Solution State Structure of pB1, the Mimotopic Peptide of Apolipoprotein B-100, by NMR

Sungran Lee, Daesung Kim, Hyojoon Kim,† Yongwoo Lee,‡ and Hoshik Won*

Department of Applied Chemistry, Hanyang University, Ansan 426-791, Korea
†Department of Biochemistry and Molecular Biology, Hanyang University, Ansan 426-791, Korea
‡R&D Center, Samsung Engineering Co., Ltd., Yongin 449-844, Korea
Received August 2, 2004

Apolipoprotein B-100 (Apo-B100) is a major protein component for low density lipoproteins (LDL). A number of mimetic peptides of Apo-B100 were screened from the phase-displayed random peptide library by utilizing monoclonal antibody (B9). Mimetic peptide for B9 epitope against apo B-100 was CRNVPPIFNDVYWIAF (pB1). From the BLAST search, the mimetic peptide pB1 had 40% homology with apo B-100. As a result of the structural determination of this mimotope using homo/hetero nuclear 2D-NMR techniques and NMR-based distance geometry (DG)/molecular dynamic (MD) computations, DG structure had low penalty value of 0.3-0.6 Å^2 and the total RMSD was 0.5-1.5 Å. Moreover, pB1 structure included a weak 3_{10}-helix from Ile^7 to Trp^{13}.

Key Words : Apolipoprotein B-100, NMR spectroscopy, Molecular dynamic computation

Introduction

Plasma lipoproteins are water-soluble particles composed of lipids and one or more specific proteins called apolipoproteins.[1] There are three main lipoprotein classes according to density : very low density lipoproteins (VLDL), low density lipoproteins (LDL), and high density lipoproteins (HDL).[2] Studies relating to the structure and metabolism of LDL are important because of the direct correlation between atherosclerosis and high LDL levels in human plasma.[3] LDL is the end product of VLDL catabolism and the major cholesterol-transporting lipoprotein in human plasma.[2] The majority of LDL particles contain a single apolipoprotein called apo B-100.[3] After the elucidation of the role of apolipoproteins in the regulation of lipoprotein metabolism, it became apparent that improvements in the characterization of apo B-100 were needed to facilitate the development of the linkage between LDL and atherosclerosis.[2]

Apolipoprotein B is the largest and one of the most important proteins that cover the lipid surfaces of lipoproteins. Apo B exists in two forms, apo B-100 and apo B-48.[4] Apolipoprotein B (apo B) is the major protein component of plasma LDL. It plays functional roles in lipoprotein bio-synthesis in liver and intestine, and is the ligand recognized by the LDL receptor during receptor-mediated endocytosis.[5] Apo B-100 which consists of 4536 amino acids has a molecular mass of 513 kDa and its levels of both LDL-cholesterol and plasma apo B are correlated with coronary heart.[6]

The profile is different from that of a typical apolipoprotein which has a high α-helical content and almost no β-sheet. In most apolipoproteins lipid binding occurs through amphipathic α-helical segment.[7-9] The structure of human apo B has been analysed in term of its functions in lipid binding, lipoprotein assembly and as the ligand responsible for LDL clearance by the LDL receptor pathway.[9] In apo B-100 few of the predicted α-helices are truly amphipathic in terms of charge distribution on the polar surface[7] except for one extended region (residues 2,000-2,600) which contains good examples of amphipathic α-helices, and may contribute to lipid binding. The secondary structure of apo B-100 has been suggested to consist of 43% α-helical, 21% β-sheet, 16% β-turn and 20% random structure.[10-12] The β-structure of apo B-100 is thought to be responsible for its interaction with lipids, due to its high hydrophobicity,[13] but is not confined to a particular region and various sections of the protein are buried in the lipid moiety.[14,15]

Recent computer modeling studies based upon biochemical analyses[16] have shown that large segment of the apo B backbone have a high amphipathic structure predicted to bind lipid. If these amphipathic β-sheets and α-helices are not folded and associated with lipid in the proper temporal sequence, the structural model predicts that the hydrophobic surfaces would become unstable in the aqueous environment of the ER lumen, leading to improper folding of nascent apo B and eventual degradation.[17]

Conformational studies for mimetic peptide CRNVPPIFNDVYWIAF (pB1) recognized by the monoclonal antibodies will be discussed from computer simulation analysis in this study. ^{1}H, ^{13}C, DEPT and 2D NMR (COSY, TOCSY, NOESY) experiments were performed using peptide, signal assignments were accomplished from experimental spectra. On the basis of these distance data from NOESY experiments, Distance Geometry (DG) and Molecular Dynamics (MD) were carried out to obtain the tertiary structure of mimetic peptide pB1.

*Corresponding Author. Fax: +82-31-407-3863; e-mail: hswon@hanyang.ac.kr

Experimental Section

Preparation of sample. Similar sequence of inherence antigen, peptide was obtained from Bio-Synthesis, Inc. Peptide was synthesized using solid-phase method (FMOC chemistry). For NMR experiments. peptide was dissolved in 350 μL DMSO-d_6.

NMR measurements. All NMR experiments were performed by using the Varian Mercury 300 MHz and Unity 500 MHz NMR spectrometer. Two-dimensional NMR experiments included correlated spectroscopy (COSY) and such as the phase sensitive total coherence spectroscopy (TOCSY). nuclear Overhauser enhancement spectroscopy (NOESY) experiments were performed with a 256 × 2048 data matrix size with 16 scans per t1 increment and spectra were zero filled of 2048 × 2048 data points. TOCSY spectrum was collected with a mixing time of 50 msec. MLEV-17 spin lock pulse sequence. Data were processed and analyzed on a SGI Octane workstation using Felix and NMRView 4.0.[18] NMRView was also used for sequential assignment of each amino acid.

Determination of solution state structure. Structure determinations were carried out using HYGEOM^{TM}, HYNMR^{TM}. Sequential assignments of amino acid spin systems were made using COSY, NOESY and TOCSY. Most important, direct way for secondary structure determination based on qualitative analysis of NOESY spectrum.[19] The structures were calculated from the NMR data according to the standard HYGEOM^{TM} simulated annealing and refinement protocols with minor modifications. NOE cross peaks were grouped according to their intensity into four categories: strong (2.0-3.0 Å), medium (2.0-3.5 Å), weak (3.0-4.5 Å). and very weak (3.5-5.0 Å).

Most of distance geometry (DG) algorithm accepts the input of distance constraints from NOE measurements.[20] (1) The set of distance restraints or bounds obtained from NOE data are determined by planarity restraints derived from the primary structure. This involves the selection of the possible intervals between lower and upper bounds consistent. (2) Embedding. The values of the distances from within the bounds obtained by bound smoothing are guessed at random. and the atomic coordinates are generated which represent the best-fit to this guess. (3) Optimization. The deviations of the coordinates from the distance bounds, as well as the stereospecific assignments. are minimized. When additional conjugate gradient minimization (CGM) was unable to further reduce the penalty for a particular structure. 2D NOESY back calculation were performed. and new distance restraints dictated by discrepancies between the experimental and back-calculated spectra were added to the experimental restraint list. Freshly embedded DG structures minimized with the modified restraints list generally exhibited penalty values lower than those of the previously refined structures and the new DG structures generally gave back-calculated NOESY spectra that were more consistent with experimental data.[21] The structure was calculated using the DG algorithm HYGEOM^{TM}, and 10 separated structures were generated using all the constrains and random input. No further refinement by energy minimization was carried out on the output of the DG calculations. RMSD (root-mean-square distances) deviations between the NMR structures were 0.4 Å for the backbone.[20] Back-calculation were assigned to GENNOE calculation in order to generate the theoretical NOEs. A consecutive serial files. obtained from

Table 1. ^{1}H-NMR chemical shift of pB1

Residue	NH	αH	βH	γH	Others
Cys^1	8.278	4.479	2.420, 2.335		
Arg^2	7.917	4.251	1.64	1.475	δH : 3.015 εH : 7.453 NH_2 : 4.297 NH : 7.535
Asn^3	8.350	4.502	2.386 2.356		NH_2 : 6.856, 7.324
Val^4	7.597	4.259	1.855	0.825, 0.750	
Pro^5		4.489	2.061	1.839, 1.735	δH : 3.456
Pro^6		4.489	2.061	1.839, 1.735	δH : 3.618
Ile^7	7.619	4.04	1.563	0.924	γCH_3, δCH_3 : 0.632
Phe^8	8.132	4.502	2.455, 2.376		Ring : *7.131*
Asn^9	7.874	4.502	2.722, 2.610		NH_2 : 7.131, 6.867
Asp^{10}	8.114	4.502	2.594, 2.455		
Val^{11}	7.452	3.996	1.854	0.629, 0.616	
Tyr^{12}	7.814	4.353	2.760, 2.590		2, 6H : 6.523 3, 5H : 6.879
Trp^{13}	7.917	4.536	3.015, 2.923		NH : 10.650 2H : 7.02 Ring : *7.234*
Ile^{14}	7.774	4.113	1.598	0.935	γCH_3, δCH_3 : 0.704
Ala^{15}	7.874	4.251	1.097		
Phe^{16}	7.949	4.357	2.949, 2.851		Ring : *6.879*

GENNOE calculation. were incorporated into HYNMR™ to generate NOE back-calculation spectra which can be directly compared with experimental NOESY spectra.[22]

Results and Discussion

Assignments using sequential NOEs can be obtained for proteins with natural isotope distribution. The ^{1}H-^{1}H connectivities that identify the different amino acid type are established via scalar spin-spin coupling, using COSY and TOCSY. Relations between protons in sequentially neighboring amino acid residues i and i+1 are established by NOEs manifesting close approach among $d_{\alpha N}$, d_{NN}, $d_{\beta N}$. Table 1 lists the correlating signals of adjacent residues on the basis of dipolar connectivities obtained from 2D spectra. Dipolar connectivities from amide protons to α- and amide

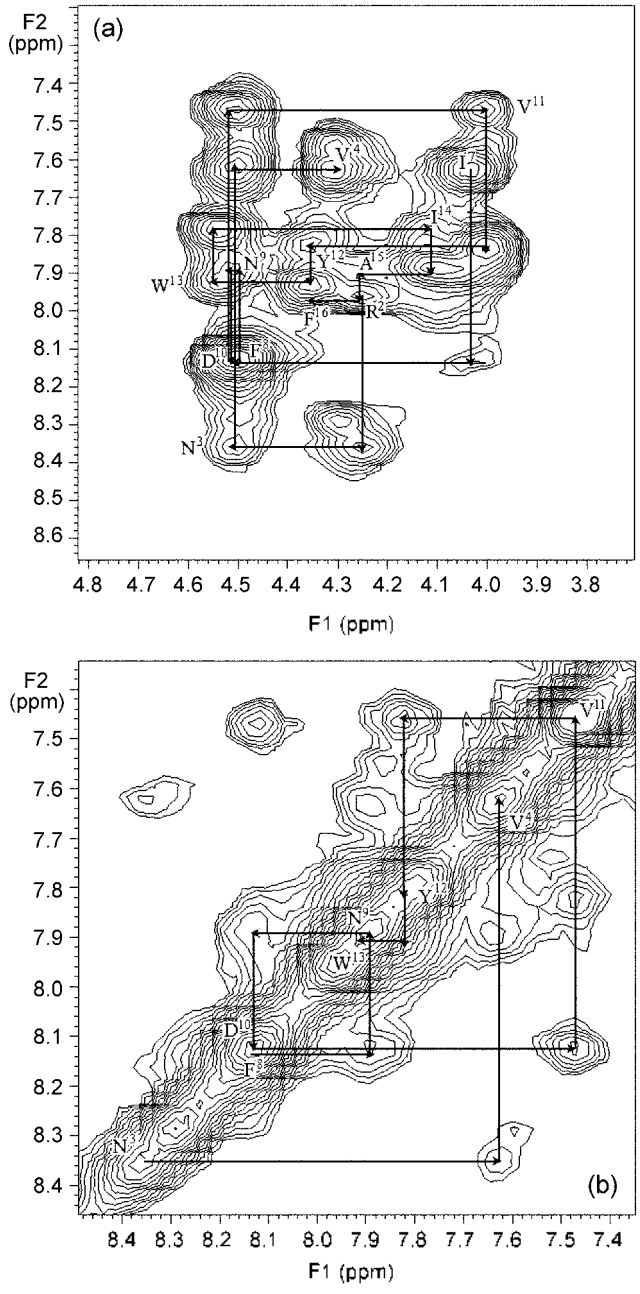

Figure 1. Intraresidue cross peaks labled in the NH-C$_\alpha$H region (upper), NH-NH region (lower) of the NOESY spectra of pB1 (τ_m = 300 ms).

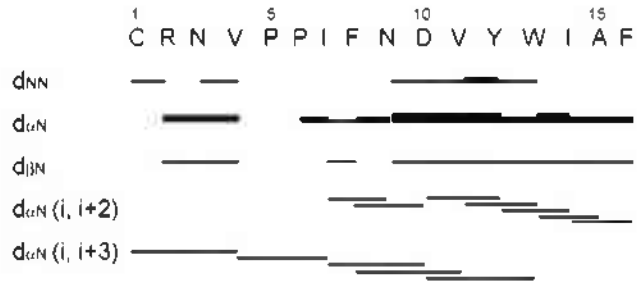

Figure 2. NOESY connectivities involving backbone protons for amino acids i and j. The height of the bars symbolizes the relative strength (strong, medium, weak) of the cross peaks in a qualitative way.

Table 2. Important NOE connectivities used for the structure determination of pB1

Residue	δ(ppm)	NOE connectivities
Cys^1_{NH}	8.278	$C^1{}_\alpha$(w), $V^4{}_{NH}$(w), $R^2{}_{NH}$(vw)
$Cys^1{}_\alpha$	4.479	$C^1{}_\beta$(s), $V^4{}_{NH}$(w)
$Arg^2{}_{NH}$	7.917	$R^2{}_\alpha$(w), $R^2{}_\beta$(w), $R^2{}_\gamma$(w)
$Arg^2{}_\alpha$	4.251	$R^2{}_\beta$(m), $R^2{}_\gamma$(m), $R^2{}_\delta$(w) , $N^3{}_\alpha$(w), $N^3{}_{NH}$(s)
$Arg^2{}_\delta$	3.015	$R^2{}_\beta$(m), $R^2{}_\gamma$(m), $R^2{}_\varepsilon$(w)
$Asn^3{}_{NH}$	8.350	$N^3{}_\alpha$(w), $N^3{}_\beta$(w), $R^2{}_\beta$(w), $R^2{}_\gamma$(vw), $V^4{}_{NH}$(w)
$Asn^3{}_\alpha$	4.502	$V^4{}_{NH}$(s), $N^3{}_\beta$(s)
$Val^4{}_{NH}$	7.597	$N^3{}_\beta$(w), $V^4{}_\alpha$(s), $V^4{}_\beta$(m), $V^4{}_\gamma$(w), $R^2{}_\beta$(w)
$Val^4{}_\alpha$	4.259	$P^5{}_\delta$(m), $V^4{}_\beta$(m), $V^4{}_\gamma$(m), $I^7{}_{NH}$(w)
$Pro^5{}_\alpha$	4.489	$P^5{}_\beta$(m), $P^5{}_\gamma$(m), $P^5{}_\delta$(m)
$Pro^5{}_\delta$	3.456	$P^5{}_\beta$(m), $P^5{}_\gamma$(m), $V^4{}_\gamma$(w)
$Pro^6{}_\alpha$	4.489	$P^6{}_\beta$(m), $P^6{}_\gamma$(m), $P^6{}_\delta$(w)
$Pro^6{}_\delta$	3.618	$P^6{}_\beta$(w), $P^6{}_\gamma$(m)
$Ile^7{}_{NH}$	7.619	$P^6{}_\alpha$(m), $N^9{}_{NH}$(w), $I^7{}_\alpha$(s), $I^7{}_\beta$(m), $I^7{}_\gamma$(m), $I^7{}_\delta$(m)
$Ile^7{}_\alpha$	4.040	$F^8{}_{NH}$(w), $I^7{}_\beta$(m), $I^7{}_{\gamma CH2}$(w), $I^7{}_\gamma$(w), $I^7{}_\delta$(w), $N^9{}_{NH}$(w)
$Ile^7{}_\beta$	1.563	$I^7{}_{\gamma CH2}$(m)
$Phe^8{}_{NH}$	8.132	$F^8{}_\alpha$(m), $F^8{}_\beta$(m), $I^7{}_\beta$(vw)
$Phe^8{}_\alpha$	4.502	$N^9{}_{NH}$(m), $F^8{}_\beta$(s), $D^{10}{}_{NH}$(m), $V^{11}{}_{NH}$(vw)
$Asn^9{}_{NH}$	7.874	$D^{10}{}_{NH}$(w), $N^9{}_\alpha$(m), $N^9{}_\beta$(m), $I^7{}_\beta$(w)
$Asn^9{}_\alpha$	4.502	$D^{10}{}_{NH}$(s), $N^9{}_\beta$(s)
$Asp^{10}{}_{NH}$	8.114	$V^{11}{}_{NH}$(m), $D^{10}{}_\alpha$(s), $D^{10}{}_\beta$(m), $I^7{}_\gamma$(w), $N^9{}_\beta$(w)
$Asp^{10}{}_\alpha$	4.502	$V^{11}{}_{NH}$(s), $D^{10}{}_\beta$(s), $V^{11}{}_\alpha$(w), $Y^{12}{}_{NH}$(w), $W^{13}{}_{NH}$(w)
$Val^{11}{}_{NH}$	7.452	$Y^{12}{}_{NH}$(m), $V^{11}{}_\alpha$(m), $V^{11}{}_\beta$(m), $V^{11}{}_\gamma$(m), $D^{10}{}_\beta$(w)
$Val^{11}{}_\alpha$	3.996	$Y^{12}{}_{NH}$(s), $V^{11}{}_\alpha$(s), $V^{11}{}_\gamma$(s), $W^{13}{}_{NH}$(w)
$Tyr^{12}{}_{NH}$	7.814	$W^{13}{}_{NH}$(vw), $Y^{12}{}_\alpha$(w), $Y^{12}{}_\beta$(m), $V^{11}{}_\beta$(m), $V^{11}{}_\gamma$(w)
$Tyr^{12}{}_\alpha$	4.353	$W^{13}{}_{NH}$(m), $Y^{12}{}_\beta$(s), $I^{14}{}_{NH}$(w)
$Trp^{13}{}_{NH}$	7.917	$W^{13}{}_\alpha$(m), $W^{13}{}_\beta$(m), $Y^{12}{}_\beta$(w)
$Trp^{13}{}_\alpha$	4.536	$W^{13}{}_\beta$(m), $I^{14}{}_{NH}$(s), $A^{15}{}_{NH}$(w)
$Ile^{14}{}_{NH}$	7.774	$I^{14}{}_\alpha$(m), $I^{14}{}_\beta$(w), $W^{13}{}_\beta$(w), $I^{14}{}_\gamma$(w), $I^{14}{}_\delta$(w)
$Ile^{14}{}_\alpha$	4.113	$I^{14}{}_\beta$(w), $I^{14}{}_\gamma$(w), $A^{15}{}_{NH}$(m), $F^{16}{}_{NH}$(w)
$Ala^{15}{}_{NH}$	7.874	$A^{15}{}_\alpha$(w), $A^{15}{}_\beta$(m), $I^{14}{}_\beta$(w), $I^{14}{}_\gamma$(w), $I^{14}{}_\delta$(w)
$Ala^{15}{}_\alpha$	4.251	$A^{15}{}_\beta$(m), $F^{16}{}_{NH}$(m)
$Phe^{16}{}_{NH}$	7.949	$F^{16}{}_\alpha$(w), $F^{16}{}_\beta$(w), $A^{15}{}_\beta$(w)
$Phe^{16}{}_\alpha$	4.357	$F^{16}{}_\beta$(w)

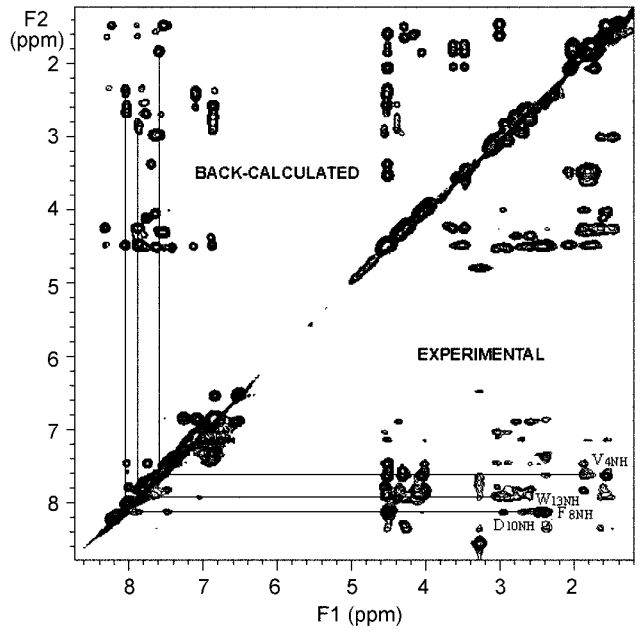

Figure 3. Comparisons of back-calculation and experimental NOESY spectrum of pB1 collected at 300 ms mixing time.

protons were also used for sequential signal assignments, and the fingerprint region of the NOESY spectrum that contains these connectivities is shown in Figure 1. Due to the ambiguity arising from degeneracy or signal overlap, the chemical shifts of ring protons were marked with italic font. Although there are some identical chemical shifts, but the ambiguity arising from severe signal overlap may be overcome by advanced NMR signal enhancement techniques.[23,24] From the NOESY connectivity table of Figure 2, it was apparent that the protein had a tendency to form weak 3_{10} helices from amino acid Ile^7 to Trp^{13} as throughout the core and basic regions.

For helical structures, one needs to observe only a small proportion of all potentially available short distances to characterize the main body of the helix, whereas accurate identification of the two ends of a helix can be more difficult and may require additional data. In the 3_{10} helix, short distances prevail between residues i and i+3, and between residues i and i+2. Observed medium-range NOEs connectivities, such as $d_{\alpha N}(i, i+2)$, $d_{\alpha N}(i, i+3)$, are indicated by lines connecting the one and two residues that are related by the NOE. $d_{\alpha N}(i, i+2)$ connectivities observed from Ile^7 to Asp^{10} and from Asp^{10} to Phe^{16}. Also, $d_{\alpha N}(i, i+3)$ connectivities are found from Ile^7 to Trp^{13}. Numerous sequential and $d_{\alpha N}(i, i+3)$ NOEs provide additional information for resolving possible ambiguities in the interpretation of the d_{NN} interactions. Structure calculation was used to many inter-residue and intra-residue NOE connectivities. Table 2 indicates the important NOE connectivities used for the structure determination.

In order to determine the DG structure, several variable-velocity simulated annealing and conjugate gradient minimization steps were used in the refinement scheme. Addition of restraints to account for minor differences between experimental and back-calculated spectra enabled the generation of new DG structures with substantially reduced penalties. To determine which of the DG structures most accurately reflect the experimental NOESY data, 2D NOESY back calculations were carried out. As illustrated in Figure 3, back-calculated spectrum of the pB1 was generally consistent with the experimental NOESY data. Ten final superpositioned DG structures are shown in Figure 4. Pairwise RMSDs obtained upon superposition of all atoms were in the range 0.5-1.5 Å. The final result of a structure determination is presented as a superposition of a group of conformers for pairwise minimum root mean square deviation (RMSD) relative to a predetermined conformer. Although there is no clear evidence for the possible dimerization arising from the N-terminal cysteine in the NMR signal assignment and through space NOE connectivities, but the possibility of dimmer formation cannot be exempted.

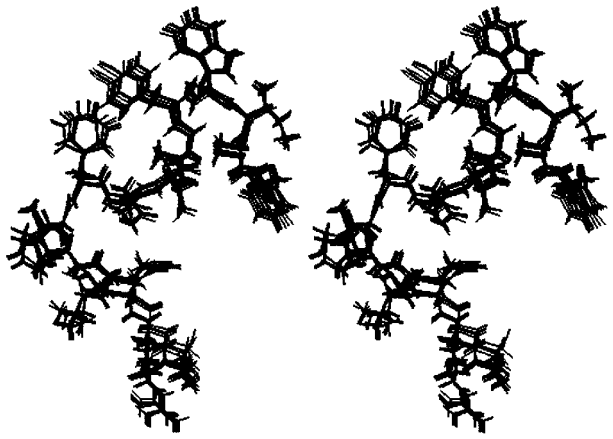

Figure 4. Stereoviews of pB1 showing best-fit superpositions of all atoms (excluding protons).

Recently, monoclonal antibodies have been used to prove the structure of apo B. In most apolipoproteins lipid binding occurs through amphipathic α-helical segment. In apo B-100, few of the predicted α-helices are truly amphipathic in terms of charge distribution on the polar surface. Simply, as Compared to first arrangement of amino acids using BLAST search, the homology of 40% was identified between apo B-100 and pB1. Generally, the conformation of apo B-100 in apo B-100-containing particles is crucial to its recognition by LDL receptor. Extensive studies have been carried out to determine the conformation of apo B-100 in LDL using monoclonal antibodies.

References

1. Cruzado, I. D.; Cockrill, S. L.; McNeal, C. J.; Macfarlane, R. D. *J. Lipid Res.* **1998**, *39*, 205.
2. Mahley, R. W.; Innerarity, T. L.; Jr. Rall, S. C.; Wesgraber, K. H. *J. Lipid Res.* **1984**, *25*, 1277.
3. Davis, R. A. *Biochemistry of Lipids, Lipoproteins and Membranes*. Elsevier Science Publishers: Canada, 1991.
4. Browman, B. H. *Hepatic Plasma Proteins Mechanisms of Function and Regulation*. Academic press, Inc; London, 1993.
5. Schumaker, V. N.; Phillips, M. L.; Chatterton, J. E. *Adv. Protein*

Chem. **1994**, *45*, 205.
6. Chen, S. H.; Yang, C. Y.; Chen, P. F.; Setzer, D.; Tamimura, M.; Li, W. H.; Gotto, A. M.; Chan, L. *J. Biol. Chem.* **1986**, *261*, 12918.
7. Segrest, J. P.; Jackson, R. L.; Morrisett, J. D.; Gotto, A. M. *FEBS Lett.* **1974**, *38*, 347.
8. Kaiser, E. T.; Kezdy, F. J. *Science* **1984**, *223*, 249.
9. Knott, T. J.; Pease, R. J.; Powell, L. M.; Wallis, S. C.; Jr. Rall, S. C.; Innerarity, R. L.; Blackhart, B.; Taylor, W. H.; Marcel, Y.; Milne, Y.; Johnson, D.; Fuller, M.; Lusis, A. J.; McCarthy, B. J.; Mahley, R. W.; Levy-Wilson, B.; Scott, J. *Nature* **1986**, *323*, 734.
10. Phillips, M. L.; Schumaker, V. N. *J. Lipid Res.* **1989**, *30*, 415.
11. Cardin, A. D.; Jackson, R. L. *Biochem. Biophys. Acta* **1986**, *877*, 366.
12. Chen, G. C.; Zhu, S.; Hardman, D. A.; Schilling, J. W.; Lau, K.; Kane, J. P. *J. Biol. Chem.* **1989**, *264*, 14369.
13. Davis, R. A.; Thrift, R. N.; Wu, C. C.; Howell, K. E. *J. Biol. Chem.* **1990**, *265*, 10005.
14. Chen, G. C.; Hardman, D. A.; Hamilton, R. L.; Mendel, C. M.; Schilling, J. W.; Thu, S.; Lau, K.; Wong, J. S.; Kane, J. P. *Biochemistry* **1989**, *28*, 2477.
15. Ettelaie, C.; Haris, P. I.; James, N. J.; Wilbourn, B.; Adam, J. M.; Bruckdorfer, K. R. *Biochimica et Biophysica Acta* **1997**, *1345*, 237.
16. Yang, C. Y.; Chen, S. H.; Gianturco, S. H.; Bradley, W. A.; Sparrow, J. T.; Tanimura, M.; Li, W. H.; Sparrow, D. A.; Deloof, S.; Rosseneu, M.; Lee, F. S.; Gu, Z. W.; Jr. Gotto, A. M.; Chan, L. *Nature* **1986**, *323*, 738.
17. Jamila, H.; Chua, C. H.; Dickson, J. K.; Chenb, Y.; Yana, M.; Billerb, S. A.; Gregga, R. C.; Wetteraua, J. R.; Gordana, D. A. *J. Lipid Res.* **1998**, *39*, 1448.
18. Johnson, B. A.; Blevins, R. A. *J. Biomol. NMR* **1994**, *4*, 603.
19. Wüthrich, K. *NMR of Proteins and Nucleic Acids*; Wiley: New York, 1986.
20. Evans, J. N. S. *Biomolecular NMR Spectroscopy*; Oxford Univ. Press: 1995.
21. South, T. L.; Blake, P. R.; Hare, D. R.; Summers, M. F. *Biochemistry* **1991**, *30*, 6342.
22. Kim, D.; Rho, J.; Won, H. *Journal of the Korean Magnetic Resonance Society* **1999**, *3*, 44.
23. Kim, D.; Lee, H.; Won, Y.; Kim, D.; Lee, Y.; Won, H. *Bull. Korean Chem. Soc.* **2003**, *24*, 967.
24. Kim, D.; Kim, D.; Lee, Y.; Won, H. *Bull. Korean Chem. Soc.* **2003**, *24*, 971.