PNAS
www.pnas.org

Supplementary Information for

Tertiary structure of apolipoprotein A-I in nascent high-density lipoproteins

Mohsen Pourmousa, Hyun D. Song, Yi He, Jay W. Heinecke, Jere P. Segrest, Richard W. Pastor

Richard W. Pastor
Email: pastorr@nhlbi.nih.gov
Jere P. Segrest
Email: j.segrest@vanderbilt.edu

This PDF file includes:

Supplementary text
Figs. S1 to S9
Tables S1 to S3
References for SI reference citations

Other supplementary materials for this manuscript include the following:

Trajectories and final configurations of Simulations 1 and 2 are available at http://doi.org/10.5281/zenodo.897027.

Methods Details

MD Simulations of Nascent HDL. Lipid nanodiscs were built using CHARMM-GUI (1). Simulations were performed at constant pressure and temperature with a constant number of particles. TIP3P water model (2) as modified for CHARMM (3) was used to describe water molecules. The Lennard-Jones (LJ) parameters of Na^+ and Cl^- as well as Na^+ and select oxygens of lipids and proteins were taken from the CHARMM C36 ion parameters (NBFIX) (4–6). All His residues were protonated, as they are in the same position as basic residues on helical repeats (7). Anton trajectories were generated with a *multigrator* (8), which separates barostat, thermostat, and Newtonian particle motion updates, with a time step of 2 fs. Temperature and pressure were kept constant at 310 K and 1 bar, respectively, using a variant (8) of the Nosé–Hoover (9) and the Martyna–Tobias–Klein algorithm (10). Electrostatic forces were calculated using the u-series method (11) on a 64 × 64 × 64 mesh for distant calculations. Water molecules and all bond lengths to hydrogen atoms were constrained using M-SHAKE algorithm (12).

Distance maps were to be compared with zero-order cross-linking data with spacer-arms of 0 Å. Since most of the cross-links were between Lys and Glu, a cutoff radius of 15.1 Å was used between C_α atoms (r_{cutoff} = spacer arm length + (Lys length from C_α + Glu length from C_α) + motion averaging factor = 0 Å + (7.1 Å + 5.0 Å) + 3 Å = 15.1 Å).

Secondary structures were calculated using the DSSP code on VMD. Simulation snapshots were generated using the VMD software (13).

Disc diameters were obtained by orienting the disc normal along the *z* axis and comparing the *z*-component of moment of inertia of the disc, $I_{z,\text{rHDL}}$, with that of a homogeneous disc, $I_{z,\text{disc}} = \frac{1}{2} N R^2$, yielding $d = 2R = 2\sqrt{(2 \times I_{z,\text{rHDL}}/N)}$, where *d*, *R* and *N* are diameter, radius, and number of atoms, respectively.

Numerical values from simulation are reported as mean ± standard error, where standard error was obtained by dividing the production runs into 10 blocks, calculating the average for each block, finding the standard deviation of 10 averages, and dividing it by $\sqrt{10}$.

Rosetta Modeling. Models were analyzed using the cluster analysis tools in Gromacs (14). First, a pairwise RMSD matrix for C_α of residues 17 to 43 was generated using the tool *gmx rms*. Secondly, the clustering was performed with an RMSD cutoff of 2.5 Å using the tool *gmx cluster*, which uses the clustering algorithm described in Ref. 15. The first 3 clusters generated included 53% of the 10,000 models. The structure at the centroid of the most populated cluster (22% of 10,000 models) was selected as the representative structure. PyMOL (16) was used for visualization.

In vacuo MD Simulation of Lipid-Free Dimeric NTDs. The pair of APOA1 residues 1–43 of the planar double belt (Fig. 2a) was simulated for 15 ns using NAMD2 (17) with the CHARMM 36 protein parameters (18–20). The time step was 2 fs and the temperature was maintained at 310 K with the Berendsen thermostat (21).

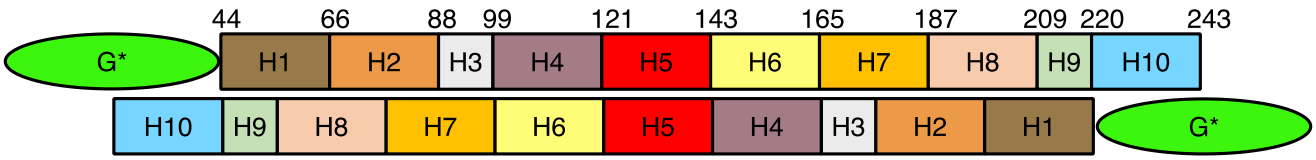

Fig. S1. Secondary structure of APOA1 in the LL5/5 double belt dimer. Residues 44 to 242 of APOA1 form 10 tandem amphipathic helices (H1–H10). These helices are primarily Type A, where the hydrophobic and hydrophilic faces are comparable in size, positive residues are on the hydrophobic/hydrophilic boundary and negative residues are on the middle of the hydrophilic face. Residues 1 to 43 are Type G* amphipathic α-helices, which differ from Type A in that positive and negative residues are more uniformly distributed on the hydrophilic face. The two proteins are arranged in an antiparallel fashion with H5 adjacent to each other. Extension of the G* domains indicates antiparallel overlap of the NTD.

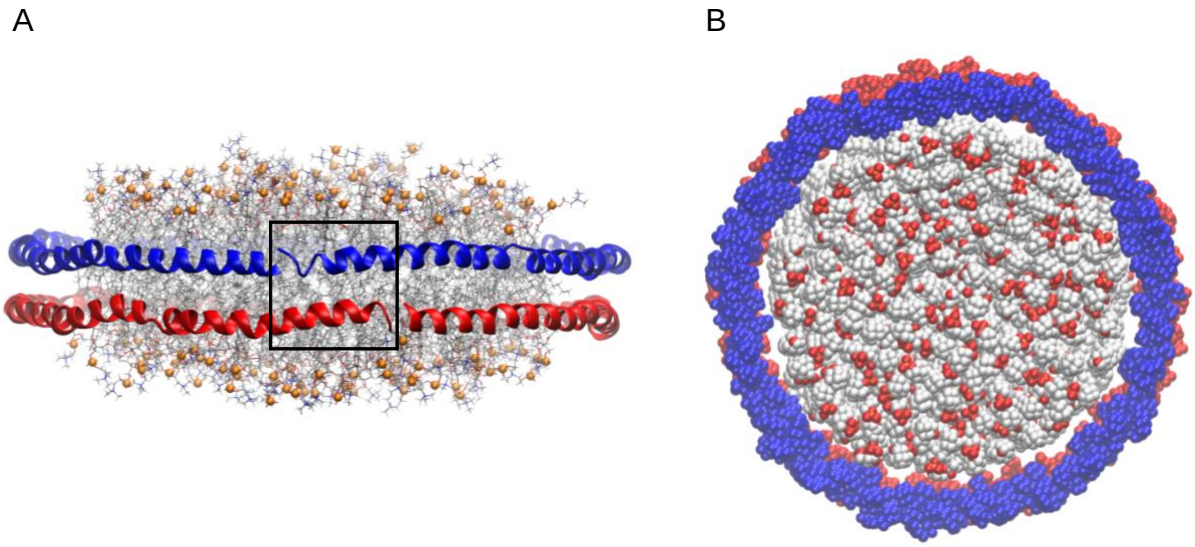

Fig. S2. Double belt model for nascent HDL of a disc containing 200:20:2 POPC:UC:APOA1 in an antiparallel arrangement with LL5/5 registry. (*A*) Side view. Box highlights partial registry of N-terminals. Coloring is the same as Fig. 1. (*B*) Top-down view in space filling, with side chains of the proteins included.

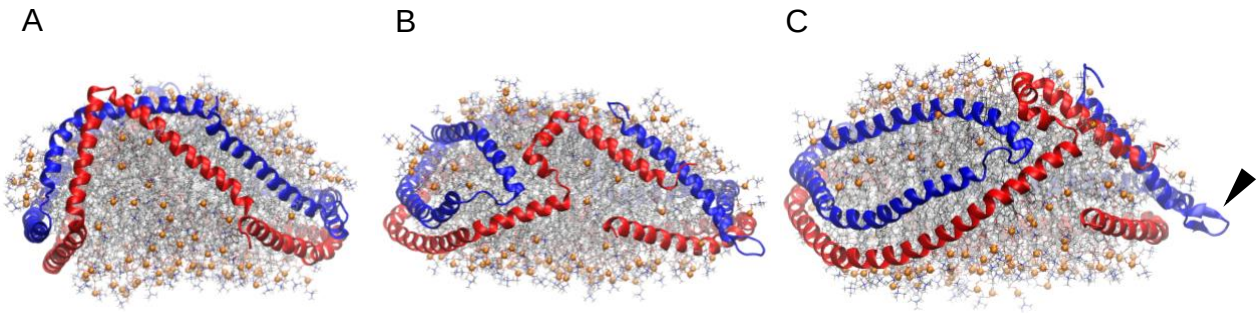

Fig. S3. Snapshots of Simulation 1. (*A*) t = 1 μs. (*B*) t = 16 μs. (*C*) t = 17 μs. Arrowhead in (*C*) points a transient β-turn. Coloring is the same as Fig. 1.

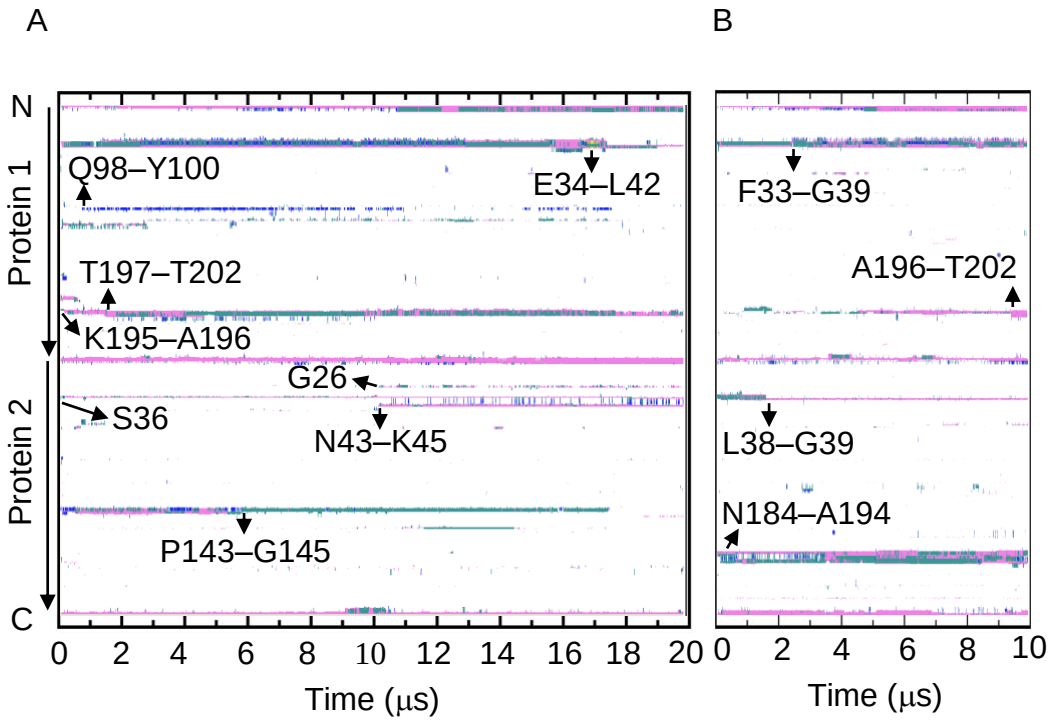

Fig. S4. Time series of secondary structure of two APOA1 molecules (Proteins 1 and 2). (*A*) Simulation 1. There is a β-turn for about 400 ns at E34–L42 of Protein 1 which is shown in Fig. S3*C*. (*B*) Simulation 2. Vertical axis from top to bottom represents residues from N- to C-terminus of each APOA1. Color codes: α-helix, white; π-helix, blue; β-sheet, yellow; turn, green; coil, pink.

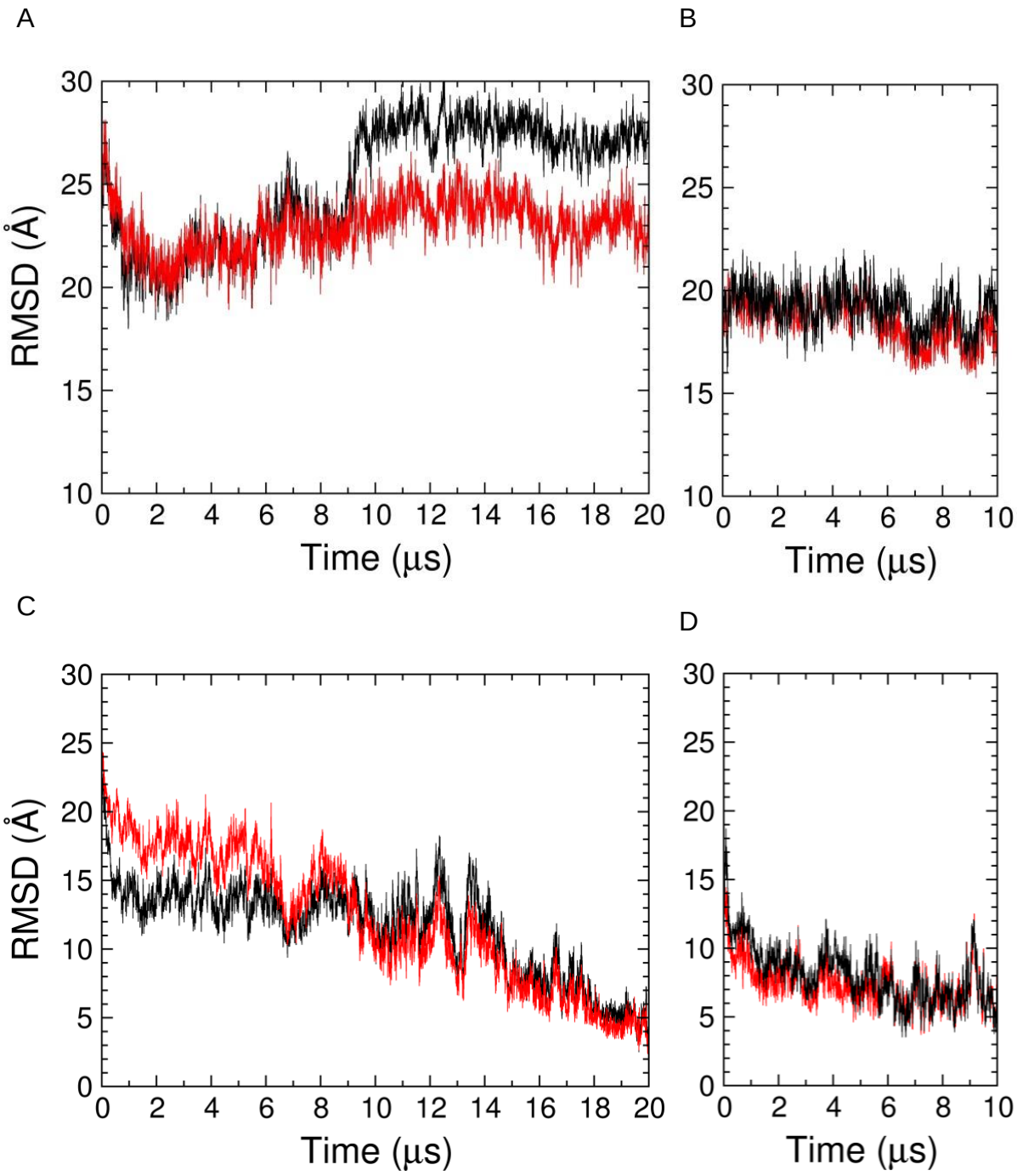

Fig. S5. RMSD of C_α of Proteins 1 (black) and 2 (red) in Simulations 1 and 2 from different reference structures. (*A*) Simulation 1, reference: planar double belt. (*B*) Simulation 2, reference: planar double belt. (*C*) Simulation 1, reference: 20 μs frame. (*D*) Simulation 2, reference: 10 μs frame.

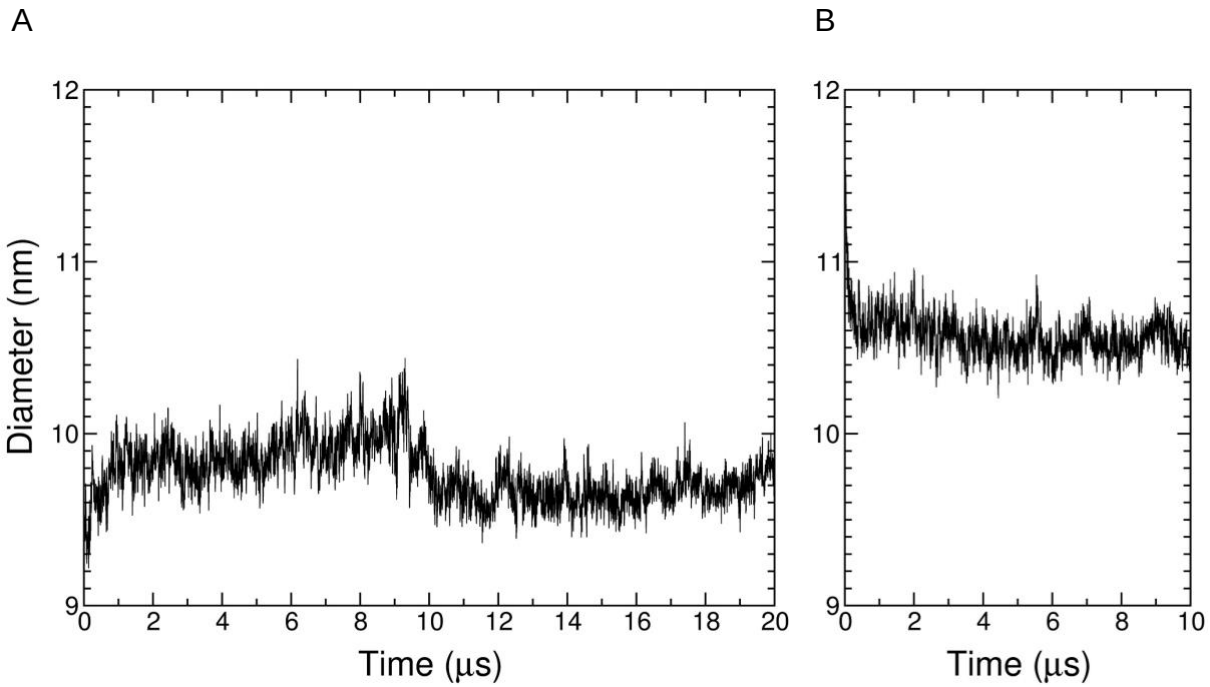

Fig. S6. Time series of disc diameters in Simulation 1 (*A*) and Simulation 2 (*B*).

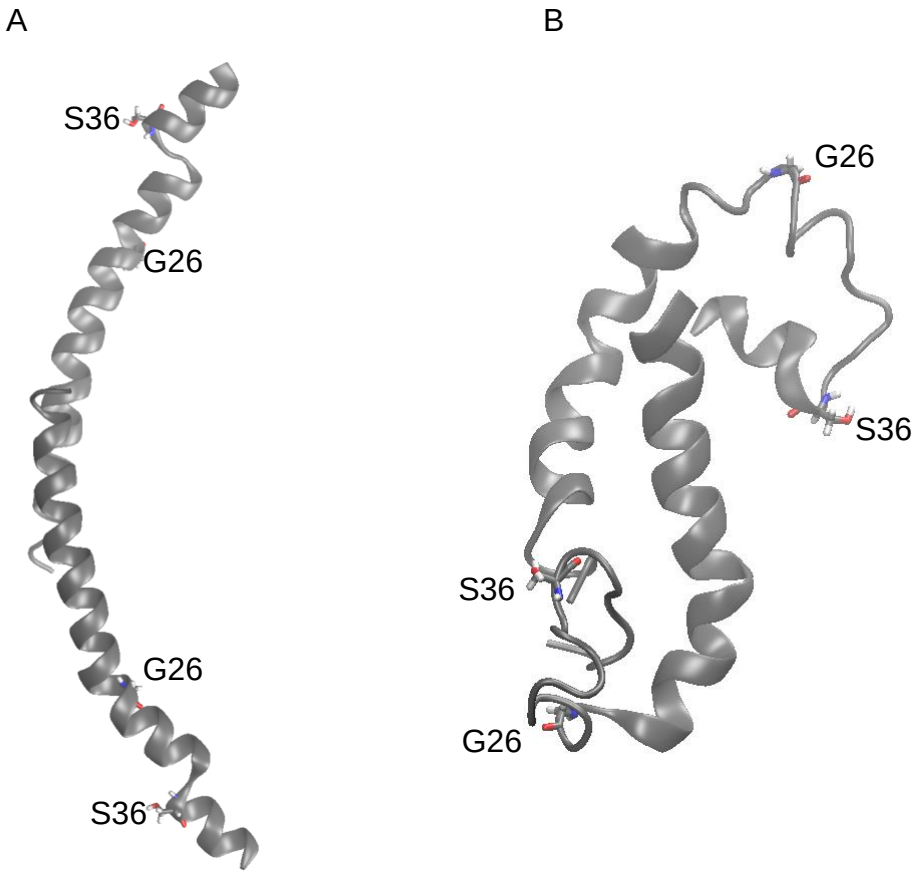

Fig. S7. In vacuo simulation of residues 1–43. (*A*) Initial condition taken from the planar double belt shown in Fig. S2. (*B*) 15 ns snapshot. Figures 2*E* and *F* in the main text show alignment of the two structures in (*B*) with that obtained from Simulation 1.

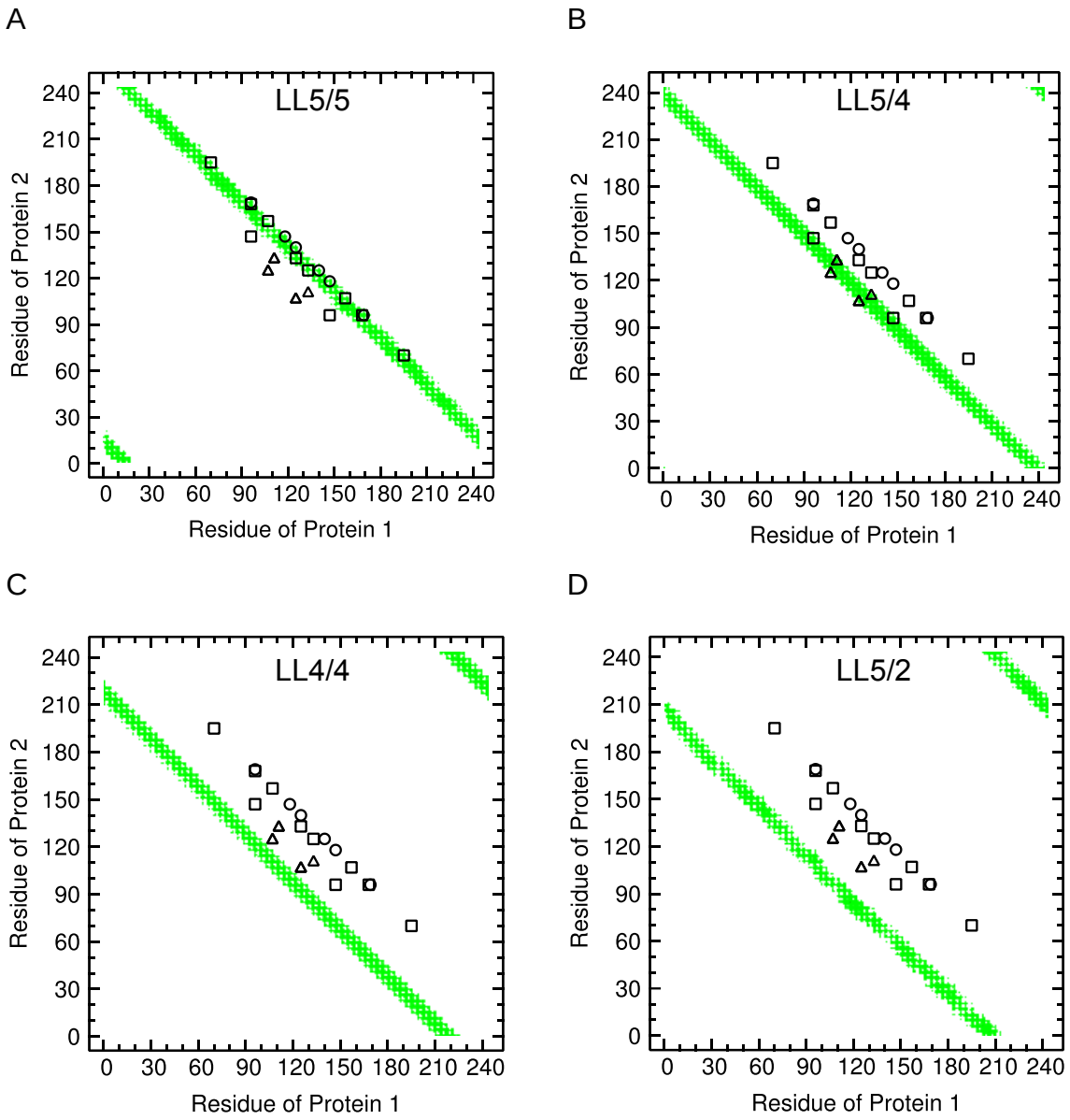

Fig. S8. Comparison of cross-link experiments with three qualitative signal strengths (circle, strong; square, medium; triangle, weak) and intermolecular distance maps of C_α from ideal belt models for LL5/5 (*A*), LL5/4 (*B*), LL4/4 (*C*), and LL5/2 (*D*) registries (green points, 15.1 Å cutoff). The distance map in panel *A* was computed from the initial condition used in the simulation of rHDL-2-100. The distance map in Figure 3*A* in the main text was computed from the 16–20 μs average of this simulation, and therefore differs slightly from that of panel *A*.

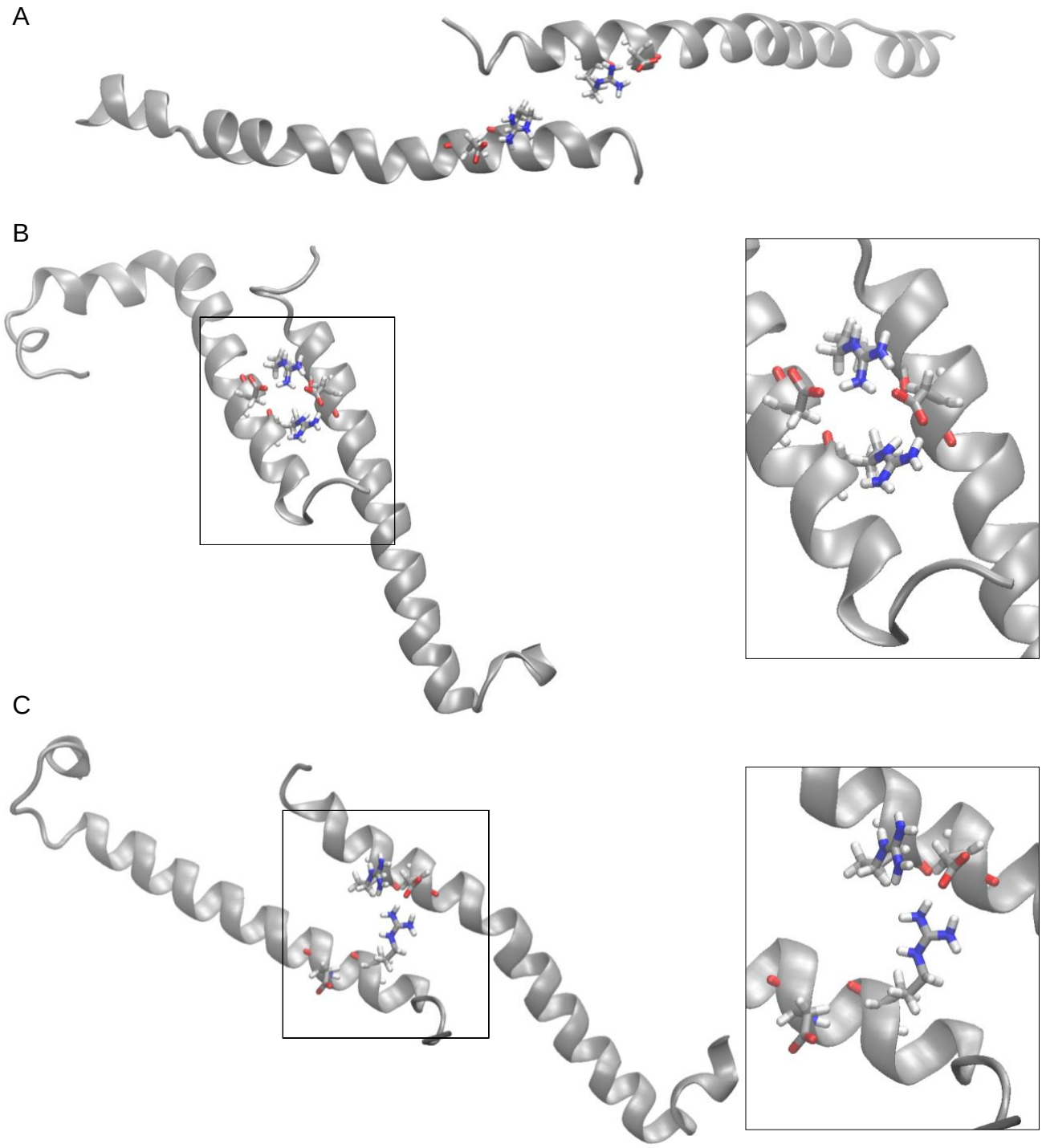

Fig. S9. Intermolecular salt bridges between Arg10 and Asp13 underlying strong interaction between NTDs (residues 1–43). (*A*) Initial condition for both Simulations 1 and 2. (*B*) Simulation 1 at 20 μs. (*C*) Simulation 2 at 10 μs. The salt bridges are magnified in the right boxed panels. Nitrogen of guanidinium group of Arg is in blue; oxygen of Asp sidechain is in red.

Table S1. Intermolecular distances between cross-linked residues of Proteins 1 and 2 averaged over 16–20 μs of Simulation 1 (rHDL-2-100). Distances that are less than cutoff (15.1 Å) are in bold typeface and demonstrate the consistency of structure with cross-links. The cross-links that are observed in Simulation 1 are also consistent with LL5/5 registry. Conversely, the cross-links that are not observed in Simulation 1 are consistent with LL5/4 registry. The last column lists the relative peak intensity of MS/MS signals. Peak intensities are only suggestive of the abundance of cross-links because ionization and cross-linking efficiencies are highly variable in different peptides.

Cross-link	Protein 1–Protein 2, C_α–C_α (Å)	Protein 2–Protein 1, C_α–C_α (Å)	Observed in Simulation 1	Consistent with LL5/4	Relative Intensity
K96–E147	25.9 (19.3,32.4)	29.7 (22.9,33.3)	No	Yes	Medium
K96–D168	**13.7** (**10.4**,17.8)	15.6 (**10.3**,21.0)	Yes	No	Medium
K96–E169	**11.7** (**8.2**,16.2)	**14.2** (**7.8**,19.9)	Yes	No	Strong
K107–E125	42.5 (32.5,48.9)	44.6 (39.3,49.1)	No	Yes	Weak
K107–D157	17.0 (**12.0**,22.8)	15.3 (**12.7**,20.4)	Yes	No	Medium
K118–E147	**13.5** (**8.0**,19.2)	**11.2** (**7.5**,**15.1**)	Yes	No	Strong
K133–E111	27.1 (24.0,31.6)	26.5 (21.0,30.1)	No	Yes	Weak
K133–E125	**10.7** (**6.3**,**15.0**)	**9.9** (**6.1**,15.5)	Yes	No	Medium
K140–E125	**10.8** (**7.1**,16.1)	**11.0** (**6.7**,16.2)	Yes	No	Strong
K195–E70	**13.6** (**7.9**,19.4)	**10.3** (**6.9**,16.0)	Yes	No	Medium

Table S2. Intramolecular distances between cross-linked residues of Proteins 1 and 2 averaged over 16–20 μs of Simulation 1 (rHDL-2-100). Distances that are less than cutoff (15.1 Å) are in bold typeface and determine the consistency of structure with cross-links. The last column lists the relative peak intensity of MS/MS signal (see comment regarding relative peak intensities in the caption of Table S1).

Cross-link	Protein 1, C_α–C_α (Å)	Protein 2, C_α–C_α (Å)	Observed in Simulation 1	Relative Intensity
D48–K23	21.1 (**13.6**,39.0)	19.7 (**13.6**,35.9)	Yes	Medium
K45–E34	**12.4** (**7.6**,20.5)	15.5 (**11.7**,20.1)	Yes	Medium
K12–E76	53.8 (40.2,77.5)	47.9 (37.9,80.0)	No	Medium
E198–K208	**11.9** (**6.9**,19.4)	16.3 (**13.1**,20.6)	Yes	Medium
K208–E223	22.9 (20.3,25.2)	23.0 (20.2,25.7)	No	Medium
K238–E34	87.6 (56.0,104.3)	43.1 (25.7,43.1)	No	Weak

Table S3. Distances calculated for rHDL-2-110 in Simulation 2 for three critical cross-links that were observed experimentally in rHDL-2-100 and were consistent with Simulation 1. Distances that are less than cutoff (15.1 Å) are in bold typeface and determine the consistency of structure with cross-links.

Cross-link	Protein 1, C_α–C_α (Å)	Protein 2, C_α–C_α (Å)
D48–K23	24.6 (22.5,27.4)	27.5 (21.4,32.5)
K45–E34	**11.9** (**9.0**,16.2)	**13.3** (**10.8**,**15.0**)
E198–K208	18.8 (15.2,24.7)	15.7 (**14.4**,19.1)

References

1. Jo S, Kim T, Iyer VG, Im W (2008) CHARMM-GUI: A web-based graphical user interface for CHARMM. *J Comput Chem* 29:1859–1865.
2. Jorgensen WL, Chandrasekhar J, Madura JD, Impey RW, Klein ML (1983) Comparison of simple potential functions for simulating liquid water. *J Chem Phys* 79:926–935.
3. Durell SR, Brooks BR, Bennaim A (1994) Solvent-induced forces between two hydrophilic groups. *J Phys Chem* 98:2198–2202.
4. Noskov SY, Roux B (2008) Control of ion selectivity in LeuT: Two Na+ binding sites with two different mechanisms. *J Mol Biol* 377:804–818.
5. Luo Y, Roux B (2009) Simulation of osmotic pressure in concentrated aqueous salt solutions. *J Phys Chem Lett* 1:183–189.
6. Venable RM, Luo Y, Gawrisch K, Roux B, Pastor RW (2013) Simulations of anionic lipid membranes: Development of interaction-specific ion parameters and validation using NMR data. *J Phys Chem B* 117:10183–10192.
7. Bashtovyy D, Jones MK, Anantharamaiah G, Segrest JP (2011) Sequence conservation of apolipoprotein AI affords novel insights into HDL structure-function. *J Lipid Res* 52:435–450.
8. Lippert RA, et al. (2013) Accurate and efficient integration for molecular dynamics simulations at constant temperature and pressure. *J Chem Phys* 139:164106.
9. Hoover WG (1985) Canonical dynamics: Equilibrium phase-space distributions. *Phys Rev A* 31:1695–1697.
10. Martyna GJ, Tobias DJ, Klein ML (1994) Constant pressure molecular dynamics algorithms. *J Chem Phys* 101:4177–4189.
11. Shaw DE, et al. (2014) Anton 2: Raising the bar for performance and programmability in a special-purpose molecular dynamics supercomputer. *Proceedings of the international conference for high performance computing, networking, storage and analysis*, (IEEE Press, Piscataway, NJ), pp 41–53.
12. Kräutler V, Van Gunsteren WF, Hünenberger PH (2001) A fast SHAKE algorithm to solve distance constraint equations for small molecules in molecular dynamics simulations. *J Comput Chem* 22:501–508.
13. Humphrey W, Dalke A, Schulten K (1996) VMD: Visual molecular dynamics. *J Mol Graph* 14:33–38.
14. Abraham MJ, et al. (2015) GROMACS: High performance molecular simulations through multi-level parallelism from laptops to supercomputers. *SoftwareX* 1:19–25.
15. Daura X, et al. (1999) Peptide folding: When simulation meets experiment. *Angew Chem Int Ed Engl* 38:236–240.
16. Schrodinger, LLC (2015) The PyMOL Molecular Graphics System, Version 1.8.
17. Phillips JC, et al. (2005) Scalable molecular dynamics with NAMD. *J Comput Chem* 26:1781–1802.

18. MacKerell Jr AD, et al. (1998) All-atom empirical potential for molecular modeling and dynamics studies of proteins. *J Phys Chem B* 102:3586–3616.
19. MacKerell Jr AD, Feig M, Brooks CL (2003) Improved treatment of the protein backbone in empirical force fields. *J Am Chem Soc* 126:698–699.
20. Best RB, et al. (2012) Optimization of the additive CHARMM all-atom protein force field targeting improved sampling of the backbone φ, ψ and side-chain χ1 and χ2 dihedral angles. *J Chem Theory Comput* 8:3257–3273.
21. Ryckaert J-P, Ciccotti G, Berendsen HJ (1977) Numerical integration of the cartesian equations of motion of a system with constraints: Molecular dynamics of n-alkanes. *J Comput Phys* 23:327–341.