

Research article

Open Access

## The N-terminal domain of apolipoprotein B-100: structural characterization by homology modeling

Hassan Al-Ali<sup>1</sup> and Hassan M Khachfe\*<sup>1,2</sup>

Address: <sup>1</sup>Computational Sciences and Bioinformatics Laboratory, American University of Beirut, Beirut, Lebanon and <sup>2</sup>Department of Physiology, Faculty of Medicine and Medical Center, American University of Beirut, Beirut, Lebanon

Email: Hassan Al-Ali - bioinfo@aub.edu.lb; Hassan M Khachfe\* - hassan.khachfe@balamand.edu.lb

\* Corresponding author

Published: 22 July 2007

Received: 9 October 2006

BMC Biochemistry 2007, 8:12 doi:10.1186/1471-2091-8-12

Accepted: 22 July 2007

This article is available from: <http://www.biomedcentral.com/1471-2091/8/12>

© 2007 Al-Ali and Khachfe; licensee BioMed Central Ltd.

This is an Open Access article distributed under the terms of the Creative Commons Attribution License (<http://creativecommons.org/licenses/by/2.0>), which permits unrestricted use, distribution, and reproduction in any medium, provided the original work is properly cited.

### Abstract

**Background:** Apolipoprotein B-100 (apo B-100) stands as one of the largest proteins in humans. Its large size of 4536 amino acids hampers the production of X-ray diffraction quality crystals and hinders in-solution NMR analysis, and thus necessitates a domain-based approach for the structural characterization of the multi-domain full-length apo B.

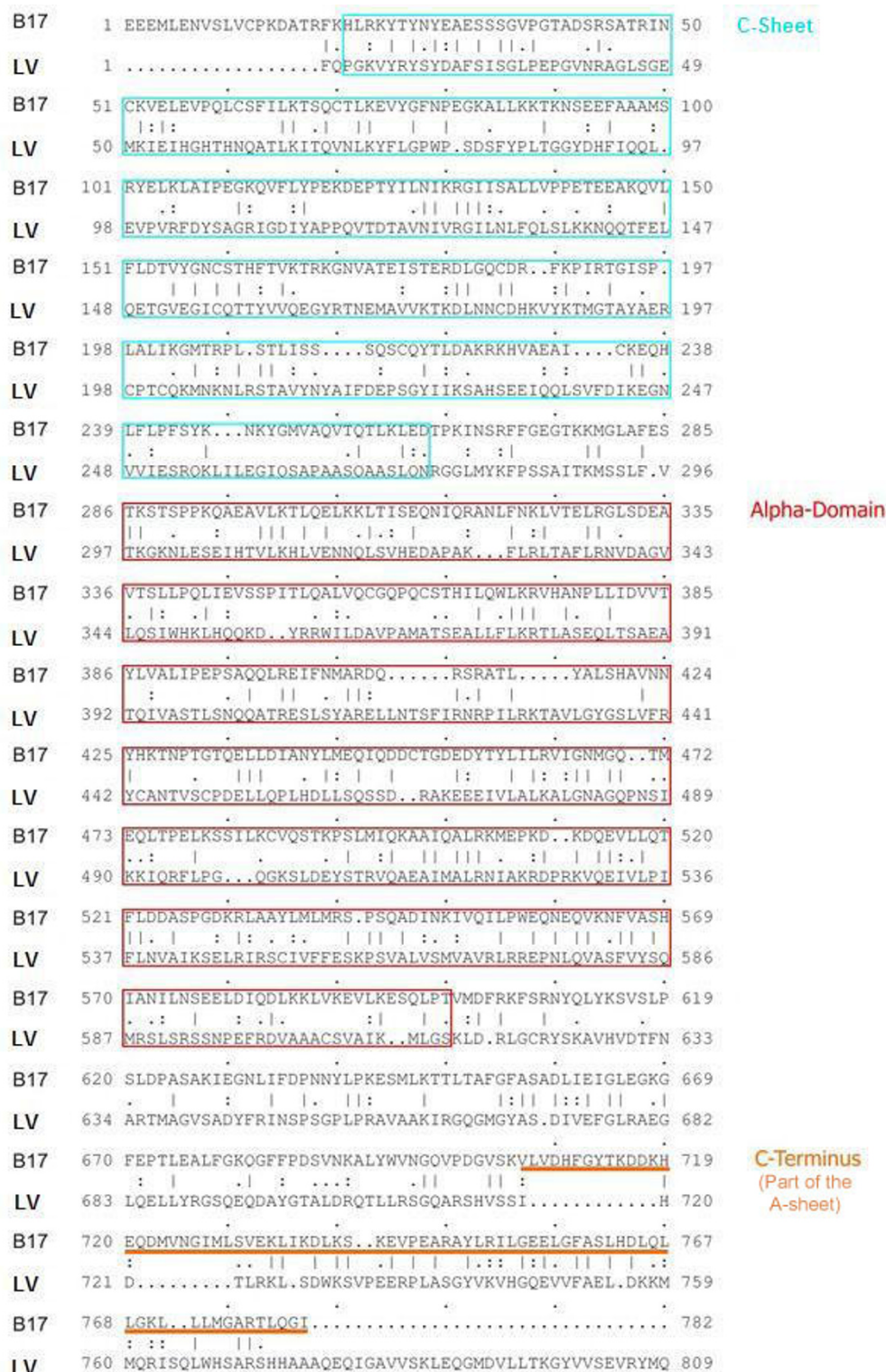
**Results:** The structure of apo B-17 (the N-terminal 17% of apolipoprotein B-100) was predicted by homology modeling based on the structure of the N-terminal domain of lipovitellin (LV), a protein that shares not only sequence similarity with B17, but also a functional aspect of lipid binding and transport. The model structure was first induced to accommodate the six disulfide bonds found in that region, and then optimized using simulated annealing.

**Conclusion:** The content of secondary structural elements in this model structure correlates well with the reported data from other biophysical probes. The overall topology of the model conforms with the structural outline corresponding to the apo B-17 domain as seen in the EM representation of the complete LDL structure.

### Background

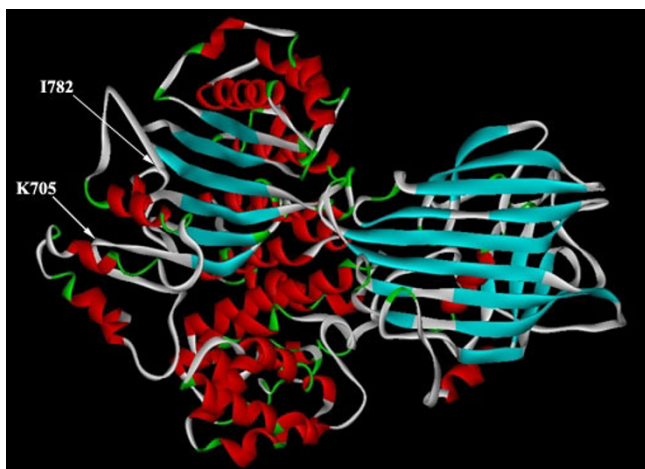
Atherosclerosis is a complex disease that has been linked to many risk factors, including hyperlipidemia, dyslipidemia, high blood pressure, and endothelial dysfunction [1]. Oxidative modification to the small low-density lipoprotein (LDL) has been dubbed the central event that initiates and propagates coronary artery diseases [2,3], and therefore, LDL is considered a major risk factor for atherosclerosis [4]. It was also shown that systemic inflammatory mechanisms may underlie the pathogenesis of atherosclerosis [5-7]. However, the specific structural interactions implicated in these mechanisms have not yet been elucidated.

Apolipoprotein B-100 (apo B) is the sole protein component of LDL [8]; however, its large size (4536 a.a.) and the limitation of current experimental techniques require that the structures of its multiple domains be analyzed separately [9,10]. Biochemical [10], calorimetric [11], computational [12-15], and spectroscopic [16] approaches were used to probe the domain arrangement and characterization of the protein, but no molecular structure has ever been assigned to any of the different domains. These techniques, however, helped in the understanding of the overall arrangement of apo B on the LDL particle and the interactions that the various secondary structures have with both the lipid and aqueous phases, and in the ability



**Figure 1**  
**The sequence alignment of B17 with the N-terminus of LV.** The figure also shows the different structural domains (Cyan boxes are for the  $\beta$ C-sheet and the magenta boxes are for the  $\alpha$ -domain. The underlined stretch indicates the region of no electron density in the template structure.





**Figure 4**  
**The structure of B17 modeled after the N-terminus of LV.**  $\beta$ -strands are colored blue, whereas  $\alpha$ -helices are colored red. Arrows and labels indicate the beginning and end of the 78-aa stretch modeled independent of the LV structure.

One disulfide pair came within binding distance after a series of minimization runs, and two pairs approached binding distance through directed (constrained) minimization, adding up to 6 disulfide bridges. However, those that were subject to constrained minimization were located within flexible loops at the surface of the protein, and thus did not cause the overall fold to change. Minimization was done in a step-wise fashion in order to explore bonding space between the sulfur groups without distorting secondary formations. Finally, a molecular dynamics simulation at 25–27 degrees Celsius was performed on the B17 molecule to allow its side chains to explore allowed conformational space.

A 78-residue stretch in the A-sheet of LV has no resolving electron density, and therefore, no coordinate assignments in the crystal structure. The correspondingly aligned amino acids in B17 (residues 706 – 782) had to be modeled separately. The secondary structures of this stretch were predicted using a variety of algorithms,

including the Chou-Fasman algorithm [26], the PROF methods of PredictProtein [27,28], and the SPDBV modality of Deep View. All of these modalities suggested an all-helical structure of the stretch, with a helical content around 65 % (Figure 3) and a reliability index approaching 90%. Several rounds of energy minimization and simulation – first in vacuum and later in water as a solvent – were performed allowing the previously-unstructured region to adopt a stable fold while its ends were fixed in space at coordinates corresponding to the crystal structure amino acids immediately preceding and succeeding the beginning and end residues in the primary sequence, respectively. Then, using the LIGATE modality in HOMOLGY, the structure of this part was pinned to the corresponding extremities in the LV-modeled B17, and the energy of the whole molecule was minimized again (Figure 4).

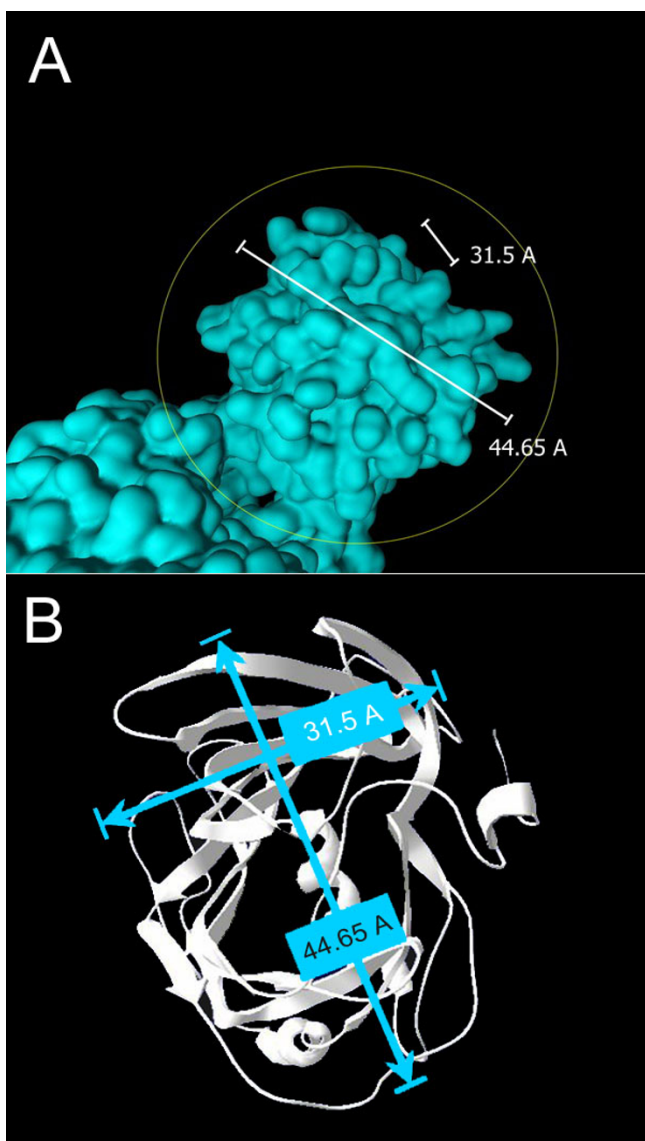
It should be noted here that a well-ordered structure is to be expected in this stretch owing to the fact that earlier biophysical studies suggested the presence of secondary structural elements that cannot be accounted for by what is reported in the crystal structure of LV only [19,20]. The secondary structural content in this complete model correlates excellently with the data reported previously using those biophysical probes. The structure also confirms the exposure of several coil-bound histidine residues that may be implicated in some helical rearrangement upon their protonation due to a slight decrease in the solvent pH [19,20]. The accessibility of these residues to the aqueous solvent was tested (Table 1), and their protonation upon the decrease in pH was confirmed.

The structure of LV has been reported to contain a completely buried salt bridge formed between R547 and E574 [21], which ties together the two "helical sheets" in the  $\alpha$ -domain, thereby increasing the stability of the local fold. A careful inspection of the B17 model structure revealed that a very similar salt bridge is formed between K530 and E557, which align – sequentially – with the above-mentioned residues in LV. Moreover, the solvent accessibility analysis illustrates that the involved side chains are well shielded from the aqueous medium and can therefore account for an extra stability in the  $\alpha$ -domain of B17 that has been previously reported [19,20].

**Table 1: Solvent accessibility for the buried salt bridge.**

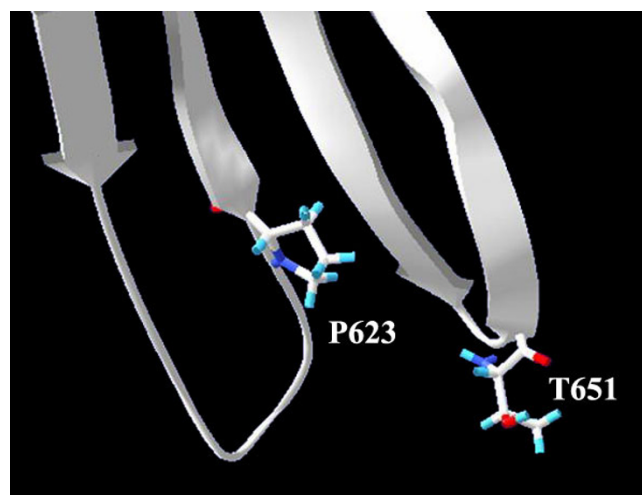
Residue	Atom	Solvent Accessibility Surface Area ( $\text{\AA}^2$ )	
		Before Solvation	After Solvation
K530	NZ	75.20	0.89
E557	OE2	1.54	0.00

Solvent accessibility surface areas for individual atoms calculated according to the hydration states of the corresponding atoms



**Figure 5**  
**The spatial dimensions of the  $\beta$ -domain of B17.** (A) Side view and (B) top view of the N-terminus of B17 with the corresponding dimensional values. These values correlate perfectly with the "knob" structure seen on the LDL particle and mapped to the N-terminus of B100.

Electron microscopy studies of intact LDL particles [29,30] showed that the N-terminus of apo B has a knob-shaped electron density with dimensions 30 – 45 Å. These dimensions approximate perfectly with the  $\beta$ -domain in the B17 model (Figure 5). These dimensions, along with the positions of the disulfide bonds and the buried, conserved salt bridge in the helical region, give credibility to the model. The lipid pocket surface accessibility – for potential lipid recruitment – towards the inside of the  $\alpha$ -domain also makes the structure trustworthy.

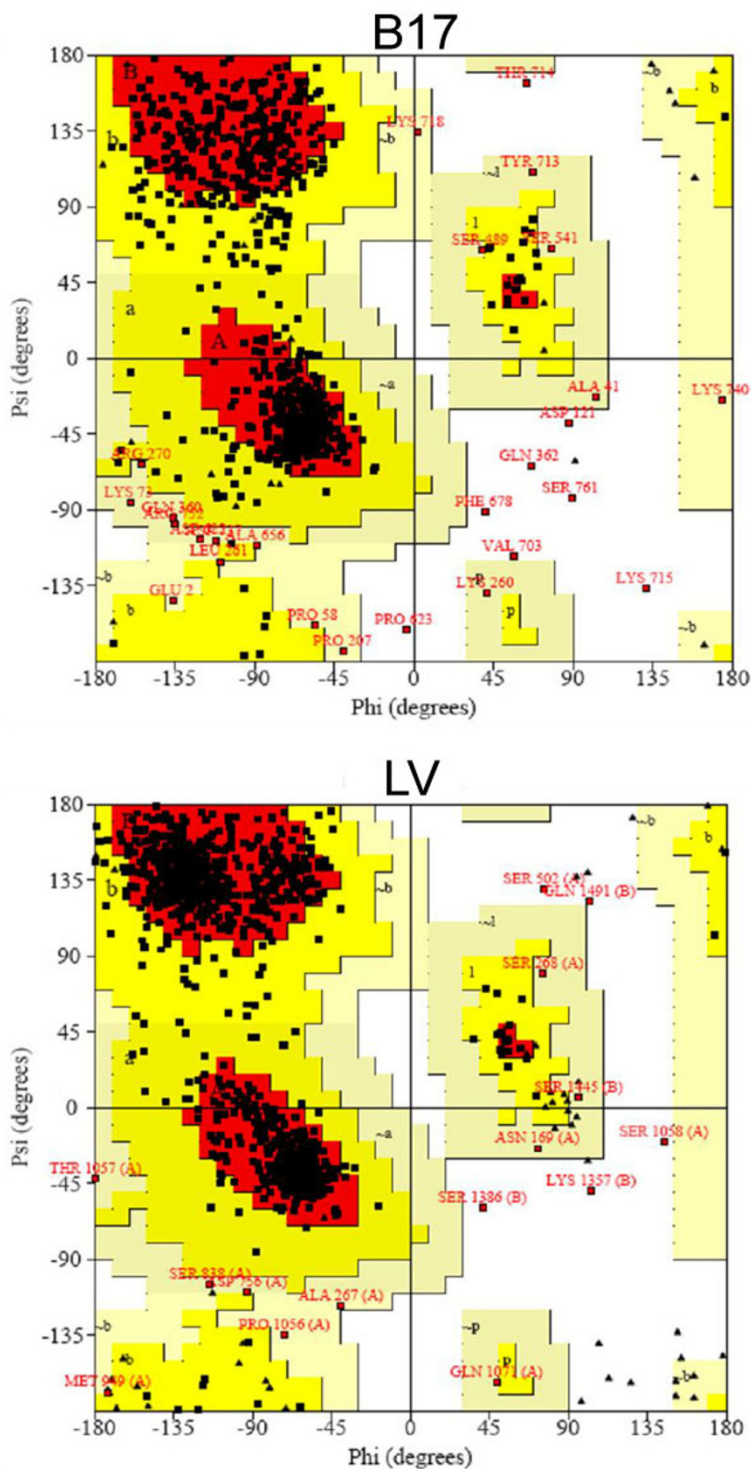


**Figure 6**  
**Fold versus angle stability:** P623 and T651, with Z-scores around -3.0, fall in turns involved in a mini sheet between the helical region and the C-terminus. The elevated torsion energy is compensated by an overall fold energy reduction.

A comprehensive structure validation test was carried out to check the physical elements of the model. Bond angles were found to deviate normally from the reported mean standard values [31]. Moreover, the RMS Z-score for bond angles in this model structure is within 9 % change with respect to that in the template structure. Bond lengths were found to have normal variability. The contact distances of all atom pairs have been checked. Among the 31 reported abnormally short interatomic distances in B17 (more than 200 in the corresponding LV template), 23 are either representations of hydrogen bonds or predictions of atoms with B-factors higher than 80, indicating that the atoms potentially implicated in these bumps are not there anyway. The evaluation of the model torsion angles did indeed show some unusual residues; however, the two amino acids, P623 and T651, with Z-scores around -3.0 (the worrying limit), actually fall in the region joining the  $\alpha$ -domain with the C-terminus of the protein. P623 is at the end of a  $\beta$ -strand and T651 is the second of a two-residue turn between two strands as well, all three of these strands are involved in a mini sheet between the helical region and the C-terminus (Figure 6), and, therefore, the slight increase in their torsional energy is compensated by the overall fold stability. Finally, the Ramachandran plot of the backbone psi-phi angles of the B17 model showed comparative results to those obtained from the crystal structure of LV (Figure 7).

### Conclusion

This model provides further insight into the structural basis for the functional attributes of B-17, and constitutes a step towards the full elucidation of the multi-domain



**Figure 7**  
**The Ramachandran plots of B17 and LV.** More than 97% of the backbone angle-pairs in B17 fall within the favorable regions, whereas 2.3% fall within the additionally allowed regions, and less than 1% is in the disallowed regions. This is very similar to the data from the LV system (98% in the favorable regions, 1% in the additionally allowed regions, and 1% in the disallowed regions).

structure of full-length Apo B-100. While the current structure ensures the globular topology of the domain and its poor lipidation state, as it does not show lipid binding pockets, the biological implications of this protein – independent of its role in apo B-100 – remain to be tested *in vitro* and, later, *in vivo*, since B17 is not a naturally occurring plasma apolipoprotein. Knowing the importance of this domain in the secretion and assembly of the full-length apo B-100, we anticipate that the current structure and subsequent physiological experiments will assist in the development of novel drugs for the treatment of and protection against diseases correlated with elevated blood LDL.

## Methods

### Molecular modeling

#### Sequence alignment

Multiple sequence alignments were done using BLAST [32] and the alignment module of the Discovery Studio suite (Accelrys Inc., Discovery Studio 1.5, San Diego: Accelrys Inc., 2004)

#### Structure prediction

The structure of B17 (residues 1–704) was modeled using MODELLER [33] of HOMOLOGY in insight II (Accelrys Inc., Insight Modeling Environment, Release 2000.1, San Diego: Accelrys Inc., 2002), based on the crystal structure of lipovitellin (LV), an egg yolk protein that shares over 30% sequence homology (in over 700 amino acid overlap) with B17. The secondary structure of the unstructured region was predicted using the Chou-Fasman Algorithm [26], the PROF methods [27,28] and the Deep View modality [34]. The calculation was performed using the Accelrys SeqWeb server of the GCG Wisconsin Package.

### Simulation

#### Energy calculations

EC's were performed using DISCOVER (Accelrys Inc., CDISCOVER Molecular Simulator, Release 2000.1, San Diego: Accelrys Inc., 2002) and CHARMM (Version c28b) [35] modules in Insight II. Energy minimizations were performed using the Steepest Descent method followed by Conjugate Gradients.

#### Molecular dynamics

MD Simulations were carried out with periodic boundary conditions using a cubic box (of appropriate size), in the Insight II package. Solvent water molecules were represented by the three-site TIP3P water model [36], in the NVT ensemble.

#### Force fields

Calculations were performed using the DISCOVER force-fields CVFF and CFF91. The CHARMM force-field used in the solvation simulation was CHARMM27.

### Analysis

#### Solvent accessibility

Solvent Accessible Surface Area (SASA) was calculated for individual atoms using the Structural Biology at NIH server (Structools), with a probe radius of 1.4 Å [37].

#### Potential maps

Solvation energy and hydrophobic interactions were calculated using the Delphi module in Insight II (Accelrys Inc., Delphi Module, Release 2000.1, San Diego: Accelrys Inc., 2002), using the CFF91 force-field. Potential maps were constructed using a grid. The dielectric value was assigned as 4 for the protein and 80 for the solvent.

### Structure validation

Structure validation tests were carried out using the PROCHECK [38] and WHATIF [39] modalities. Calculated values were referenced to the reported mean standard values [31].

### Authors' contributions

HAA carried out all structural prediction and optimization exercises and participated in the analysis of the results. HMK conceived of the study, designed the experimental approach, coordinated the work, analyzed the results, and drafted the manuscript. All authors read and approved the final manuscript.

### Acknowledgements

The authors wish to thank Professor David Atkinson for his insightful comments and Professor Sawsan Khouri for her helpful remarks. This work has been supported in part by the American University of Beirut's Medical Practice Plan (AUB-MPP) and University Research Board (AUB-URB) funds.

### References

1. Soltero-Perez IF: **Thinking intelligently about therapy of atherosclerosis.** *American Journal of Therapeutics* 2003, **10(6)**:429-437.
2. Yla-Herttuala S, Palinski W, Rosenfeld ME, Parthasarathy S, Carew TE, Butler S, Witztum JL, Steinberg D: **Evidence for the presence of oxidatively modified low density lipoprotein in atherosclerotic lesions of rabbit and man.** *Journal of Clinical Investigation* 1989, **84(4)**:1086-1095.
3. Yla-Herttuala S, Palinski W, Rosenfeld ME, Steinberg D, Witztum JL: **Lipoproteins in normal and atherosclerotic aorta.** *European heart journal* 1990, **11(Suppl E)**:88-99.
4. Archbold RA, Timmis AD: **Modification of coronary artery disease progression by cholesterol-lowering therapy: the angiographic studies.** *Current opinion in lipidology* 1999, **10(6)**:527-534.
5. Bach-Ngohou K, Nazih H, Nazih-Sanderson F, Zair Y, Le Carrer D, Krempf M, Bard JM: **Negative and independent influence of apolipoprotein E on C-reactive protein (CRP) concentration in obese adults. Potential anti-inflammatory role of apoE in vivo.** *International Journal of Obesity & Related Metabolic Disorders: Journal of the International Association for the Study of Obesity* 2001, **25(12)**:1752-1758.
6. Hulthe J, Fagerberg B: **Circulating oxidized LDL is associated with increased levels of cell-adhesion molecules in clinically healthy 58-year old men (AIR study).** *Medical Science Monitor* 2002, **8(3)**:CR148-52.
7. Titov VN: **The functional role of arterial intima. Endogenous and exogenous pathogens and specificity of atheromatosis as an inflammation.** *Klinicheskaia Laboratornaia Diagnostika* 2003:23-24.

8. Mahley RW, Angelin B: **Type III hyperlipoproteinemia: recent insights into the genetic defect of familial dysbetalipoproteinemia.** *Advances in Internal Medicine* 1984, **29**:385-411.
9. Cladaras C, Hadzopoulou-Cladaras M, Nolte RT, Atkinson D, Zannis VI: **The complete sequence and structural analysis of human apolipoprotein B-100: relationship between apoB-100 and apoB-48 forms.** *EMBO Journal* 1986, **5**(13):3495-3507.
10. Yang CY, Gu ZW, Weng SA, Kim TW, Chen SH, Pownall HJ, Sharp PM, Liu SW, Li WH, Gotto AM Jr: **Structure of apolipoprotein B-100 of human low density lipoproteins.** *Arteriosclerosis* 1989, **9**(1):96-108.
11. Walsh MT, Atkinson D: **Calorimetric and spectroscopic investigation of the unfolding of human apolipoprotein B.** *Journal of lipid research* 1990, **31**(6):1051-1062.
12. Nolte RT: **Structural analysis of the human apolipoproteins: An integrated approach utilizing physical and computational methods.** In *PhD Dissertation Boston University, Department of Biophysics*; 1994.
13. Segrest JP, Garber DW, Brouillette CG, Harvey SC, Anantharamaiah GM: **The amphipathic alpha helix: a multifunctional structural motif in plasma apolipoproteins.** *Advances in Protein Chemistry* 1994, **45**:303-369.
14. Segrest JP, Jones MK, Mishra VK, Pierotti V, Young SH, Boren J, Innerarity TL, Dashti N: **Apolipoprotein B-100: conservation of lipid-associating amphipathic secondary structural motifs in nine species of vertebrates.** *Journal of lipid research* 1998, **39**(1):85-102.
15. Segrest JP, Jones MK, De Loof H, Dashti N: **Structure of apolipoprotein B-100 in low density lipoproteins.** *Journal of lipid research* 2001, **42**(9):1346-1367.
16. Walsh MT, Atkinson D: **Physical properties of apoprotein B mixed micelles with sodium deoxycholate and in a vesicle with dimyristoyl phosphatidylcholine.** *Journal of lipid research* 1986, **27**(3):316-325.
17. Herscovitz H, Hadzopoulou-Cladaras M, Walsh MT, Cladaras C, Zannis VI, Small DM: **Expression, secretion, and lipid-binding characterization of the N-terminal 17% of apolipoprotein B.** *Proceedings of the National Academy of Sciences of the United States of America* 1991, **88**(16):7313-7317. [erratum appears in Proc Natl Acad Sci U S A 1991 Oct 15;88(20):9375]
18. Herscovitz H, Kritis A, Talianidis I, Zanni E, Zannis V, Small DM: **Murine mammary-derived cells secrete the N-terminal 41% of human apolipoprotein B on high density lipoprotein-sized lipoproteins containing a triacylglycerol-rich core.** *Proceedings of the National Academy of Sciences of the United States of America* 1995, **92**(3):659-663.
19. Khachfe HM, Atkinson D: **Structural Analysis and Characterization of the 17% N-terminal Domain of Apolipoprotein B-100 Using CD Spectroscopy [abstract].** *Biophys J* 2001, **80**(1):62a.
20. Khachfe HM: **Spectroscopic and Calorimetric Studies of the 17% N-terminal Domain of Apolipoprotein B-100.** In *PhD Dissertation Boston University School of Medicine, Department of Physiology and Biophysics*; 2002.
21. Mann CJ, Anderson TA, Read J, Chester SA, Harrison GB, Kochl S, Ritchie PJ, Bradbury P, Hussain FS, Amey J, Vanloo B, Rosseneu M, Infante R, Hancock JM, Levitt DG, Banaszak LJ, Scott J, Shoulders CC: **The structure of vitellogenin provides a molecular model for the assembly and secretion of atherogenic lipoproteins.** *Journal of Molecular Biology* 1999, **285**(1):391-408.
22. Raag R, Appelt K, Xuong NH, Banaszak L: **Structure of the lamprey yolk lipid-protein complex lipovitellin-phosvitin at 2.8 Å resolution.** *Journal of Molecular Biology* 1988, **200**(3):553-569.
23. Segrest JP, Jones MK, Dashti N: **N-terminal domain of apolipoprotein B has structural homology to lipovitellin and microsomal triglyceride transfer protein: a "lipid pocket" model for self-assembly of apoB-containing lipoprotein particles.** *Journal of lipid research* 1999, **40**(8):1401-1416.
24. Shelness GS, Thornburg JT: **Role of intramolecular disulfide bond formation in the assembly and secretion of apolipoprotein B-100-containing lipoproteins.** *Journal of lipid research* 1996, **37**(2):408-419.
25. Yang CY, Kim TW, Weng SA, Lee BR, Yang ML, Gotto AM Jr: **Isolation and characterization of sulfhydryl and disulfide peptides of human apolipoprotein B-100.** *Proceedings of the National Academy of Sciences of the United States of America* 1990, **87**(14):5523-5527.
26. Chou PY, Fasman GD: **Prediction of protein conformation.** *Biochemistry* 1974, **13**(2):222-245.
27. Rost B, Sander C: **PROF.** *J Mol Biol* 1993, **232**:584-599.
28. Rost B, Fariselli P, Casadio R: **PROFhtm.** *Prot Science* 1996, **7**:1704-1718.
29. Orlova EV, Sherman MB, Chiu W, Mowri H, Smith LC, Gotto AM Jr: **Three-dimensional structure of low density lipoproteins by electron cryomicroscopy.** *Proceedings of the National Academy of Sciences of the United States of America* 1999, **96**(15):8420-8425.
30. Poulos GV: **The three dimensional structure of low density lipoprotein via cryoelectron microscopy.** In *PhD Dissertation Boston University, Department of Biophysics*; 2001.
31. Engh R, Huber R: **Accurate Bond and Angle Parameters for X-ray Protein Structure Refinement.** *Acta Crystallogr* 1991, **A47**:392-400.
32. Altschul SF, Lipman DJ: **Protein database searches for multiple alignments.** *Proc Natl Acad Sci USA* 1990, **87**(14):5509-13. 31 Sali A, Blundell TL: **Comparative protein modelling by satisfaction of spatial restraints.** *Journal of Molecular Biology* 1993, **234**(3):779-815.
33. Sali A, Potterton L, Yuan F, van Vlijmen H, Karplus M: **Evaluation of comparative protein modeling by MODELLER.** *Proteins* 1995, **23**(3):318-326.
34. **Swiss-PDB viewer** [<http://www.expasy.org/spdbv/>]
35. Brooks BR, Brucoleri RE, Olafson BD, States DJ, Swaminathan S, et al: **CHARMM:A program for Macromolecular Energy, Minimization, and Dynamics Calculations.** *J Comp Chem* 1983, **4**:187-217.
36. Jorgensen WL, Chandrasekhar J, Buckner JK, Madura JD: **Computer simulations of organic reactions in solution.** *Annals of the New York Academy of Sciences* 1986, **482**:198-209.
37. Gerstein M: **A Resolution-Sensitive Procedure for Comparing Protein Surfaces and its Application to the Comparison of Antigen-Combining Sites.** 1992 [<http://www.ncbi.nlm.nih.gov/structools.htm>].
38. Laskowski RA, MacArthur MW, Moss DS, Thornton JM: **PRO-CHECK: a program to check the stereochemical quality of protein structures.** *J Appl Cryst* 1993, **26**:283-291.
39. Vriend G: **WHAT IF: a molecular modelling and drug design program.** *J Mol Graph* 1990, **8**:52-56.

Publish with **BioMed Central** and every scientist can read your work free of charge

"BioMed Central will be the most significant development for disseminating the results of biomedical research in our lifetime."

Sir Paul Nurse, Cancer Research UK

Your research papers will be:

- available free of charge to the entire biomedical community
- peer reviewed and published immediately upon acceptance
- cited in PubMed and archived on PubMed Central
- yours — you keep the copyright

Submit your manuscript here:  
[http://www.biomedcentral.com/info/publishing\\_adv.asp](http://www.biomedcentral.com/info/publishing_adv.asp)

