http://dx.doi.org/10.5012/bkcs.2011.32.9.3425

# Solution State Structure of pA1, the Mimotopic Peptide of Apolipoprotein A-I, by NMR Spectroscopy

**Hyojoon Kim[†] and Hoshik Won[*]**

*Department of Applied Chemistry, Hanyang University, Ansan 426-791, Korea. [*]E-mail: hswon@hanyang.ac.kr*
*[†]Department of Biochemistry and Molecular Biology, Hanyang University, Ansan 426-791, Korea*
*Received June 9, 2011, Accepted July 31, 2011*

Apolipoprotein A-I (Apo A-I) is a major component for high density lipoproteins (HDL). A number of mimetic peptides of Apo A-I were screened from the phase-displayed random peptide library by utilizing monoclonal antibodies (A12). Mimetic peptide for A12 epitope against Apo A-I was selected as CPFARLPVEHHDVVGL (pA1). From the BLAST search, the mimetic peptide pA1 had 40% homology with Apo A-I. 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 0.3-0.7 Å$^2$ and the total RMSD was 0.6-1.6 Å. The mimotope pA1 exhibited characteristic conformation including a β-turn from Pro[7] to His[11].

**Key Words :** Apolipoprotein A-I, NMR spectroscopy, Molecular dynamic computation

## Introduction

Plasma lipoproteins are water-soluble particles composed of lipids and one or more specific proteins called apolipoproteins.[1] Apolipoproteins playing an important role in lipid transport and metabolic process in blood stream are components of LDL and HDL. Apolipoproteins are classified into A-I, A-II, B, C-I, C-II, D, E, respectively, based on the ABC nomenclature. There are three main lipoprotein classes according to density: very low density lipoproteins (VLDL), low density lipoproteins (LDL, d = 1.019-1.063 g/mL), and high density lipoproteins (HDL, d = 1.063-1.21 g/mL).[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 analyzed 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 á-helices, and may contribute to lipid binding.

The structural modifications of apolipoprotein are known to be more important than that of blood serum and lipid-protein in the disease diagnosis of circulatory system. Therefore the change of apolipoprotein became an important estimation factor. As an example, the ratio apo A-I/apo B-100 is now often used as index causing the disease of coronary arteries.

HDL is a heterogeneous mixture of lipoprotein particles. The major protein component (~70%) of HDL is apo A-I carrying 243 amino acids synthesized in the small intestine and liver. The structure of truncated human apolipoprotein A-I, the major protein component of high density lipoprotein, has been determined at 4-Å resolution.[10] Although the function of the apo A-II (*ca.* 30% of apo A) is not clearly known, but some clinical results indicated that the apo A-I activates the lecithin cholesterol acyltransferase (LCAT) enzyme to remove the content of cholesterol in peripheral liver tissue. The level of apo A-I in HDL of cell plasma is responsible for the level of cholesterol. The results of clinical trials clearly have demonstrated that elevated high density lipoprotein levels strongly correlated with reduced risks of HDL to prevention of atherosclerosis is consonant with its role as the mediator of reverse transport of cholesterol for peripheral to the liver for catabolism.[11] Hydro-

dynamics, electron microscopy and small angle X-ray scattering methods showed that spherical HDL consists of an ~84 Å diameter, core of neutral lipids and phospholipids acyl chains, surrounded by ~12 Å thickness of phospholipid head groups and protein.[12] The C-terminal of apo A-I consists of vertically repeating units of amphiphilic helical structures, 11 or 2 × 11 amino acids in length, nestled between the phospholipid head groups exhibiting a strongly binding affinity to lipid, and this enhance the conversion of nascent discoidal HDL to spherical HDL particle by action of lecithin cholesterol acyltransferase.[13]

Conformational studies for mimetic peptide CPFARLPV-EHHDVVGL (pA1) recognized by the monoclonal antibodies providing the regulation of surface conformational change are designed based on the screening from the phase-displayed random peptide library by utilizing monoclonal antibodies (A12). For NMR based structure determination and complete NMR signal assignments of this mimetic peptide, $^{1}$H, $^{13}$C, DEPT and 2D NMR experiments were performed. On the basis of distance data from NOESY experiments, distance geometry (DG) and molecular dynamics (MD) were carried out to obtain the tertiary structure of pA1 of the mimotope apolipoprotein A-I.

## Experimental Section

**Preparation of Sample.** Mimetic peptide CPFARLPV-EHHDVVGL (pA1)e was obtained from Bio-Synthesis, Inc. Peptide was synthesized using solid-phase method (FMOC chemistry). Because pA1 peptide is designed to interact with high density lipoprotein, the mimotopic pA1 peptide was water insoluble. The 3.0 mM concentration of peptide was prepared in 350 μL DMSO-$d_6$ carrying relatively high dielectric constant close to that of water.

**NMR Measurements.** All NMR experiments were performed by using the Varian Mercury 300 MHz and Unity INOVA 500 MHz NMR spectrometer at 25 °C. 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. Although NOESY experiments were recorded at three different 200 ms, 300 ms and 400 ms mixing time, the NOE spectrum at 300 ms mixing time was used for signal assignment and 2D-NOE back-calculation. TOCSY spectrum was collected with a mixing time of 50 ms, MLEV-17 spin lock pulse sequence. All of NMR data were processed by Varian VNMRJ and analyzed by using NMRView.[14]

**Solution State Structure.** Structure determinations were carried out by using HYGEOM™, HYNMR™. 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.[15] Structures were calculated from the NMR data according to the standard HYGEOM™ 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).

General ways of distance geometry (DG) algorithm accepts the input of distance constraints from NOE measurements.[16] 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. After internal coordinates are embedded in space, the values of the distances within the bounds obtained by bound smoothing are guessed at random, and the atomic coordinates representing the best-fit to this guess are generated. The deviations of the coordinates from the distance bounds, as well as the stereospecific assignments, are minimized by lowering penalty values. 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.[17] The structure was calculated using the DG algorithm HYGEOM™, 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 0.6-1.6 Å for the entire backbone.[16] Back-calculation were assigned to GENNOE calculation in order to generate the theoretical NOEs. A consecutive serial files, obtained from GENNOE calculation, were incorporated into HYNMR™ to generate NOE back-calculation spectra which can be directly compared with experimental NOESY spectra.[18]

## Results and Discussion

Connectivities derived from through bond J-coupling and through space coupling are important in NMR signal assignment and solution state structure determination. The $^{1}$H-$^{1}$H connectivities that identify the different amino acid type are established via scalar spin-spin coupling, using COSY and TOCSY. Assignments using sequential NOEs can be obtained for proteins with natural isotope distribution in space. Relations between protons in sequentially neighboring amino acid residues i and i+1 are established by NOEs manifesting close approach among $d_{\alpha N}$, $d_{NN}$ and $d_{\beta N}$. Except for flexible terminal amino acid residues, complete NMR signal assignments listed in Table 1 were accomplished by utilizing homonuclear J-resolved 2D NMR experiments. Table 2 lists the correlating signals of adjacent residues on

**Table 1.** $^1$H-NMR signal assignment of mimotopic peptide pA1

| Residue | NH | αH | βH | γH | Others |
|---|---|---|---|---|---|
| Pro[2] | | 4.170 | 2.042 | 1.827 | δH:3.440 |
| Phe[3] | 7.954 | 4.393 | 2.934, 2.752 | | ring:7.187 |
| Ala[4] | 8.030 | 4.204 | 1.130 | | |
| Arg[5] | 7.850 | 4.206 | 1.593 | 1.396 | δH:3.006<br>εH:7.623 |
| Leu[6] | 7.921 | 4.450 | 1.541 | 1.343 | δH:0.793 |
| Pro[7] | | 4.341 | 1.961 | 1.786 | δH:3.562 |
| Val[8] | 7.757 | 4.011 | 1.855 | 0.746 | |
| Glu[9] | 7.938 | 4.197 | 1.636, 1.778 | 2.136 | |
| His[10] | 8.052 | 4.492 | 2.947, 2.903 | | 2H:8.713<br>4H:7.199 |
| His[11] | 8.237 | 4.519 | 2.985, 2.925 | | 2H:8.738<br>4H:7.237 |
| Asp[12] | 8.547 | 4.552 | 2.619, 2.472 | | |
| Val[13] | 7.901 | 4.164 | 1.914 | 0.744 | |
| Val[14] | 7.723 | 4.153 | 1.904 | 0.757 | |
| Gly[15] | 7.980 | 3.924, 3.728 | | | |

**Table 2.** Important NOE connectivities used for the structure determination of pA1

| Residue | Chemical shift (ppm) | NOE connectivities |
|---|---|---|
| $Pro[2]_{\alpha}$ | 4.170 | $P_{2\beta}$(m), $P_{2\gamma}$(m), $F_{3NH}$(m) |
| $Pro[2]_{\gamma}$ | 1.827 | $P_{2\beta}$(m), $P_{2\delta}$(m) |
| $Pro[2]_{\delta}$ | 3.440 | $L_{6\beta}$(w), $L_{6\gamma}$(m) |
| $Phe[3]_{\alpha}$ | 4.393 | $F_{3NH}$(m), $F_{3\beta}$(m), $A_{4NH}$(m) |
| $Phe[3]_{\beta}$ | 2.752, 2.934 | $F_{3NH}$(m) |
| $Ala[4]_{NH}$ | 8.030 | $F_{3\beta}$(w), $A_{4\beta}$(s), $R_{5\gamma}$(w) |
| $Ala[4]_{\alpha}$ | 4.204 | $A_{4\beta}$(s), $A_{4NH}$(s), $R_{5NH}$(m) |
| $Arg[5]_{NH}$ | 7.850 | $R_{5\beta}$(m), $R_{5\gamma}$(m), $A_{4\beta}$(m) |
| $Arg[5]_{\alpha}$ | 4.206 | $R_{5NH}$(s), $R_{5\beta}$(s), $R_{5\gamma}$(s), $L_{6NH}$(s) |
| $Arg[5]_{\beta}$ | 1.593 | $R_{5\gamma}$(s), $R_{5\delta}$(w) |
| $Arg[5]_{\gamma}$ | 1.396 | $R_{5\delta}$(s), $R_{5\varepsilon}$(w) |
| $Leu[6]_{NH}$ | 7.921 | $L_{6\beta}$(w), $L_{6\gamma}$(m) |
| $Leu[6]_{\alpha}$ | 4.450 | $L_{6\beta}$(w), $L_{6\gamma}$(m), $P_{7\delta}$(m), $P_{2\delta}$(m), $L_{6NH}$(m) |
| $Pro[7]_{\alpha}$ | 4.341 | $P_{7\beta}$(s), $P_{7\gamma}$(s), $V_{8NH}$(s), $E_{9NH}$(m) |
| $Pro[7]_{\delta}$ | 3.562 | $P_{7\beta}$(w), $P_{7\delta}$(m), $L_{6\gamma}$(m), $L_{6\beta}$(w) |
| $Val[8]_{NH}$ | 7.757 | $P_{7\gamma}$(m), $V_{8\beta}$(m), $V_{8\gamma}$(m) |
| $Val[8]_{\alpha}$ | 4.011 | $E_{9NH}$(s), $V_{8NH}$(s), $V_{8\beta}$(s), $V_{8\gamma}$(m), $E_{9\beta}$(w) |
| $Glu[9]_{NH}$ | 7.938 | $H_{10NH}$(w), $E_{9\beta}$(w), $E_{9\gamma}$(m), $H_{11NH}$(vw) |
| $Glu[9]_{\alpha}$ | 4.197 | $H_{10NH}$(s), $E_{9NH}$(s), $E_{9\beta}$(m), $E_{9\gamma}$(m) |
| $Glu[9]_{\beta}$ | 1.636, 1.778 | $E_{9\gamma}$(s) |
| $His[10]_{NH}$ | 8.052 | $H_{11NH}$(w), $V_{8\alpha}$(w), $H_{10\beta}$(m), $E_{9\beta}$(w), $E_{9\gamma}$(w) |
| $His[10]_{\alpha}$ | 4.492 | $H_{10NH}$(s), $H_{11NH}$(s), $H_{10\beta}$(w) |
| $His[11]_{NH}$ | 8.237 | $D_{12NH}$(w), $E_{9\alpha}$(w), $H_{11\beta}$(m), $A_{4\beta}$(w) |
| $His[11]_{\alpha}$ | 4.519 | $H_{11NH}$(s), $D_{12NH}$(s), $H_{11\beta}$(m) |
| $Asp[12]_{NH}$ | 8.547 | $V_{13NH}$(w), $H_{11\beta}$(w), $D_{12\beta}$(m) |
| $Asp[12]_{\alpha}$ | 4.552 | $D_{12NH}$(s), $V_{13NH}$(s), $D_{12\beta}$(s), $V_{14NH}$(w) |
| $Val[13]_{NH}$ | 7.901 | $V_{14NH}$(m), $D_{12\beta}$(w), $V_{13\beta}$(m), $V_{13\gamma}$(m) |
| $Val[13]_{\alpha}$ | 4.164 | $V_{13NH}$(s), $V_{14NH}$(s), $V_{13\beta}$(m), $V_{13\gamma}$(m) |
| $Val[14]_{NH}$ | 7.723 | $V_{14\beta}$(m), $V_{14\gamma}$(m), $G_{15NH}$(w) |
| $Val[14]_{\alpha}$ | 4.153 | $V_{14NH}$(s), $V_{14\beta}$(m), $V_{14\gamma}$(m), $G_{15NH}$(m) |
| $Gly[15]_{NH}$ | 7.980 | $G_{15\alpha}$(s), $V_{14\beta}$(w) |

*Important NOEs are classified as strong(s), medium(m), weak(w), respectively, for NOE restraints in structure determination.

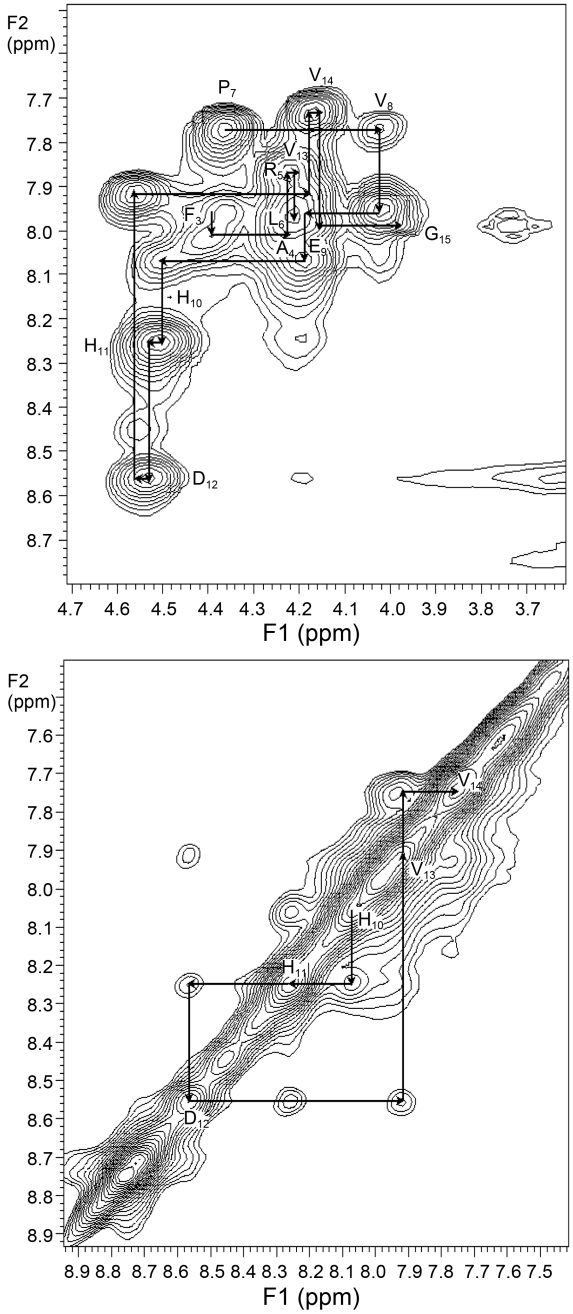


**Figure 1.** Intraresidue cross peaks labled in the NH-$C_{\alpha}$H region (upper), NH-NH region (lower) of the NOESY spectra of pA1 ($\tau_m$ = 300 ms).

the basis of dipolar connectivities obtained from 2D NOE spectra. Dipolar connectivities from amide protons to α- and amide protons were also used for sequential signal assignments, and the fingerprint region of the NOESY spectra recorded at 300 ms are shown in Figure 1. Although relatively weak $d_{\alpha N}$ (i, i+2) dipolar connectivities from residue number 7 to 11 were observed as shown in Figure 2, it was apparent that the mimotopic peptide pA1 has a characteristic conformation including a β-turn from Pro[7] to His[11].

Although the first residue Cys[1] and the last residue Leu[16] were not clearly showing inter-residual NOE connectivities, but important NOEs exhibiting peptide skeleton

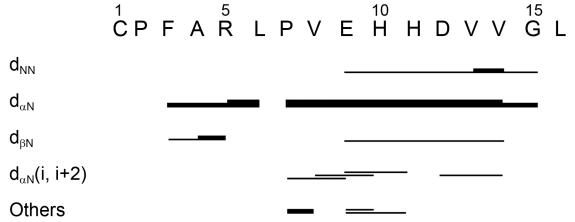


**Figure 2.** NOESY connectivity table of pA1: NOESY connectivities involving backbone protons for amino acids i and j. The height of the bars symbolizes the relative strength (strong, medium, weak) of the cross peaks in a qualitative way.

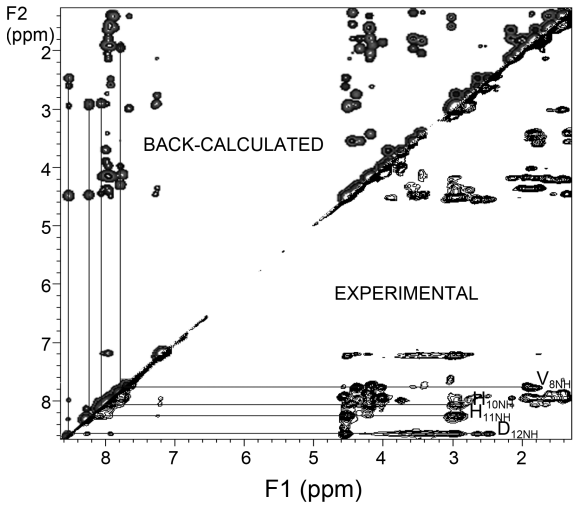


**Figure 3.** Comparisons of back-calculation and experimental NOESY spectrum of pA1 recorded at 300 ms mixing time are made.

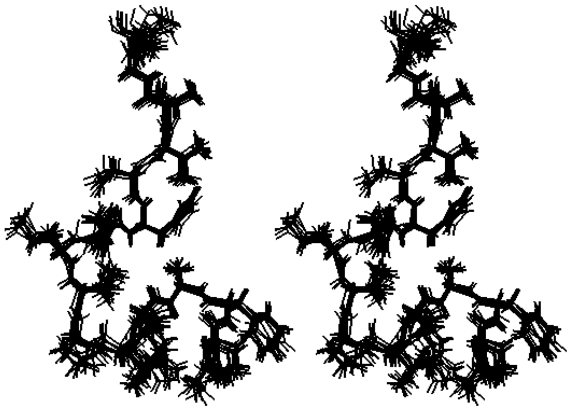

**Figure 4.** Stereoviews of pA1 showing best-fit superpositions of all atoms (excluding protons).

and specific conformation of mimetic peptide pA1 are observed to be Ala[4]$_{\beta}$-His[11]$_{NH}$, Pro[7]$_{\alpha}$-Glu[9]$_{NH}$, Val[8]$_{\alpha}$-His[10]$_{NH}$, Glu[9]$_{NH}$-His[11]$_{NH}$, Asp[12]$_{\alpha}$-Val[14]$_{NH}$. In α-helical peptide conformation, proline enhances to form turn structure. Out of mimotope pA1 CPFARLPVEHHDVVGL residues Ala[4]$_{\beta}$-His[11]$_{NH}$ demonstrates such a general tendency. By utilizing distance geometry and molecular dynamic computations incorporated with 2D NOE back-calculation enable to confirm β-turn conformation.

In order to determine the DG structure, several variable velocities 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 pA1 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 0.6-1.6 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.

## 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.; Rall, S. C., Jr.; 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.; Rall, S. C., Jr.; 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. Borhani, D. W.; Rogers, D. P.; Engler, J. A.; Brouillette, C. G. *Proc. Natl. Acad. Sci. U. S. A.* **1997**, *94*, 12291.
11. Castelli, W. P.; Garrison, R. J.; Wilson, P. W. F.; Abbott, R. D.; Kalousdian, S.; Kannel, W. B. *J. Am. Med. Assoc.* **1986**, *256*, 2835.
12. Atkinson, D.; Davis, M. A. F.; Leslie, R. B. *Proc. R. Soc. London* **1974**, *186*, 165.
13. Segrest, J. P.; Jones, M. K.; De Loof, H.; Brouillette, C. G.; Venkatachalapathi, Y. V.; Anantharamaiah, G. M. *J. Lipid Res.* **1992**, *33*, 141.
14. Johnson, B. A.; Blevins, R. A. *J. Biomol. NMR* **1994**, *4*, 603.
15. Wüthrich, K. *NMR of Proteins and Nucleic Acids*; Wiley: New York, 1986.
16. Evans, J. N. S. *Biomolecular NMR Spectroscopy*, Oxford Univ. Press, 1995.
17. South, T. L.; Blake, P. R.; Hare, D. R.; Summers, M. F. *Biochemistry* **1991**, *30*, 6342.
18. Kim, D.; Rho, J.; Won, H. *J. Korean Mag. Reson. Soc.* **1999**, *3*, 44.