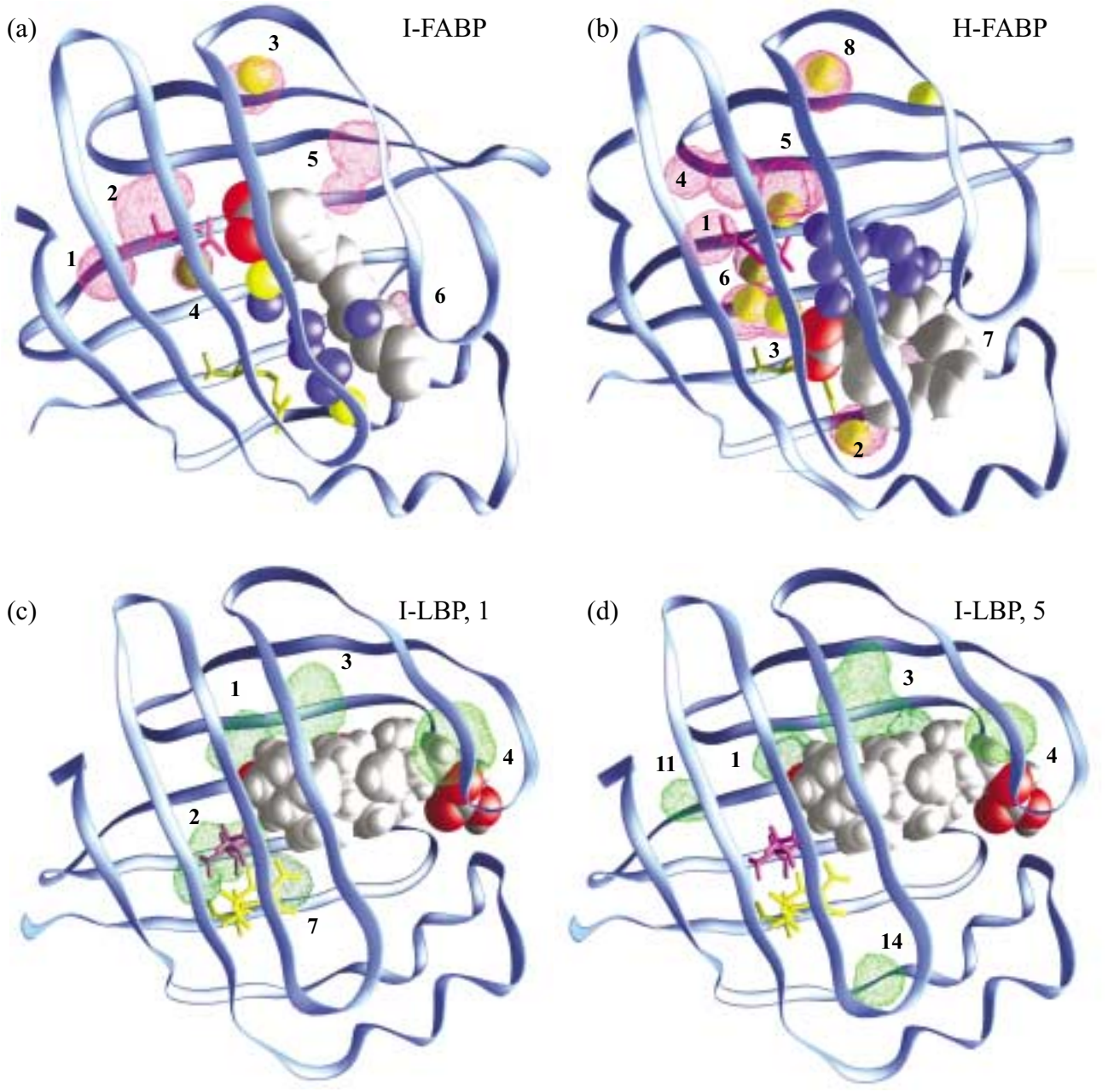
family of fatty acid binding proteins. *Protein Sci.* **9**, 497–504.
23. Lassen, D., Lücke, C., Kveder, M., Mesgarzadeh, A., Schmidt, J. M., Specht, B. *et al.* (1995). Three-dimensional structure of bovine heart fatty-acid-binding protein with bound palmitic acid, determined by multidimensional NMR spectroscopy. *Eur. J. Biochem.* **230**, 266–280.
24. Lücke, C., Zhang, F., Rüterjans, H., Hamilton, J. A. & Sacchettini, J. C. (1996). Flexibility is a likely determinant of binding specificity in the case of ileal lipid binding protein. *Structure*, **4**, 785–800.
25. Lücke, C., Zhang, F., Hamilton, J. A., Sacchettini, J. C. & Rüterjans, H. (2000). Solution structure of ileal lipid binding protein in complex with glycocholate. *Eur. J. Biochem.* **267**, 2929–2938.
26. Lücke, C., Rademacher, M., Zimmerman, A. W., van Moerkerk, H. T. B., Veerkamp, J. H. & Ruterjans, H. (2001). Spin-system heterogeneities indicate a selected-fit mechanism in fatty acid binding to heart-type fatty acid-binding protein (H-FABP). *Biochem. J.* **354**, 259–266.
27. Kramer, W., Sauber, K., Baringhaus, K.-H., Kurz, M., Stengelin, S., Lange, G. *et al.* (2001). Identification of the bile acid-binding site of ileal lipid-binding protein by photoaffinity labeling, matrix-assisted laser desorption ionization-mass spectrometry, and NMR structure. *J. Biol. Chem.* **276**, 7291–7301.
28. Zhu, L., Kurian, E., Prendergast, F. G. & Kemple, M. D. (1999). Dynamics of palmitic acid complexed with rat intestinal fatty acid binding protein. *Biochemistry*, **38**, 1554–1561.
29. Denisov, V. P. & Halle, B. (1996). Protein hydration dynamics in aqueous solution. *Faraday Discuss.* **103**, 227–244.
30. Halle, B. (1998). Water in biological systems: the NMR picture. In *Hydration Processes in Biology* (Bellisent-Funel, M.-C., ed.), pp. 233–249, IOS Press, Dordrecht.
31. Denisov, V. P. & Halle, B. (1995). Protein hydration dynamics in aqueous solution: a comparison of bovine pancreatic trypsin inhibitor and ubiquitin by oxygen-17 spin relaxation dispersion. *J. Mol. Biol.* **245**, 682–697.
32. Halle, B., Denisov, V. P. & Venu, K. (1999). Multinuclear relaxation dispersion studies of protein hydration. In *Biological Magnetic Resonance* (Krishna, N. R. & Berliner, L. J., eds), vol. 17, pp. 419–484, Kluwer/Plenum, New York.
33. Halle, B. & Denisov, V. P. (2001). Magnetic relaxation dispersion studies of biomolecular solutions. *Methods Enzymol.* **338**, 178–201.
34. Press, W. H., Flannery, B. P., Teukolsky, S. A. & Vetterling, W. T. (1992). *Numerical Recipes in C. The Art of Scientific Computing*, 2nd edit., Cambridge University Press, Cambridge.
35. Gottschalk, M., Nilsson, H., Roos, H., Halle, B. (2003). Protein self-association in solution: the bovine β-lactoglobulin dimer and octamer. *Protein Sci*. In the press.
36. Kleywegt, G. J. & Jones, A. T. (1993). Detection, dilineation, measurement and display of cavities in macromolecular structures. *Acta Crystallog. sect. D*, **50**, 178–185.
37. Likic, V. A. & Prendergast, F. G. (1999). Structure and dynamics of the fatty acid binding cavity in apo rat intestinal fatty acid binding protein. *Protein Sci.* **8**, 1649–1657.
38. Thompson, J., Ory, J., Reese-Wagoner, A. & Banaszak, L. (1999). The liver fatty acid binding protein—comparison of cavity properties of intracellular lipid-binding proteins. *Mol. Cell. Biochem.* **192**, 9–16.
39. Gerstein, M. & Chothia, C. (1996). Packing at the protein–water interface. *Proc. Natl Acad. Sci. USA*, **93**, 10167–10172.
40. Tochtrop, G. P., Richter, K., Tang, C. G., 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.
41. Kramer, W., Corsiero, D., Friedrich, M., Girbig, F., Stengelin, S. & Weyland, C. (1998). Intestinal absorption of bile acids: paradoxical behaviour of the 14 kDa ileal lipid-binding protein in differential photoaffinity labelling. *Biochem. J.* **333**, 335–341.
42. Impey, R. W., Madden, P. A. & McDonald, I. R. (1983). Hydration and mobility of ions in solution. *J. Phys. Chem.* **27**, 5071–5083.
43. Bakowies, D. & van Gunsteren, W. F. (2002). Water in protein cavities: a procedure to identify internal water and exchange pathways and application to fatty acid-binding protein. *Proteins: Struct. Funct. Genet.* **47**, 534–545.
44. Lücke, C., Huang, S., Rademacher, M. & Rüterjans, H. (2002). New insights into intracellular lipid binding proteins: the role of buried water. *Protein Sci.* **11**, 2382–2392.
45. Otting, G., Liepinsh, E. & Wüthrich, K. (1991). Protein hydration in aqueous solution. *Science*, **254**, 974–980.
46. Otting, G. (1997). NMR studies of water bound to biological molecules. *Prog. Nucl. Magn. Reson. Spectrosc.* **31**, 259–285.
47. Sobol, A. G., Wider, G., Iwai, H. & Wüthrich, K. (1998). Solvent magnetization artefacts in high-field NMR studies of macromoleular hydration. *J. Magn. Reson.* **130**, 262–271.
48. Richieri, G. V., Ogata, R. T. & Kleinfeld, A. M. (1999). Fatty acid interactions with native and mutant fatty acid binding proteins. *Mol. Cell. Biochem.* **192**, 77–85.
49. Glatz, J. F. C. & Veerkamp, J. H. (1983). A radiochemical procedure for the assay of fatty acid binding by proteins. *Anal. Biochem.* **132**, 89–95.
50. Nicholls, A., Bharadwaj, R. & Honig, B. (1993). GRASP: graphical representation and analysis of surface properties. *Biophys. J.* **64**, 166–170.
51. Garcia de la Torre, J., Huertas, L. & Carrasco, B. (2000). HYDRONMR: prediction of NMR relaxation of globular proteins from atomic-level structures and hydrodynamic calculations. *J. Magn. Reson.* **147**, 138–146.
52. Lücke, C., Fushman, D., Ludwig, C., Hamilton, J. A., Sacchettini, J. C. & Rüterjans, H. (1999). A comparative study of the backbone dynamics of two closely related lipid binding proteins: bovine heart fatty acid binding protein and procine ileal lipid binding protein. *Mol. Cell. Biochem.* **192**, 109–121.
53. Constantine, K. L., Friedrichs, M. S., Wittekind, M., Jamil, H., Chu, C.-H., Parker, R. A. *et al.* (1998). Backbone and side chain dynamics of uncomplexed human adipocyte and muscle fatty acid-binding proteins. *Biochemistry*, **37**, 7965–7980.
54. Weiner, S. J., Kollman, P. A., Case, D. A., Singh, C. U., Ghio, C., Alagona, G. *et al.* (1984). A new force filed

for molecular mechanical simulation of nucleic acids and proteins. *J. Am. Chem. Soc.* **106**, 765–784.
55. Connolly, M. (1983). Analytical molecular surface calculation. *J. Appl. Crystallog.* **16**, 548–558.
56. Brooks, B. R., Bruccoleri, R. E., Olafson, B. D., States, D. J., Swaminathan, S. & Karplus, M. (1983). CHARMM: a program for macromolecular energy, minimization, and dynamics calculations. *J. Comput. Chem.* **4**, 187–217.

Edited by P. Wright

(Received 7 April 2003; received in revised form 21 July 2003; accepted 21 July 2003)

Supplementary Material for this paper is available on Science Direct


(a)
I-FABP
3
5
2
1
4
6
(b)
H-FABP
8
5
4
1
6
3
7
2
(c)
I-LBP, 1
3
1
4
2
7
(d)
I-LBP, 5
3
11
1
4
14

Supplementary Material to "Water Dynamics in the Large Cavity of Three Lipid-Binding Proteins Monitored by ^{17}O Magnetic Relaxation Dispersion"

K. Modig, M. Rademacher, C. Lücke och B. Halle

Spectral density for the dynamic cluster model

We derive here the quadrupolar (^{2}H or ^{17}O) spectral density function for water molecules exchanging among hydration sites within a cavity in a protein undergoing rotational diffusion. The reduced time correlation function (TCF) for the water molecules in the cavity is

$$g^{\mathrm{C}}(t) = \left\langle U_0^{\mathrm{L}*}(0)\, U_0^{\mathrm{L}}(t) \right\rangle \tag{S1}$$

where $U_0^{\mathrm{L}} = V_0^{\mathrm{L}} / \langle |V_0^{\mathrm{L}}|^2 \rangle^{1/2}$ and V_0^{L} is the $m = 0$ spherical component of the irreducible electric field gradient (EFG) tensor in a lab-fixed frame.[1] The fluctuations in $U_0^{\mathrm{L}}(t)$ are due to four processes: (1) local water motion (mainly librations and C_2 flips) in each hydration site, (2) water exchange among hydration sites within the cavity, (3) water escape from the cavity into the bulk solution, and (4) rotational diffusion of the protein molecule.

If local water motions (loc) are much faster than site exchange (X) and protein rotation (R), the TCF can be split in two independent parts:

$$g^{\mathrm{C}}(t) = g^{\mathrm{C}}_{\mathrm{loc}}(t) + g^{\mathrm{C}}_{\mathrm{RX}}(t) \tag{S2}$$

Next, we assume that site exchange and protein rotation are statistically independent (but not necessarily time-scale separated) processes. Furthermore, we assume that protein rotational diffusion is isotropic and that the cavity escape kinetics are site-independent. After some algebra, the second term in equation (S2) can then be expressed as

$$g^{\mathrm{C}}_{\mathrm{RX}}(t) = S^2_{\mathrm{C}}[A^2_{\mathrm{C}} + (1 - A^2_{\mathrm{C}})\, g_{\mathrm{X}}(t)] \exp(-t/\tau_{\mathrm{RC}}) \tag{S3}$$

with

$$\frac{1}{\tau_{\mathrm{RC}}} = \frac{1}{\tau_{\mathrm{R}}} + \frac{1}{\tau_{\mathrm{C}}} \tag{S4}$$

Here, τ_R is the usual rank-2 rotational correlation time of the protein and τ_C is the mean residence time of a water molecule inside the cavity.

The two orientational order parameters in equation (S3) are defined through

$$S_{\mathrm{C}}^2 = (1+\eta^2/3)^{-1}\left\langle \sum_m \left| \sum_n \left\langle D_{mn}^2(\Omega_{\mathrm{SM}})\right\rangle_{\mathrm{S}} \zeta_n \right|^2 \right\rangle_{\mathrm{C}} \tag{S5}$$

$$A_{\mathrm{C}}^2 = \sum_m \left| \sum_n \left\langle D_{mn}^2(\Omega_{\mathrm{CS}})\zeta_n \right\rangle_{\mathrm{C}} \right|^2 \tag{S6}$$

Here, $D_{mn}^2(\Omega)$ is the rank-2 Wigner function of the Euler angles Ω.[2] The angles Ω_{SM} specify the orientation of the water molecule relative to the principal frame of the EFG tensor averaged over local motions in a particular hydration site. The angles Ω_{CS} relate this local EFG frame to the principal frame of the EFG tensor averaged over all hydration sites within the cavity. The coefficients ζ_n and ζ_n are given by

$$\zeta_n = \frac{\sum_p \left\langle D_{np}^2(\Omega_{\mathrm{SM}})\right\rangle_{\mathrm{S}} \zeta_p}{\left(\sum_m \left| \sum_p D_{mp}^2(\Omega_{\mathrm{SM}})\zeta_p \right|^2 \right)^{1/2}} \tag{S7}$$

$$\zeta_n = D_{n0}^2(\Omega_{\mathrm{MF}}) + \frac{\eta}{\sqrt{6}}[D_{n2}^2(\Omega_{\mathrm{MF}}) + D_{n,-2}^2(\Omega_{\mathrm{MF}})] \tag{S8}$$

The fixed angles Ω_{MF} relate the principal frame of the instantaneous EFG tensor to another frame fixed in the water molecule[3] and η is the asymmetry parameter of the instantaneous EFG tensor.[1] In equations (S5) – (S7), $\langle\cdots\rangle_{\mathrm{S}}$ denotes an average over local motions in a hydration site and $\langle\cdots\rangle_{\mathrm{C}}$ an average over all hydration sites in the cavity.

The reduced TCF $g_{\mathrm{X}}(t)$ in equation (S3) describes the effect of water exchange among the hydration sites in the cavity. We adopt a Markovian jump model with N_{C} equally populated sites. For simplicity, we assume that the mean residence time τ_{S} of a water molecule in a hydration site is the same for all sites and that exchange from a given site leads to any of the other $N_{\mathrm{C}} - 1$ sites with equal probability. For this exchange model,

$$g_{\mathrm{X}}(t) = \exp(-t/\tau_{\mathrm{X}}) \tag{S9}$$

with

$$\tau_{\mathrm{X}} = \left(1 - \frac{1}{N_{\mathrm{C}}}\right)\tau_{\mathrm{S}} \tag{S10}$$

The TCF for the cavity waters is now obtained by combining equations (S2), (S3) and (S9):

$$g^{\mathrm{C}}(t) = (1 - S_{\mathrm{C}}^2)\tau_{\mathrm{C}}^{\mathrm{loc}} + S_{\mathrm{C}}^2 A_{\mathrm{C}}^2 \exp(-t/\tau_{\mathrm{RC}}) + S_{\mathrm{C}}^2(1 - A_{\mathrm{C}}^2)\exp(-t/\tau_{\mathrm{RCX}}) \tag{S11}$$

with

$$\tau_{\mathrm{C}}^{\mathrm{loc}} = \int_0^\infty dt\, g_{\mathrm{loc}}^{\mathrm{C}}(t) / g_{\mathrm{loc}}^{\mathrm{C}}(0) \tag{S12}$$

and

$$\frac{1}{\tau_{\mathrm{RCX}}} = \frac{1}{\tau_{\mathrm{RC}}} + \frac{1}{\tau_{\mathrm{X}}} \tag{S13}$$

In the limit of slow intersite exchange, where $\tau_X >> \tau_{RC}$, equation (S13) yields $\tau_{RCX} = \tau_{RC}$. In this limit, which includes the case of a singly buried internal water molecule, equation (S11) reduces to the familiar result[3-5]

$$g^{\mathrm{C}}(t) = (1 - S_{\mathrm{C}}^2)\tau_{\mathrm{C}}^{\mathrm{loc}} + S_{\mathrm{C}}^2 \exp(-t/\tau_{\mathrm{RC}}) \tag{S14}$$

Consider now a protein containing N_C mutually interchanging cavity waters and N_I other internal water molecules that exchange directly with the bulk solution. We assume that the exchange of these water molecules with the external solvent is fast compared to the spin relaxation time scale, but slow compared to the rotational diffusion of the protein. The spectral density function $J(\omega_0)$ that enters equation (1) of the main text can then be obtained as the cosine transform of the population-weighted average of the TCFs in equations (S11) and (S14), the latter with C replaced by I. The result can then be expressed on the form of equation (2), with

$$\tau = \frac{N_{\mathrm{S}}}{N_{\mathrm{T}}}\left(\frac{\tau_{\mathrm{S}}}{\tau_{\mathrm{bulk}}} - 1\right)R_{\mathrm{bulk}} + \frac{\Delta_{\mathrm{Q}}^2}{N_{\mathrm{T}}}\left[N_{\mathrm{I}}(1 - S_{\mathrm{I}}^2)\tau_{\mathrm{loc}}^{\mathrm{I}} + N_{\mathrm{C}}(1 - S_{\mathrm{C}}^2)\tau_{\mathrm{loc}}^{\mathrm{C}}\right] \tag{S15}$$

$$\tau = \frac{\Delta_{\mathrm{Q}}^2}{N_{\mathrm{T}}}(N_{\mathrm{I}}S_{\mathrm{I}}^2 + N_{\mathrm{C}}S_{\mathrm{C}}^2 A_{\mathrm{C}}^2) \tag{S16}$$

$$\tau = \frac{\Delta_{\mathrm{Q}}^2}{N_{\mathrm{T}}} N_{\mathrm{C}}S_{\mathrm{C}}^2(1 - A_{\mathrm{C}}^2) \tag{S17}$$

$$\tau_{\tau} = \tau_{\mathrm{R}} \tag{S18}$$

$$\tau_{\perp} = \left(\frac{1}{\tau_R} + \frac{N_C}{(N_C - 1)\tau_S} \right)^{-1} \quad \text{(S19)}$$

The second term in equation (S15) can usually be neglected because $N_I + N_C << N_S$ (the number of water molecules in contact with the external protein surface). Equation (2) with the parameters given by equations (S15)-(S19) defines the dynamic cluster model (DCM) discussed in the main text. The static cluster model (SCM) is recovered from these expressions by letting $\tau_S \to \infty$ and $A_C = 0$.

Cavity calculations – small cavities

As a quantitative measure of the polarity of a cavity, we calculate the hydrophilic atom surface density (HSD)

$$HSD = \frac{N_P}{A_C} \approx \frac{N_P}{(6\sqrt{\pi} V_C)^{2/3}} \quad \text{(S20)}$$

where A_C is the probe-accessible surface area of the cavity, V_C the corresponding cavity volume and N_P the number of polar atoms lining the cavity. Because VOIDOO only calculates V_C (not A_C), we invoke the approximation $A_C = (6\sqrt{\pi} V_C)^{2/3}$, rigorously valid for a spherical cavity. The small cavities detected in I-FABP, H-FABP and I-LBP are shown in Figure S1. The residual part of the binding cavity not occupied by ligand and not included in the small cavities, is approximately 300 ± 50 Å^3 for each of the three proteins.

I-FABP

No water-occupied cavities are created by the ligand (Table S1). The only cavity created by the ligand appears to be empty (cavity 5 in Table S1 and Figure S1), even though it has a fairly high HSD.

Table S1. Small cavities in I-FABP with bound myristate

Cavity	V (Å^3)	HSD (Å^{-2})	Water	FA lining	Comment
1	10.3	0.00			Only found in 1 out of 10 translations
2	53.4 ± 2.8	0.00			
3	20.3 ± 2.3	0.25	134 (135)		
4	13.7 ± 0.8	0.29	158 (217)		
5	34.0 ± 2.5	0.20		CA, C3-6	
6	12.6 ± 1.2	0.19			

PDB file 1ICM was used. Note that cavity 5 is created by the ligand. Water numbering refers to the PDB file 1ICM (1IFC). The quoted error in the volume is the standard deviation for all translations (see Materials and Methods).

H-FABP

The cavities and their water molecules are described in Tables S2 and S3, respectively. The coding refers to PDB file 1HMS. All water molecules outside the binding cavity, except W166, are slightly solvent-exposed in at least one of the three investigated holo proteins. Only W166 and W222 have close contact with both hydrogen bond donors and acceptors and only for W166 all these are backbone atoms. This indicates that W166 has the longest residence time.

In addition to these water molecules, four water molecules are found in small sub-cavities within or in the vicinity of the binding cavity: W167 (cavity 5), W174, W175 (cavity 3) and W181 (cavity 2) These are created directly by the ligand, or by amino acid side-chains interacting with it. Thus, it is unlikely that these water molecules are isolated also in the apo protein. Among these water molecules, W167 may be responsible for some of the NOE cross-peaks with surrounding protein protons.[6] W167 is shielded from the water cluster by the side-chain of Arg106. W174 and W175 bridge the fatty acid head-group to two of the nitrogen atoms in the side-chain of Arg106. W181 is situated between the fatty acid, the C-terminal end of helix II and is within hydrogen-bonding distance of the side-chain of Arg126, which is directly involved in binding of the fatty acid head-group.

Four other water molecules in the binding cavity are potentially perturbed by the ligand, but are not found in separate sub-cavities: W138, W152, W178 and W183. W138 is located on the periphery of the water cluster, in the vicinity of Arg106, and contacts the cluster via W153. W152 has two other water contacts and coordinates the fatty acid head-group. W178 is located close to the portal region and interacts with W142. Finally, W183 is located between β-strands C and D and the rest of the water cluster. It is connected to the water cluster through a single water molecule. The total volume of the ligand-created cavities (2-3 and 5) is 140 $Å^3$.

Two cavities, 1 and 7, are small and appear hydrophilic, but are not occupied by crystallographically visible water molecules.

Table S2. Small cavities in H-FABP with bound oleate

Cavity	V ($Å^3$)	HSD ($Å^{-2}$)	Water	FA lining	Comment
1	17.5 ± 1.0	0.25			Outside binding cavity
2	23.5 ± 1.1	0.20	181	C13-16	Close to Arg126.N
3	42.8 ± 1.9	0.10	174, 175	C, O, OXT	
4	60.6 ± 3.0	0.12			Above Arg106, close to cavity 2 in holo I-FABP
5	71.2 ± 4.9	0.11	167		Above Arg106
6	14.6 ± 1.2	0.31	191 (217)		Corresponds to cavity 4 in holo I-FABP
7	13.7 ± 1.6	0.22			Between helix I and strand K
8	11.9 ± 0.2	0.28	166 (135)		Corresponds to cavity 3 in holo I-FABP

PDB file 1HMS was used. Note that cavities 2 and 3 are created by the ligand. Cavities 6 and 8 have counter parts in I-FABP. Water numbering refers to the PDB file 1HMS (1IFC). The quoted error in the volume is the standard deviation for all translations.

I-LBP

The small cavities are listed in Tables S4a and S4b. In the NMR structures, the ligand is modeled so that the hydroxyl group in ring A appears to be hydrogen bonded to the carboxyl group of Glu110, which in turn is hydrogen bonded to Arg121.[7] If the docking is correct, cavity 1 has room for two water molecules, which could make hydrogen bonds to the hydroxyl group of bile acid ring A and the protein. Another small cavity (2) is shielded from the binding cavity by Arg121 and Glu110 and the bile acid. This cavity is not very hydrophilic, but comparable to the cavity (3) of H-FABP, which is modeled as containing two water molecules in the crystal structure. Water molecules trapped in this cavity can escape only if the hydrogen bond network bile acid – Glu110 – Arg121 is broken.

The ligand creates two other fairly hydrophilic cavities (4 and 5) that may contain water molecules, and one more hydrophobic cavity (6). In addition, cavity 10 may accommodate a water molecule that needs to pass the bile acid to sneak out into “the gap” or ligand cavity. The total volume of the ligand-created cavities (1-6 and 8) is 170-270 $Å^3$, corresponding to 6-14 water molecules.[8]

Turning to cavities that are expected to be present also in the apo form, VOIDOO finds seven cavities not directly in contact with the binding cavity. Of these, only one (11) is seen in more than one of the five NMR models and only one (10) is hydrophilic enough to accommodate a water molecule. A cavity corresponding to W166 of H-FABP is not found, but this cavity should be very small and may therefore escape detection unless the coordinates are highly accurate.

Table S3. Potentially long-lived water molecules in human holo H-FABP

Water [a]	Cavity [b]	B (Å^2) [c]	A_S (Å^2) [d]	H-bond [e]	R_{OX} (Å) [f]	NOE [g]
166	8	8.3-10.2	0.0	Lys65:O	2.8	Gly67
143				Val68:O	2.8	Val68
(135)				Val84:N	2.9	Val84
171	–	13.8-16.2	2.1-2.8	Asp12:O□	2.7	
187				Thr125:O	3.0	
				Thr127:O□	2.9	
191	6	14.9-16.3	0.7-1.2	Val105:O	3.2	
224				Glu107:O□	2.8	
(217)				Ile114:O	3.0	
				Thr116: O□	2.7	
222	–	19.8-21.0	0.0-2.6	Lys81:O	2.8	Ile83
243				Leu94:O	3.1	
				Lys96:N□	2.7	
138	C	16.8-21.1	0.9-1.4	Arg106:N□1	2.7	
137				Arg106:N□2	3.2	
				W153	2.7	
152	C	11.1-14.6	0.0	Thr53:O□	2.9	
139				FA:O	2.7	
				W153	2.8	
				W179	2.9	
167	C (5)	10.2-13.0	1.2-1.4	Leu91:O	2.8	Arg106
134				Leu104:O	2.9	
				Arg106:N□	3.0	
174	C (3)	11.6-13.4	0.0	Thr40:O□	2.8	
140				Arg106:N□	2.9	
				FA:O	2.8	
				W175	2.7	

175	C (3)	16.8-17.1	0.0	Arg106:N_{η}	2.9	
142				W174	2.7	
178	C	11.2-13.1	0.0	Tyr19:O_{η}	3.1	
147				Arg78:N_{η}	2.9	
				Gln95:O_{ε}	2.7	
				W142	3.0	
181	C (2)	18.7-23.3	0-1.0	Ala33:O	2.7	
135				Thr36:O	2.7	
				Arg126:N_{η}	2.9	
183	C	11.2-13.6	0.0	Leu51:O	2.9	Lys52
145				Thr53:N	3.3	Ile62
				Thr53:O_{γ}	3.0	
				Thr60:O	3.4	
				Thr60: O_{γ}	2.8	
				W149	2.7	

[a] The first number refers to 1HMS and 1HMR labeling, the second to 1HMT. Number in parenthesis refers to homologous water molecules in apo I-FABP (PDB file 1IFC).

[b] The numbering refers to Table S2. C indicates that the water is found in the binding cavity. A number within parenthesis indicates that the water is situated in the ligand cavity, but is sealed off from it by the ligand or a residue perturbed by the ligand.

[c] Crystallographic B-factor for water oxygen in PDB files, 1HMS, 1HMR and 1HMT.

[d] Solvent-accessible surface area of water molecule (probe radius 1.2 Å) calculated with GRASP, using PDB files 1HMS, 1HMR and 1HMT.

[e] Water residue numbers refer to labeling in 1HMS.

[f] Only distances in 1HMS are shown. The same potential HB partners were found in 1HMR and 1HMT.

[g] Observed NOEs in the study by Mesgarzadeh *et al*.[6] The water molecules were assigned using the PDB file 1HMS.

Table S4a. Small cavities in I-LBP with bound glycocholate

Cavity	Model 1		Model 2		Model 3		Model 4		Model 5	
	V (Å^3)	HSD (Å^{-2})	V (Å^3)	HSD (Å^{-2})	V (Å^3)	HSD (Å^{-2})	V (Å^3)	HSD (Å^{-2})	V (Å^3)	HSD (Å^{-2})
1	41.5 ± 2.2	0.14	34.3 ± 4.9	0.18	21.8 ± 3.0	0.21	53.6 ± 1.3	0.15	42.9 ± 1.6	0.13
2	67.6 ± 1.5	0.19	81.3 ± 1.2	0.07	76.8 ± 0.2	0.08	75.4 ± 0.2	0.13	9.4	
3	45.5 ±1.8	0.00			79.0 ±2.0	0.01	19.0 ±2.3	0.00	87.5 ±0.4	0.07
4	71.0 ± 1.8	0.16			39.1 ± 2.0	0.18			32.3 ± 2.9	0.18
5			19.8 ± 0.7	0.23	28.6 ± 4.0	0.18	26.0 ± 1.0	0.16		
6					26.2 ± 0.9	0.02	14.9 ± 0.8	0.14		
7	21.7 ± 1.2	0.21								
8			59.4 ± 0.3	0.08						
9			18.7 ± 4.9	0.18						
10			10.4 ± 0.4	0.30						
11			13.6 ± 1.2	0.14					15.6 ± 2.0	0.07
12							18.9 ± 2.3	0.15		
13							29.9 ± 0.3	0.00		
14									24.3 ± 0.7	0.15

PDB file 1EIO was used. Note that cavities 1, 2, 3, 4, 5, 6 and 8 are created by the ligand. Also note that several cavities are not found in all five models, and that the cavity volume varies considerably among the models. Further information is found in Table S4b. The quoted error in the volume is the standard deviation for all translations. Some of the cavities are visualised in Figure S1.

Table S4b. More information on the cavities listed in Table S4a

Cavity	Lining bile atoms	Comment
1	C, C4-5, O	
2	C,C1-2,C7,O	Shielded from ligand bile acid by Glu110 and Arg121in some models
3	C3-5, C10-13, C18-19, O3	
4	C13, C16-23, O2	
5	C, C1-7	In front of ligand
6	C6-11, C23	In the middle of "the gap"
7		
8	CA, C12, C18-22, C24, N, O3, O5	Close to cavity 4, but on the other side of ligand
9		Outside cavity, near strands G/H
10		Behind ligand
11		
12		In front of/under ligand, between strands C/D
13		Close to cavity 1, but closer to N-terminal
14		May correspond to cavity 2 in H-FABP

An atom is lining if it is lining in any of the five models. For instance, O3 of glycocholate is lining cavity 3 in all models except model 1. The positions of cavities that are not shown in Figure S1 are indicated.

Cavity calculations – the binding cavity

The results of our calculations and others found in the literature are compiled in Table S5. In addition to the LBPs investigated by MRD, we also calculated the cavity volumes of apo and holo adipocyte lipid-binding protein (A-LBP). As is evident from the table, the volumes depend strongly on the algorithm, probe radius and atom parameter set. Nevertheless, it is also evident that the cavity expands substantially upon ligand binding (see I-FABP, I-LBP and A-LBP). As pointed out by Likic and Prendergast,[9] the I-FABP cavity expands its hydrophobic surface upon ligand binding, and some side chains need to change their conformation to accommodate the fatty acid. For I-LBP, most of the cavity expansion seems to arise from the large change in position of strands E and F, required to accommodate the bulky bile acid.

The common cavity detection algorithms only identifies cavities that are closed in the sense that they cannot be accessed by an external probe sphere. Because the binding cavity in the LBPs has several small holes that can be penetrated by a 1.2 Å probe, the cavity volume was calculated after these holes had been sealed by artificial blocking atoms. The positions of these blocking atoms are given in Table S6.

Table S5. Volume of binding cavity in four lipid-binding proteins.

Protein	Type	Probe	Program / param	V (Å^3)	Reference
I-FABP, apo	PO	1.2	GR / Am	521	This work
I-FABP, holo	PO	1.2	GR / Am	926	This work
H-FABP, holo	PO	1.2	GR / Am	911	This work
I-LBP, apo	PO	1.2	GR / Am	401	This work
I-LBP, holo	PO	1.2	GR / Am	1013	This work
A-LBP, apo [a]	PO	1.2	GR / Am	743	This work
A-LBP, holo [a]	PO	1.2	GR / Am	893	This work
I-FABP	PO?	1.0	Qo / Co	850	Scapin *et al.* (1992)[10]
H-FABP	PO?	1.0	Qo / Co	850	Scapin *et al.* (1993)[11]
I-FABP apo	PO	1.2	VO / CH22	609 ± 4	Liki_ & Prendergast[9]
I-FABP apo	PA	1.2	VO / CH22	109 ± 1	Liki_ & Prendergast[9]
I-FABP apo	PO	1.4	VO / CH22	539 ± 8	Liki_ & Prendergast[9]
I-FABP apo	PA	1.4	VO / CH22	61 ± 1	Liki_ & Prendergast[9]
I-FABP apo	PO	3.0	LB	451	Banaszak *et al.*[12]
I-LBP, holo	PO	?	GR / Co	1100	Lücke *et al.*[7]
I-FABP, holo	PA	1.4	?	234	Thompson *et al.*[13]
H-FABP	PA	1.4	?	323	Thompson *et al.*[13]

For the NMR structures, hydrogen atoms were removed prior to the volume calculation.

Abbreviations: PO, probe-occupied volume; PA, probe-accessible volume; GR, GRASP; VO, VOIDOO; Am, Amber; CH, original united-atom CHARMM; CH22, CHARMM22;[14] Qu, Quanta; Co, Connolly's parameter set;[15] LB, method of Levitt and Banaszak.[16]

[a] For adipocyte lipid-binding protein, the PDB files were 1LIB (apo) and 1LID (holo, oleic acid).[17]

Table S6. Positions of blocking atoms inserted to close the FA binding cavity. The blocking atom was placed halfway between the atoms in each atom pair in the list. All blocking atoms were given the size of an oxygen atom.

PDB file	Blocking atoms
1EIO, mod 1	The glycine part of glycocholate; Asp26.O–Arg121.NH2; Ile59.CG1–Tyr53.CE2; Ile74.CG1–Asn61.ND2; Leu34.O–Tyr14.CB; Leu23.CG–Ile59.CG1; Glu72.O–Tyr97.OH; Thr62.O–Asp70.O.
1EAL, mod 1	Met18.CE–Ile74.CG2 (Blk128); Met18.CE–Tyr53.OH; Tyr53.CE2–Ala31.CB; Glu11.OE2–Asn33.O; Tyr119.OH–Arg32.O; Lys35.CG–Glu11.OE1; Leu21.CD2–Tyr97.CD1; Blk128–Tyr53.CE1; Phe17.CE1–Asn96.O; Thr62.N–Asp70.O; Ile71.CD1–Phe63.CD2
1IFC	Gly31.CA–Asp74.CA (Blk384); Met21.CA–Asp74.CA; Val60.CA–Tyr70.CA; Asp74.CB–Phe55.CE2; Gly31.CA–Ile23.CD1; Blk384–Phe55.CZ; Blk384–Ile23.CD1; Glu63.O–Phe68.CD1
1ICM	Ile23.CG2–Asp74.N; Lys27.CE–Ala73.CB; Ile23.CD1–Leu72.CD1
1HMS	Thr29.CG2–Asp76.CB; Thr29.CG2–Lys58.CD; Glu72.CG–Thr60.OG1; Thr73.O–Ile62.CG2
1LIB	Thr60.CB–Glu72:OE2
1LID	Fatty acid carbons 16-18; Ala75.O–Thr29.CG2; Ola133.C17–Thr29.CG2; Ola133.C17–Ala75.CB; Thr60.CG2–Ile73.O; Thr60.CG2–Glu72.OE2; Thr60.CB–Ile62.CG1; Val32.CG1–Thr29.CG2; Gly34.O–Arg126.CZ; Gly34.O–Ala36.O; Gly34.N–Phe16.CD2

References

1. Abragam, A. (1961). *The Principles of Nuclear Magnetism*, Clarendon Press, Oxford.
2. Brink, D. M. & Satchler, G. R. (1968). *Angular Momentum*. 2nd ed., Clarendon Press, Oxford.
3. Halle, B., Denisov, V. P. & Venu, K. (1999). Multinuclear relaxation dispersion studies of protein hydration. In *Biological Magnetic Resonance* (Krishna, N. R. & Berliner, L. J., eds.), Vol. **17**, pp. 419-484. Kluwer/Plenum, New York.
4. Halle, B. & Denisov, V. P. (2001). Magnetic relaxation dispersion studies of biomolecular solutions. *Methods Enzymol.* **338**, 178-201.
5. Halle, B. & Wennerström, H. (1981). Interpretation of magnetic resonance data from water nuclei in heterogenous systems. *J. Chem. Phys.* **75**, 1928-1943.
6. Mesgarzadeh, A., Pfeiffer, S., Engelke, J., Lassen, D. & Rüterjans, H. (1998). Bound water in apo and holo bovine heart fatty-acid-binding protein determined by heteronuclear NMR spectroscopy. *Eur. J. Biochem.* **251**, 781-786.
7. Lücke, C., Zhang, F., Hamilton, J. A., Sacchettini, J. C. & Rüterjans, H. (2000). Solution structure of ileal lipid binding protein in complex with glycocholate. *Eur. J. Biochem.* **267**, 2929-2938.
8. Gerstein, M. & Chothia, C. (1996). Packing at the protein-water interface. *Proc. Natl. Acad. Sci. U. S. A.* **93**, 10167-10172.
9. Likic, V. A. & Prendergast, F. G. (1999). Structure and dynamics of the fatty acid binding cavity in apo rat intestinal fatty acid binding protein. *Protein Sci.* **8**, 1649-1657.
10. Scapin, G., Gordon, J. I. & Sacchettini, J. C. (1992). Refinement of the structure of recombinant rat intestinal fatty acid-binding apoprotein at 1.2-Å resolution. *J. Biol. Chem.* **267**, 4253-4269.
11. Scapin, G., Young, A. C. M., Kromminga, A., Veerkamp, J. H., Gordon, J. I. & Sacchettini, J. C. (1993). High resolution X-ray studies of mammalian intestinal and muscle fatty acid-binding proteins provide an opportunity for defining the chemical nature of fatty acid: protein interactions. *Mol. Cell. Biochem.* **123**, 3-13.
12. Banaszak, L., Winter, N., Xu, Z., Bernlohr, D. A., Cowan, S. & Jones, A. T. (1994). Lipid-binding proteins: A family of fatty acid and retinoid transport proteins. *Adv. Protein Chem.* **45**, 89-151.

13. Thompson, J., Ory, J., Reese-Wagoner, A. & Banaszak, L. (1999). The liver fatty acid binding protein - comparison of cavity properties of intracellular lipid-binding proteins. *Mol. Cell. Biochem.* **192**, 9-16.

14. MacKerell, J., A. D., Bashford, D., Bellott, M., Dunbrack, J., R. L., Evanseck, J. D., Field, M. J., Fischer, S., Gao, J., Guo, H., Ha, S., Joseph-McCarthy, D., Kuchnir, L., Kuczera, K., Lau, F. T. K., Mattos, C., Michnick, S., Ngo, T., Nguyen, D. T., Prodhom, B., Reiher, W. E., Roux, B., Schlenkrich, M., Smith, J. C., Stote, R., Straub, J., Watanabe, M., Wiórkiewicz-Kuczera, J., Yin, D. & Karplus, M. (1998). All-atom empirical potential for molecular modeling and dynamics studies of proteins. *J. Phys. Chem. B* **102**, 3586-3616, doi: 10.1021/jp973084f.

15. Connolly, M. (1983). Analytical molecular surface calculation. *J. Appl. Crystallogr.* **16**, 548-558.

16. Levitt, D. G. & Banaszak, L. (1992). Pocket - A computer-graphics method for identifying and displaying protein cavities and their surrounding amino-acids. *J. Molec. Graph.* **10**, 229-234.

17. Xu, Z., Bernlohr, D. A. & Banaszak, L. (1993). The adipocyte lipid-binding protein at 1.6-Å resolution: Crystal structures of the apoprotein and with bound saturated and unsaturated fatty acids. *J. Biol. Chem.* **268**, 7874-7884.

Figure S1. Small cavities found by VOIDOO, visualized with GRASP as described in Materials and methods. The cavities are numbered according to Table S1, S2 and S4. The panels show cavities in (a) holo (myristate) I-FABP (PDB code 1ICM), (b) holo (oleate) H-FABP (1HMS), (c) holo I-LBP (1EIO, model 1), and (d) holo I-LBP (1EIO, model 5). Additional cavities were found in 1EIO models 2-4 (not shown). Potentially long-lived water molecules are shown in yellow, other cavity waters in blue. The yellow waters of I-FABP are, from top to bottom, W134, W158, W170 and W183. The yellow waters in H-FABP are W166, W222, W167, W191, W175, W174 and W181. The coloured amino acids in I-FABP and H-FABP are Arg106 (magenta) and Arg126 (yellow), and in I-LBP they are Glu110 (magenta) and Arg121 (yellow). Ligands are colored grey with red oxygen atoms.