July 18, 2021

Dr. Nikolay Dokholyan

G. Thomas Passananti Professor and Vice Chair for Research, Department of Pharmacology

Professor, Department of Biochemistry and Molecular Biology

Penn State Neuroscience Institute

30 Hope Dr, Unit 1300

Hershey, PA 17033 JOURNAL

Dear Editor:

Thank you for your consideration of our manuscript entitled "Evolutionary and Structural Constraints Influencing Apolipoprotein A-I Amyloid Behaviour" (Prot-00199-2021) and request for a revised version. We have copied the reviewers' comments and criticisms in the 'Reply to Reviewers' file, and addressed each one individually (blue characters). As you will see, we have made every attempt to incorporate all the suggestions as thoroughly as possible in the revised manuscript. All the changes in the manuscript, including minor errors, can be seen in the version of the revised manuscript highlighted with red font.

We hope that you find the new revised version satisfactory.

Sincerely, Dr. Prof. María Alejandra Tricerri

Reviewer's comments

Reviewer 1

1. The authors create homology models for several apoA-I sequences followed by energy minimization. They report that different sequences showed very similar helix bundle conformation and similar intrinsic dynamics profiles (as represented by MSF), including low dynamics in APR1 (page 5). Since the models for each sequence were not subjected to simulated annealing and equilibration, it is likely that these models were stuck in the initial free energy minimum corresponding to the starting helix-bundle structure. If so, it is not surprising that the dynamic profiles in well-ordered regions did not change in these simulations. The authors may

want to either rebut this criticism or tone down the conclusions regarding the local dynamics.

"We agree with the reviewer opinion that the our modelling pipeline doesn't provide an extensive conformational sampling to ensure that local energy minima are avoided for each structural model. We have included a sentence in the results section mentioning this concern."

2. There are nearly 100 known variants of human apoA-I, most of them single point substitutions. Less that ¼ variants are associated with amyloidosis; many others are linked to dyslipidemia and cardiovascular disease. Does the analysis of the current study reveal any unusual properties of these disease-associated single substitution mutants as compared to WT apoA-I?

"The question posed by the reviewer is really interesting given that several of this mutations are linked to pathological conditions. During the preparation of the manuscript, an analysis of all the variants available for apoA-I was performed, including those associate with HDL diseases. No significant evidence of a destabilizing effect was found for these HDL variants (attached the reviewer will find a figure with this preliminar data). However, it has to be noted that our study used a lipid-free structure for apoA-I, and it could be possible that these variants present a prominent impact on HDL structure rather than monomeric apoA-I."

3. As the authors mention, proteolysis in specific apoA-I regions is important for amyloid formation. In vivo, proteolytic degradation and misfolding probably compete, and the initial cleavage site may determine the fate of the generated fragments (misfolding vs degradation). Do the authors observe increased dynamics in proteolysis-prone regions of some variants, e.g. in the central h5 region or the C-terminal h8 region?

"As the reviewer mention, specific regions on apoA-I sequence have been described as targets for proteolysis, in particular helices H5 and H8. We searched for increased mobility during molecular dynamics simulations, but we didn't evidence a significant change on the dynamics of these residues in the variants analyzed."

4. Fig 2B what are the units of MSF? In all main Figures, it would be more instructive
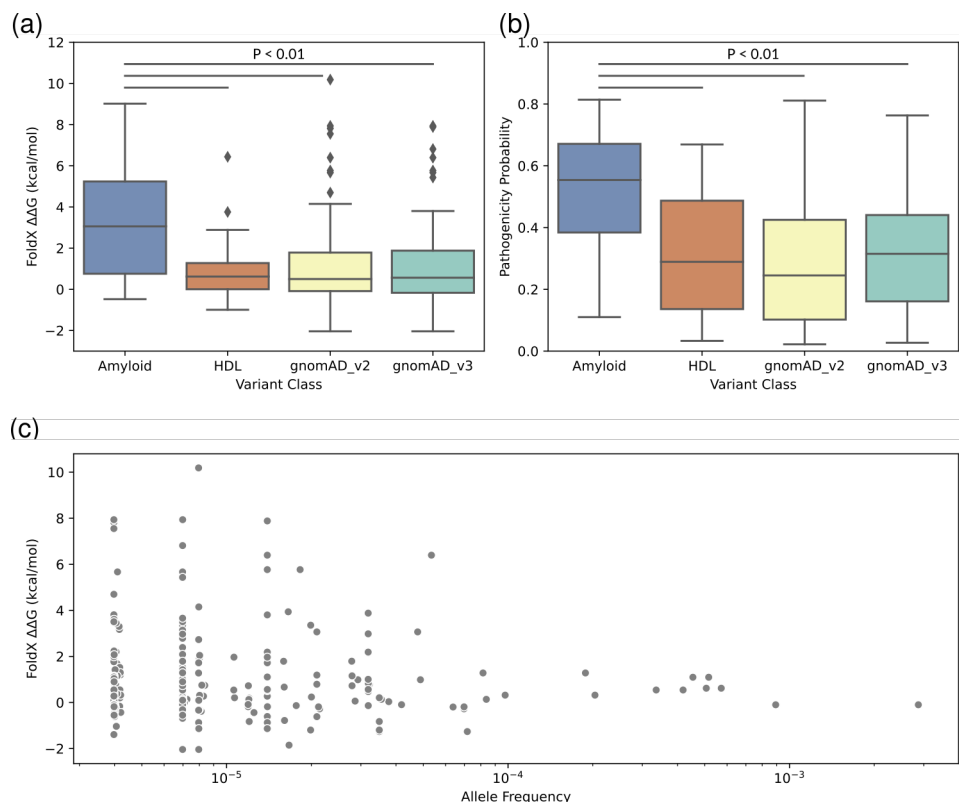
Figure 1: Impact of HDL-associated variants on apoA-I stability.

to show the results for each of the four APR separately, as in Fig 1A, rather than combining them together.

"As implemented in the Prody package, mean square fluctuations are calculated as arbitrary or relative units, and this detail has been added to the legend of Figure 2. For panels 1B, 2B and 2C, APRs plots have been modified to show the results of each individual APR in a separate manner."

5. "Page 3, line 24 "short C-terminus domain (residues 170-178)." This is a residue segment in the C-terminal half of the protein, not a C-terminal domain.

"This error has been corrected according to reviewer's suggestion."

6. The paper may benefit from editing for brevity.

"Whenever it possible, we tried to be more concise and brief along the manuscript."

Reviewer 2

1. One of the best indicators of 3D model quality is the identity percent between target sequences and protein template. These number should be show in the manuscript

"As suggested by the reviewer, the identity percent shared between the template and each target sequence have been added to the Materials and Methods section of the manuscript."

2. The homology modeling is not a method to predict conformational diversity of a protein so I would avoid phrases like "apoA-I structures have conserved their overall intrinsic dynamics" and "Dynamic and structural properties of apoA-I homologous structures." In this case MSF may reflect structural divergence (or just errors due to problems in the alignment and/or 3D modelling) more than dynamic properties of each protein. If these conclusions were derived from normal mode analysis please mention this in the text.

"Following reviewers suggestion, we modified the manuscript to address these inconsistencies. Also, we added a mention in the manuscript for the results that are derived from gaussian network model (GNM) analysis."

3. Site-specific evolutionary rates correlates with the number of contacts for a given position. Is there any correlation between WCN, H and dN/dS for the APRs?

"As the reviewer suggest, this correlation between evolutionary rate and biophysical properties in proteins has been reported previously for several protein families. Unfortunately, we couldn't detect a clear correlations between any of these variables in the case of APRs. Attached you will find a figure showing these results."
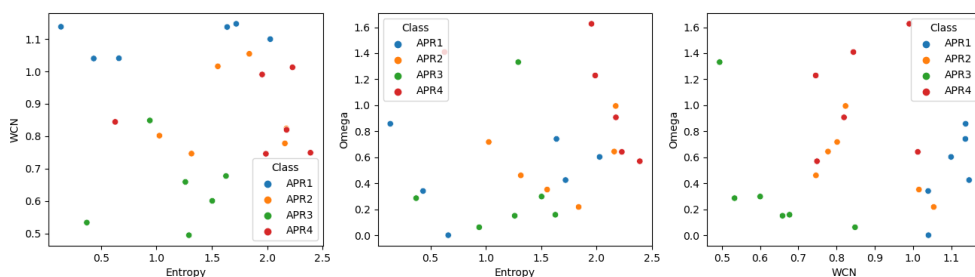

Figure 2: Lack of correlations for WCN, H and Omega (dN/dS) for APRS.)

4. Figure 3B and C are very interesting. Is there any other cluster/region with similar stability in the protein? Is the APR1 region the more stable considering all the

protein?

"Although we evidence several position which are highly susceptible to mutations, in particular the proline residues delimiting the helical segments of apoA-I, we didn't evidence another region that displays a susceptibility profile so pronounced as the APR1."

5. It is not clear to me where "variants" in supplementary Table 1 come from. I understand that four are from previous studies, but the others?

"We agree with the reviewer that it is ambiguous the source of information for the amyloidogenic variants listed in supplementary table 1. These variants were collated from the reference webpage www.amyloidosismutations.com, together with the two novel variants described by (Horike et al. 2018). A brief sentence mentioning these references has been added to the results section of the manuscript to clarify this point. Additionally, we would like to point out an error in the supplementary table 1: the mutation G35V was included in an older version of this table during our preliminary analysis of apoA-I amyloid behavior but it was later discarded because we found ambiguous information about its link with amyloidosis. We have remove this mutation from the supplementary table 1 in this revised version."

6. It is not clear to me the purpose of the use of normal mode analysis and MD. If the main objective was to explore protein flexibility and/or effect of mutations, I would certainly trust in the MD results more than in the normal mode analysis. The normal mode analysis relies on "equilibrium" state dynamics and MD (if it is longer enough) could explore much more conformational space. "We shared the opinion of the reviewer about the power of MD simulations to sample a larger fraction of the conformational space of apoA-I. Our motivation to explore protein dynamics with gaussian network models, specifically for the case of homology-based models, was based on the computational cost of running several simulation for different apoA-I structures (in this case, eight different models). Methods based on normal mode analysis are much more cheaper on the computational side and can provide an approximation to the dynamics of the system. However, for the analysis of the impact of different variants on apoA-I structure we wanted to have more comprehensive data about the impact of a given mutation, so we decided to invest on the MD simulations."

7. As I mentioned before, the finding using computational techniques that APR1 plays a key role in contrast to the other APRs is for me very interesting and I think that this result should be highlighted in the manuscript

"We are glad that the reviewer finds this result interesting. Modifications have been added to the discussion section to highlight this result"

TODO

- ☒ Moderated modelling conclusions
- ☒ Attach HDL variants plot
- ☒ Show data for proteolytic sites
- ☒ MSF units and split APRs figures
- ☒ Little error in page 3
- • [] Brevity?
- ☒ Incorporate % identity for the homology models
- ☒ Moderate homology modeling and MSF conclusions
- ☒ Correlations WCN, H and dN/dS
- ☒ Look for other conserved regions besides APRs
- ☒ Detail where do variants come from
- ☒ Explain why we used GNM in first place (limited computational resources)
- ☒ Emphasize APR1 results in contrast to the other APRs

References

Horike, Keiji, Asami Takeda, Makoto Tsujita, Norihiko Goto, Yoshihiko Watarai, Kazuharu Uchida, Akio Katayama, et al. 2018. "Two Novel Apoa1 Gene Mutations in a Japanese Renal Transplant Recipient with Recurrent Apolipoprotein a-i Related Amyloidosis." *Nephrology* 23 (S2): 17–21. https://doi.org/https://doi.org/10.1111/nep.13278.