Simulations of Fatty Acid-Binding Proteins. II. Sites for Discrimination of Monounsaturated Ligands

Thomas B. Woolf and Michael Tychko

Department of Physiology, Johns Hopkins University School of Medicine, Baltimore, Maryland 21205 USA

ABSTRACT Fatty acid binding proteins (FABPs) can discriminate between saturated and unsaturated fatty acids via molecular mechanisms that are not understood. Molecular dynamics computer calculations are used to suggest the relationship between tertiary structure and binding specificity. Three separate 1-ns simulations, with explicit solvent, are presented: 1) oleic acid (C18:1 *cis*) bound to adipocyte FABP, 2) oleic acid bound to human muscle FABP, and 3) elaidic acid (C18:1 *trans*) bound to human muscle FABP. The average structural, dynamic, and energetic properties of the trajectory were analyzed, as were the motional correlations. The molecular dynamics trajectories reveal intriguing differences among all three systems. For example, the two proteins have different strengths of interaction energy with the ligand and different motional coupling, as seen with covariance analysis. This suggests distinctive molecular behavior of monounsaturated fatty acids in the two similar proteins. An importance scale, based on motional correlation and interaction energy between protein and ligand, is proposed, to help identify amino acids involved with the discrimination of ligand saturation state or geometric isomerization.

INTRODUCTION

Fatty acid binding proteins (FABPs) are 12–15-kDa proteins that are abundant in many tissue types and have 20–70% sequence identity (see Banaszak et al., 1994; Veerkamp et al., 1991, for overviews). The members of this family are capable of discriminating among fatty acids on the basis of chain length and saturation state, but there is no detailed molecular understanding of this process. Highresolution structures of several FABPs have been determined by x-ray diffraction methods (see Banaszak et al., 1994; LaLonde et al., 1994, for overviews). In each case, a β -clam structure is formed by two faces composed of five β -strands that enclose an \sim 500-Å³ internal binding cavity. This pocket is lined with hydrophobic as well as hydrophilic residues and contains water, even when the ligand is present. Recent binding measurements performed with acrylodan-derivatized intestinal FABP (ADIFAB) suggest that fatty acid affinities lie in the nanomolar range and vary with tissue origin of the protein (Richieri et al., 1994, 1995, 1996). In particular, a 5–10-fold difference in oleic acid binding affinity was measured, favoring human muscle FABP (M-FABP) over adipocyte FABP (A-FABP) (Richieri et al., 1994).

The specific physiological roles of FABPs have not been established with certainty. Although it is probable that these proteins are involved with the regulation and transport of free fatty acids, there may be other important functional roles (for a review see Glatz and van der Vusse, 1990; Bass, 1993; Veerkamp, 1995). For example, FABPs have been

© 1998 by the Biophysical Society 0006-3495/98/02/694/14 \$2.00

implicated in the regulation of fatty acid metabolism, signal transduction, and adipose cell differentiation (Glatz et al., 1995; Glatz and van der Vusse, 1996; Amri et al., 1995; Teboul et al., 1995).

High-resolution x-ray structures of M-FABP and A-FABP with monounsaturated fatty acids are available at 1.4 Å and 1.6 Å, respectively (Xu et al., 1993; Young et al., 1994). The protein structures are similar, with a backbone root mean square (RMS) difference of 0.7 Å. The conformation of the fatty acid ligand is similar in the area near the headgroup, but diverges toward the end of the alkane chain (Xu et al., 1993; Young et al., 1994). The x-ray-resolved waters near the headgroup are found in a similar pattern in both structures.

Given the structural homology of all three systems, the origins of ligand specificity are not immediately apparent. Interestingly, the bound conformations of monounsaturated ligands in M-FABP are similar to that of stearic acid (C18:0) (Xu et al., 1993; Young et al., 1994). Therefore, the original x-ray paper suggested that the relative binding affinity could be partly rationalized on the basis of conformational energies for the three ligands relative to an extended state (Young et al., 1994). However, this analysis does not account for the enthalpic and entropic effects of the protein environment. Although the current calculations do not directly determine the free energy of binding, they do unveil the details of molecular motion. These dynamic phenomena must be considered to gain a full understanding of the specificity inherent in this family of proteins.

Molecular dynamics calculations have been used to provide insight into the relationship of protein structure and function by exploring atomic motion (e.g., Brooks et al., 1988). The scope of this approach is still limited by modern computer speeds, but simulations are now able to approach nanosecond time scales. The trajectories contain a wealth of information about molecular behavior that is not obvious

Received for publication 26 June 1997 and in final form 3 November 1997. Address reprint requests to Dr. Thomas B. Woolf, Department of Physiology, Johns Hopkins University School of Medicine, 725 N. Wolfe St., Baltimore, MD 21205. Tel.: 410-614-2643; Fax: 410-955-0461; E-mail: twoolf@welchlink.welch.jhu.edu.

from a static structural view. In particular, correlated movements and enthalpic interactions between subsets of atoms can be determined from the calculations.

Conclusions drawn from these simulations can illuminate general questions beyond this particular system. For example, the molecular details of ligand interaction within a protein-binding site are of obvious importance to the field of rational drug design (see Ajay and Murcko, 1995, and Gilson et al., 1997, for recent reviews). A better understanding of hydrophobic ligands within a buried cavity may provide general insights that could help to improve ligand binding design. Furthermore, very little is currently known about the effect of the lipid bilayer environment on membrane proteins. The protein-lipid interactions evident in FABPs are suggestive of an inside-out version of those involved with membrane proteins. Therefore, an improved understanding of this system may help to unravel the general effect of the membrane on protein structure and function.

Recent studies have suggested that oleic and elaidic acids (*cis* and *trans* C18:1) have distinct metabolic fates in humans (e.g., Mensink and Katan, 1993). The *trans* geometry of the single double bond may contribute to eventual arteriosclerosis. The current simulations uncover subtle differences in the behavior of these two fatty acids within FABPs. This knowledge could provide further guidance for understanding the molecular physiology of fatty acids.

MATERIALS AND METHODS

The CHARMm program was used to calculate the detailed motions of holo FABPs in solution. Three separate systems were examined: A-FABP:oleic acid, M-FABP:oleic acid, and M-FABP:elaidic acid. For the initial conformation, the current simulations used high-resolution x-ray crystallographic data collected by two groups (Xu et al., 1993; Young et al., 1994). The coordinates provide both a starting point for calculation and a basis for comparison during analysis. Although the calculations were performed by a method identical to that of a previous paper (Woolf, 1998), the methods are repeated here for completeness.

The PDB coordinates (1hmr, 1hms, 1lid) were oriented with the center of mass at the origin and the largest principal axis along the *x* direction. The structure was then gradually relaxed into the CHARMm potential surface (Schlenkrich et al., 1996). This was achieved by first fixing in place all heavy atoms and then relaxing the positions of added hydrogens. After this step, a series of harmonic restraints was added. Two loops were used to gradually decrease the strength of the restraints. The first loop used a set of 10 decreasing harmonic restraints. The force constant was initially 100.0 kcal/mol- \AA^2 and then decreased in steps of 10 kcal/mol- \AA^2 to 10.0 kcal/ mol- \AA^2 . A second loop used decreasing harmonic restraints from 10.0 to 1.0 kcal/mol-Å² . After the second loop, a set of steepest descent and adopted basis Newton-Raphson minimizations was performed. This resulted in a starting point with small deviation from the x-ray structure that was at a minimum in the CHARMm potential function.

The protonated form of the fatty acid was used for all three simulations. This is consistent with our previous simulations (Woolf, 1998) and is based on both Poisson-Boltzmann DELPHI and CHARMm calculations. The charged form of the fatty acids was found to lead to large deviations from the initial crystal structures. The protonated form generated stable trajectories in the current three simulations and the previous two. Similar calculations for the I-FABP system suggest a charged ligand for that environment. This is consistent with the only available experimental evidence for the pK of the carboxylate group (Cistola et al., 1989). Further analysis is planned of other alternative hydrogen placements in the vicinity of the headgroup.

Solvation of the protein used a box of preequilibrated waters overlaid on the relaxed structure. Any waters with oxygens closer than 2.6 Å to atoms from the original coordinates were deleted. A spherical droplet was established by deleting any waters with oxygens outside of a 27.0-Å radius. The systems consisted of \sim 7500 atoms. A variant of the stochastic method of Brooks and Karplus was used to restrain the droplet of overlaid waters during the simulation (Brooks and Karplus, 1983). The MMFP routine in CHARMm was used so that solvent atoms near the periphery experience small restraining forces that prevent their exit from the simulation shell. The restraining solvent potential is set to zero from the center of the sphere to within a radial distance of 0.7 Å from the well minima. The potential function increases rapidly from -0.25 kcal/mol minima at 27.0 Å to 1.9 kcal/mol at 28.0 Å. The goal is to moderately perturb those waters near the boundary with a minimal effect on the inside of the system. This boundary condition may influence the motions of surface residues and affect the rate of water transfer inside and outside of the protein. Thus the protein surface motions are not considered in detail for the analysis below. Future work will compare simulations that use particle mesh Ewald with constant pressure and temperature with the current results (Tychko and Woolf, manuscript in preparation). However, the current choice of boundary conditions should lead to reasonable dynamic behavior for atoms within the binding cavity of FABP.

Equilibration consisted of several steps. First, the added water was relaxed while all other atoms were fixed. This lowered the gradient to a point where dynamics could be initiated. A series of gradually decreasing harmonic restraints was imposed to prevent large protein motions during this stage. After this relaxation of the overlaid waters, two dynamic equilibration stages were performed. The first 25 ps of dynamics used Langevin dynamics with coupling to a 300 K heat bath (friction coefficient of 25 ps^{-1}). After this run, 25 ps of dynamics was performed with velocity scaling. The approach rescaled velocities every 2.5 ps if the temperature fluctuated more than 5° from 300 K. After this equilibration, 1-ns trajectories were produced. Conformations were saved every 25 steps (50 fs) for analysis.

Analysis was performed with CHARMm and custom scripts. The covariance analysis utilized a combination of CHARMm routines, and a perl script to create a file that INSIGHT (Biosym) could display graphically. In particular, the CHARMm correl subroutine was used to produce a covariance time series for each residue, with either ligand heavy atoms or those water molecules that had a high contact frequency with the ligand.

Potentially interesting sites for mutagenesis were identified from a combination of the positional covariance and individual amino acid-ligand interaction energy. This was performed by normalizing both quantities (negative covariance values were multiplied by -1). Critical residues were suggested by multiplying the two normalized quantities and summing over all fatty acid atoms. This procedure selects amino acids that have both a

FIGURE 1 Conformations of oleic and elaidic acids in A-FABP and M-FABP crystal structures. The conformation of oleic acid in A-FABP is extended relative to the curved structure of the ligand in M-FABP.

FIGURE 2 (*A*) RMS deviations from the trajectory averaged structure. *Top two panels*, M-FABP; *bottom*, A-FABP. The top of each panel represents C-^a deviations, and the bottom deviations are for all heavy atoms. In all cases, the *x* axis represents the residue number. Larger deviations are seen for the turns, compared to smaller deviations for β -sheets. The overall RMS deviations are small, indicative of a well-defined trajectory. (*B*) Mapping of the RMS deviations onto a color representation. The scale runs from blue to dark red $(0-3 \text{ Å})$. This emphasizes the differences between the turns and β -sheets. (*C*) Fatty acid RMS deviations. The *x* axis moves along the heavy atoms of the ligand from the oxygens of the headgroup to the terminal methyl group. (*D*) Color mapping of the RMS deviations seen in *C*. The scale runs from blue to red (0.6–3.0 Å).

consistently strong interaction energy and a strong correlated motion with the ligand.

RESULTS

Three 1-ns simulations of FABPs were performed using CHARMm: M-FABP:elaidic acid, M-FABP:oleic acid, and A-FABP:oleic acid. The calculations were initiated from separate high-resolution x-ray coordinates that are homologous. Although the ligands adopt similar structures near the headgroup, the conformation of the fatty acid alkane chain differs. Oleic and elaidic acids are both curved in M-FABP, whereas oleic acid is extended in the A-FABP complex. These initial conformations for the fatty acids are shown in Fig. 1. The starting points for the current simulations were very similar to previous calculations on stearic acid bound to M-FABP and A-FABP (Woolf, 1998).

Analysis of the trajectories focused on several issues: average structural properties, dynamic properties, interaction energies, water behavior, and covariance analysis. This set of analyses can enhance molecular insight into the systems and help identify the residues most involved in specific binding.

Structural properties

The time-averaged structural properties of simulated systems can be used to help ascertain the validity of the model, as well as to provide insight into motionally averaged measurements on the nanosecond time scale. For example, large RMS deviations from the original structure would imply that the simulation was not well defined and stable, whereas variations in the pattern of deviations give insight into the relative mobilities of different parts of the system. The current trajectory showed small RMS deviations throughout the simulation, as can be seen in Fig. 2. There was no evidence of large changes within subsets of the structure or over the whole structure. Fig. 2, *A* and *B*, shows the protein RMS deviations from the trajectory-averaged structures. Similar deviations from the x-ray structures are seen. The finding of larger fluctuations within turns is consistent with other molecular dynamics calculations (e.g., Brooks et al., 1988). For example, the turn between β -strands B and C has relatively large RMS deviations, as seen in Fig. 2, *A* and *B*.

Average, isotropic motions of the bound ligand are provided by crystallographic B-factors. A comparison of the RMS deviations for the ligand heavy atoms and the Bfactors showed excellent agreement. Fig. 2, *C* and *D*, shows the RMS deviations of the ligands. In particular, the largest RMS deviations occurred at the terminal methyl group of the A-FABP:oleic acid simulation, consistent with the xray-determined B-factors (Xu et al., 1993). Thus this figure suggests a characteristic difference in the average motions of oleic acid in A-FABP and M-FABP.

Backbone dihedral angles define the overall fold of the protein. Therefore, large changes in these angles would not be consistent with the RMS deviations. However, small fluctuations in the dihedral values may have a functional

role. A sense of the regions of secondary structure with the most flexibility is important for determining the types of motion present in the system. Fig. 3 shows the average values of backbone dihedrals as well as RMS deviations from those averages. The most flexible regions are immediately apparent from this analysis. Thus, consistent with the RMS deviations of Fig. 2, the dihedral fluctuations are larger in the turns.

Dynamic properties

The ensemble of conformations generated during a molecular dynamics trajectory has a temporal sense. This allows the time-dependent motion of sets of atoms to be considered in detail. In particular, the motion of the fatty acid ligands can be closely examined. It may have been argued that the tight binding of protein and ligand would result in a largely immobile fatty acid. Contrary to this thought, Fig. 4 shows that there are time-dependent changes in the alkane chain dihedral angles. That is, the fatty acid chain was not locked into one conformation during the trajectory, but explored a range of possible states. In all cases, the transitions conserved the overall structure of the ligand. It was interesting to further consider the pairs of transitions that occurred. However, to test for concerted pairs of transitions, better statistics than are available from our 1-ns simulation of a single alkane chain are required (e.g., Brown et al., 1995).

FIGURE 3 Average and RMS deviation for backbone dihedral angles. Gray bars represent the trajectory-averaged values; the black tips represent the RMS deviations from the average. The fluctuations are largest for the turn regions. The psi values are projected down and the phi values are projected up to enable both sets of data to be presented in a single figure.

FIGURE 4 Time series of ligand dihedral angles. The bottom trace is from the headgroup, and the top trace is for the end of the alkane chain. Each trace was shifted by 360° to make it possible to present all of the time series in a single figure. A 720° scale bar is provided for the magnitude of changes in each trace. Tables 1–3 should be consulted for the numbers of transitions and observed pairs of dihedral changes.

Nonetheless, some data on the transitions seen in the present case are collected in Tables 1–3. In each table the total number of transitions and numbers of observed paired transitions are listed. The pairs are scored as occurring within 3.0 ps of each other. This is the optimal window size seen in the analysis of mean-field DPPC, bilayer DPPC, and hexadecane (Brown et al., 1995). A full estimate of a concerted, energetic transition pair requires an ability to separate the random number of observed pairs from the concerted number (Brown et al., 1995). Despite the statistical inablity to meaningfully perform this test for the current data set, it is intriguing that $i:i + 2$ transitions seem to occur at a greater rate than other pair types. A preference for $i:i + 2$ transitions has been observed in pure bilayer simulations (Venable et al., 1993; Brown et al., 1995).

Dihedral number 3 is defined by the C1-C2-C3-C4 atoms. Dihedral number 17 is defined by the C15-C16-C17-C18 atoms. Note that the number of paired transitions does not sum to the total number of transitions because more than a pair of transitions may occur within a single window. The window size was 3.0 ps (from Brown et al., 1995, Table II).

Interaction energies

The interaction energy between the ligand and the rest of the system can be used to qualitatively estimate the relative free energy of binding via a linear response approximation (e.g., Aqvist et al., 1994; Aqvist and Hansson, 1996). Both the electrostatic and vdW energies are used for this estimate. This implies that an analysis of the differences in interaction energy between various components of the system can suggest residues that might be important for the binding function.

Dihedral number 3 is defined by the C1-C2-C3-C4 atoms. Dihedral number 17 is defined by the C15-C16-C17-C18 atoms. Note that the number of paired transitions does not sum to the total number of transitions because more than a pair of transitions may occur within a single window. The window size was 3.0 ps (from Brown et al., 1995, Table II).

Dihedral number 3 is defined by the C1-C2-C3-C4 atoms. Dihedral number 17 is defined by the C15-C16-C17-C18 atoms. Note that the number of paired transitions does not sum to the total number of transitions because more than a pair of transitions may occur within a single window. The window size was 3.0 ps (from Brown et al., 1995, Table II).

The interactions of the ligand with the rest of the system were computed for each saved conformation of the trajectory. The resulting set of numbers was binned and normalized to produce a probability distribution separated into headgroup and acyl chain components. This is shown in Fig. 5 *A*. Differences were seen among all three simulations. First, the headgroup interaction of the A-FABP system was relatively weaker and more broadly distributed than either M-FABP system. Second, the tail interaction of the A-FABP:oleic acid system was weaker than that of the M-FABP systems. If the reference system (water/bilayer/vesicle/micelle) provides a similar interaction energy, then the differences in oleic acid interactions between M-FABP and A-FABP simulations may be directly related to the relative difference in binding free energy. It is also intriguing to notice the slight but clear differences between the interactions of oleic and elaidic acids in M-FABP. The oleic acid distributions are narrower and shifted slightly relative to the elaidic acid results.

The alkane chain component of the ligand interaction energy can be analyzed at the level of individual methylene groups, as shown in Fig. 5 *B*. In particular, the differences in interaction energy along the chain are related to the local vdW and electrostatic interactions provided by the environment of the binding cavity. Again, there were differences between all three simulations. The A-FABP results show a tighter distribution, with the majority of interaction energies clustered about -2.5 kcal/mol with a width of \sim 1.0 kcal/ mol. Two distributions are stronger and wider: those of the C2 and C18 methylene groups. Furthermore, the C3 methylene group is weaker and less symmetrical than the average distribution. The M-FABP simulations show interaction energies that are more broadly distributed over the length of the alkane chain. In particular, two interactions are markedly stronger than other interactions. The C18 and C2 methylene units have interaction energies between -4 and -5 kcal/mol, with a broader distribution for the C18 group than for the C2 group. The range of interaction energies for the other units varies from -2 to -3.5 kcal/mol. It is especially interesting that the C3 methylene group has a much weaker interaction in the elaidic acid simulation than in the oleic acid simulation.

Further insight can be gained by analyzing the interaction energy of individual amino acids with the ligand. This is seen in Fig. 6. The calculation is presented in terms of the total and the vdW and electrostatic components. In principle, this provides the connection for comparing differences in relative free energy of binding between the systems, by identifying those residues with the strongest enthalpic contributions. Residues with strong interaction energies are presumed to have an appreciable effect on binding. The figure shows differences and similarities between the three systems. For example, R126 had the largest interaction energy for the M-FABP systems. In contrast, R106 had the largest interaction for the A-FABP simulation.

Water behavior

The water molecules inside FABPs are intrinsically involved with the binding process. For example, a hydrogenbonded water is observed in the crystal structure bridging R106 to the fatty acid headgroup. The motional behavior of this water has not been measured directly. Thus a detailed description of the motion of water and its interaction with the ligand could help to clarify its functional role.

The analysis focused on a subset of waters that had the most frequent contacts with the fatty acid. Waters that passed within 4.0 \AA of the ligand were tabulated for every fourth conformation in the trajectory. Interestingly, the top 40 waters on this list exhibited a range of effective diffusion constants and effective lifetimes for regions near the ligand. Some stayed near the ligand throughout the simulation, whereas others ventured out of the binding pocket. An estimate of the diffusion constant for each water was determined from the slope of the mean square displacement correlation function. The estimates suggest that the effective diffusion constants varied by two orders of magnitude. The waters with the most motion were roughly half as mobile as waters in bulk molecular dynamics simulations (Brooks et al., 1988). More restricted water was roughly 100 times less mobile than water observed entering and leaving the interior.

Tables 4, 5, and 6 describe the 40 waters that had the largest set of contacts with the ligand averaged over each trajectory. The table also includes a listing of the most frequent neighbors for those waters. For example, Table 6 (for A-FABP:oleic acid) lists the water bridging R106 and the headgroup first. Its most frequent neighbors were the headgroup atoms C1, O1, O2, C2, and C3. The second water has a set of contacts with the headgroup (O1, O2, C1, C2) and another set of contacts with the chain end (C10, C12,

FIGURE 5 (*A*) Interaction energy probability distribution functions for the headgroup and the tail regions of the fatty acid. The tail region was roughly twice as strong in energetic interaction as the headgroup region. (*B*) Interaction energy probability distribution function for the methylene groups along the fatty acid acyl chain. Each system has a characteristic set of distributions. The set of distributions is wider for M-FABP than for A-FABP.

C18). The most motionally restricted water was not the water most frequently in contact with the fatty acid. This most slowly diffusing water had 125 contacts within 4.0 Å and had an estimated diffusion constant that was five times less than the water with the most contacts.

Covariance analysis

It has been suggested that an essential configurational subspace of a protein can be defined by diagonalization of the pairwise matrix of atomic positional fluctuations (Garcia, 1992; Amadei et al., 1993). This subspace contains a small fraction of the overall degrees of freedom, which may, in turn, capture the most functionally relevant motions. The use of covariance analysis has also been fruitful for determining the dynamic connections between regions of structure (e.g., Ichiye and Karplus, 1991). The current calculations employ covariance analysis to reveal patterns of motional coupling between ligand and protein or ligand and water in FABPs. Strong positive or negative zero-time positional covariances imply a nearly direct coupling between subsets of atoms.

The covariance analysis was performed between the ligand and either protein heavy atoms or the 40 nearby waters determined above. This is shown in Fig. 7. Although the current findings show clear differences among the three systems, the contrasts were not as striking as those of stearic acid in M-FABP versus A-FABP (Woolf, 1998). The previous simulations showed strong alternating patterns of correlation along the stearic acid for various regions of M-FABP, similar to the current M-FABP-elaidic acid data. The current simulations also show some coupling for the M-FABP-elaidic acid trajectory that was not present in M-FABP-oleic acid. However, the pattern of protein-ligand covariance was clearly stronger in the previous M-FABPstearic acid simulation. Furthermore, the current A-FABPoleic acid data are quite different from the previous A-FABP-stearic acid calculations. In contrast, it is interesting to note that certain regions of the protein correlate similarly in all three current simulations, such as the patch of positive correlation between the second α -helix and the middle of the ligand tail.

The water-ligand covariances of the current trajectories were much milder than in the previous simulations of stearic

FIGURE 6 Interaction energy between each amino acid and the ligand. The total interaction energy is also separated into electrostatic and vdW components. Gray bars describe the average interaction energy, and the black tips illustrate the additional increase in magnitude that is possible with RMS fluctuations. Note that the values for R106 and R126 are off scale. The electrostatic and vdW breakdowns give an indication of their relative contributions.

acid in M-FABP and A-FABP (Woolf, 1998). However, unlike the stearic acid simulations, the current data show a strong difference between A-FABP and M-FABP. The water in the A-FABP simulation showed both strong correlated and anticorrelated motions compared to the scattered sets of weak correlations for both M-FABP systems. For example, the most frequently contacted water (*left side of the x axis*) has a strong anticorrelated set of motions with respect to the ligand headgroup. The next three waters have a positive correlation with the chain of the fatty acid. Waters 10, 12, 15, and 16 have strong anticorrelations with the ligand terminus. Although the M-FABP data suggest weaker coupling to water, it is still interesting to note differences like the band of stronger correlations for C11–C13 of M-FABPelaidic acid versus M-FABP-oleic acid.

DISCUSSION

Protein-lipid interactions play a crucial role in a variety of biological systems, but there is currently little understanding of the nuances of these important effects. Differences in the molecular behavior of saturated versus unsaturated alkane chains are expected to give rise to distinct protein-lipid interactions (e.g., Gennis, 1989). The majority of computer simulations have addressed fully saturated fatty acids such as DMPC or DPPC (e.g., Merz and Roux, 1996, for overview). However, this study examined a series of saturated and unsaturated fatty acid ligands bound to several FABPs. Thus the current simulations provide an ideal setting for probing the differences between saturated and monounsaturated fatty acids interacting with binding proteins. It is interesting to note that the high-resolution crystal structures of these systems did not immediately reveal any mechanism

Estimated

TABLE 6 Nearest water neighbors to the oleate in A-FABP

Contacts (no. /tri)	esumated diffusion constant $(\mu m^2/ms)$	Most frequent near neighbors
9761	0.02	C1 O2 O1 C2 C3 C4 C5
8912	0.02	01 C1 C2 02 C3 C13 C12 C18
6459	0.08	02 01 C11 C13 C3 C15 C11
3909	0.1	C13 C5 C6 O1 C11 C3 C8
3818	0.02	C7 C6 C8 C9 C5
3437	0.03	C14 C16 C13 C15 C18 C17 C12
3416	0.07	C5 C6 C8 C7 C4 C3 O2 C9
3117	0.03	C8 C7 C6 C5 C11 C9 C10 C4
3053	0.18	C16 C5 C3 O1 C13 O2 C15
1926	0.02	01 02 C2 C1 C4 C3 C6 C5
1522	0.07	02 C1 01 C2 C3 C11 C13 C6
1104	0.33	02 C5 C3 C1 O1 C6 C2 C4
664	0.02	C15 C18 C13 C12 C16 C17 C14
544	0.9	C ₁₆ C ₁₄ C ₁₅ C ₁₈ C ₁₇ C ₁₃
479	0.8	C11 O2 C13 O1 C15 C2 C14
440	0.15	02 C3 C1 01 C5 C4 C6 C8
238	0.007	C5 C6 C4 C3 O2 C2
218	0.7	C14 C16 C15 C17 C18 C13 C12
164	0.7	C18 C16 C17 C14 C15
161	0.82	C18 C17 C16
159	0.58	C18 C16 C17 C14 C15
127	0.88	C14 C16 C15 C13
125	0.003	C ₂ O ₁ C ₃ O ₂ C ₅ C ₄
119	0.85	C ₁₄ C ₁₆ C ₁₅ C ₁₇ C ₁₈
115	0.78	C18 C16 C17 C14
109	0.58	C ₂ C ₁ C ₃ O ₁ C ₄ O ₂ C ₆
107	0.6	C18 C17 C16
81	0.75	C18 C17 C16 C14 C15
80	0.85	C14 C15 C16 C17 C13
78	0.88	C18 C17 C16 C14
69	1.0	C18 C17 C16 C14
65	0.58	C ₁₈ C ₁₇ C ₁₆
62	0.77	C ₁₅ C ₁₈ C ₁₆ C ₁₇ C ₁₃
60	0.63	C18 C17 C16 C14
59	0.67	C15 C14 C16 C17
58	0.77	C18 C16 C17
51	0.57	C18
51	0.67	C18 C17 C16
49	0.76	C18 C16
43	0.78	C18 C17 C16 C14

for discrimination among ligands with varying length or saturation state. The molecular dynamics method may reveal particular aspects of these systems that are related to specificity.

In fact, intriguing differences were found among the molecular dynamics simulations of C18 fatty acids bound to M-FABP and A-FABP. The results suggest that differences in the motion and interaction energy of the systems could play a significant role in the binding function of FABPs. For example, water was more strongly coupled with the ligand motion in A-FABP than M-FABP. The total protein-ligand interaction energy was strong in all three simulations, which is consistent with tight binding. However, the ligand was not locked into a single conformation by the shape of the binding site. Instead, dihedral transitions were observed along the length of the alkane chain. Furthermore, the interaction energy of the methylene groups along the chain varied from one system to another. This implies that the shape of the binding pocket is optimized for a particular ligand, or set of ligands, within a specific FABP. In particular, some methylene groups had strong interaction energies of -4 to -5 kcal/mol, whereas others had weaker interaction energies of -1 to -2 kcal/mol. A further breakdown of interaction energies in terms of individual protein residues suggested that the two proteins are coupled differently to their ligands, despite very similar tertiary structures. For example, R126 had the strongest interaction in M-FABP, whereas R106 was stronger for A-FABP.

Mutagenesis suggestions

A combination of the covariance and interaction energy analyses was used to establish an importance scale that suggested key residues involved in overall binding and discrimination of the fatty acid saturation state. A joint weighted comparison was made by normalizing both sets of information. The interaction energy was normalized on a scale from 0 to 1, with the largest value being the greatest interaction energy. The covariance analysis was normalized on the same type of 0 to 1 scale, with the largest values being either a large correlated or anticorrelated motion. Thus the predictions are based purely on the basis of amino acid residues that are both strongly interacting and strongly coupled in their motion with the ligand. This type of analysis could have several sources of error. For instance, the analysis could be flawed if the motions and energetic interactions are not representative of that found in the real system. This scenario is possible if the computer calculations improperly sample the relevant set of conformations accessible to the protein in solution. The predictions could also be suspect if strong enthalpic interactions and covariances are not true markers of functional significance. Despite these caveats, the approach may provide a useful estimate of those residues most important for binding function. The predictions are experimentally testable, and the validity of the approach can be checked.

Pairs of the five molecular dynamics runs (three in the current paper and two from the previous paper) were compared. Fig. 8 shows sums and 10 pairwise differences of combinations. The information is best summarized by three categories: 1) sites that may be important for binding in general; 2) sites that may discriminate monounsaturated fatty acids from saturated fatty acids; and 3) sites that could be important for *cis* versus *trans* discrimination.

Sites for general binding

The most important residues for binding in general can be estimated from looking for amino acids with high scores in all five simulations. Separate sums of three M-FABP systems, two A-FABP systems, and all five simulations are presented in Fig. 8 *A*. Six conserved residues were identified

FIGURE 7 Covariance analysis. (*A*) Ligand and protein motions. The zero-time covariance function was computed as an average over the full trajectory. The color scale bar represents the range of averaged motional correlation between protein atoms and fatty acid atoms. Blue regions are anticorrelated, and red regions are positively correlated. The 20 heavy atoms of the fatty acid ligand are represented along the vertical axis (*headgroup at the bottom*), and the horizontal axis represents the protein sequence (*N-teminus left*). The regions of secondary structure are indicated along the bottom. (*B*) Ligand and 40 nearest water molecules. The 20 atoms of the fatty acid ligand are represented vertically (*headgroup at the bottom*), and the water is ordered from left to right (most to least contacts). Tables 4–6 should be consulted for the most frequent neighbors of the water and a calculated estimate of the diffusion constant. A-FABP has a stronger set of covariance connections with the water than M-FABP.

by this test—three near the headgroup and three near the midpoint of the alkane chain.

The importance of the first three was predicted from the x-ray structure (e.g., Jakoby et al., 1993; Xu et al., 1993; Young et al., 1994). In addition, the simulation results suggested that R106 is more important for A-FABP and R126 is more important for M-FABP, whereas Y128 is somewhat important in both FABPs. The other three residues suggested by this analysis were not obvious from consideration of the crystal structure. The three are A75, D76, and F16. The A75 and D76 residues are in the mobile E-F turn. The location is near the bend in the M-FABP ligands and the midpoint of the A-FABP ligands. There is no x-ray-resolved water near the site. The location is also near the presumed portal region, and thus mutations at this site may effect the kinetics of binding. The F16 site is in the beginning of the first α -helix. Its side chain occupies part of the space in the bend of the ligand in M-FABP and along the midpoint of the ligand in A-FABP. Mutations that change the size of this side chain may have a large impact on the binding site, by changing the local packing arrangements that are possible for the ligand. Finally, a seventh potentially

FIGURE 8 Plots based on the proposed method for identifying functionally important amino acids for mutagenesis. (*A*) Separate sums of three M-FABP systems, two A-FABP systems, and all five simulations. (*B*) The ten pairs of differences. This provides suggestions for amino acids that might be involved with general ligand binding and selectivity.

interesting site is P38, which seems more important in the M-FABP simulations, even though it is conserved in A-FABP.

Sites for saturation state discrimination

Determining the sources of discrimination between saturated and monounsaturated ligands is difficult. As has been pointed out (Richieri et al., 1994; Cistola et al., 1988), the solubility of the fatty acids in water and their phase diagrams depend on saturation state. Therefore, part of the measured difference in binding affinity between stearic and oleic acids is possibly due to differences in their respective physiological reference states. Without a clear understanding of what the reference state actually is, it is difficult to estimate the contribution of the reference to differences in binding affinity from a single simulation.

Given these difficulties, a qualitative prediction might still be possible, because the change in binding affinity for oleic versus stearic acid is opposite in the two proteins. In M-FABP:oleic acid, the affinity decreases by 2.5 (Richieri et al., 1994). In contrast, the affinity for A-FABP:oleic acid increases by 1.4 (Richieri et al., 1994). Because the changes are in opposite directions, it was possible to look for changes in the scoring for amino acids that are unique to each of the two systems. The full set of differences is shown in Fig. 8 *B*. Those residues with large differences in score (and therefore, presumably, which are important in discrimination) were counted for each pair. Then those residues that were present in both sets were discounted (under the assumption that residues are not acting positively in one system and negatively in the other; this assumption assumes a type of independence of each residue that may not be true). The resulting set of amino acids was then postulated

to be contributing to stearic acid recognition over oleic acid in M-FABP and to oleic acid over stearic acid in A-FABP.

The M-FABP residues selected as targets by this approach were M20, R78, Q95, and L117. The A-FABP residues selected are all present in α -helix II: V25, T29, and V32. These three helical residues may thus be involved with the kinetics of fatty acid transfer in addition to defining the binding site. The M-FABP residues are located at different sites around the ligand: M20 in β -I, R78 in the E-F turn, Q95 in β -G, and L117 in β -I. Interestingly, L117 is the only nonconserved residue in this set of seven C117 in A-FABP.

Sites for *cis/trans* **selection**

By considering a larger set of pairwise combinations, two amino acids that could be important for *cis*/*trans* discrimination are suggested. The M-FABP:oleic acid and M-FABP:elaidic acid systems were each compared with M-FABP:stearic acid. Furthermore, the oleic and elaidic acid M-FABP runs were also compared against A-FABP: stearic acid as a second screen. Last, the elaidic and oleic acid M-FABP runs were compared directly against themselves. Two residues were identified based on these comparisons, that is, M20 and K58. The M20 contributes vdW connections to the midpoint of the fatty acid alkane chain. It is closer to the *cis-trans* region than to the F16 side chain. It was also selected as being important for the discrimination of saturation state in the M-FABP. The K58 side chain is in the C-D turn region and may have a larger degree of mobility than M20. In particular, correlations with the ligand C3 region are strong for M20. This was consistent with the finding of the C3 methylene group having a greater *cis*/*trans* difference in interaction energy than other units along the chain.

Finally, the current simulations are also important in light of the recent evidence suggesting that *trans* monounsaturated fatty acids may be metabolized differently from the *cis* isomers (e.g., Mensink and Katan, 1993). The increasing prevalence of *trans* monounsaturated fatty acids in the human diet calls for an increased understanding of the details involved in discrimination between geometric isomers of a given fatty acid. In particular, a better molecular understanding of the changes in behavior of an endogenous protein interacting with a substrate less commonly found during evolution could suggest mechanisms for the differential metabolism. Furthermore, it may eventually be possible to design a modified FABP that has high affinity for *trans* monounsaturated fatty acids.

SUMMARY

Three 1-ns molecular dynamics simulations are presented that examine the details of monounsaturated fatty acids in complex with M-FABP and A-FABP. The results suggest differences in correlated motion and interaction energy between oleic acid in M-FABP and A-FABP. The simulations also reveal differences between elaidic and oleic acids in

M-FABP. These calculations are related to a long-term goal of connecting the FABP protein tertiary structure to ligandbinding affinities and the eventual rational design of new FABPs capable of binding specific fatty acids. In particular, the results suggest sites for mutagenesis of residues that may be important in the discrimination of saturation state.

Alan Grossfield is thanked for comments on the manuscript and help with the figure production and presentation.

The Biomedical Engineering Department is thanked for use of the Whitaker Foundation Computer Center.

The American Heart Association is gratefully acknowledged for a grantin-aid that supported this work. The Bard Foundation is acknowledged for additional funding at important stages of this work.

REFERENCES

- Ajay and M. A. Murcko. 1995. Computational methods to predict binding free energy in ligand-receptor complexes. *J. Med. Chem.* 38:4953–4967.
- Amadei, A., B. M. Linssen, and H. J. C. Berendsen. 1993. Essential dynamics of proteins. *Proteins.* 17:412–425.
- Amri, E. Z., F. Bonino, G. Ailhaud, N. A. Abumrad, and P. A. Grimaldi. 1995. Cloning of a protein that mediates transcriptional effects of fatty acids in preadipocytes. *J. Biol. Chem.* 270:2367–2371.
- Aqvist, J., and T. Hansson. 1996. On the validity of electrostatic linear response in polar solvents. *J. Phys. Chem.* 100:9512–9521.
- Aqvist, J., C. Medina, and J.-E. Samuelsson. 1994. A new method for predicting binding affinity in computer-aided drug design. *Protein Eng.* 7:385–391.
- Banaszak, L., N. Winter, Z. Xu, D. A. Bernlohr, S. Cowan, and T. A. Jones. 1994. Lipid-binding proteins: a family of fatty acid and retinoid transport proteins. *Adv. Protein Chem.* 45:89–151.
- Bass, N. M. 1993. Cellular binding proteins for fatty acids and retinoids: similar or specialized functions? *Mol. Cell. Biochem.* 123:191–202.
- Brooks, C. L., III, and M. Karplus. 1983. Deformable stochastic boundaries in molecular dynamics. *J. Chem. Phys.* 79:6312–6325.
- Brooks, C. L. I., M. Karplus, and B. M. Pettitt. 1988. Proteins: A Theoretical Perspecitive of Dynamics, Structure, and Thermodynamics. John Wiley and Sons, New York.
- Brown, M. L., R. M. Venable, and R. W. Pastor. 1995. A method for characterizing transition concertedness from polymer dynamics computer simulations. *Biopolymers.* 35:31–46.
- Cistola, D. P., J. A. Hamilton, D. Jackson, and D. M. Small. 1988. Ionization and phase behavior of fatty acids in water: application of the Biggs phase rule. *Biochemistry.* 27:1881–1888.
- Cistola, D. P., J. C. Sacchettini, L. J. Banaszak, M. T. Walsh, and J. I. Gordon. 1989. Fatty acid interactions with rat intestinal and liver fatty acid-binding proteins expressed in *Escherichia coli. J. Biol. Chem.* 264:2700–2710.
- Garcia, A. E. 1992. Large-amplitude nonlinear motions in proteins. *Phys. Rev. Lett.* 68:2696–2699.
- Gennis, R. B. 1989. Biomembranes: Molecular Structure and Function. Springer Verlag, New York.
- Gilson, M. K., J. A. Given, B. L. Bush, and J. A. McCammon. 1997. The statistical-thermodynamic basis for computation of binding affinities: a critical review. *Biophys. J.* 72:1047–1069.
- Glatz, J. F., T. Borchers, F. Spener, and G. J. van der Vusse. 1995. Fatty acids in cell signalling: modulation by lipid binding proteins. *Prostaglandins Leukot. Essent. Fatty Acids.* 52:121–7.
- Glatz, J. F. C., and G. J. van der Vusse. 1990. Cellular fatty acid-binding proteins: current concepts and future directions. *Mol. Cell. Biochem.* 98:237–251.
- Glatz, J. F., and G. J. van der Vusse. 1996. Cellular fatty acid binding proteins: their function and physiological significance. *Prog. Lipid Res.* 35:243–282.
- Ichiye, T., and M. Karplus. 1991. Collective motions in proteins: a covariance analysis of atomic fluctuations in molecular dynamics and normal mode simulations. *Proteins Struct. Funct. Genet.* 11:205–217.
- Jakoby, M. G., K. R. Miller, J. J. Toner, A. Bauman, L. Cheng, E. Li, and D. P. Cistola. 1993. Ligand-protein electrostatic interactions govern the specificity of retinol- and fatty acid-binding proteins. *Biochemistry.* 32:872–878.
- LaLonde, J. M., D. A. Bernlohr, and L. J. Banaszak. 1994. The up-anddown beta-barrel proteins. *FASEB J.* 8:1240–1247.
- Mensink, R. P., and M. B. Katan. 1993. Trans monounsaturated fatty acids in nutrition and their impact on serum lipoprotein levels in man. *Prog. Lipid Res.* 32:111–122.
- Merz, K. M., Jr., and B. Roux. 1996. Biological membranes: a molecular perspective from computation and experiment. Birkhaüser Press, Boston, MA.
- Richieri, G. V., R. T. Ogata, and A. M. Kleinfeld. 1994. Equilibrium constants for the binding of fatty acids with fatty acid-binding proteins from adipocyte, intenstine, heart, and liver measured with the fluorescent probe ADIFAB. *J. Biol. Chem.* 269:23918–23930.
- Richieri, G. V., R. T. Ogata, and A. M. Kleinfeld. 1995. Thermodynamics of fatty acid binding to fatty acid-binding proteins and fatty acid partition between water and membranes measured using the fluorescent probe ADIFAB. *J. Biol. Chem.* 270:15076–15084.
- Richieri, G. V., R. T. Ogata, and A. M. Kleinfeld. 1996. Kinetics of fatty acid interactions with fatty acid binding proteins from adipocyte, heart, and intestine. *J. Biol. Chem.* 271:11291–11300.
- Schlenkrich, M., J. Brickmann, A. D. J. MacKerell, and M. Karplus. 1996. Empirical potential energy function for phospholipids: criteria for parameter optimization and applications. *In* Biological Membranes: A Molecular Perspective from Computation and Experiment. K. M. J. Merz and B. Roux, editors. Boston, Birkhäuser. 31-81.
- Teboul, L., D. Gaillard, L. Staccini, H. Inadera, E. Z. Amri, and P. A. Grimaldi. 1995. Thiazolidinediones and fatty acids convert myogenic cells into adipose-like cells. *J. Biol. Chem.* 270:28183–28187.
- Veerkamp, J. H. 1995. Fatty acid transport and fatty acid-binding proteins. *Proc. Nutr. Soc.* 54:23–37.
- Veerkamp, J. H., R. A. Peeters, and R. G. H. J. Maatman. 1991. Structural and functional features of different types of cytoplasmic fatty acidbinding proteins. *Biochim. Biophys. Acta.* 108:1–24.
- Venable, R. M., Y. Zhang, B. J. Hardy, and R. W. Pastor. 1993. Molecular dynamics simulations of a lipid bilayer and of hexadecane: an investigation of membrane fluidity. *Science.* 262:223–226.
- Woolf, T. B. 1998. Simulations of fatty acid-binding proteins suggest sites important for function. I. stearic acid. *Biophys. J.* 74:681–693.
- Xu, Z., Bernlohr, and L. J. Banaszak. 1993. The adipocyte lipid-binding protein at 1.6 Å resolution. *J. Biol. Chem.* 268:7874–7884.
- Young, A. C., G. Scapin, A. Kromminga, S. B. Patel, J. H. Veerkamp, and J. C. Sacchettini. 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.