Structural Studies of High Density Lipoprotein by Electron Microscopy and Flexible Fitting

Lin Zhu1, Jitka Petrlova2, Peter Gysbers3, Hans Hebert1, Stefan Wallin3, Caroline Jegerschöld4 and Jens O. Lagerstedt2

1. School of Technology and Health, Royal Institute of Technology, Huddinge, Sweden
2. Department of Experimental Medical Science, Lund University, Lund, Sweden
3. Department of Physics and Physical Oceanography, Memorial University of Newfoundland, St. John's, NL A1B 3X7, Canada
4. Department of Biosciences and Nutrition, Karolinska Institute, Huddinge, Sweden

Abstract

The structure of high density lipoprotein, HDL, was investigated in vitreous samples, frozen at cryogenic temperatures, as well as in negatively stained samples by transmission electron microscopy. The HDL was assembled in vitro from apolipoprotein A-I and 1-palmitoyl-2-oleoyl-sn-glycero-3-phosphocholine at a 1:50 molar ratio yielding a particle of 132 kDa and 9.6 nm diameter. Low resolution electron density maps were obtained showing a nearly circular, flat particle with an uneven particle thickness. Using these maps as restraints, flexible fitting of apolipoprotein A-I structures, derived from atomic resolution X-ray structures, was applied. Two models for the full-length structure of apolipoprotein A-I dimer in the lipid bound conformation are presented.

Keywords: Apolipoprotein, HDL, cryoEM, negative stain, flexible fitting

Introduction

High density lipoprotein (HDL, or “good cholesterol”) and its main protein component apolipoprotein A-I (apoA-I) are central to the reverse cholesterol transport (RCT) which is an integral part of the cholesterol transport and metabolism of the human body (1, 2). Both cholesterol and other lipids are transported in a soluble mode by a variety of high to low density lipoproteins and chylomicrons in the blood (3-6). HDL is negatively correlated to risk of cardiovascular disease and atherosclerosis, and has been shown to have beneficial effects in dia-
betes, obesity and metabolic disease as well as in inflammation (1, 4-11). The apoA-I protein is primarily produced by the liver, and to a lesser extent in the intestines. In the liver, the newly synthesized apoA-I protein interacts with the hepatic membranes to recruit lipids and form nascent HDL particles (12-14) that are secreted into the circulation. The formation of the nascent HDL particles is believed to be catalyzed by the cellular transmembrane ATP-binding cassette (ABC) receptors ABCA1 and ABCG1 in an energy driven process (15, 16). In this process, the interaction between the apoA-I protein and the receptors leads to an extraction of a patch of plasma membrane from the cell surface by the apoA-I protein. The formed protein-lipid complexes are composed of two apoA-I proteins that are arranged in a way that each copy of the apoA-I covers the acyl chains of each leaflet in the extracted bilayer patch. The shape of the nascent HDL particles would therefore be expected to be discoidal or elliptic.

The apoA-I molecule is composed of 243 amino acids and has a molecular weight of 28 kDa. The N-terminus (amino acid residues 1-43) and remaining sequence (amino acid residues 44-243) are encoded by two regions/exons of the gene, respectively (17). The 44-243-region comprises 10 tandem 11-mer/22-mer α-helical segments which are denoted as H1-H10 (18). Each α-helical segment contains one or two of the 11-residue long consensus sequences: homology A and homology B (19). The α-helical segments are thus, denoted in accordance with their locations in the apoA-I sequence as well as with the type of 11-mer/22-mer repeats. For instance, H2 contains two homology A, which are denoted as H1A1 and H1A2; H3 contains only one homology B, which is denoted as H3B; H4 contains one homology A and one homology B, which are denoted as H4A and H4B.

Several experimental approaches have been employed with the goal to describe the structure of HDL particles, and specifically the structural fold of apoA-I in these particles. Transmission electron microscopy (TEM) was one of the first methods used where HDL discs (14, 20) as well as HDL spheres (14, 21) were visualized by the use of sodium phosphotungstate (NaPT) for negative staining. However, it became evident that the use of NaPT induced a specific aggregation of phospholipid surfaces resulting in extended stacks of particles, or discs, that then formed a “rouleaux” structure (14). This phenomenon was not observed with the use of another common stain, uranyl formate (UF) where each HDL particle instead was well separated from its neighbor, oriented with its flat plane adhering to the support (22). The early TEM studies thus provided a first understanding of the general shape of the particles. Other methodologies such as X-ray crystallography (18, 23, 24), electron paramagnetic spectroscopy (EPR)(24, 25) covalent cross-linking combined with mass-spectroscopy (26-30),
FRET analyses (31), hydrogen/deuterium (H/D) exchange (32), or combination of NMR, EPR and TEM (33) have later been successfully used to further our understanding of the apoA-I protein in HDL. These important advances have enabled researchers to propose structure models of the apoA-I protein in association with phospholipids. The models generally agrees on that two apoA-I proteins circulate the phospholipid bilayer in an antiparallel fashion in the nascent HDL particle. However, detailed structure knowledge, including how the N- and C-termini are organized, still remains to be clearly elucidated.

Cryo-EM for biomolecule structure determination is an approach that has gained significant recognition over the last years and detailed structures of large macromolecules such as the ribosome (reviewed in (34)), the mammalian mitochondrial complex I (35) or the γ-secretase (36) have been determined with this technology. However, cryo-EM has been less used for structure analyses of biomolecules of smaller sizes such as the HDL particle (about 130 kDa and 8-12 nm in diameter depending on protein-to-lipid ratio) as they are relatively hard to detect in the vitreous ice (37). To overcome this limitation, Murray and colleagues used a high phospholipid-to-apoA-I protein ratio to create large HDL-like particles suitable for cryo-EM analysis, and presented a circular disc with 36 nm diameter and 4.5 nm thick as expected for a DMPC bilayer (38). Zhang and Ren used another approach, electron cryotomography, to study a HDL particle with a diameter of 11 nm and showed a circular disc of intermediate resolution (39).

We here used an approach where cryo-EM structure analysis was combined with in silico molecular dynamics simulation to study the structure of the apoA-I protein in 9.6 nm rHDL particles, reconstituted from apoA-I and the phospholipid POPC (1-palmitoyl-2-oleoyl-sn-glycerol-3-phosphatidylcholine). We applied cryo-EM on unstained vitrified preparations of rHDL, negative stain EM (ns-EM) on the rHDL as well as statistics to characterize the dynamic structure of rHDL. To interpret the obtained densities derived from the cryo-EM particles, we developed and applied a flexible docking procedure based on biased molecular simulations of apoA-I by using a “cross-correlation” function between calculated and experimental density maps. The underlying computational model combines Monte Carlo simulation techniques and an implicit-solvent energy function, and has been used previously in various applications, such as protein-peptide binding (40), mechanical unfolding (41) and disordered proteins including the 1-93 fragment of apoA-I (42). Initial structures for the flexible docking procedure were constructed using the determined crystal structure of fragment 3-184 of apoA-I (18). Two final structure models are presented depicting two apoA-I proteins arranged in an
anti-parallel double belt fashion that covers the edges of the acyl chains of the lipids in the bilayer. Specifically, the structure models illuminate important aspects of the fold of the N- and C-termini in the HDL particles.

Materials and Methods

Materials

POPC (1-palmitoyl-2-oleoyl-3-sn-glycero-phosphatidylcholine) was bought from Avanti polar lipids. UF was from Polysciences, Inc., and NaPT was from Sigma-Aldrich.

Preparation of homogenous 9.6 nm rHDL

A bacterial expression system consisting of pEXP-5 plasmid in *Escherichia coli* strain BL21(DE3) pLysS cells (Invitrogen) was used to produce the apoA-I WT proteins (Uniprot code of the human apoA-I: P02647), as previously described (43, 44). After purification of apoA-I proteins on Ni\(^{2+}\)- chelated columns (GE Healthcare) and desalting to remove imidazole, TEV protease treatment was employed to cleave the His-tag. This was followed by a second Ni\(^{2+}\)- column passage where TEV protease and the cleaved His-tag were retained on the column. The flow-through containing cleaved apoA-I proteins were desalted into TBS buffer, (pH 7.4), concentrated with 10 kDa molecular weight cut-off Amicon Ultra centrifugal filter devices (Millipore) and stored at 4°C prior to use. Protein purity was confirmed by sodium dodecyl sulfate (SDS)-polyacrylamide gel electrophoresis with Coomassie blue staining and protein concentrations determined by Nanodrop, using molecular weight and extinction coefficients (45, 46). A solution of 22 mM sodium deoxycholate (in 0.5 ml) was added to an equal volume of 16.3 mM 1-palmitoyl-2-oleoyl-sn-3-glycero-phosphocholine (POPC) in Tris-buffered saline (TBS), pH 7.4. The mixture was vortexed and incubated at 37 °C until clear. ApoA-I (3 mg) was added to the solution (1:4:50 (w/w) apoA-I : sodium deoxycholate : POPC), followed by incubation at 37 °C for 1 hour. Reconstitution was done with the POPC : apoA1 molar ratio of 50:1 (instead of the commonly used 80:1) as the lower POPC content promotes the formation of a highly homogenous population of 9.6 nm particles. Sodium deoxycholate was removed by extensive dialysis against TBS, pH 7.4. rHDL particles were recovered by FPLC gelfiltration. Fractions were collected and the size of rHDL was determined by native blue gel electrophoresis (Invitrogen)(31). The 9.6 nm particles from the FPLC peak
fractions resulted in concentrations of typically 0.05 – 0.1 mg/ml that could be used directly for EM. The preparations were stored in TBS at 4 °C and were used within a month for electron microscopy although negative stain specimens appeared homogenous for up to three months when kept at 4 °C (47).

**Preparation of grids**

For negative stain, 200-mesh copper grids (Ted Pella Inc. USA) were used with manually coated continuous carbon film, whereas for cryo-EM R2/4 holey carbon film-coated 400-mesh copper grids (Quantifoil) were used. Both types of grids were glow-discharged for 20 seconds and 30 mA before use.

**Negative stain EM and data processing**

A drop of 3.5µl sample was added per grid. After 30s incubation the surplus liquid was blotted off on Whatman filter paper. Immediately a drop of 1% uranyl formate (UF) was applied to the grid. After 30 s the surplus stain was blotted off and the grid was allowed to dry. Uranyl formate was prepared in the dark as in (48), aliquotted and frozen until use. Data was collected at 200 kV on a Jeol SF2100F equipped with a 4K x 4K CCD camera (Tiez Video and Imaging Systems GmbH, Germany) at 69,500 x magnification. The pixel size of the CCD camera is 15 µm which gives a corresponding value of 2.08 Å/pix on the specimen level for the EM images. From 5 micrographs, 1358 particles were manually boxed with BOXER in the EMAN2 software (see below) (49) with box size 84 pixels corresponding to 17.5 nm, based on which 98 2D projections were generated through 2D Class-averaging.

**Cryo-EM and data processing**

Grids were rapidly frozen in the controlled environment of a Vitrobot (FEI, model: FP 5350/60). 3.5µl was added on the grid for 60 s incubation under the 22 °C temperature and 100% humidity environment. After blotting, the grid was rapidly plunged into liquid ethane for freezing and then transferred into liquid nitrogen for storage. Data was collected at 200 kV on a Jeol SF2100F electron microscope equipped with a 4K x 4K CCD camera (Tiez Video and Imaging Systems GmbH, Germany) at 69500x magnification. The pixel size of the CCD camera is 15 µm which gives a corresponding value of 2.08 Å/pix on the specimen level for the EM images. For cryo-EM, 223 images were collected for processing by EMAN2 and
RELION (50). 6050 particles were manually boxed with a box size 26.6 nm. Processing for 2D class averages was done in EMAN2 using “no-shrink” to avoid loss of structure information, followed by the reconstruction of 3D initial map. The original particle set and 3D initial map obtained from EMAN2 were used for the subsequent 3D refinement in RELION.

Monte Carlo simulations

Simulations of apoA-I dimers are performed with the PROFASI package (51). Two new energy terms were implemented, such that simulations could be carried out with a modified potential energy function,

\[ E = E_0 + E_{\text{disc}} + E_{\text{cc}} , \]

where \( E_0 \) is the standard energy function of the protein model (52), \( E_{\text{disc}} \) is a term for protein-lipid interactions through an implicit representation of the central bilayer disc, and \( E_{\text{cc}} \) is a “cross-correlation” term score the difference between experimental and calculated electron density maps. The two terms, \( E_{\text{disc}} \) and \( E_{\text{cc}} \), are described in detail below.

Simulations were carried out using fixed-temperature Metropolis Monte Carlo. Three different Monte Carlo update types were included: (1) pivot moves, which turn individual backbone torsional angles (\( \varphi, \psi \)), (2) sidechain rotations that turn individual \( \chi \)-angles, and (3) a semilocal move that turns up to 8 consecutive \( \varphi, \psi \)-angles in a coordinated way such that conformational changes are approximately local (53). Backbone torsional angles of residues 110-153, corresponding to alpha helices H4-H6, are held fixed. All other angles, including all sidechain angles, are free to change. All simulations are carried out 330 K.

Implicit lipid bilayer model

To take into account the interaction between apoA-I and lipid molecules in a computationally convenient way, we introduce an implicit representation of the lipid bilayer disc of HDL particles. The disc is geometrically represented as a cylinder with dimensions realistic for 9.5 nm reconstituted HDL particles. It is meant to represent the hydrophobic core region of the bilayer, i.e., the lipid tails, and has a thickness of 30 Å. The cylinder is made impenetrable to all non-hydrogen backbone and sidechain atoms of the protein, except some of the sidechain atoms of amino acids Pro, Tyr, Val, Ile, Leu, Met, Phe, Trp, Arg, and Lys. The non-H atoms
that can penetrate the cylinder are those classified as hydrophobic in the model, e.g., C$_\beta$, C$_{\gamma 1}$, C$_{\gamma 2}$ and C$_\delta$ of Ile and C$_\beta$, C$_\gamma$ and C$_\delta$ of Lys, (see Table 4 in Ref. PMC Biophysics 2, 2 (2009). This allows the sidechains of amino acids with hydrophobic character to partially penetrate the lipid bilayer, as has been observed in previous studies (54).

To mimic the effective hydrophobic attraction between exposed lipid tails (cylinder side) and hydrophobic amino acids, we used the energy term

$$E_{\text{disc}} = -k_{\text{disc}} \sum_i e^{-(d_i - \sigma)^2/2},$$

where the sum goes over all C$_\beta$ atoms of Pro, Tyr, Val, Ile, Leu, Met, Phe, Trp, Arg, and Lys, and $d_i$ is the minimal distance between atom $i$ and the cylinder side surface, and $\sigma = 3$ Å. The strength of the interaction was chosen to be $k_{\text{disc}} = 1$.

Simulations with two different cylinder base area geometries were carried out: (1) a circular shape with radius 35 Å, and (2) a “stadium” shape with width 70 Å and length 78 Å, i.e., two half circles (radii 35 Å) separated by 8 Å and an 8 x 70 Å$^2$ rectangle in the space between the half circles. These two shapes are referred to as circular and elongated in the text, respectively.

**Calculated electron density maps and correlation coefficient**

Simulations were biased using the experimental electron density maps with the term

$$E_{\text{cc}} = k_{\text{cc}} (1-r^2),$$

where

$$r^2 = \frac{\left(\sum_{ijk} \rho_{i,j,k}^{\text{exp}} \rho_{i,j,k}^{\text{calc}}\right)^2}{\sum_{ijk} \left(\rho_{i,j,k}^{\text{exp}}\right)^2 \sum_{ijk} \left(\rho_{i,j,k}^{\text{calc}}\right)^2}$$

is the “cross-correlation” between experimental $\rho^{\text{exp}}$ and calculated $\rho^{\text{calc}}$ electron densities, and $k_{\text{cc}}=50$ controls the strength of the bias. The sums are taken over all voxels $ijk$. In determining
the calculated electron density $\rho^{\text{calc}}$, we essentially followed the approach of Tama et al (55). Thus, we calculated the electron density of a model conformation $C$ of voxel $ijk$ using

$$
\rho^{\text{calc}}_{i,j,k} = \sum_{n=1}^{N} \int_{V_{ijk}} d\mathbf{r} e^{-3|r-r_n|^2/2\sigma},
$$

where the sum goes over all non-H atoms $n$, $r^n$ is the position of atom $n$, the integration is taken over the voxel volume $V_{ijk}$, and $\sigma$ represents the resolution of the calculated density. Because the resolution of density map calculated this way is roughly $2\sigma$ (56), we set $\sigma = 10$ Å.

Because of the relatively low resolution of our EM data, and for computational reasons, we chose a voxel side length of 6 Å in calculations of the $\rho^{\text{calc}}$ and $r^2$. This side length is larger than the 2 Å side obtained from the experimental data. Before simulations, we therefore coarse-grained the EM-derived density maps by averaging over 3x3x3 voxels to obtain the value in a 6 Å-voxel. Cross-correlations calculations with varying voxel sizes have been carried out by Zheng (57), and revealed only minor impact on the quality of fits for voxel sizes up to around 5 Å for maps with higher resolution than ours.

**Results**

**rHDL reconstitution**

HDL particles were prepared by an established reconstitution described in materials and methods. A lipid-to-protein ratio of 80:1 (POPC:apoA-I) was first used to generate the rHDL particles with an expected diameter of 9.6 nm. While the 9.6 nm rHDL particles were formed at this lipid-to-protein ratio, inspection by TEM of the negative stained samples of the rHDL preparations indicated that an unsatisfactory high level of large membranous aggregates was also present (not shown). A lower relative level of lipids was therefore tested to overcome this. Indeed, a ratio of 50:1 for POPC:apoA-I resulted in minimum presence of the lipid aggregates and a maintained 9.6 nm rHDL particle integrity as judged by native PAGE gel-electrophoresis (supplementary Fig. SI_1).
Negative staining EM of rHDL particles

Negative staining is frequently used to verify homogeneity and to identify potential aggregation problems before performing the cryo-EM analyses. However, when carefully made, negative stain may reveal useful structural information (58). For example, negative staining of rHDL particle has provided information on deviations from circularity, in particular with the use of a uranyl salt as the stain (22), which was also used here for initial assessment of the formed 9.6 nm rHDL particles. We did not observe aggregation or “stacking” with the use of uranyl acetate or uranyl formate (UF). However, the rHDL particles appeared to have a preferred orientation after staining by any of the uranyl salts (Fig. 1A; UF is used as the stain in the figure), with an apparent adherence by the membrane side (or phospholipids head group) to the carbon film support. Consequently, essentially only views perpendicular to the membrane surface were observed (top-views) (Fig. 1A). Based on these particles, class-average projections obtained from image processing (Fig. 1B) showed objects that deviated from perfect circularity and that also presented some edginess, suggesting an elliptical shape of the particles. Next, cryo-EM was used to provide additional details of the rHDL particle structures.

Fig. 1. The reconstituted HDL appears to deviate from circularity as some contain “edges” or “one corner”. (A) rHDL were stained with UF where the raw image (left) shows a homogenous population without aggregates. Scale bar 100 nm in the raw image (left) and 50 nm in the magnified cutout (right). (B) From the raw images 1358 particles were selected and processed to give 2D class-averages. The panel in (B) shows the result for 98 class-average projections with a box size of 17.5 nm.
Cryo-EM

High resolution structural information is possible to achieve by cryo-EM (59, 60), however, the smaller the particles are, the harder the detection is and as a consequence low resolution is typically achieved. The buffer composition may also affect the quality of the contrast. In our analyses tris-buffer provided somewhat better contrast of the rHDL particles than phosphate-buffer (not shown), hence tris-buffer was used. The fact that the snap-freezing leading to rapid immobilization of the rHDL particles in solution removed the risk for the biased orientation obtained in the negative stain EM analyses. rHDL particles could thus be viewed in different orientations by conventional cryo-EM (Fig. 2A) and were selected for processing (Fig. 2B).

Fig. 2 In vitreous ice, rHDL discs may be viewed from several angles although side views provide more contrast and are therefore easier to identify. (A) A raw cryo-EM image obtained at liquid nitrogen temperatures and 200 keV by a CCD camera. Scale bar is 200 nm. (B) A magnified section from the raw cryo-EM image where a selection of rHDL particles has been boxed (White boxes: tilted or side views. Green boxes: Top views. Box size: 26.6 nm). (C) Two-dimensional class-average projections of rHDL. Box size: 26.6 nm.
A total number of 6050 particles were used for generation of 2D class-averages (Fig. 2C). According to the generated projections, most of them appeared to show side views of rHDL particles, whereas fewer showed clearly tilted views or top views.

To obtain a 3-dimensional reconstruction of the particle, usually an initial 3D map is first reconstructed from the 2D class-average projections. This 3D initial map provides a preliminary notion of the particle structure and can be used as a reference map in the following refinement step as. Alternatively, in order to avoid the impact on the final refinement results from unexpected structure features in the early stages of 3D map building, the usage of a featureless initial map is sometimes preferred as the reference for the refinement.

For the nascent HDL particle, although detailed structural information is still not clear, a disc-like overall shape is the consensus of most researchers (12) (28) (26). Both C64 and C8 (octagon) symmetries, two levels of symmetrization, were applied in the construction of initial map, which resulted in two near circular disc-like maps with few structural features (suppl SI_2). These two highly symmetrized initial maps were used as the reference maps for the following 3D refinement.

The 3D refinement was processed in Relion (50). Apart from the reference map, a set of raw particles are likewise required as the input. Before processed in the refinement, the same particle set (6050 particles) was pre-evaluated through 2D classification step in Relion. A large number of bad particles were manually removed due to low signal-to-noise ratio (SNR), unmatched size and shape, and other unexpected structural features. The remaining particle set contained 3870 particles and was used for the 3D refinement, in which each particle was separately compared to the initial map and contributed to the building of a new and refined 3D map. In addition, the initial maps were low-pass filtered to 60 Å before usage. D1 symmetry was applied in the refinement. The two refined maps – D1C64 and D1C8, are shown in Fig. 3.
Fig. 3 Three-dimensional refinements of rHDL based on two different reference maps shows deviations from circularity. The nascent HDL was proposed to assume a disc shape where two copies of apoA-I encircle the acyl chain layer of the phospholipid bilayer patch that has been extracted from the plasma membrane (18). The two maps are very similar as expected with a shape that deviates from circularity viewed perpendicular to the membrane. Viewed along the membrane a thicker side and 180° to this a thin side where a zig zag wave is apparent. In the refinement, both maps were symmetrized in D1. A) Refined map D1C64. Resolution: 17 Å. B) Refined map D1C8. Resolution: 16 Å. Scale bar is 10 nm.

The two refined maps share similarities as expected but also present some differences. Both have an elongated shape, which is different from the circular shape of class-average projections of particle top-views obtained by negative staining EM (Fig. 1B), and show an uneven distribution of density - thicker on one side and thinner on the other side (Fig.3, middle right). By comparing the two refined maps, D1C64 (Fig. 3A) appears slightly less elliptic than D1C8 (Fig. 3B) with somewhat more pronounced features. The cryo-EM derived maps together with published structural information (18) were next used in a flexible docking procedure to obtain a model for how the two apoA-I molecules are arranged in the 9.6 nm high density lipoprotein.
Flexible fitting procedure

In order to structurally interpret the above data, we developed a Monte Carlo-based procedure for fitting apoA-I dimer structures into the cryo-EM derived electron density maps. The overall aim of this procedure was to generate a collection of physically sound structures with good fits to the EM data that could be further assessed using additional experimental data. To this end, simulations of apoA-I dimers were carried out using an implicit-solvent, all-atom protein model (52) with a modified potential energy function, $E = E_0 + E_{cc} + E_{disc}$, to take into account the fit between calculated and experimental electron density maps ($E_{cc}$), and the interaction between protein and lipids via an implicit representation of the lipid bilayer disc ($E_{disc}$). Details are given below and in Methods. To account for the overall shape of the obtained electron density maps, we consider two different shapes of these discs, circular and elongated cylinders. The elongated disc has a ratio of around 1.11 between the long and short axes, chosen to be similar to the ratio of the EM density maps. We carried out the flexible fitting procedure on the D1C64 and D1C8 density maps, for which the D1 symmetry has been applied.

Initial structures of apoA-I dimers

The starting structures for our simulations are shown in Figure 4. From the $\Delta$(185-243) X-ray dimer structure of Mei and Atkinson (18) we made the following modifications. First, to create a full-length version of the apoA-I protein, residues 1-2 and 185-243 were added to both chains. The C terminal tail regions (residues 185-243 = 58 aa), containing the helical segments H8, H9 and H10, were constructed in an $\alpha$-helical conformation by modeling them on the helices H4-H7. Second, the N terminal helical bundle was “unhinged” by pivoting around the flexible linker region Gly65/Pro66. This unhinging has been suggested to occur during the lipid loading process of HDL upon the transition from small (~8.4 nm) to mid-sized (~9.6nm) double-belt conformations (18, 61). In this way, we created an “open” structure which is poised to wrap around a lipid bilayer disc without assuming any particular tertiary structure in the N terminal regions. In the final step, we positioned the H4-H6 regions of the initial dimer manually on the thin side of the EM-derived density maps, D1C64 and D1C8 (see Figure 4). Our motivation for this was twofold. First the “pivoting” of the first 1-66 segment on the N terminal bundle creates an overall structure with a thin (helices H3-H7 = 88-187) and a thick (H8-H10 and the 1-66 segment) side, which thus fits well with the overall shape of the EM-derived density maps. Second, the undulating character of the H4-H6 segment (a feature of
the Mei and Atkinson structure) makes it fit well into the thin side of both the D1C64 and D1C8 density maps, as can be seen in Figure 4.

Fig. 4 The initial apoA-I structure used for our simulations of rHDL, superimposed on the two D1 symmetry maps D1C64 and D1C8. Scale bar is 10 nm.
**Biased simulations and structure selections**

To achieve our final structural models for disc-shaped rHDL, we carried out the following steps:

1) A first round of biased simulations was carried out from the initial structures in Figure 4, using the simulation protocol described in Methods. Eight independent runs with 400,000 MC steps were performed for both the circular and elongated lipid discs. The results are shown in Figure 5. We found that, in all cases, the apoA-I chains wrapped around the lipid disc to form double-belt like structures. Although there are substantial variations in the structures obtained, we noted some common structural features. A particularly robust feature in our simulations is that the C-terminal helices of the two apoA-I chains (residues 186-243) tended to wrap tightly around the lipid disc and pack against in each (19 out of 32 runs). Another common feature is that a short alpha helical segment at around residues 55-65 tended to pack against the side of the lipid disc, such that it closes the “gap” produced by the termination of the H10 helix at the C terminus. Close interactions between C terminal alpha helices have been proposed to be present in HDL particles of various sizes (61) and a H10-H10 interaction is present in the X-ray structure of Borhani et al (23). Based on the above considerations, and on their fit with the electron density, we disregarded all structures that lacked a close packing between the C-terminal helices H9 or H10.
Fig. 5 A total of 32 structures generated from the first round of steered simulations. A) Results based on the EM-derived density map D1C64 and the implicit circular lipid bilayer disc. B) Results based on the EM-derived density map D1C64 and the implicit elongated lipid bilayer disc. C) Results based on the EM-derived density map D1C8 and the implicit circular lipid bilayer disc. D) Results based on the EM-derived density map D1C8 and the implicit elongated lipid bilayer disc. The structures selected for the second round simulation are boxed in blue color.
2) The remaining 19 structures were used as starting points for a second round of simulations, with the same number of MC steps. The minimum-energy conformations occurring in this second round of simulations were collected (Fig. 6).

Fig. 6 The results of second round simulations on the selected 19 structures. A) Results based on D1C64 and circular disc. B) Results based on D1C64 and elongated disc. C) Results based on D1C8 and circular disc. D) Results based on D1C8 and elongated disc. The two best structures are boxed in blue color (see text below).
3) These minimum-energy conformations were thereafter assessed for consistency with data from cross-correlation experiments (26, 28, 29, 31, 62). To this end, we collected all intermolecular pairs of amino acid positions that have been identified through cross-linking or fluorescence energy resonance transfer (FRET) spectroscopy in 9.6nm HDL particles (see Table S1) and that involved either the N terminal or C terminal regions. A score for each minimum-energy structure was obtained by calculating the distance $d$ for all experimentally verified cross-linked residue-residue pairs, and counting the number of pairs that were consistent with experiments (criterion $d < 26 \text{ Å}$) (Table S2-S5). We selected the top scoring structures for each of the D1C64 and D1C8 map which satisfy 17 and 16 distance restraints, respectively (out of 25 possible). Both top-scoring structural models have elongated lipid discs, although some structures with circular discs did also score well.

4) The structures obtained from simulations are generally not symmetrical. In particular, we found that they differ in the N-terminal region. Because of the dynamic nature of rHDL, some asymmetry in the apoA-I structures is not unexpected. However, presumably the apoA-I homodimer is symmetric in an averaged sense. We therefore “symmetrize” the two top-scoring structures by making the two monomers structurally similar in regions where they differ. Which of the two monomers was held fixed (and thus used as a template for the other monomer) for each part of the structure, was chosen such that the number of restraints with the cross-correlation data was maximized. After this symmetrization, the two selected structures satisfy 17 (D1C64) and 15 (D1C8) restraints, respectively.

5) Finally, the symmetrized structures were relaxed by an additional simulation at 298 K, carried out with only small step MC updates included (no pivot moves included).

The two final structural models of apoA-I restrained by cryo-EM and other known experimental data, are shown in Fig. 7.
Fig. 7 Two final structures of the apoA-I pair obtained from three independent simulations started from the initial structures in Figure 4 and fitted into corresponding EM density maps (parameter values, disc radius $R_{\text{disc}}=35$ Å, disc height $H=30$ Å, cross-correlation strength $k_{\text{EM}} = 100$, and disc interaction parameter $k_{\text{disc}} = 1$). A) The protein structure (magenta) based on the EM map D1C64 (cyan). B) The protein structure (green) based on the EM map D1C8 (pale lilac). Scale bar for both A and B is 10 nm.

There are some significant differences between the two models. In particular, the organization of the N terminal helical region is different in the two cases. In the D1C64 map, the segment 55-65 contacts the lipid disc. By contract, in the D1C8 map, the corresponding lipid region on the disc side is protected by the 5-23 residue segment. Interestingly, both models have some similarities to the mid-sized double-belt model proposed by Gursky (61).
Discussion

EM-derived density maps and flexible fitting template structure

The HDL particle size (≈10 nm) is near the detection limit of cryo-EM data collection and processing. Nevertheless, an attempt was made to obtain a three dimensional map of the apoA-I protein in discoidal HDL particles by cryo-EM on preparations of reconstituted HDL assembled in test tube from purified apoA-I protein and POPC phospholipids. The electron density obtained allowed for refinement by use of flexible docking of reference structures obtained from modifications of X-ray structures.

Both cryo-EM maps, D1C64 and D1C8, showed an uneven distribution of density and had elongated shapes. The uneven thickness of the particles observed by cryo-EM, could not be detected in UF stained samples but this was most likely due to the lack of side-views. The elongated shape is likely to have originated from specific structural features of the raw particle since the reference maps used for the cryo-EM data processing had more or less round shapes (Suppl SI_2). However, as several models propose a nearly circular shape of HDL (12) (26) (28), the deviations and uneven thickness raised some questions regarding whether this was a result of an uneven distribution of particle views biased towards a high number of side-views and a consequent steering of the 3D map towards an elliptic shape or if the rHDL particle actually has an oval shape as suggested by some results (63, 64).

Along the lines of the work by Murray et al (38) we therefore attempted to tilt the sample (65). However, the small size of these particles made it impossible to follow the rotation of a single rHDL particle. A clear side view, observed at 0° tilt, was lost after as little as a 30° tilt (not shown). Another approach to evaluate this was the use of carbon film in the cryo-EM sample preparation (65). This was used as the presence of carbon film could potentially have the effect to guide the orientation of the discs, so that more discs would lie flat on the carbon film, hence presenting a higher number of top-views. Indeed, the side-views disappeared, however top-views were not clearly observed and possibly the presence of the carbon film disturbed a clear view of particles at this orientation (not shown).

Despite the limited success in applying these two additional approaches, we concluded that the two 3D maps (Fig. 3) presented similar structural features and both had a shape deviating from circularity, which is in accordance with a number of class-average projections of top-
views obtained by negative stain (Fig. 1B). In support of our conclusion, the structure of the deletion mutant of apoA-I (33) also showed a somewhat elongated shape.

The resulting structure models with the two encircling apoA-I proteins arranged in an anti-parallel fashion with the two 43-residue N-terminal α-helices folding back so that there is a stretch with three helices on this side of the disc (i.e., an extra density on this side of the disc rim as compared to the opposite side where the H5-H5 registry is located).

**X-ray crystal structures of apoA-I as initial reference for flexible fitting to cryo-EM density maps**

Another critical step in the generation of HDL structure models was the selection of initial reference for the flexible fitting. Among the discoidal models for nascent HDL, most support the notion that the two apoA-I encircles the lipid bilayer disc in an anti-parallel fashion so that the N- and C-termini overlap on one side (61) (12) (66) (67). This suggested us to apply D1 symmetry in the refinement of the two EM density. We initially also considered the so called double superhelix model, which is a structure model of full-length apoA-I in lipid-bound form derived from negative stain EM and SANS data (64). However, this structure has been questioned using molecular dynamics (62), which in addition to the convincing published data on the discoidal, anti-parallel organization, led us to not further evaluate this model as reference in our fitting procedure. Instead we focused on apoA-I crystal structures.

Two high-resolution apoA-I structures have been determined by X-ray crystallography. Both structures are obtained from truncated apoA-I variants that either lack the N-terminal domain ([44-243]apoA-I) (18), this variant crystallized as a tetramer, or lack the C-terminal domain ([1-184]apoA-I) (23), this variant crystallized as a dimer. The differences in fold and oligomerization between these apoA-I structures, lacking either the first 43 residues on the N-terminus or the last 59 residues on the C-terminus, are striking. In the N-terminal truncation, the structure is open and nearly circular. In this structure, two apoA-I molecules form one anti-parallel pair and two such pairs dimerize into a tetrameric organization in a manner to cover all hydrophobic surfaces (23). It was speculated that the presence of a high salt concentration during the crystallization process possibly induced this structural arrangement. The potentially salt-driven interaction could also be the reason for the unexpectedly high (> 90 %) helicity observed in this structure (23). Thereeto, as has recently been shown (44), high concentrations of apoA-I in solution induces oligomerization of the apoA-I protein and increased helical structure content, possibly due to coiled-coil interactions. This would also suggest a
general caution when attempting to deduce the apoA-I structure from crystals which by nature have very high densities of protein.

Compared to the [44-243]apoA-I structure, the [1-184]apoA-I has a more compact structure, in particular in the region of where the N-terminal 66 residues folds back on the H2-H4 segment (18). In addition, in this structure, two copies form a dimer with H5 in register. Hence, the [1-184]apoA-I dimer was used for modeling by manipulations of the structure similar to those reported by Mei and Atkinson (18) and discussed by Gursky (61). The thereby derived initial structure, shown in Fig. 4, was thus deemed suitable to use for flexible fitting in our cryo-EM map in several steps.

Interestingly, both cryo-EM maps had a thin side displaying an undulating shape (Fig. 7, rightmost). The thickness and shape corresponded to amino acid stretch 110-154 (H4B-H5-H6A with the Atkinson depiction (68)) in the [1-184]apoA-I dimer and the initial structure was manually placed here to make a good fit and kept static during the flexible fitting procedure. A potential limitation of the static arrangement is that the locked positions of amino acids 110-154 contributed to a reduced freedom in the flexibility of residues 99-109 and 155-161, which were indeed minimally altered in the flexible fitting procedure.

**Organization of the N- and C-termini on the rHDL discs**

The consensus regarding the N-terminal half of the protein is that it stabilizes the lipid free conformation of apoA-I whereas the C-terminal part is highly involved in initial lipid binding. The C-terminal region was not present in the crystal structure used in our flexible fitting and hence residues 185 to 243 were manually added and freely flexible before fitting into the cryo-EM maps. Looking at the position of hydrophobic residues in the structure model D1C64 residues G186-Q243, a Leu189 is the first residue to point inwards toward the lipid bilayer acyl chains. As every third-fourth residue is hydrophobic and the helicity is not interrupted there is a hydrophobic surface along this stretch of residues (11 Leu residues altogether) (Fig. 8). Helix 9-10 (P209-Q243) has been proposed to have the highest lipid affinity of all apoA-I helices, and Helix 10 (P220-Q243) to be directly involved in interactions with ABCA1 and the lipid bilayer in the formation of the nascent HDL disc (69, 70).
Sections within residues 1-109 were also observed to move in the fitting. Helical segments H4A (P99-109), H3 (P99-109) and H2A2 (G78-87) moved only slightly relative to the reference structure, likely due to stabilization by interaction of hydrophobic side-chains to the lipidphilic core. In contrast, residues upstream to domain H2A2 exhibited larger dynamic changes. In the [1-184]apoA-I crystal structure a large number of stabilizing salt bridges, aromatic clusters and pi-cation connections stabilize the interaction between the N-terminal 35-residue kinked α-helix (residues 1-35) and the kinked α-helical segment 67-110, and also the coil/helix that connects these two (18). Interestingly, although the entire segment 1-66 had been manually displaced outwards in the starting structure (Fig. 4) it folded around the central “elongated cylinder” to provide a structure (Fig. 7) similar to predicted structural organizations (61, 71) and experimentally supported (24). In the latter, EPR data proposes that 43 N-terminal amino acids “fold back” and interact with α-helices downstream of the first 43 residues on the same apoA-I molecule and with the C-terminal α-helix on the other apoA-I molecule in the dimeric HDL (24). Collectively, this implies that the 9.6 nm HDL-disc would be thicker on one side, which is consistent with our two structure models. Indeed, the similar “folding back” conformation of the N-terminus can be observed in the cryo-EM D1C64 model (Fig. SI_4). In this model, the H1B1 (residues 44-55) forms a coil and the second part of H1, the H1B2 (residues 56-65) forms an α-helix. This is similar to the EPR derived secondary structure (25) which generally agrees with the cryo-EM structure models presented here. Both experimental approaches indicate residues 6-34 as α-helix and residues 36 to about 50 as random coil or β-strand structure. A major difference between the two is that the EPR data indicates β-strand structure at residues 149-157, which is not seen in the cryo-EM models. However, it should be noted that this stretch of amino acid residues falls within the locked posi-
tions in the molecular fitting procedure (as described above) which has very little or no degree of freedom motion. The notion that the N-terminal domain has a unique fold and is separated from the regions that circulate the lipid bilayer was further supported by Denisov et al (72). The authors reported that deletion of either the first 11 residues (H1B1) or the first 22 (H1B1 plus H1B2) of the apoA-I protein still allowed for reconstitution of 9.6 nm rHDL particles as shown by size-exclusion chromatography and SAXS analysis. Recently the three-dimensional structure of a shortened apoA-I and was determined by use of a combination of NMR, EPR and TEM data (33). The shortened version of apoA-I lacked residues 1-54 and H5 (residues 121-142). Two copies of the shortened apoA-I protein was shown to form a double-belt structure with apparent edges. This is in line with our current negative stain results (Fig. 1B) and with earlier images of negatively stained HDL (22). It was also concluded that the deletion of H5 did not affect the capability to form HDL particles. However, to maintain a hydrophobic interior, a specific rotation of the helices was necessary. A molecular model of discoidal HDL based on hydrogen-deuterium exchange mass spectrometry has been proposed that supports the antiparallel double-belt structure (73). In this model, residues 159–180 of each apoA-I form a so-called “solar flare” region, which is a highly open and solvent-exposed loop. The current cryoEM structures did not suggest this flexibility, however it cannot be excluded that this region of the model is biased by the [1-184]apoA-I crystal reference structure.

**Conclusion**

The low-resolution EM density maps are the first 3D maps of 9.6nm rHDL discs determined by cryo-EM. The overall shape and unevenly-distributed density of the maps provide structural insights on the conformation of apoA-I molecules in nascent HDL. Based on the EM maps and a published X-ray crystal structure, structure models of apoA-I were generated through a multi-step biased simulation procedure. The novel full-length structures of apoA-I dimers provide support to previous structural studies of apoA-I and HDL, and deepens the understanding to the structure-function relationship of nascent HDL with significance for the prevention of lipoprotein-related disease. In addition, the biased simulation method used in this study provide a powerful and convenient modelling tool with applicability for structural studies and modelling of other proteins and protein complexes.
References


Supplementary Materials

Fig. SI_1. Reconstitution of rHDL with a ratio of 50:1 POPC:apoA-I yields fractions with a disc size of 9.6 nm despite the lower amount of POPC available compared to the frequently used ratio of 80:1.

Fig. SI_2. Reference maps of rHDL obtained by symmetrization of initial maps.
Fig. SI_3. Fourier Shell Curve plots for assessment of resolution. A. FSC for the EM map D1C64, the resolution is 17.4 Å. B. FSC for the EM map D1C8, the resolution is 15.6 Å. Both resolutions were estimated at FSC=0.143.

Fig. SI_4. Top view of structure model from D1C64 cryoEM map where helices were coloured according to the scheme in Atkinson and later Gursky.
<table>
<thead>
<tr>
<th>Linked residues</th>
<th>Method</th>
<th>Type and spacer arm length</th>
<th>Estimated backbone-to-backbone distance</th>
<th>Reference</th>
</tr>
</thead>
<tbody>
<tr>
<td>Silva et al Biochemistry 2005</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>K208-K208</td>
<td>X-link + MS</td>
<td>BS3, 11.3Å</td>
<td>25Å</td>
<td>Silva et al 2005</td>
</tr>
<tr>
<td>K88-K118</td>
<td>X-link + MS</td>
<td>BS3, 11.3Å</td>
<td>25Å</td>
<td>Silva et al 2005</td>
</tr>
<tr>
<td>K96-K118</td>
<td>X-link + MS</td>
<td>BS3, 11.3Å</td>
<td>25Å</td>
<td>Silva et al 2005</td>
</tr>
<tr>
<td>K40-K239</td>
<td>X-link + MS</td>
<td>BS3, 11.3Å</td>
<td>25Å</td>
<td>Silva et al 2005</td>
</tr>
<tr>
<td>K118-K140</td>
<td>X-link + MS</td>
<td>BS3, 11.3Å</td>
<td>25Å</td>
<td>Silva et al 2005</td>
</tr>
<tr>
<td>K59-K208</td>
<td>X-link + MS</td>
<td>BS3, 11.3Å</td>
<td>25Å</td>
<td>Silva et al 2005</td>
</tr>
<tr>
<td>K226-K239</td>
<td>X-link + MS</td>
<td>BS3, 11.3Å</td>
<td>25Å</td>
<td>Silva et al 2005</td>
</tr>
<tr>
<td>K59-K195</td>
<td>X-link + MS</td>
<td>BS3, 11.3Å</td>
<td>25Å</td>
<td>Silva et al 2005</td>
</tr>
<tr>
<td>K77-K195</td>
<td>X-link + MS</td>
<td>BS3, 11.3Å</td>
<td>25Å</td>
<td>Silva et al 2005</td>
</tr>
<tr>
<td>Bhat et al JBC 2005</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>K45-K208</td>
<td>X-link + MS</td>
<td>DSP, 12Å</td>
<td>25Å</td>
<td>Bhat et al 2005</td>
</tr>
<tr>
<td>K59–K195</td>
<td>X-link + MS</td>
<td>DSP, 12Å</td>
<td>25Å</td>
<td>Bhat et al 2005</td>
</tr>
<tr>
<td>K59-K206</td>
<td>X-link + MS</td>
<td>DSP, 12Å</td>
<td>25Å</td>
<td>Bhat et al 2005</td>
</tr>
<tr>
<td>K59-K208</td>
<td>X-link + MS</td>
<td>DSP, 12Å</td>
<td>25Å</td>
<td>Bhat et al 2005</td>
</tr>
<tr>
<td>K77-K182</td>
<td>X-link + MS</td>
<td>DSP, 12Å</td>
<td>25Å</td>
<td>Bhat et al 2005</td>
</tr>
<tr>
<td>K7–K195</td>
<td>X-link + MS</td>
<td>DSP, 12Å</td>
<td>25Å</td>
<td>Bhat et al 2005</td>
</tr>
<tr>
<td>K88-K182</td>
<td>X-link + MS</td>
<td>DSP, 12Å</td>
<td>25Å</td>
<td>Bhat et al 2005</td>
</tr>
<tr>
<td>K226-K238</td>
<td>X-link + MS</td>
<td>DSP, 12Å</td>
<td>25Å</td>
<td>Bhat et al 2005</td>
</tr>
<tr>
<td>K226-K239</td>
<td>X-link + MS</td>
<td>DSP, 12Å</td>
<td>25Å</td>
<td>Bhat et al 2005</td>
</tr>
<tr>
<td>Bhat et al Biochemistry 2007</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>K12-K182</td>
<td>X-link + MS</td>
<td>DSP, 12Å</td>
<td>DSP, 26.3Å</td>
<td>Bhat et al 2007</td>
</tr>
<tr>
<td>K40-K239</td>
<td>X-link + MS</td>
<td>DSG, 7.7Å; DSP, 12Å</td>
<td>DSG, 22.3Å; DSP, 26.6Å</td>
<td>Bhat et al 2007</td>
</tr>
<tr>
<td>Martin et al JBC 2006</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>W50-A230</td>
<td>FRET, Trp to AEDANS</td>
<td></td>
<td>22.7Å</td>
<td>Martin et al 2006</td>
</tr>
<tr>
<td>W72-A210</td>
<td>FRET, Trp to AEDANS</td>
<td></td>
<td>23.5Å</td>
<td>Martin et al 2006</td>
</tr>
<tr>
<td>W90-A190</td>
<td>FRET, Trp to AEDANS</td>
<td></td>
<td>24Å</td>
<td>Martin et al 2006</td>
</tr>
<tr>
<td>Jones et al JBC 2010</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>W40-L240</td>
<td>FRET, Trp to AEDANS</td>
<td></td>
<td>28.34Å</td>
<td>Jones et al 2010</td>
</tr>
</tbody>
</table>

**Table S1.** Published interpeptide interactions in 9.6 nm rHDL identified by cross-linking or FRET spectroscopy, and used for structure model evaluation in this study.
Table S2. Number of inter-molecular distance restraints (see Table S1) satisfied in the obtained minimum-energy structures from the second round of simulations based on the D1C64 map and a circular disc. The numbering of structures (Struct 1 to 5) corresponds structures from left to right in Fig. 6A. Any amino acids pair \(ij\) identified as a constraint in these datasets can, in principle, be satisfied in two ways for non-symmetrical dimer structures. We therefore treat the two cases \(i<j\) and \(j>i\) as separate constraints for the purpose of scoring our structures.

<table>
<thead>
<tr>
<th>Data set</th>
<th>Struct 1</th>
<th>Struct 2</th>
<th>Struct 3</th>
<th>Struct 4</th>
<th>Struct 5</th>
</tr>
</thead>
<tbody>
<tr>
<td>Silva et al 2005</td>
<td>8</td>
<td>7</td>
<td>8</td>
<td>8</td>
<td>7</td>
</tr>
<tr>
<td>Bhat et al 2005</td>
<td>13</td>
<td>9</td>
<td>13</td>
<td>14</td>
<td>14</td>
</tr>
<tr>
<td>Bhat et al 2007</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>Martin et al 2006</td>
<td>0</td>
<td>1</td>
<td>3</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>Jones et al 2010</td>
<td>2</td>
<td>0</td>
<td>2</td>
<td>0</td>
<td>1</td>
</tr>
<tr>
<td><strong>TOTAL</strong></td>
<td><strong>23</strong></td>
<td><strong>17</strong></td>
<td><strong>26</strong></td>
<td><strong>22</strong></td>
<td><strong>22</strong></td>
</tr>
</tbody>
</table>

Table S3. Number of inter-molecular distance restraints (see Table S1) satisfied in the obtained minimum-energy structures from the second round of simulations based on the D1C64 map and an elongated disc. The numbering of structures (Struct 1 to 5) corresponds structures from left to right in Fig. 6B.

<table>
<thead>
<tr>
<th>Data set</th>
<th>Struct 1</th>
<th>Struct 2</th>
<th>Struct 3</th>
<th>Struct 4</th>
<th>Struct 5</th>
</tr>
</thead>
<tbody>
<tr>
<td>Silva et al 2005</td>
<td>11</td>
<td>6</td>
<td>8</td>
<td>9</td>
<td>7</td>
</tr>
<tr>
<td>Bhat et al 2005</td>
<td>16</td>
<td>7</td>
<td>12</td>
<td>13</td>
<td>12</td>
</tr>
<tr>
<td>Bhat et al 2007</td>
<td>1</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>Martin et al 2006</td>
<td>1</td>
<td>1</td>
<td>3</td>
<td>1</td>
<td>1</td>
</tr>
<tr>
<td>Jones et al 2010</td>
<td>1</td>
<td>1</td>
<td>0</td>
<td>2</td>
<td>0</td>
</tr>
<tr>
<td><strong>TOTAL</strong></td>
<td><strong>30</strong></td>
<td><strong>15</strong></td>
<td><strong>23</strong></td>
<td><strong>25</strong></td>
<td><strong>20</strong></td>
</tr>
</tbody>
</table>

Table S4. Number of inter-molecular distance restraints (see Table S1) satisfied in the obtained minimum-energy structures from the second round of simulations based on the D1C8 map and a circular disc. The numbering of structures (Struct 1 to 4) corresponds structures from left to right in Fig. 6C.

<table>
<thead>
<tr>
<th>Data set</th>
<th>Struct 1</th>
<th>Struct 2</th>
<th>Struct 3</th>
<th>Struct 4</th>
</tr>
</thead>
<tbody>
<tr>
<td>Silva et al 2005</td>
<td>8</td>
<td>9</td>
<td>4</td>
<td>9</td>
</tr>
<tr>
<td>Bhat et al 2005</td>
<td>9</td>
<td>13</td>
<td>6</td>
<td>12</td>
</tr>
<tr>
<td>Bhat et al 2007</td>
<td>1</td>
<td>1</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>Martin et al 2006</td>
<td>2</td>
<td>1</td>
<td>2</td>
<td>1</td>
</tr>
<tr>
<td>Jones et al 2010</td>
<td>1</td>
<td>1</td>
<td>1</td>
<td>1</td>
</tr>
<tr>
<td><strong>TOTAL</strong></td>
<td><strong>19</strong></td>
<td><strong>25</strong></td>
<td><strong>13</strong></td>
<td><strong>23</strong></td>
</tr>
</tbody>
</table>

Table S5. Number of inter-molecular distance restraints (see Table S1) satisfied in the obtained minimum-energy structures from the second round of simulations based on the D1C8 map and an elongated disc. The numbering of structures (Struct 1 to 5) correspond structures from left to right in Fig. 6D.

<table>
<thead>
<tr>
<th>Data set</th>
<th>Struct 1</th>
<th>Struct 2</th>
<th>Struct 3</th>
<th>Struct 4</th>
<th>Struct 5</th>
</tr>
</thead>
<tbody>
<tr>
<td>Silva et al 2005</td>
<td>8</td>
<td>9</td>
<td>8</td>
<td>8</td>
<td>7</td>
</tr>
<tr>
<td>Bhat et al 2005</td>
<td>11</td>
<td>15</td>
<td>12</td>
<td>13</td>
<td>10</td>
</tr>
<tr>
<td>Bhat et al 2007</td>
<td>0</td>
<td>1</td>
<td>0</td>
<td>0</td>
<td>2</td>
</tr>
<tr>
<td>Martin et al 2006</td>
<td>1</td>
<td>1</td>
<td>1</td>
<td>0</td>
<td>1</td>
</tr>
<tr>
<td>Jones et al 2010</td>
<td>1</td>
<td>2</td>
<td>2</td>
<td>1</td>
<td>2</td>
</tr>
<tr>
<td><strong>TOTAL</strong></td>
<td><strong>21</strong></td>
<td><strong>28</strong></td>
<td><strong>23</strong></td>
<td><strong>22</strong></td>
<td><strong>22</strong></td>
</tr>
</tbody>
</table>