Amyloidogenic mutations in human apolipoprotein A-I are not necessarily destabilizing: A common mechanism of apoA-I misfolding in familial amyloidosis and atherosclerosis

Madhurima Das, Xiaohu Mei, Shobini Jayaraman, David Atkinson, and Olga Gursky

SUPPLEMENT

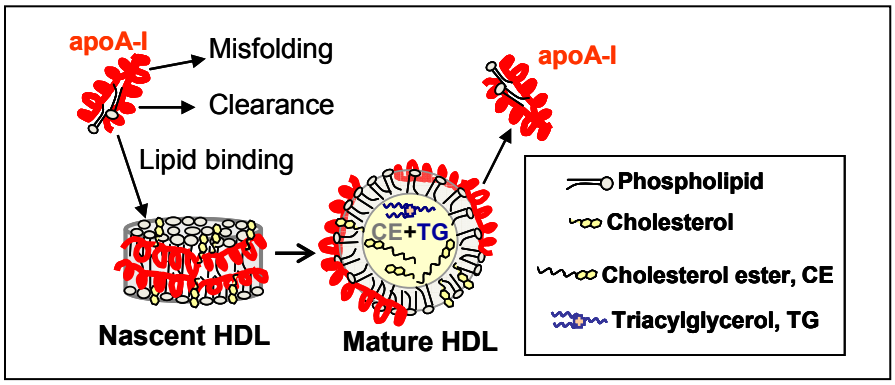

Figure S1 Cartoon representation of high-density lipoprotein (HDL) and free apoA-I (in red). Free protein, which constitutes ~5% of all circulating apoA-I, can bind lipids and form nascent HDL. Nascent HDL is envisioned as a "discoidal" nanoparticle (d~10 nm) comprized of a cholesterol-containing phospholipid bilayer with apoA-I α-helices wrapped around the circumference, although other models have also been proposed ([1] and references therein). In vivo, nascent HDLs are rapidly remodeled into mature "spherical" particles that contain a core of apolar lipids, mainly cholesterol esters with a small amount of triacylglycerides.

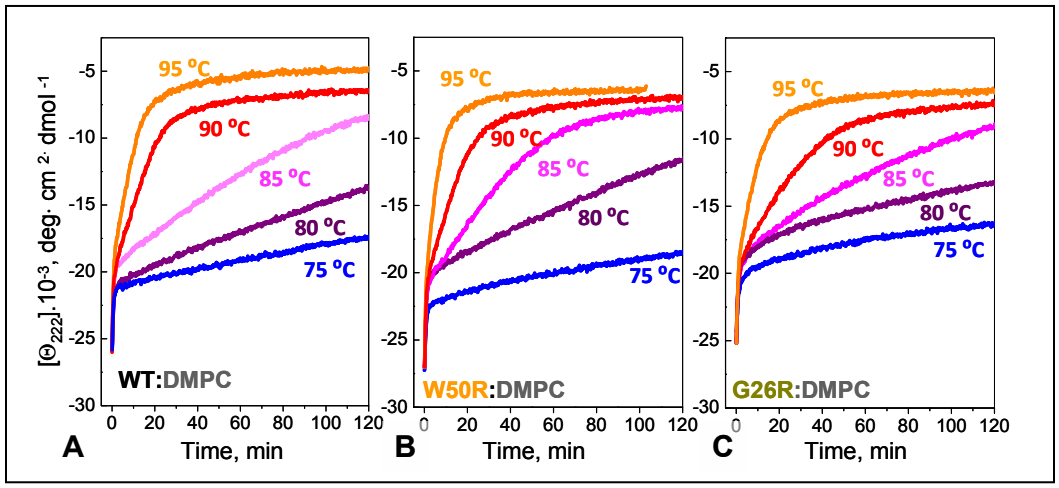

Figure S2. Comparison of kinetic stability of DMPC complexes with recombinant human wild type (WT) and mutant apoA-I. Thermal denaturation of these complexes was triggered by a rapid increase in temperature from 25 °C to a higher constant value ranging from 75 to 95 °C (indicated on the lines). The time course of protein unfolding, which is coupled to protein dissociation from the lipid and lipoprotein fusion, was monitored by CD signal at 222 nm, $[\Theta_{222}](t)$. The unfolding rate increased with increasing temperature, indicating reduced kinetic stability, ΔG^*, at high temperature, and a high enthalpic component of this stability, ΔH^* [2]. The results show no detectable difference in the kinetic stability of DMPC complexes with WT apoA-I or its two amyloidogenic mutants, W50R and G26R. Similarly, no differences in thermal stability were detected in the complexes of these proteins with POPC and cholesterol (data not shown). Consistent with these temperature-jump studies, limited kinetic studies of chemical denaturation using 4-8M Urea at 37 °C showed no significant differences in the denaturation time course of DMPC complexes containing WT, G25R or W50R apoA-I (data not shown).

Doc. S1. Amyloid prediction algorithms and consensus sequence analysis

We used consensus method AmylPred [3] and its expanded version, AmylPred2 [4]. AmylPred2 combines 11 sequence-based prediction algorithms (five of which were included in AmylPred) to search for the amyloidogenic segments that:

1. Have average packing density above threshold [5];
2. Contain conformational switches from α-helix to β-sheet [6];
3. Match the known amyloidogenic patterns in hexapeptides [7];
4. Have TANGO scores above threshold, which reflects average physical properties important for β-sheet formation, such as hydrophobicity, charge, and secondary structure [8];
5. Have hexapeptide conformational energy below threshold, which is derived from conformational modeling using microcrystal structures [9];
6. Use aggregation propensity scale derived for amino acids based on *in vivo* studies, AGGRESCAN [10];
7. Search for amyloidogenic mutants by using energy calculations to characterize mutational landscapes, AmyloidMutants [11];
8. Search for β-strand contiguity by using simple algorithm for sliding averages, SALSA [12];
9. Search for regions that appear natively α-helical but have hidden β-propensity, NetCSSP [13];
10. Use a phenomenological method based on support vector machines to identify hexapeptides that form amyloid, Pafig [14];
11. Use position-specific scoring matrices to identify amyloid-forming segments and distinguish them from non-fibrillar aggregates, WALTZ [15].

More details on specific algorithms and their use can be found in [4] and references therein.

In addition, we used sequence-based algorithm PASTA [16] that searches for potential β-strand pairings and predicts their relative orientation. The results are shown in Tables S1 and S2.

Residue #	10 20 30 40 50 60 70 80 90 100 110 120
1. AGGRESCAN	DEPPQSPWDR VKDLATVYVD VLKDSGRDYV SQFEGSALGK QLN*LKLLDNW DSVTSTFSKL* REQLGPVTQE FWDN*LEKETE GL*RQEMSKDL EEVKAKVQPY LDDFQKKWQE EMELYRQKVE
2. Amyl. Mutants	DEPPQSPWDR VKDLATVYVD VLKDSGRDYV SQFEGSALGK QLN*LKLLDNW DSVTSTFSKL* REQLGPVTQE FWDN*LEKETE GL*RQEMSKDL EEVKAKVQPY LDDFQKKWQE EMELYRQKVE
3. Amyl. Pattern	DEPPQSPWDR VKDLATVYVD VLKDSGRDYV SQFEGSALGK QLN*LKLLDNW DSVTSTFSKL* REQLGPVTQE FWDN*LEKETE GL*RQEMSKDL EEVKAKVQPY LDDFQKKWQE EMELYRQKVE
4. Av. Pack. Dens.	DEPPQSPWDR VKDLATVYVD VLKDSGRDYV SQFEGSALGK QLN*LKLLDNW DSVTSTFSKL* REQLGPVTQE FWDN*LEKETE GL*RQEMSKDL EEVKAKVQPY LDDFQKKWQE EMELYRQKVE
5. SALSA	DEPPQSPWDR VKDLATVYVD VLKDSGRDYV SQFEGSALGK QLN*LKLLDNW DSVTSTFSKL* REQLGPVTQE FWDN*LEKETE GL*RQEMSKDL EEVKAKVQPY LDDFQKKWQE EMELYRQKVE
6. Hexapept. Eng.	DEPPQSPWDR VKDLATVYVD VLKDSGRDYV SQFEGSALGK QLN*LKLLDNW DSVTSTFSKL* REQLGPVTQE FWDN*LEKETE GL*RQEMSKDL EEVKAKVQPY LDDFQKKWQE EMELYRQKVE
7. NetCSSP	DEPPQSPWDR VKDLATVYVD VLKDSGRDYV SQFEGSALGK QLN*LKLLDNW DSVTSTFSKL* REQLGPVTQE FWDN*LEKETE GL*RQEMSKDL EEVKAKVQPY LDDFQKKWQE EMELYRQKVE
8. Pafig	DEPPQSPWDR VKDLATVYVD VLKDSGRDYV SQFEGSALGK QLN*LKLLDNW DSVTSTFSKL* REQLGPVTQE FWDN*LEKETE GL*RQEMSKDL EEVKAKVQPY LDDFQKKWQE EMELYRQKVE
9. SecStr	DEPPQSPWDR VKDLATVYVD VLKDSGRDYV SQFEGSALGK QLN*LKLLDNW DSVTSTFSKL* REQLGPVTQE FWDN*LEKETE GL*RQEMSKDL EEVKAKVQPY LDDFQKKWQE EMELYRQKVE
10. TANGO	DEPPQSPWDR VKDLATVYVD VLKDSGRDYV SQFEGSALGK QLN*LKLLDNW DSVTSTFSKL* REQLGPVTQE FWDN*LEKETE GL*RQEMSKDL EEVKAKVQPY LDDFQKKWQE EMELYRQKVE
11. WALTZ	DEPPQSPWDR VKDLATVYVD VLKDSGRDYV SQFEGSALGK QLN*LKLLDNW DSVTSTFSKL* REQLGPVTQE FWDN*LEKETE GL*RQEMSKDL EEVKAKVQPY LDDFQKKWQE EMELYRQKVE
12. PASTA	DEPPQSPWDR VKDLATVYVD VLKDSGRDYV SQFEGSALGK QLN*LKLLDNW DSVTSTFSKL* REQLGPVTQE FWDN*LEKETE GL*RQEMSKDL EEVKAKVQPY LDDFQKKWQE EMELYRQKVE

Residue #	130 140 150 160 170 180 190 200 210 220 230 240
1. AGGRESCAN	PLRAELQEGA RQKLHELQEK LSPLGEEMRD RARAHVDALR THLAPYSDEL RQRLAARLEA LKENGGARLA EYHAKATEHL STLSEKAKPA LEDLRQGLLP VLESFKVSFL SALEEYTKKL NTQ
2. Am. Mutants	PLRAELQEGA RQKLHELQEK LSPLGEEMRD RARAHVDALR THLAPYSDEL RQRLAARLEA LKENGGARLA EYHAKATEHL STLSEKAKPA LEDLRQGLLP VLESFKVSFL SALEEYTKKL NTQ
3. Amyl. Pattern	PLRAELQEGA RQKLHELQEK LSPLGEEMRD RARAHVDALR THLAPYSDEL RQRLAARLEA LKENGGARLA EYHAKATEHL STLSEKAKPA LEDLRQGLLP VLESFKVSFL SALEEYTKKL NTQ
4. Av. Pack. Dens.	PLRAELQEGA RQKLHELQEK LSPLGEEMRD RARAHVDALR THLAPYSDEL RQRLAARLEA LKENGGARLA EYHAKATEHL STLSEKAKPA LEDLRQGLLP VLESFKVSFL SALEEYTKKL NTQ
5. SALSA	PLRAELQEGA RQKLHELQEK LSPLGEEMRD RARAHVDALR THLAPYSDEL RQRLAARLEA LKENGGARLA EYHAKATEHL STLSEKAKPA LEDLRQGLLP VLESFKVSFL SALEEYTKKL NTQ
6. Hexapept. Eng.	PLRAELQEGA RQKLHELQEK LSPLGEEMRD RARAHVDALR THLAPYSDEL RQRLAARLEA LKENGGARLA EYHAKATEHL STLSEKAKPA LEDLRQGLLP VLESFKVSFL SALEEYTKKL NTQ
7. NetCSSP	PLRAELQEGA RQKLHELQEK LSPLGEEMRD RARAHVDALR THLAPYSDEL RQRLAARLEA LKENGGARLA EYHAKATEHL STLSEKAKPA LEDLRQGLLP VLESFKVSFL SALEEYTKKL NTQ
8. Pafig	PLRAELQEGA RQKLHELQEK LSPLGEEMRD RARAHVDALR THLAPYSDEL RQRLAARLEA LKENGGARLA EYHAKATEHL STLSEKAKPA LEDLRQGLLP VLESFKVSFL SALEEYTKKL NTQ
9. SecStr	PLRAELQEGA RQKLHELQEK LSPLGEEMRD RARAHVDALR THLAPYSDEL RQRLAARLEA LKENGGARLA EYHAKATEHL STLSEKAKPA LEDLRQGLLP VLESFKVSFL SALEEYTKKL NTQ
10. TANGO	PLRAELQEGA RQKLHELQEK LSPLGEEMRD RARAHVDALR THLAPYSDEL RQRLAARLEA LKENGGARLA EYHAKATEHL STLSEKAKPA LEDLRQGLLP VLESFKVSFL SALEEYTKKL NTQ
11. WALTZ	PLRAELQEGA RQKLHELQEK LSPLGEEMRD RARAHVDALR THLAPYSDEL RQRLAARLEA LKENGGARLA EYHAKATEHL STLSEKAKPA LEDLRQGLLP VLESFKVSFL SALEEYTKKL NTQ
12. PASTA	PLRAELQEGA RQKLHELQEK LSPLGEEMRD RARAHVDALR THLAPYSDEL RQRLAARLEA LKENGGARLA EYHAKATEHL STLSEKAKPA LEDLRQGLLP VLESFKVSFL SALEEYTKKL NTQ

Table S1. Summary of apoA-I sequence analysis by using AmylPred 2 [4] (11 algorithms) and PASTA [16]. Boxed regions show amyloid hot spots predicted by consensus of 4 or more methods. Residues *44-55* (extended strand in the crystal structure of Δ(185-243)apoA-I) and *76-81* (expected β-breaking motif) are in *italics*.

Doc. S2. Parallel in-register β-sheets are predicted for the major amylopid hot spots in apoA-I

PASTA server [16] was used to search apoA-I sequence for possible β-sheet pairings and to assess the likely β-sheet orientation (parallel or antiparallel). The output is shown in Table S1 and Figure S4.

pairing 1	PASTA	energy -5.32	length	6	between segments	17-22	and 17-22	parallel
pairing 2	PASTA	energy -5.14	length	9	between segments	14-22	and 14-22	parallel
pairing 3	PASTA	energy -5.00	length	7	between segments	16-22	and 16-22	parallel
pairing 4	PASTA	energy -4.63	length	10	between segments	221-230	and 221-230	parallel
pairing 5	PASTA	energy -4.54	length	5	between segments	17-21	and 17-21	parallel
pairing 6	PASTA	energy -4.46	length	13	between segments	221-233	and 221-233	parallel
pairing 7	PASTA	energy -4.37	length	8	between segments	14-21	and 14-21	parallel
pairing 8	PASTA	energy -4.37	length	8	between segments	15-22	and 15-22	parallel
pairing 9	PASTA	energy -4.31	length	11	between segments	221-231	and 221-231	parallel
pairing 10	PASTA	energy -4.23	length	12	between segments	11-22	and 11-22	parallel
pairing 11	PASTA	energy -4.22	length	6	between segments	16-21	and 16-21	parallel
pairing 12	PASTA	energy -4.10	length	7	between segments	17-23	and 17-23	parallel
pairing 13	PASTA	energy -4.07	length	6	between segments	14-19	and 14-19	parallel
pairing 14	PASTA	energy -4.07	length	6	between segments	17-22	and 225-23	parallel

Table S2 Sequence analysis of human apoA-I by using amyloid prediction software PASTA [16] to assess the major amyloidogenic "hot spots" and their preferred orientation in fibrils (parallel or antiparallel β-strands). In fibrils, β-pairings generally have PASTA energies of -4.0 or lower [16]. PASTA output for top apoA-I pairings between residue segments with energies below -4 is listed. These include two major amyloidogenic hot spots in the N-terminal (blue) and C-terminal (red) regions. Similar major hot spots have been predicted by other methods including those in Amylpred2. Importantly, PASTA predicts that both major hot spots in apoA-I form parallel β-sheet.

Our sequence analyses using PASTA predicts parallel in-register β-sheet in most other apolipoproteins, including apoC-II and α-synuclein. PASTA energies for the best pairings are comparable (-5.37 for apoC-II) or lower (-6.2 for α-synuclein) than those of apoA-I (-5.3, Table S1). EPR and solid-state NMR show parallel in-register β-sheets in fibrils of apoC-II and α-synuclein [17-19], directly verifying these predictions. Further, FTIR spectra of fibrils formed by apoA-I fragment 1-43 that contains the major hot spot 14-11 [20] lack the peak near 1690 cm^{-1}, which indicates parallel β-sheet in these fibrils.

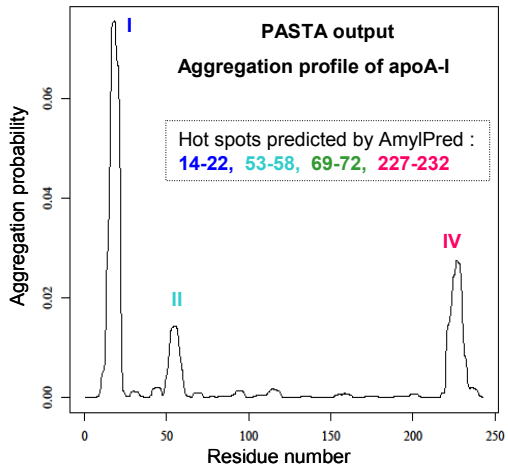

Figure S3. Amyloidogenic profile of apoA-I predicted by PASTA shows three (I, II, and IV) out of four hot spots predicted by AmylPred in the same rank order, **I**>**IV**>**II**.

The prediction that the two major N- and C-terminal amyloidogenic hot spots in apoA-I form parallel in-register β-sheets in amyloid has two important implications. First, it suggests that apoA-I has high propensity to form parallel β-sheet in which each hot spot segment is paired with its counterparts from the adjacent molecules and hence, different hot spots act in synergy. Second, in the well-established "double-belt" conformation of apoA-I molecular dimer on HDL and free in solution ([21-23] and references therein), the two highly helical dimer-forming apoA-I molecules are antiparallel (Figure 1A). This difference between the antiparallel molecular orientation in apoA-I helical dimer and the parallel intermolecular β-sheet in fibrils supports the importance of apoA-I dissociation into monomers as an obligatory step in fibril formation.

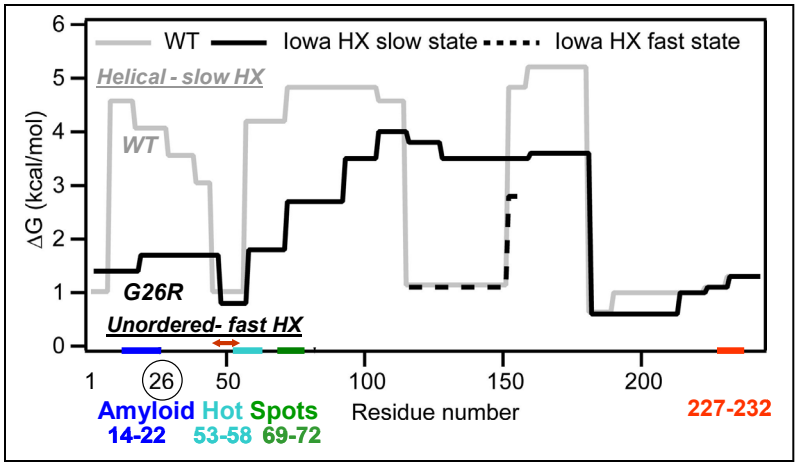

Fig. S4 Hydrogen-deuterium exchange (HX) – mass spectrometry (MS) profile of monomeric lipid-free human WT and G26R apoA-I. Free energy of secondary structural stability ΔG, which is proportional to –RT ln *k,* where *k* is the experimentally determined HX rate, is plotted as a function of residue number. Slow exchange corresponding to high ΔG~4-5 kcal/mol is characteristic of hydrogen bonding in a well-ordered α-helical structure, while fast exchange correspondding to low ΔG~1 kcal/mol indicates unstructured regions. The results of HX/MS secondary structural assignment in WT apoA-I (grey line) are in excellent agreement with the crystal structure of Δ(185-243)apoA-I [21]. For example, the extended segment 44-55 seen in the crystal structure shows low protection in HX (double arrow), while α-helical segments show high protection (grey line). Comparison with HX data for G26R (black line) indicates that G26R mutation causes loss of α-helical structure in residues 1-90 [21]. We note that these residues encompass three predicted N-terminal amyloid hot spots (indicated by colored bars and residue numbers on the X-axis). (Adapted with permission from [22]; copyright 2012 American Chemical Society. Our modifications to the figure include text in italics or in color.)

Residue	Atom	Distance
Arg 177'	C_γ	3.8 Å
	NH_2^+	4.6 Å
	N_ε	4.6 Å
	C_γ	4.8 Å
Leu 174'	C_α	4.0 Å
	C_β	4.2 Å
Arg 173'	O	4.4 Å
	C_α	4.4 Å
	C_β	4.8 Å
Asp 89	O^-_δ	4.9 Å
Tyr 18	C_β	4.9 Å

Table S3. Nearest neighbors of Met86 sulfur in the helix bundle structure of free apoA-I. Atoms located within 5 Å from Met86 S_δ in the crystal structure of lipid-free Δ(185-243)apoA-I (PDB ID 3R2P) are listed. Hyphens indicate residues from the segment 170-178 that contains all known "outside" point mutations in AApoA-I. Apolar atoms are in black, polar in green, charged in blue (positive) and red (negative). Packing of Met86 sulfur in this predominantly apolar environment suggests strongly that polar Met sulfoxide will disrupt this packing. This will perturb the packing of Tyr18 in the middle of the major amyloidogenic hot spot (residues 14-22), and thereby promote β-aggregation. The closest contacts between Met86 and Tyr18 include 4.0 Å (M86 C_β to Y18 C_γ) and 4.9 Å (M86 S to Y18 C_β).

Doc. S3. References

1. Gursky O (2013). Crystal structure of Δ(185-243)apoA-I suggests a mechanistic framework for the protein adaptation to the changing lipid load in Good Cholesterol: From Flatland to Sphereland via Double Belt, Belt Buckle, Double Hairpin and Trefoil/Tetrafoil. *J Mol Biol* **425(1)**, 1-16.
2. Gursky O, Ranjana & Gantz DL (2002). Complex of human apolipoprotein C-1 with phospholipid: thermodynamic or kinetic stability? *Biochemistry* **41(23)**, 7373-7384.
3. Frousios KK, Iconomidou VA, Karletidi CM & Hamodrakas SJ (2009). Amyloidogenic determinants are usually not buried. *BMC Struct, Biol*, **9**, 44.
4. Tsolis AC, Papandreou NC, Iconomidou VA, Hamodrakas SJ (2013) A consensus method for the prediction of 'aggregation-prone' peptides in globular proteins. PLoS One **8(1)**, e54175.
5. Galzitskaya OV, Garbuzynskiy SG & Lobanov MV (2006) Prediction of amyloidogenic and disordered regions in protein chains. *PloS Comput Biol* **2**, 1639-1648.
6. Hamodrakas SJ, Liappa C & Iconomidou,VA (2007) Consensus prediction of amyloidogenic determinants in amyloid-forming proteins. *Int J Biol Macromol* **41**,295-300.
7. López de la Paz M & Serrano L (2004) Sequence determinants of amyloid fibril formation. *Proc Natl. Acad. Sci. USA* **101**, 87-92.
8. Fernandez-Escamilla AM, Rousseaux F, Schymkowitz J & Serrano L (2004) Prediction of sequence-dependent and mutational effects on the aggregation of peptides and proteins. *Nature Biotech* **22**, 1302-1306.
9. Zhang Z, Chen H & Lai L (2007) Identification of amyloid fibril-forming segments based on structure and residue-based statistical potential. *Bioinformatics* **23**, 2218-2225.
10. Conchillo-Solé O, de Groot NS, Avilés FX, Vendrell J, Daura X & Ventura S (2007) AGGRESCAN: A server for the prediction and evaluation of “hot spots” of aggregation in polypeptides. *BMC Bioinformatics* **8**, 65.
11. O'Donnell CW, Waldispühl J, Lis M, Halfmann R, Devadas S, Lindquist S & Berger B (2011) A method for probing the mutational landscape of amyloid structure. *Bioinformatics* **27(13)**, i34-42.
12. Zibaee S, Makin OS, Goedert M & Serpell LC (2007) A simple algorithm locates beta-strands in the amyloid fibril core of alpha-synuclein, Abeta, and tau using the amino acid sequence alone. *Protein Sci* **16**, 906–918.
13. Kim C, Choi J, Lee SJ, Welsh WJ, Yoon S (2009) NetCSSP: Web application for predicting chameleon sequences and amyloid fibril formation. *Nucleic Acids Res* **37**, W469-473.
14. Tian J, Wu N, Guo J & Fan Y (2009) Prediction of amyloid fibril-forming segments based on a support vector machine. *BMC Bioinformatics* **10**, Suppl 1:S45.
15. Maurer-Stroh S, Debulpaep M, Kuemmerer N, Lopez de la Paz M, Martins IC, Reumers J, Morris KL, Copland A, Serpell L, Serrano L, Schymkowitz JW & Rousseau F (2010) Exploring the sequence determinants of amyloid structure using position-specific scoring matrices. *Nat Methods* **7(3)**, 237-242.
16. Trovato A, Seno F, Tosatto SC (2007) The PASTA server for protein aggregation prediction. *Protein Eng Des Sel* **20(10)**, 521-523.
17. Teoh CL, Pham CL, Todorova N, Hung A, Lincoln CN, Lees E, Lam YH, Binger KJ, Thomson NH, Radford SE, Smith TA, Müller SA, Engel A, Griffin, MD, Yarovsky I, Gooley PR & Howlett GJ (2011) A structural model for apolipoprotein C-II amyloid fibrils: experimental characterization and molecular dynamics simulations. *J Mol Biol* **405(5),** 1246-1266.
18. Der-Sarkissian A, Jao CC, Chen J, Langen R (2003) Structural organization of α-synuclein fibrils studied by site-directed spin labeling. *J Biol Chem* **278**, 37530–37535.
19. Lv G, Kumar A, Giller K, Orcellet ML, Riedel D, Fernández CO, Becker S & Lange A (2012) Structural comparison of mouse and human α synuclein amyloid fibrils by solid-state NMR. *J Mol Biol* **420(1-2)**, 99-111.
20. Adachi E, Kosaka A, Tsuji K, Mizuguchi C, Kawashima H, Shigenaga A, Nagao K, Akaji K, Otaka A & Saito H (2014) The extreme N-terminal region of human apolipoprotein A-I has a strong propensity to form amyloid fibrils. *FEBS Lett* **588(3)**, 389-394.
21. Mei X & Atkinson D (2011) Crystal structure of C-terminal truncated apolipoprotein A-I reveals the assembly of high density lipoprotein (HDL) by dimerization. *J Biol Chem* **286(44)**, 38570-38582.
22. Chetty PS, Ohshiro M, Saito H, Dhanasekaran P, Lund-Katz S, Mayne L, Englander W & Phillips MC (2012) Effects of the Iowa and Milano mutations on apolipoprotein A-I structure and dynamics determined by hydrogen exchange and mass spectrometry. *Biochemistry,* **51(44)**, 8993-9001.
23. Borhani DW, Rogers DP, Engler JA & Brouillette CG (1997) Crystal structure of truncated human apolipoprotein A-I suggests a lipid-bound conformation. *Proc Natl Acad Sci USA* **94**, 12291-12296.