BIOINFORMATION Discovery at the interface of physical and biological sciences

open access

www.bioinformation.net

Volume 9(20)

Hypothesis

Model of β-Sheet of Muscle Fatty Acid Binding Protein of *Locusta migratoria* Displays Characteristic Topology

Nadeem A Kizilbash*, Abdul Hai & Jamal Alruwaili

Department of Biochemistry, Faculty of Medicine & Applied Medical Sciences, Northern Border University, Arar-91431, Saudi Arabia; Nadeem A. Kizilbash - Email: fsd707@gmail.com; Phone: +966-592390610; *Corresponding author

Received November 06, 2013; Accepted December 12, 2013; Published December 27, 2013

Abstract:

The β -sheet of muscle fatty acid binding protein of *Locusta migratoria* (*Lm*-FABP) was modeled by employing 2-D NMR data and the Rigid Body Assembly method. The model shows the β -sheet to comprise ten β -strands arranged anti-parallel to each other. There is a β -bulge between Ser 13 and Gln 14 which is a difference from the published structure of β -sheet of bovine heart Fatty Acid Binding Protein. Also, a hydrophobic patch consisting of Ile 45, Phe 51, Phe 64 and Phe 66 is present on the surface which is characteristic of most Fatty Acid Binding Proteins. A "gap" is present between β_D and β_E that provides evidence for the presence of a portal or opening between the polypeptide chains which allows ligand fatty acids to enter the protein cavity and bind to the protein.

Keywords: Muscle Fatty Acid Binding Protein, Nuclear Overhauser Effect, β-sheet of Lm-FABP, Portal hypothesis, *Locusta migratoria*.

Background:

In insects, body fat plays a major role **[1]**. Insects store lipids as glycogen and triglycerides in the adipocytes. The metabolism of lipids is important in insects for growth and reproduction and provides energy during non-feeding periods. In the resting locust, the major source of energy is Trehalose (a hemolymph) **[2]**. One advantage of using lipids as a source of energy is that they weigh less than isocaloric amounts of carbohydrates **[3, 4]**. The main storage form of lipids is as triglycerides which are released as diglycerides during flight **[4, 5]**.

Based on the tissue of origin, the intracellular fatty acid-binding proteins (FABPs) are divided into three categories: (i) hepatic-type FABPS, (ii) intestinal-type FABPS and (iii) muscle/cardiac-type FABPs. These proteins are 14-15 kDa in size and are important for the uptake, metabolism and transport of long-chain fatty acids. They are also responsible for the modulation

of cell growth and proliferation. Human muscle and bovine heart fatty acid binding proteins bind long-chain fatty acids in the cytosol of muscle tissues.

According to the Portal hypothesis proposed for the uptake of fatty acids by FABPS, the fatty acid first adsorbs to the protein surface and then searches for an opening or portal by which it can enter the protein cavity. After finding the opening, the fatty acid enters the protein cavity. The next step is the protonation of the ligand, before de-solvation, which demands a large amount of energy (30 kJ mol⁻¹). Since the insertion of the negative charged ligand into the low dielectric matrix of the protein occurs before the protonation of the ligand. The insertion of the charged head-group of the ligand fatty acid inside the protein cavity requires energy in the amount of 300 kJ mol⁻¹[6]. Molecular dynamics simulations of I-FABP have

shown that the ligand first adsorbs to the protein surface. The rate limiting step is the de-solvation of the carboxylate head group of the fatty acid anion [7].

The β -sheet of human muscle fatty acid binding protein displays ten β -strands arranged anti-parallel to each other in which each strand is hydrogen bonded to its neighbor. This arrangement is completed by the formation of hydrogen bonds between the first and the last β -strand (Figure 1). The β -sheet is folded to form a β -barrel and contains an interior and an exterior surface. The interior surface serves as a cavity or a pit where the ligand binds [8]. The volume of this cavity varies between 300 and 700 Å³ [9]. In this project, 2-D NMR data, specifically information provided by the Nuclear Overhauser Effect (NOE), was used to build a model of the β -sheet of *Lm*-FABP using the Rigid Body Assembly method and bovine heart FABP as a template [10].



Figure 1: Backbone of human muscle FABP showing the arrangement of two α -helices (α_{I} and α_{II}) and ten β -strands (β_{A} , β_{B} , β_{C} , β_{D} , β_{E} , β_{F} , β_{G} , β_{H} , β_{I} and β_{I}) **[13**].

Methodology:

The methodology used in the study is summarized in schematic 1 (see supplementary material).

Protein expression and purification

Lm-FABP was expressed in *E.coli* cells using previously published protocols **[11]**. The cDNA of the protein was isolated after digestion with restriction endonucleases, NcoI and BamHI and ligated with the pET3d vector. The plasmid pET3d/*Lm*-FABP was used to transform the *E. coli* D12 strain BL21 (DE3) cells. Lm-FABP was purified by the use of previously published protocols **[11]**. The protocol included cell lysis, centrifugation and purification by a Sepharose HR 26/10 FPLC column. *Lm*-FABP eluted with 50 mM Tris/HCl (pH 8.0). The fractions were analyzed by SDS/PAGE and Western blotting. Fractions containing *Lm*-FABP were further purified by gel filtration chromatography using a Sephacryl S-100 column. ISSN 0973-2063 (online) 0973-8894 (print)

Bioinformation 9 (20):1003-1009 (2013)

NMR experiments

2-D homonuclear NMR spectra were acquired for protein sample bound to native fatty acids in 20 mM phosphate buffer (10% D₂O, 0.05% NaN₃) at pH 5.5 and 35 °C. These spectra were acquired in the "phase sensitive" mode with pre-saturation of the water signal to obtain maximum signal-to-noise ratio. All the chemical shifts were referenced with respect to sodium 3-(trimethylsilyl) [2,2,3,3-2H₄] propionate (TSP). Once the spectra had been acquired, they were analyzed to assign chemical shifts and to collect NOE constraints that consist of sequential, medium range and long range constraints.

Amino Acid Sequence Alignment

Before analysis of the NMR data, the amino acid sequence of *Lm*-FABP was analyzed. The sequence is characterized by the presence of thirty one amino acids with aliphatic side chains. It also contains three Proline residues (Pro39, Pro91 & Pro103). The sequence also contains one Cysteine residue (Cys115). There are ten aromatic amino acid residues (seven Phenylalanines and three Tyrosines) and two Histidines (His102 and His96), also present in the sequence. In addition, the sequence is characterized by two Methionines, (f-Met 0 and Met 21).

A comparison of the sequence of *Lm*-FABP with the bovine heart FABP was performed using the FSSP data-base. The alignment revealed 42% sequence homology between two proteins. There are three insertions (Ile 7, Leu 46 and Asn 92) and two deletions (Gly 99 and Ala 132) in the sequence of *Lm*-FABP as compared to bovine heart FABP. The amino acid insertions are exactly the same if the sequence of *Lm*-FABP is compared to human muscle FABP. The only amino acid deleted is Gly 99.



Figure 2: Comparison of secondary structure elements of *Lm*-FABP and bovine heart FABP showing the three insertions (Ile 7, Leu 46 and Asn 92) and the two deletions (Gly 99 and Ala 132).

Spin System Identification

2-D homonuclear COSY and TOCSY spectra were used to identify amino acid spin systems comprising J-coupled ¹H. The amide ¹H chemical shift dispersion in the 2-D TOCSY ranged from 10.18-6.78 ppm. The scalar coupled NH and $C_{\alpha}H$ ¹H were recognized in the "finger print" region of the 2-D COSY spectrum, as shown in **(Figure 2)**. To identify spin systems representing individual amino acids (comprising NH and $C_{\alpha}H$

along with the side-chains ¹H), the program AURELIA **[12]** was used to peak-pick the 2-D data. Initially, about 165 spin systems were identified. These included some doublets due to the presence of spin system heterogeneity. This list later was narrowed down to about 145 spin systems for 130 non-Proline amino acids in the protein. The three Proline residues present in the sequence were identified separately since they don't possess an amide ¹H.

NMR Chemical Shift Assignment

The strategy of sequential ¹H chemical shift assignment was followed to assign the backbone and side-chain NMR chemical shifts of *Lm*-FABP (data not shown).

Identification of Secondary Structure Elements

Secondary structure elements (Figure 2) were identified by characteristic Nuclear Overhauser Effect (NOE) cross-peaks as shown in schematic 2 (see supplementary material). The sequential NOEs observed for α -helices consist of a strong d_{NN} and a weak $d_{\alpha N}$ NOE cross-peak for each "i-1" and "i" amino acid residue. In addition to these medium range NOEs are also observed as shown in schematic. For β -strands, the characteristic NOE information is a strong $d_{\alpha N}$ and a weak d_{NN} cross-peak for "i-1" and "i" amino acid residues. For the β -turns, a strong-weak d_{NN} and strong to medium $d_{\alpha N}$ NOE crosspeaks are expected for "i-1" and "i" amino acid residues.



Figure 3: β-sheet of *Lm*-FABP modeled after bovine heart FABP by the use of 2-D NMR data and the Rigid Body Assembly method **[10]**.

Homology Modeling of β -sheet

The β -sheet of *Lm*-FABP (Figure 3) was modeled after the β sheet of Bovine Heart FABP using the Rigid Body Assembly method [10]. This method starts with the identification of the ISSN 0973-2063 (online) 0973-8894 (print) Bioinformation 9 (20):1003-1009 (2013) 1 conserved β -sheet of bovine H-FABP and then assembly of the model of β -sheet of *Lm*-FABP from the H-FABP template using the inter-strand NOE cross-peak intensities in **schematic 3**, **Table 1 (see supplementary material)**.

Discussion:

Many structures of human, bovine and insect FABPs are known via X-ray crystallography or NMR **[13-17]**. These structures have provided details about ligand-binding property of FABPs. The FABPs display a ten-stranded β -sheet structure and a large, hydrophilic, water-filled inner cavity that serves as a ligand-binding site for hydrophobic ligands such as fatty acids and retinoids **[6]**. The β -sheet of Lm-FABP shown in **(Figure 3)** closely resembles that of the other FABPs. The fatty acid is bound to FABPs by electrostatic and hydrogen bond interactions of its carboxylate head group with charged or polar residues inside the protein and by interactions of its tail with hydrophobic residues. A "gap" is observed between β D and β E in the β -sheet which can serve as a portal for entry of long chain fatty acids in the inner cavity of the protein.

A description of the ligand-binding process is given by the Portal hypothesis which describes the entry of long chain fatty acids inside the protein cavity through an opening or portal. The Portal hypothesis is supported by many experimental studies, most of which were carried out using either the I-FABP or ALBP. Crystallographic studies of both these proteins have indicated that the lipid tail is located near the suggested entry site **[18, 19]**. NMR measurements of I-FABP have suggested that the protein exhibits a pronounced backbone disorder at the portal region and is more mobile than the rest of the protein suggesting that this region may be involved in ligand insertion **[17, 20, 21, 22]**.

Conclusion:

The information provided by the Nuclear Overhauled Effect was used to build a model of the β -sheet of *Lm*-FABP using the Rigid Body Assembly method and bovine heart FABP as a template. The β -sheet structure of *Lm*-FABP displays a typical structure of 10 anti-parallel β -strands arranged to form a sheet. A "gap" is observed between β_D and β_E . A β -bulge is also present between Ser 13 and Gln 14 which is a difference from the structure of bovine heart fatty acid binding protein. The observed "gap" between β_D and β_E provides evidence for the presence of a portal between the polypeptide chains which allows long chain fatty acids to enter the protein cavity and bind to the protein.

Acknowledgement:

We would like to thank HEJ Research Institute of Chemistry (Karachi, Pakistan) for the use of their 600 MHz NMR spectrometer.

References:

- [1] Law JH & Wells MA, J Biol Chem. 1989 264: 16335 [PMID: 2674129]
- [2] Weis-Fogh T, Phil Trans R Soc (B). 1952 237:1
- [3] Beenakers AMT, Insect Physiol. 1965 11: 879
- [4] Beenakers AMT, Gen Camp Endocr. 1969 13: 12
- [5] Mayer RJ & Candy DJ, Comp Biochem Physiol. 1969 31: 409 [PMID: 5351972]

- [6] Friedman R *et al. Biochemistry* 2005 44: 4275 [PMID: 15766256]
- [7] Richieri GV et al. J Biol Chem. 1996 271: 11291 [PMID: 8626681]
- [8] Veerkamp JH et al. Biochim Biophys Acta. 1991 1081: 1 [PMID: 1991151]
- [9] Massolini G & Calleri E, J Chromatogr B Analyt Technol Biomed Life Sci. 2003 797: 255
- [10] Marti-Renom MA et al. Annu Rev Biophys Biomol Struct. 2000 29: 291 [PMID: 10940251]
- [11] Maatman RG et al. Eur J Biochem. 1994 221: 801 [PMID: 8174560]
- [12] Gorler A et al. J Magn Reson. 1999 137: 39 [PMID: 10053131]
- [13] Young ACM et al. Structure 1994 2: 523 [PMID: 7922029]

- [14] Haunerland NH et al. Biochemistry. 1994 33: 12378 [PMID: 7918460]
- [15] Lassen D et al. Eur J Biochem. 1995 230: 266 [PMID: 7601110]
- [16] Zhang F et al. J Biomol NMR. 1996 9: 213 [PMID: 9204553]
- [17] Lucke C et al. Structure 1996 4: 785 [PMID: 8805562]
- [18] Sacchettini JC et al. J Mol Biol. 1989 208: 327 [PMID: 2671390]
- [19] Xu Z et al. J Biol Chem. 1993 268: 7874 [PMID: 8463311]
- [20] Hodsdon ME & Cistola DP, *Biochemistry* 1997 36: 2278 [PMID: 9047330]
- [21] Hodsdon ME & Cistola DP, *Biochemistry* 1997 36: 1450 [PMID: 9063893]
- [22] Zhang FL et al. Biochemistry 2003 42: 7339 [PMID: 12809489]

Edited by P Kangueane

Citation: Kizilbash et al. Bioinformation 9(20): 1003-1009 (2013)

License statement: This is an open-access article, which permits unrestricted use, distribution, and reproduction in any medium, for non-commercial purposes, provided the original author and source are credited

Supplementary material:

Table 1: List of Nuclear Overhauser Effect (NOE) cross-peak intensities used for modeling of β -sheet of *Lm*-FABP after bovine heart FABP by the use of the Rigid Body Assembly method.

¹ H Nuclei of Different Amino Acid Residues exhibiting Dipolar		Intensities of NOF Crossneaks observed in 2-D NOFSV	
Interaction	and childring Dipolai	spectrum	
Ala 5 CαH	Val 45 NH	medium NOE	
Ala 5 CaH	Glu 44 CaH	strong NOE	
Gly 6 NH	Glu 44 CaH	medium NOE	
Gly 6 NH	Leu 43 NH	weak NOE	
Lys 8 CaH	Leu 43 NH	medium NOE	
Lys 8 CaH	Glu 42 CaH	strong NOE	
Ala 5 CaH	Val 45 NH	medium NOE	
Ala 5 CaH	Glu 44 CaH	strong NOE	
Gly 6 NH	Glu 44 CaH	medium NOE	
Gly 6 NH	Leu 43 NH	weak NOE	
Lys 8 NH	Gln 133 NH	weak NOE	
Tyr 9 NH	Glu 42 CaH	medium NOE	
Tyr 9 NH	Ile 41 NH	weak NOE	
Tyr 9 CaH	Ala 132 CaH	strong NOE	
Lys 8 NH	Gln 133 NH	weak NOE	
Tyr 9 NH	Glu 42 CaH	medium NOE	
Tyr 9 NH	lie 41 NH	weak NOE	
I yr 9 CαH	Ala 132 C α H	strong NOE	
Lys 10 CaH	11e 41 NH No1 40 Cov	meanum NOE	
Lys 10 CaH	val 40 CaH	strong NOE	
Ly5 10 INFL Lyc 10 NFL	A1a 152 C0H I we 121 NH		
Lys 10 1011 Lou11 NH	Lys 131 INF Val 40 Cort	medium NOF	
	Val 40 COH Lyc 131 NH	medium NOE	
Leu II Cull	Lys 131 MI Twr 120 CoH	strong NOE	
Val 40 NH	I vs 56 NH	weak NOF	
lle 41 CaH	Lys 56 NH	medium NOE	
Ile 41 Coll	Ser 55 CaH	strong NOE	
Glu 42 NH	Ser 55 CaH	medium NOE	
Glu 42 NH	Thr 54 NH	weak NOE	
Leu 43CaH	Thr 54 NH	medium NOE	
Leu 43CaH	Leu 53 CaH	strong NOE	
Glu 44 NH	Leu 53 CaH	medium NOE	
Glu 44 NH	Lys 52 NH	weak NOE	
Val 45 CaH	Lys 52 NH	medium NOE	
Val 45 CaH	Phe 51 CaH	strong NOE	
Asp 47NH	Phe 51 CaH	medium NOE	
Asp 47 NH	Lys 50 NH	weak NOE	
Lys 50 CaH	Leu 68 NH	medium NOE	
Lys 50 CaH	Lys 67 CaH	strong NOE	
Phe 51 NH	Lys 67 CaH	medium NOE	
Phe 51 NH	Phe 66 NH	weak NOE	
Lys 52 CaH	H Phe 66 NH	medium NOE	
Lys 52 CaH	Inr 65 CaH	strong NOE	
Leu 53 NH Leu 52 NH	Thr 65 CaH Pho 64 NU	medium NOE	
Leu 53 NH The 54 Could	rne 64 NH Pho 64 NH	Weak NOE	
INF 54 CCH The 54 Could		atrong NOE	
ти 34 Сип Sor 55 NH		modium NOF	
Sor 55 NH	GIU 03 CUH Thr 62 NH		
Lve 56 CoH	Thr 62 NH	medium NOF	
Lys 56 CaH	Asn 61 CoH	strong NOE	
Thr 57 NH	Asn 61 CaH	medium NOE	
Thr 57 NH	Lvs 60 NH	weak NOE	
Glu 70 CaH	Ile 86 NH	medium NOE	
Glu 71 NH	Ile 86 NH	weak NOE	
Glu 71 NH	Ile 85 CaH	medium NOE	
Phe 72 CaH	Ile 85 CaH	strong NOE	
Phe 72 CaH	Ser 84 NH	medium NOE	
Asp 73 NH	Ser 84 NH	weak NOE	
Glu 71 NH	Ile 85 CaH	medium NOE	

ISSN 0973-2063 (online) 0973-8894 (print) Bioinformation 9 (20):1003-1009 (2013)

Phe 72 CaH	Ile 85 CaH	strong NOE
Phe 72 CαH	Ser 84 NH	medium NOE
Asp 73 NH	Ser 84 NH	weak NOE
Asp 73 NH	Lvs 83 CaH	medium NOE
Glu 74 CαH	Lvs 83 CaH	strong NOE
Glu 74 CoH	Val 82 NH	medium NOE
Val 82 Call	Asp 101 NH	modium NOE
Val 62 Cull	Charles Carl	atron a NOE
Val 82 CaH	GIV 100 CaH	Strong NOE
Lys 83 NH	GIY 100 CaH	medium NOE
Lys 83 NH	Lys 99 NH	weak NOE
Ser 84 CaH	Lys 99 NH	medium NOE
Ser 84 CaH	Gln 98 CaH	strong NOE
Ile 85 NH	Gln 98 CaH	medium NOE
Ile 85 NH	Glu 97 NH	weak NOE
Ile 86 CaH	Glu 97 NH	medium NOE
Ile 86 CaH	His 96 CaH	strong NOE
Thr 87 NH	His 96 CaH	medium NOE
Thr 87 NH	Val 95 NH	weak NOE
Leu 94 NH	Glu 109 CaH	medium NOE
Leu 94 NH	Arg 108 NH	weak NOE
Val 95 CoH	Arg 108 NH	medium NOE
Val 95 Call	Ilo 107 CorH	strong NOE
His 96 NH		modium NOF
1115 90 INII 11:0 06 NUI	Ile 107 Cull	medium NOE
Glu 97 CaH		medium NOE
Glu 97 CaH	lle 105 CaH	strong NOE
GIn 98 NH	Ile 105 CaH	medium NOE
Gln 98 NH	Thr 104 NH	weak NOE
Lys 99 CaH	1 hr 104 NH	medium NOE
Lys 99 CaH	Pro 103 CaH	strong NOE
Gly 100 NH	Pro 103 CaH	medium NOE
Gly 100 NH	His 102 NH	weak NOE
Thr 87 NH	His 96 CaH	medium NOE
Thr 87 NH	Val 95 NH	weak NOE
Leu 94 NH	Glu 109 CaH	medium NOE
Leu 94 NH	Arg 108 NH	weak NOE
Thr 104 CaH	Gly 122 NH	medium NOE
Ile 105 NH	Lys 120 NH	weak NOE
Ile 106 CaH	Lys 120 NH	medium NOE
Ile 106 CαH	Ile 119 CaH	strong NOE
Ile 107 NH	Ile 119 CaH	medium NOE
Ile 107 NH	Thr 118 NH	weak NOE
Arg 108 CaH	Thr 118 NH	medium NOE
Ile 107 NH	Ile 119 CaH	medium NOE
Arg 108 CaH	Ile 117 CαH	strong NOE
Glu 109 NH	Ile 117 CaH	medium NOE
Glu 109 NH	Val 116 NH	weak NOE
Phe 110 CaH	Val 116 NH	medium NOE
Phe 110 CoH	Cys 115 CoH	strong NOF
Ser 111 NH	Cyc 115 CoH	medium NOF
Ser 111 NH	Gln 114 NH	week NOF
Clp 114 CoH	Ala 132 NH	medium NOE
Cln 114 Coll		strong NOE
GIII 114 COM	Lys 131 Carl	madium NOE
Cys 115 NII	Lys 151 Can	medium NOE
	Tyr 130 NH	weak NOE
		medium NOE
Val 116 Coch	lle 129 CaH	strong NOE
lle 117 NH	Ile 129 CaH	medium NOE
Ile 117 NH	Arg 128 NH	weak NOE
Thr 118 CaH	Arg 128 NH	medium NOE
Thr 118 CaH	Thr 127 CaH	strong NOE
Ile 119 NH	Thr 127 CaH	medium NOE
Ile 119 NH	Ala 126 NH	weak NOE
Lys 120 CaH	Ala 126 NH	medium NOE
Lys 120 CaH	Val 125 CaH	strong NOE

*Intensity<60,000 = weak; Intensity <90,000 = medium; Intensity >90,000 = strong *Intensities of NOE cross-peaks as observed in AURELIA peak-picked data.

Schematic 1: Methodology used for acquisition of NMR data and determination of the β-sheet.



Schematic 2: Nuclear Overhauser Effect (NOE) cross-peak intensities used to identify \Box -helices, β -strands and β -turns.

1 (11(11))		typ	e I β-turn
	(2.4 Å) strong NOE (3.4 Å) medium NOE (4.2 Å) weak NOE (4.0 Å) weak NOE	d _{NN} (residue 1,2/3,4) d _{NN} (residue 2,3)	(2.6 Å) strong NOE (2.4 Å) strong NOE
d _{αN} (i-3,i) d _{αN} (i-4,i) d _{βα} (i-3,i)	(3.4 Å) medium NOE (4.3 Å) weak NOE (3.5-5.0 Å) medium-weak NOE	d _{aN} (residue 1,2/3,4) d _{aN} (residue 3,4)	(3.2 Å) medium NOE (3.4 Å) medium NOE
Sequential and 1	medium range NOE cross-	type II β-turn	
peaks observed for α-helices.		d _{NN} (residue 1,2/3,4) d _{NN} (residue 2,3)	(2.4 Å) strong NOE (4.5 Å) weak NOE
d _{NN} (i,i+1/i,i-1) d _{Na} (i,i-1)	(4.5 Å) weak NOE (2.2 Å) strong NOE	d _{aN} (residue 1,2/3,4) d _{aN} (residue 2,3)	(3.2 Å) medium NOE (2.2 Å) strong NOE
Sequential (i,i+ strand.	1/i,i-1) NOEs observed for β -	Sequential NOE cro turns.	sspeaks observed for β

Schematic 3: Inter-strand Nuclear Overhauser Effect (NOE) cross-peak intensities used for modeling of the β-sheet of *Lm*-FABP.

Sequential NOEs

d _{NN} (i,i+1/i,i-1) d _{Na} (i,i-1)	4.5 Å 2.2 Å	(weak NOE) (strong NOE)				
Inter-strand NOEs						
d _{NN} (i,j)	3.3 Å (w	veak NOE)				
d _{Na} (i,j)	3.2 Å (n	nedium NOE)				
daa (i,j)	2.3 Å (s	trong NOE)				