Minireview

Rat Intestinal Fatty Acid Binding Protein

A MODEL SYSTEM FOR ANALYZING THE FORCES THAT CAN BIND FATTY ACIDS TO PROTEINS*

James C. Sacchettini†‡ and Jeffrey I. Gordon¶

From the Department of Biochemistry, Albert Einstein College of Medicine, Bronx, New York 10461 and the Department of Molecular Biology and Pharmacology, Washington University School of Medicine, St. Louis, Missouri 63110

Rat intestinal fatty acid binding protein (I-FABP)1 is produced in the columnar absorptive epithelial cells (enterocytes) of the intestine (1). Its precise function in the absorption of long chain fatty acids and their subsequent trafficking to intracellular sites of metabolic processing remain unclear. Nonetheless, there are several features of this fatty acid "receptor" that contribute to its attractiveness as a model for studying the manner in which fatty acids interact with proteins. It is a small monomeric protein composed of 131 residues that binds a single molecule of long chain fatty acid in a noncovalent fashion (2, 5). It has no known co- or post-translational modifications and can be efficiently produced in bacteria with full preservation of function (7). Escherichia coli-derived rat I-FABP can be readily crystallized with and without bound fatty acid. The structure of the apoprotein has been refined to 1.2-Å resolution (6). The structures of I-FABP with bound tetracanoate (myristate, C14:0), hexadecanoate (palmitate, C16:0), and 9Z-octadecanoate (oleate, C18:1) have been refined to 1.5, 1.9, and 1.7 Å, respectively, so that an assessment can be made of how the protein accommodates changes in the chain length/conformation of its bound ligand (7, 8). Moreover, I-FABP is a member of a family of lipid binding proteins. This family includes other members that are expressed in different cell lineages as well as proteins that have evolved to bind different types of ligands, e.g. retinoids (cellular retinol binding protein [CRBP], CRBPII, cellular retinoic acid binding protein [CRABP]) or bile acids (bile acid binding protein). The structures of several of these other family members have been determined: mouse adipocyte lipid binding protein with bound octadecanoate and 9Z-octadecanoate to 1.6 Å (9, 10), bovine peripheral nerve myelin P2 protein with endogenous fatty acids and with 9Z-octadecanoate to 2.4 and 2.7 Å (11, 12), human muscle FABP with a population fatty acid to 2.1 Å (13), bovine heart FABP with endogenous fatty acid ligands to 3.2 Å (14), chicken liver FABP to 3.0 Å (15), plus CRBP and CRBPII with bound all-trans-retinol to 2.1 and 1.9 Å, respectively (12, 16). All these proteins have similar conformations even though multiple sequence alignments reveal that they have as little as 20% amino acid sequence identity (11, 14). Sequence similarities and differences between these proteins have provided insights about the structural basis for their overlapping and distinct binding specificities. In this minireview, we focus on how x-ray crystallography combined with mutagenesis has been used to analyze the forces that bind fatty acids to I-FABP. This "mutagenesis" has involved both the protein and its ligand.

* This minireview will be reprinted in the Minireview Compendium, which will be available in December, 1993. Work for the authors' laboratories cited in this review was supported by National Institutes of Health Grants GM45866, DK30292, AI27179, and AI31888.

† To whom correspondence should be addressed.

‡ The abbreviations used are: I-FABP, intestinal fatty acid binding protein; CRBP, cellular retinol binding protein II; r.m.s., root mean square.


The Structure of Apo-I-FABP—Apo-I-FABP has two, 7-residue, right-handed α-helices (α1 and αII) located near its N terminus. Ninety-two of the protein's 131 residues form 10 anti-parallel β-strands (βA–βL). These β-strands are organized into two nearly orthogonal β-sheets, giving the protein the overall appearance of a clam shell (Fig. 1A). The first β-sheet contains βA, βB, βC, βD, βE, plus the first five amino acids of βF. The second sheet contains the remainder of βF through βI. The main chain atoms of βE and βF are not directly connected to one another by hydrogen (H–) bonds. Rather, a 5–11-Å "gap" separates the two strands (Fig. 1A). This gap is filled by the side chains of the strand's component amino acids (Fig. 1, A and B). βF functions as a structural linker between I-FABP's two β-sheets; the main chain atoms of the N-terminal portion of this strand are connected to the N-terminal atoms of βE while the main chain atoms of the C-terminal portion of the strand are H-bonded to several residues of βG. One side of the "last" β strand in the protein (βJ) shares H-bonds with main chain atoms of β while the other side interacts with the "first" strand (βA). These H-bonds serve to unite the two β-sheets.

A single solvent-accessible opening or portal is present on the surface of the protein and bounded by αII plus the two type I turns that connect βC and βD and βE and βF (Fig. 1B). The oval shaped portal is described by the van der Waals surfaces of 4 main chain atoms (Ile23, Lys27, Arg28, and Gly31) and the side chain atoms of 3 polar apolar (Ile23, Phe32, and Ala40). The side chain of Lys27 (from αII) is the only charged element in the immediate vicinity of the portal (Fig. 1B). The benzene side chain of Phe55 forms a hinged lid over the portal and exhibits two distinct conformations in the 1.85-Å structure of the apoprotein. Neither conformation appears to completely block the orifice of the portal.

Apo-I-FABP contains a very large solvent-filled cavity. Interactions between side chain atoms of nonpolar amino acids from each of the β-strands and αII serve to form a continuous, interior, U-shaped "scaffold" (Fig. 1A). The absence of a large, segregated hydrophobic core in apo-I-FABP appears to create a dependence on the interior, horseshoe-shaped scaffold plus interactions between exterior charged/polar side chains and ordered waters for stabilization of the protein's secondary structural elements. The hydrophilic domain in the apoprotein's interior cavity is composed principally of waters plus a few polar and charged residues. The total volume of the solvent cavity is ~850 Å3 (compared with an estimated volume of 17,800 Å3 for the entire apoprotein). 720 Å3 of the 850 Å3 is occupied by 24 ordered waters (Fig. 1A) while ~130 Å3 appears to be occupied by ordered water. The remaining 24 ordered waters are located in the protein's fatty acid binding site. Apo- and Holoproteins—Superimposition of the structures of apo-I-FABP and I-FABP containing bound 9Z-octadecanoate (oleate) reveals that protein makes remarkably few conformational adjustments to accommodate this fatty acid (Fig. 1C). These adjustments are principally limited to movement of main chain and side chain atoms that form the solvent-accessible portal, an ordering of the positions of the side chains of Tyr70, Tyr117, and Phe55, plus a reduction in the number of interior ordered waters. The fatty acid is linked to the protein via a series of feeble forces (H-bonds, van der Waals contacts) that involve residues from both sheets and from the helices.

The carbohydrate group of the bound 9Z-octadecanoate (oleate) is bound inside the protein while the methyl terminus of its acyl chain is positioned within the solvent-accessible portal (Fig. 1, B–D). The carboxylate group is linked to the protein via an electrostatic network that includes seven H-bonds (Fig. 1E). The 2 oxygen atoms of carboxylate (OE1 and OE2) are H-bonded to the NH1 and NH2 atoms of Arg106. OE1 also forms a H-bond with the indole nitrogen of Tyr70 while OE2 also forms a H-bond with an ordered water molecule. This water is H-bonded to another ordered water which, in turn, is H-bonded to NE2 of Gln115. Another H-bond links the OE1 atom of Gln115 and the NE atom of Arg106.

The hydrocarbon chain of 9Z-octadecanoate is bound to I-FABP...
A Model Fatty Acid Receptor

U-shaped "scaffold" that surrounds the fatty acid binding pocket. This scaffold is formed by a series of van der Waals interactions between the side chains of hydrophobic and aromatic residues (green). Eight turns connect these strands. With one exception, all are typical β-turns, i.e. either type I, II, or III with a i−1, i+2 H-bond motif (17). The two, short, right-handed helices are maintained by typical n to n+4 main chain H-bonds (17). The protein is oriented such that the gap between βD and βE is in the foreground. Panel B, the molecular surface of I-FABP-oleate, orientated as shown for the apoprotein in panel A. The electrostatic potential of the surface is depicted as follows: blue, areas of positive charge (calculated at pH 7.0), red, areas of negative charge, white, neutral. The terminal methyl group of oleate (yellow) is seen protruding from the solvent-accessible portal region. The blue density adjacent to this portal represents positive charge from the side chain of Lys27. Panel C, stereodiagram of a superimposition of the refined structures of apo-I-FABP and I-FABP-oleate. A backbone trace of apo-I-FABP plus the side chains of its Tyr70, Tyr117, and Phe55 residues are blue. The 24 interior ordered waters of the apoprotein are shown as purple spheres. The backbone trace of I-FABP-oleate, the side chains of its Tyr70, Tyr117, and Phe55 residues, plus the atoms of its bound oleate are shown in yellow. The eight interior ordered waters in the holoprotein are depicted as red spheres. Panel D, I-FABP-oleate illustrating its interior, 855-A³ cavity (outlined by dots) that contains bound oleate (yellow) and eight ordered waters (red). The side chains of Phe55, Tyr70, and Tyr117 are shown in green while the side chain of Arg106 is shown in dark blue. Panel E, stereodiagram showing the electrostatic network that binds the carboxylate group of oleate in I-FABP-oleate. Panels A–D were generated using GRASP (B. Honig and A. Nicholls).
in a bent conformation reflecting the presence of gauche bonds plus the single cis double bond between C9 and C10. This conformation is stabilized through a series of interactions with the side chain atoms of aromatic and hydrophobic residues and with several ordered interior waters (Fig. 1, C and D). The convex face of the chain is within van der Waals bonding distance of the side chains of 3 aromatic and 3 hydrophobic residues (Trp65, Trp72, Trp77, Tyr117, Leu82, Val162, Met183). Together, these side chains form a "crevice" that cradles the acyl chain. The phenolic groups of Tyr72 and Tyr117 show discrete disorder in the apoprotein because of rotation around their C-C-C-C bonds. Binding of the fatty acid restricts each side chain to a single well ordered position (Fig. 1C). This ordering reflects the contributions of both van der Waals contacts with CB of Ala72, to an open position where it interacts with Leu30, Asp134, and Ala73. The electron density appears similar in both conformations, suggesting that each set of van der Waals contacts provides similar free energies of association. In contrast, the terminal methyl of oleate is oriented toward the center of the benzene, allowing it to stabilize the position of this aromatic lid through van der Waals contacts. The torsional angles of the C-C bonds at the ω-terminus of this fatty acid produce a bent conformation that "increases" the area of contact between the van der Waals surfaces of oleate and the benzene (Fig. 1, C and D).

Eight interior ordered waters are discernible in the refined 1.75-Å structure of I-FABP:oleate (Fig. 1D). These waters are contained in an area of ~300 Å2 located between the concave surface of the chain and several polar or hydrophobic residues (Fig. 1D). The eight ordered waters in I-FABP:oleate are located in essentially the same positions occupied by 8 of the 24 internal ordered waters in apo-I-FABP; six of the other interior ordered waters in the apoprotein are "replaced" by C1, C2, C3, C12, C14, and C16 of oleate in the holoprotein (Fig. 1, C and E). These eight interior waters are remarkably well ordered, reflecting an elaborate series of interactions involving both H-bonds and water and appear to be van der Waals contacts. The oxygens of two of the eight waters are located within van der Waals distance of C1, C2, C3, and C11 (although the precise nature of the interaction of the oxygens with 9Z-octadecenoate is not known, ~85 Å2 of the 255 Å2 van der Waals surface of the acyl chain appears to be exposed to these waters). Each of the eight waters forms an average of 2.75 H-bonds. Water 170 in I-FABP:oleate is an example of a complex network of interactions involving feeble forces that stabilize the positions of these interior waters; water 170 H-bonds to three polar groups (OE2 of the fatty acid's carboxyl group, OE2 of Asp81 and a second inter nal water) and is also located within van der Waals distance of C1, C2, C3 (Fig. 1, D and E).

Comparison of the structures of I-FABP:oleate and I-FABP:palmitate, and I-FABP:myristate indicates that removal of 2 or 4 methylenes produces only modest structural adjustments in the protein. The overall r.m.s. deviation of all main chain atoms between I-FABP:myristate and I-FABP:oleate is only 0.53 Å. Most of the larger deviations occur in the turns connecting βE and βF, βF and βG, and βF and βI, turns that also have the highest temperature factors in I-FABP. The structures of each holoprotein. The components of the electrostatic network that bind the carboxylate groups of oleate, palmitate, and myristate occupy comparable positions. The locations of the side chain atoms that interact with the acyl chain of oleate are highly conserved in I-FABP:myristate and I-FABP:palmitate (r.m.s. deviation, 0.82 Å). Subtraction of 2 or 4 methylenes from oleate does not perturb the positions of the eight interior ordered waters in I-FABP:oleate; including the solvent accessible main chain and side chain atoms in the portal. The apparent movement in the acyl chain over a distance of 4 methylenes in I-FABP:oleate does not appear to be associated with movement of the positions of Trp92, Trp72, and Tyr117 or other components of the crevice that cradles the convex face of the carboxyl terminal methyl of the fatty acid chain, nor does it produce marked alterations in the positions of the seven remaining internal waters.

These findings suggest a number of conclusions. First, the movement of the acyl chain associated with removal of Arg106 by Gln emphasizes the role played by H-bond formation between the hydrogens of the fatty acid's carboxylate group and an ordered water in...
stabilizing acyl chain position. Second, the fact that the 7 remaining interior waters do not exhibit marked variations in their positions suggests that their locations are predominantly determined by H-bond interactions between themselves and with the side chain atoms of several polar residues. Third, the location of the J-shaped density suggests a potential way in which I-FABP can accommodate long chain length (C18-C22) fatty acids. Fourth, the ability of the receptor to substitute Glu for Ile in I-FABP and CRBPII to produce such dramatic changes in their binding specificities emphasizes the potentially generic nature of their lipid binding sites as well as the contribution of the H-bonding capabilities of their comparably positioned Arg<sup>106</sup>Gln<sup>115</sup> residues to ligand specificity. A total of five members of this protein family contain an Arg residue at a position equivalent to residue 106 in I-FABP-CRABP which I-FABP may represent an extreme (epitomized) example of how the forces that “drive” the binding mechanism and the determinants of entry/exit of the ligand/water into I-FABP remain unclear. I-FABP may represent an extreme (optimized) example of how the free energy of dehydration of the binding site can function as a major energetic contributor to a binding reaction, perhaps as an important contributor as the free energy of stabilization of the complex. The remarkably few apparent conformational adjustments made by I-FABP after acquisition of its ligand evokes a vision of the protein as a molecular water pump. However, we do not have sufficient insight about the binding mechanism to establish whether there is progressive movement of water and ligand into and out of the binding pocket. Nonetheless, the free energy of fatty acid desolvation should play an important role in binding. There appears to be a correlation between the solubility of fatty acids in aqueous solution and their binding affinities for I-FABP (4). I-FABP may act like a non-polar solvent, the K<sub>s</sub> for fatty acid binding paralleling the partition coefficient for the ligand between water and hexane (except for the constant favorable contribution of electrostatic interactions involving the carboxylate group and Arg<sup>106</sup>Gln<sup>115</sup>Trp<sup>82</sup>, and ordered waters). One hypothesis about the mechanism of entry of fatty acid into the protein involves a series of progressive interactions of carboxylate that direct it through the portal to the interior of the protein (7). An initial interaction would involve carboxylate and the side chains of Lys<sup>27</sup> (from oii) and/or Asp<sup>34</sup> (from the type I turn connecting bE and bF). These are the only charged elements in the immediate vicinity of the portal (Fig. 1B). A second interaction may involve the carboxylate of Arg<sup>106</sup>Gln<sup>115</sup>Asn<sup>11</sup>, and/or Asp<sup>34</sup> plus two ordered solvents form an electrostatic network that resembles the electrostatic network involving Arg<sup>106</sup>Gln<sup>115</sup>, and interior waters in the apoprotein. The methylene chain could assume a bent conformation that has its hydrophobic portion of the Val2 interactions with the protein while the carboxylate group forms an electrostatic network with Trp<sup>82</sup>Arg<sup>106</sup>Gln<sup>115</sup>, and two interior ordered waters. The mechanism and pathway for egress of water from the apoprotein during this process of ligand acquisition remain uncertain but may involve the “gap” between bD and bE. Finally, fatty acid release could occur through a reversal of this process or via an exit portal located near the “bottom” of the protein. It would appear that this exit portal would have to widen through movement of the side chains of Arg<sup>106</sup>Gln<sup>115</sup> and one or more members of the stack of three benzenes (from Phe<sup>47</sup>Phe<sup>62</sup>, and Phe<sup>92</sup>), which form a barrier between the interior and exterior of the protein. (N.B. Phe<sup>47</sup> exhibits its discrete disorder in the 1.2-Å structure of apo-I-FABP.) Determination of the solution state structure of E. coli-derived rat I-FABP using isotope-edited NMR methods should provide insights about the binding mechanism.

Finally, the structural motif exemplified by I-FABP and other family members that bind fatty acids may have broad biological applications. For example, the N-methyl-D-aspartate receptor is an ion channel whose activity is modulated by arachidonic acid. Petrou et al. (24) have noted a significant degree of primary sequence similarities between I-FABP and the receptor. Probing protein data bases with the sequences of FABP family members may help establish the representation of this motif in proteins with known, or heretofore unknown, function.

REFERENCES