Vol. 80, No. 1, 1978 BIOCHEMICAL AND BIOPHYSICAL RESEARCH COMMUNICATIONS

January 13, 1978 Pages 140-146

ON THE STRUCTURAL SIMILARITY OF APOA-I AND APOA-II OF HUMAN HIGH DENSITY LIPOPROTEINS

Ingming Jeng
Lipid Research Center
Department of Preventive Medicine
Washington University School of Medicine
St. Louis, Missouri 63110

Received November 21, 1977

SUMMARY: The possible evolutionary origin of apolipoproteins was studied by comparing the primary structures of different plasma apolipoproteins and other phospholipid-binding proteins. Apolipoprotein A-I (ApoA-I) and apolipoprotein A-II (ApoA-II) of human high density lipoprotein (HDL) are related. The resemblance of these two HDL apolipoproteins are apparently restricted to the carboxyl terminal regions suggesting that these portions of the molecules are derived from the same ancestor. The homologous carboxyl terminal segments may be involved in the regulation of HDL metabolism or in the interaction with phospholipids.

Plasma lipoproteins are transport vehicles for the otherwise insoluble lipids in the circulation. They can be classified according to their behavior in the ultracentrifuge or upon electrophoresis (1, 2). The former method enjoys greater popularity since it can be easily adapted for preparative purposes. This classification system divides plasma lipoproteins into four major classes: chylomicrons, very low density lipoproteins, low density lipoproteins, and high density lipoproteins. In addition to their function as lipid carriers, lipoproteins participate in other aspects of lipid metabolism. The structure and metabolism of lipoproteins have been the subject of extensive study with the hope of finding relevance to an understanding of lipid disorders in man. The recent observation that high plasma HDL levels correlate with a low incidence of coronary heart disease (3) has stimulated a great deal of interest in the structure and metabolism of HDL.

HDL, isolated between d 1.063 and 1.210, is the smallest lipoprotein

0006-291X/78/0801-0140$01.00/0

particle. HDL contains about equal amounts of lipid and protein by weight. ApoA-I and ApoA-II, two major protein components of HDL, have been purified, characterized, and sequenced. ApoA-I, the most abundant apoprotein of HDL, has a single polypeptide chain of 245 amino acid residues (4). Its binding to different phospholipids has been studied using the techniques of electron microscopy, circular dichroism, nuclear magnetic resonance, fluorescence spectroscopy, and radioimmunoassay (5-8). ApoA-II is smaller than ApoA-I. Human ApoA-II is a dimer of two identical polypeptide chains containing 77 amino acid residues. These two polypeptides are connected by an intrachain disulfide bond (9,10). This disulfide bond apparently is not essential for its function as a lipid carrier since reduction and alkylation of this disulfide bridge do not alter the protein's ability to reassociate with phospholipids (11). Furthermore, ApoA-II in the monkey is a monomer containing no cysteine (12).

The determinations of the primary structures of ApoA-I and ApoA-II have provided a necessary fundation in the study of protein-lipid interactions, HDL structure, and the possible evolutionary histories of these two proteins. The interrelationships between ApoA-I and ApoA-II were studied. It was found that they are similar at their carboxyl terminal regions. The resemblance is much stronger than one would expect by chance suggesting a common evolutionary origin of these carboxyl terminal regions. The resistance of these polypeptides to successful mutation might be used as evidence that they are involved in a common and essential function of HDL.

METHODS

The sequences of ApoA-I, ApoA-II, and ApoC peptides have been reported recently (13, 14). All other sequences of phospholipid-binding proteins, such as blood clotting factors, phospholipases, and neurotoxins, were taken from the Atlas of Protein Sequence and Structures (13). The method used for testing similarities and possible repeats in sequences is the diagonal method developed by Gibbs and McIntyre (15).

RESULTS AND DISCUSSION

The binding between apolipoproteins and phospholipids is the primary

interaction in the assembly of lipoproteins since cholesterol and triglycerides can associate with apolipoproteins only in the presence of phospholipids (16). Both ApoA-I and ApoA-II readily form complexes with phospholipids, hence, it may be expected that they are homologous to other phospholipid-binding proteins. The sequences of phospholipases, neurotoxins, prothrombins, and other blood clotting proteins are compared by both the visual inspection and diagonal diagram method (15). Careful analyses of these comparisons, however, revealed no significant homology. In addition to the comparison of whole proteins, an attempt was also made to align individual short segments in the hope that the regions binding to phospholipid might be similar. For example, the amino terminal region of prothrombin has been shown to interact with phospholipid. This type of comparison failed to detect any homology suggesting that their "active sites" for associating with phospholipids are not necessarily derived from the same ancestral gene.

In the study of primary structure relationships among all the apolipoproteins of known sequences, all ApoA proteins and ApoC peptides were again compared by the same techniques. It was of interest that although ApoA-I and ApoA-II are different in size, there was a significant homology between these two proteins at their carboxyl terminal regions (Fig. 1). There are eight identical residues in the last seventeen amino acid residues representing 47% identity. This is substantially higher than one might expect from random distribution. Therefore, it is highly unlikely that this similarity in sequences could have occurred by chance. Furthermore, if these two chains are related to each other, it may be expected that the non-identical amino acid residues in these regions could be derived from the change of a single nucleotide in the triplet coding for the amino acid. The nature of non-identical amino acid residues in these regions was, therefore, studied. It was found that seven out of eight pairs could be related to one another by a single

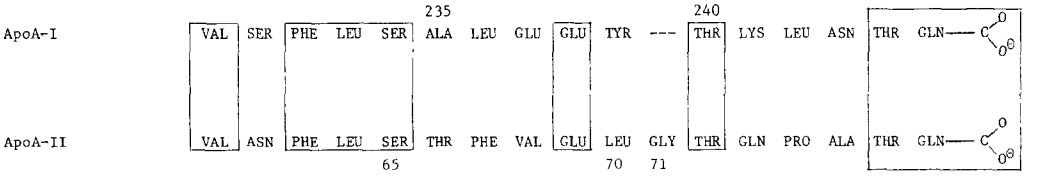

Fig. 1. Carboxyl Terminal Sequences of Human ApoA-I and A-II

base change in the nucleotides coding for the amino acid. The only exception is the pair of ASN 245 of ApoA-I and ALA 75 of ApoA-II. In this instance, they relate to each other by a two base change in the messengers. The relatedness of genomes for non-identical amino acid residues provides one more argument for the existence of structural similarity between these apolipoproteins in these regions.

The similarity between these two proteins is apparently restricted to this part of the molecules and does not extend over the remainder. It is possible that other regions of these two proteins are derived independently for their distinct functions. For example, ApoA-I is a known activator of lecithin cholesterol acyltransferase. It is likely that a part of ApoA-I, non-existent in ApoA-II, is required for this activity.

In order to maximize the homology, a gap has to be inserted in the ApoA-I sequence. This gap is assigned to a position corresponding to residue 71 of ApoA-II as shown in Fig. 1. It can, however, be introduced to residue numbers 65, 66, 67, 69 with the same number of identical residues. It should also be mentioned that the introduction of a gap probably does not alter the similarity of secondary structures of these two apolipoproteins in these regions (16).

Furthermore, three out of eight identical residues are hydroxyl amino acid residues, i.e. serine and threonine. The possible role of these two residues in apolipoproteins structure is interesting since all three ApoC peptides are also rich in these two hydroxyl amino acids, namely

ApoC-I, C-II, C-III have 18%, 22%, 20% of these two residues, respectively. Although serine and threonine sometimes are essential for the activity of proteins such as "serine proteases" and some glycoproteins, they can be successfully replaced by several other amino acids of similar properties. It has been reported that the mutability of serine is the highest among all 20 amino acids (13). With the exception of phenylalanine, the other seven identical residues in these regions are of high mutabilities. The conservation of amino acids with high mutabilities lends additional support to the contention that the resemblance is significant.

Finally, the position of the homologous region is intriguing. Since ApoA-I is much larger than ApoA-II in size, it may not be coincidental that the homology is at the carboxyl terminal regions. If the position of this region is not important, the homologous region in ApoA-I might be internal rather than external. The possibility exists that the function of these peptides is such that they must be at the carboxyl-termini to express their common functions. The fact that two of the identical residues, THR-GLN, are at the carboxyl termini is consistent with the idea that the position of the homology is critical.

The observation that these two apolipoproteins in HDL are partially homologous suggests that these carboxyl terminal regions of the proteins are the end products of a common ancestor. However, it is difficult to rule out the possibility that they are the result of convergent evolution due to their similar function. Regardless of their origin the significant homology in these terminal portions does suggest that they may be essential for a common and important function of these two HDL proteins. The structural relatedness has prompted us to propose that these regions may be important for either the integrity of HDL structure or the metabolism of these two HDL proteins. It should be emphasized, however, that these possibilities are not mutually exclusive.

One of the possibilities is that these regions are directly involved

in the lipid-protein interaction. An amphipathic helical region has been hypothesized to be the lipid-binding domain of the plasma apolipoproteins (17, 18). This model, derived mainly from model building studies, has not been conclusively proved. Also, it is unknown whether regions without amphipathic helix can bind to phospholipid or not. Since these two homologous terminal regions do not contain amphipathic helical regions (17), it would be interesting to test the abilities of these regions to associate with phospholipids.

Besides the binding to phospholipid, the possible involvement of these regions in the metabolism of HDL has to be considered. It is now clear that apolipoprotein levels in the circulation are under rigid control. We have found that the levels of these two apolipoproteins do not change significantly over a 24-hour period. Furthermore, the ratio of ApoA-I and ApoA-II remains constant in subfractions of HDL (19). This indicates that the metabolism of ApoA-I and ApoA-II are coordinatedly regulated and the possibility exists that their biosynthesis and/or degradation are controlled by common processes. If a common protein or enzyme indeed participates in their biosynthesis, secretion, or degradation, it is possible that these homologous segments of ApoA-I and ApoA-II are the recognition sites for a common regulatory protein. Of interest in this regard is the observation that the carboxyl terminal region of ApoA-I in HDL is more reactive immunologically than the amino terminal part (6) suggesting that the carboxyl terminal region is more exposed. This is consistent with the notion that the carboxyl terminal regions interact with other proteins or enzymes to determine the fate of HDL particles.

The amino acid sequences of these two HDL apolipoproteins have been known for the last few years, however, the relatedness of these two proteins has not previously been recognized. The documentation of this similarity opens up many experimental approaches to investigate the function of these homologous carboxyl terminal regions. The size and position of these

regions make them readily accessible to experimental manipulation. Derivatives with specific modifications in these regions may be compared with the native proteins in regard to lipid-binding ability and biological half life. Information of this type should prove useful in elucidating the biological function of these proteins.

ACKNOWLEDGEMENTS: I am grateful to Drs. Gustav Schonfeld, Ralph A. Bradshaw, Stuart Weidman, Michicael Sutton, and Robert E. Silverman for their helpful discussions. I acknowledge the assistance of Mrs. Lillie Beal in typing the manuscript. This work was supported by NIH Grant 5-R01-HL-15308-05 from National Institutes of Health.

REFERENCES

1. Lindgren, F. T., Jense, L. C., and Hatch, F. T. (1972) Blood Lipids and Lipoproteins: Quantitation, Composition, and Metabolism (Nelson, G. J., ed.), Wiley-Interscience, New York, p. 181-274.
2. Lees, R. S., and Hatch, F. T. (1963) J. Lab. Clin. Med. 61, 518-527.
3. Miller, G. J., and Miller, N. E. (1975) Lancet 1, 16-19.
4. Baker, N. H., Delahunty, T., Gotto, A. M., Jr., and Jackson, R. L. (1974) Proc. Natl. Acad. Sci. USA 71, 3631-3634.
5. Schonfeld, G., Pfleger, B., and Roy, R. (1975) J. Biol. Chem. 250, 7943-7950.
6. Schonfeld, G., Bradshaw, R. A., and Chen, J. S. J. Biol. Chem. 251, 3921-3926.
7. Lux, S. E., John, K. M., Fleischer, S., Jackson, R. L., and Gotto, A. M. (1972) Biochem. Biophys. Res. Commun. 49, 23-29.
8. Scanu, A. M. (1970) Biochem. Biophys. Acta 200, 570-572.
9. Lux, S. E., John, K. M., Ronan, R., and Brewer, H. B. (1972) J. Biol. Chem. 247, 7519-7527.
10. Brewer, H. B., Jr., Lux, S. E., Ronan, R., John, K. M. (1972) Proc. Natl. Acad. Sci. 69, 1304-1308.
11. Edelstein, C., Lim, C. T., and Scanu, A. M. (1973) J. Biol. Chem. 248, 7653-7660.
12. Edelstein, C., Noyes, C., and Scanu, A. M. (1974) FEBS Lett. 38, 166-170.
13. Dayhoff, M. O. (1976) Atlas of Protein Sequence and Structure, 5, National Biomedical Research Foundation, Washington, D.C.
14. Jackson, R. L., Baker, H. N., Gilliam, E. G., and Gotto, A. M. (1977) Proc. Natl. Acad. Sci. USA 74, 1942-1945.
15. Gibbs, A. J., and McIntyre, G. A. (1970) Europ. J. Biochem. 16, 1-11.
16. Barry, D., and Jeng, I. Manuscript in preparation.
17. Segrest, J. P., Jackson, R. L., Morrisett, J. O., and Gotto, A. M. (1974) FEBS Lett. 38, 247-253.
18. Segrest, J. P., Pownall, H. J., Jackson, R. L., Glenner, G. G., and Pollock, P. S. (1976) Biochem. 15, 3187-3191.
19. Cheung, M. C., and Albers, J. J. (1977) J. Clin. Invest. 60, 43-50.