A robust all-atom model for LCAT generated by homology modeling[S]

Jere P. Segrest,[1],*,† Martin K. Jones,*,† Andrea Catte,[2],*,† and Saravana P. Thirumuruganandham[3],*,†

Center for Computational and Structural Dynamics* and Department of Medicine,† University of Alabama at Birmingham, Birmingham, AL 35294-0012

Abstract   LCAT is activated by apoA-I to form cholesteryl ester. We combined two structures, phospholipase A₂ (PLA₂) that hydrolyzes the ester bond at the *sn-2* position of oxidized (short) acyl chains of phospholipid, and bacteriophage tubulin PhuZ, as C- and N-terminal templates, respectively, to create a novel homology model for human LCAT. The juxtaposition of multiple structural motifs matching experimental data is compelling evidence for the general correctness of many features of the model: *i*) The N-terminal 10 residues of the model, required for LCAT activity, extend the hydrophobic binding trough for the *sn-2* chain 15–20 Å relative to PLA₂. *ii*) The topography of the trough places the ester bond of the *sn-2* chain less than 5 Å from the hydroxyl of the catalytic nucleophile, S181. *iii*) A β-hairpin resembling a lipase lid separates S181 from solvent. *iv*) S181 interacts with three functionally critical residues: E149, that regulates *sn-2* chain specificity, and K128 and R147, whose mutations cause LCAT deficiency.[Jlr] Because the model provides a novel explanation for the complicated thermodynamic problem of the transfer of hydrophobic substrates from HDL to the catalytic triad of LCAT, it is an important step toward understanding the antiatherogenic role of HDL in reverse cholesterol transport.—Segrest, J. P., M. K. Jones, A. Catte, and S. P. Thirumuruganandham. **A robust all-atom model for LCAT generated by homology modeling.** *J. Lipid Res.* 2015. **56:** 620–634.

Supplementary key words apolipoproteins • cholesterol/efflux • high density lipoprotein • phospholipases/A2 • lecithin:cholesterol acyltransferase

Of the possible mechanisms suggested to explain the atheroprotective role of apoA-I, a process called reverse cholesterol transport is most completely understood at the molecular level (1). Activation of the plasma enzyme LCAT (2–4) by apoA-I on nascent discoidal HDL (dHDL) is necessary for esterification of the unesterified cholesterol (UC) molecules of HDL to cholesteryl ester (CE)

and leads to the conversion of phospholipid-rich dHDL to spheroidal CE-rich (circulating) HDL, a central step of reverse cholesterol transport (1).

Unique among enzymes, LCAT is responsible for catalyzing two distinct steps in the formation of plasma-derived CE (4). In the first step, apoA-I presents the *sn-2* acyl chain of medium- to long-chain phosphatidylcholine (PC) with about 90% fidelity (5, 6) to the hydrolytic active site of LCAT to allow fatty acid hydrolysis of the *sn-2* acyl chain, formation of an acyl ester enzyme intermediate with the catalytic residue S181, and release of lysolecithin (2). This is the rate limiting step in the process (7); short-chain (soluble) phospholipids can be hydrolyzed by LCAT with no requirement for apoA-I to activate the enzyme. In the second step, the acyl ester enzyme intermediate reacts with the 3β-hydroxyl group of UC to form the insoluble product CE. Functionally, LCAT is the major source of plasma-derived CE and is responsible for transforming nascent or lipid-poor HDL into spherical HDL containing a core of CE with an exterior composed of apoA-I, phospholipid, and UC. If no LCAT is present, little to no spherical HDL is formed (8, 9).

The lack of a crystal structure of full length LCAT complicates any attempt at understanding the molecular mechanisms of its activation by apoA-I. A homology model

of a portion of the structure of the 416 residue long LCAT chain, residues 73–210, was published 15 years ago by Peelman et al. (3, 4, 10). More recently, Sensi et al. (11) published a homology model for full length LCAT that is composed of an amalgam of homology and ab initio fragments developed by a step-by-step selection of best fit models. A key criterion used for best fit was spatial clustering of the three residues of the catalytic triad proposed by Peelman et al. (10). Although S181 is known to be the nucleophile in LCAT (12), the other residues of the proposed triad, D345 and H377, have not been established.

Here we explore the creation of a robust homology model for LCAT using the multi-alignment routine of the program MODELLER and two recently determined crystal structures homologous to portions of LCAT as templates. Derivation of this model required many fewer assumptions than employed by either Peelman et al. (10) or Sensi et al. (11). A single dual alignment generates multiple models essentially all possessing structural attributes matching LCAT features previously determined by biochemical and biophysical methods (3, 4). We then examined the best models produced through molecular dynamics (MD) simulations and normal-mode analysis (NMA) and explored the molecular mechanisms whereby LCAT associates with HDL to create a tunnel for transfer of lipids from HDL to the LCAT active site.

MATERIALS AND METHODS

Homology modeling calculations

Homology models search of the Research Collaboratory for Structural Bioinformatics (RCSB) Protein Data Bank (PDB) was carried out on the Basic Local Alignment Search Tool (BLAST) server by selecting the protein BLAST program and by using the position-specific iterated BLAST (PSI-BLAST) algorithm. We found the best template for the C-terminal domain of human LCAT (residues 51–416) to be the X-ray crystal structure of human lipoprotein-associated phospholipase A_2 (Lp-PLA$_2$) (RCSB PDB identification (ID): 3D59) (13). The best template for the N-terminal domain of human LCAT (residues 1–50) was the X-ray crystal structure of tubulin-like protein, PhuZ, from bacteriophage 201φ2-1 (RCSB PDB ID: 3RB8) (14).

In order to assemble the best templates for the N-terminal and the C-terminal domains of LCAT, the model of human LCAT was built using the homology modeling software MODELLER (15–18) version 9.11. The structural model of the human LCAT target sequence was based on the alignment against these two templates, which was performed using a Python script of MODELLER, and as a graphical user interface the molecular graphics program UCSF Chimera (19) version 1.8. Finally, we used the discrete optimized protein energy (DOPE) potential to evaluate the best 10 models (defined as those with minimal DOPE values) out of the 100 structures generated by MODELLER, using the "loopmodel" function.

In a parallel experiment, the disulfide bonds were specified by the "self.patch" function within the "special_patches" routine in a modification to the "loopmodel" function, and then another 100 structures were generated by MODELLER and the best 10 models were selected as above.

Multiple sequence alignment

The sampling of the sequence similarity of human LCAT with protein sequences from other species was performed using as a seed the sequence of human LCAT and the PSI-BLAST algorithm. Non-redundant protein sequence databases were searched, and general parameters were set using values for the expected threshold (20) of 10 and a word size of 3; the scoring parameters used were BLOcks SUbstitution Matrix 62 (BLOSUM62) protein substitution matrix (21) with costs of gap opening and extension of 12 and 1, respectively. The search generated more than 40 sequences from fish, amphibians, birds, and mammal species.

The sequences similar to LCAT were aligned using a combination of UCSF Chimera and Clustal 2 (22, 23). As far as regards the software Clustal 2, we used the graphical user interface ClustalX (24–26) version 2.1 in order to produce Fig. 4. Sequences were aligned by performing a complete alignment in the multiple alignment mode.

NMA

NMA is a direct way to analyze the vibrational motion of proteins based on a structural analysis (27–30). In NMA calculations, the potential function is approximated by a sum of quadratic terms in the displacements around a minimum energy conformation.

We utilized the Normal Mode Analysis, Deformation, and Refinement (NOMAD-REF) web server for all atomic NMA (31). NMA was performed on the homology model of LCAT coordinates consisting of 3,331 atoms as reference structures in combination with the following parameters for the elastic network model for the first 36 modes. All interactions were weighted by a distance-weight parameter of 5 Å, and for the nonbonding interactions a cut-off of 10 Å was chosen. We obtained modes in PDB formats and inspected them using visual molecular dynamics (VMD) (32) for the specific motions that are described in results.

MD

All-atom minimization and MD simulations were used to explore the dynamic motion and measure the stability of the models we developed. The MD was performed using NAMD 2.9 (33) as described (34, 35).

To obtain m5 from m5*, 10,000 steps of conjugate-gradient energy minimization were performed to remove possible close contacts between atoms after specifying harmonic constraints of 1 kcal/mol on all backbone atoms except for residues within five of the disulfide-linked cysteines. In addition, secondary-structure restraints (as determined by the initial structure from MODELLER) were also specified by using the VMD (32) plug-in "SSRestraints" to ameliorate disruption during minimization. Hydrogen bonds were not restrained. These restraints were then used by NAMD's "extrabonds" command.

To obtain m5p, a POPC molecule was manually docked into the hydrophobic trough of m5, then 10,000 steps of conjugate-gradient energy minimization were performed after freezing all protein atoms in place.

To obtain m5d, residues 150–153 were removed from m5p, then E149-E154 were joined to form a continuous protein backbone. Then the structure was minimized for 10,000 steps while freezing residues 103–112, 122, 149, 181, 208, 288, and 342 to prevent disruption of the lipid. Next, while applying harmonic constraints of 1 kcal/mol on all residues >10 Å from N, O, and P atoms of the lipid, as well as the terminal methyls of the lipid, the system was subjected to 10,000 steps of conjugate-gradient energy minimization, heated to 310 K, and then simulated for 5.5 ns, giving m5d. In addition, this final structure was also solvated, minimized for 10,000 steps, heated to 310 K, and simulated for

another 5 ns without harmonic constraints. Finally, m5d was also simulated for another 1 ns without solvent and without any harmonic constraints.

For the simulation of the four residue peptide, consisting of residues 179–182, the initial unsolvated structure was minimized for 1,000 steps, heated to 310 K, and then simulated for 1 ns.

RESULTS

Creation of the homology model

When we performed a homology search against human LCAT (**Fig. 1A**), the top scoring homology match of residues 1–50 of human LCAT was to residues 2–51 of tubulin PhuZ, bacteriophage 201phi2-1 [Fig. 1B; RCSB PDB ID: 3R4V (14)], a segment that forms a β-sheet/α-helix motif (Fig. 1C). In spite of a similarity/identity of only 30/14%, the alignment was good with only two gaps of two residues each (**Fig. 2**).

We found a best match of residues 51–416 of LCAT to residues 54–427 of Lp-PLA$_2$ [Fig. 1D–F; RCSB PDB ID: 3D59 (13)], a member of the α/β-hydrolase fold family characterized by a β-sheet core of eight or so strands connected by α-helixes to form an α/β/α sandwich (36). Lp-PLA$_2$, also known as platelet-activating factor acetylhydrolase,

is similar to LCAT in that it tightly associates with HDL to hydrolyze the ester bond at the *sn-2* position of oxidized (short) acyl chains of phospholipid. Further, LCAT can also transesterify and hydrolyze platelet-activating factor-2 or oxidized PC, each of which has shorter chains than POPC in the *sn-2* position. The similarities make Lp-PLA$_2$ a credible template for LCAT.

The crystal structure of Lp-PLA$_2$ has a catalytic triad of S273, H351, and D296 (Fig. 1F), a short (∼7 Å) aromatic-rich binding pocket for the short *sn-2* oxidized acyl chain (Fig. 1D), and an acyl trough for binding the longer *sn-1* chain (Fig. 1E). Again, in spite of an even lower similarity/identity than tubulin PhuZ (20/7%), the alignment was good with only eight gaps, six of one to three residues, and two of four and five residues (Fig. 2). Importantly, there is only a single gap of one residue in the alignment of the core sequence (37) of LCAT (residues 106–260, green box in Fig. 2).

The chief inaccuracies in homology modeling derive from errors in sequence alignment and improper template selection (38). In spite of weak identity, because of a core alignment with only one gap, core regions tend to have conserved structural folds and are more amenable to homology modeling than peripheral loop regions (39), we used the multi-alignment routine in the homology program, MODELLER, to combine the Lp-PLA$_2$ structure
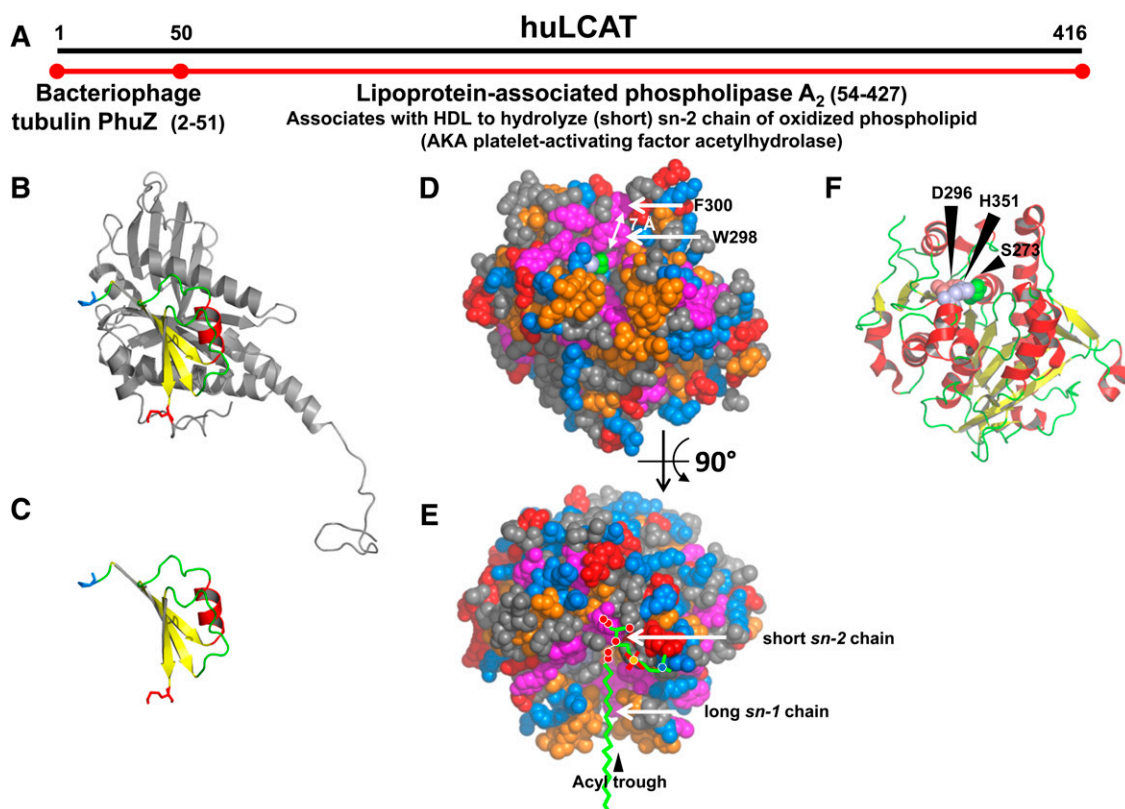
Fig. 1. Templates used to model LCAT. A: Multi-alignment used with MODELLER, 30% similarity/14% identity for PhuZ and 20% similarity/7% identity for Lp-PLA$_2$. B: Ribbons image of tubulin PhuZ. C: Ribbons image of residues 2-51 of tubulin PhuZ. D: Space filling image of PLA$_2$. Color code: Aromatic residues (magenta); hydrophobic residues (gold); basic residues (blue); acidic residues (red); binding site for *sn-2* chain (white double-headed arrow); S273 of catalytic triad (green). E: (D) rotated 90° showing proposed binding of human plasma platelet-activating factor based upon a crystal structure of the organophosphate compound paraoxon bound to PLA$_2$. F: Ribbons image of (D) showing active site triad.

Template `PVKVCLI FAGGTGMNVATKL VD LGEA VHCFD TC D K N VVD -` 39
LCAT `FWLLNVL FPPHT--TPKAEL S N HTRP VILVPG CL G N QLEA` 38

N-Terminal Domain

Template `-VHRSVN VTLTK-NGPYSVGC TDLMFDHTNKGT FLR---L` 74
LCAT `KLDKPDV VNWMCYRKTEDFFT IWLDLNMFLPLGVD CWIDN` 78

PhuZ ← → **Lp-PLA₂**

Template `YYPSQDN DRLDTLWIPNKEYFWGLS---KFLGTHWLMGNI` 111
LCAT `TRVVYNR SSGLVSNAPGVQIRVPGFGKTYSVEYLDSSKLA` 118

Template `LRLLFGS MTTPANWNSPLRPGEKYPLVVFSHGLGAFRTLY` 151
LCAT `GYLHTLV QNLVNNGYVRDETVRAAPYDWRLEPGQQEEYYR` 158

LID Domain

Template `SAIGIDLASHGFIVAAVEHRDRSASATYYFKDQSAAEIGD` 191
LCAT `KLAGLVEE MHAAYGKPVFLIGHSLGCLHLLYFLLRQ-PQA` 197

TRIAD Domain

Template `KSWLYLRTLKQEEETHIRNEQVRQRAKECSQALSLILDID` 231
LCAT `WKDRFIDGFISLGAPWGGSIKPMLVLASGDNQGIPIMSSI` 237

Template `HGKPVKNALDLKFDMEQLKDS ID REKIAVIGHSFGGATVI` 271
LCAT `KLKEEQR ITTTSPWMFPSRMAWP---EDHVFISTPSFN--` 272

Template `QTLSEDQRFRCGIALDA--WMFPLGDEVYSRIPQPLFFIN` 309
LCAT `---YTGRDFQRFFADLHFEEGWYMWLQSRDLLAGLPAPGV` 309

Template `SEYFQYPANIIKMKKCYSPDKERKMITIRGSVHQNFADFT` 349
LCAT `EVYCLYGVG----LPTPRTYIYDHGFPYTDPVGVLYEDGD` 345

Template `FATGKIIGHMLKLKGDIDSNVAIDLSNKASLAFLQKHLGL` 389
LCAT `DTVATRS TELCGLWQGR-QPQPVHLLPLHGIQHLNMVFSN` 384

Template `HKDFDQWDCLIEGDDENLIPGTNINTT------` 416
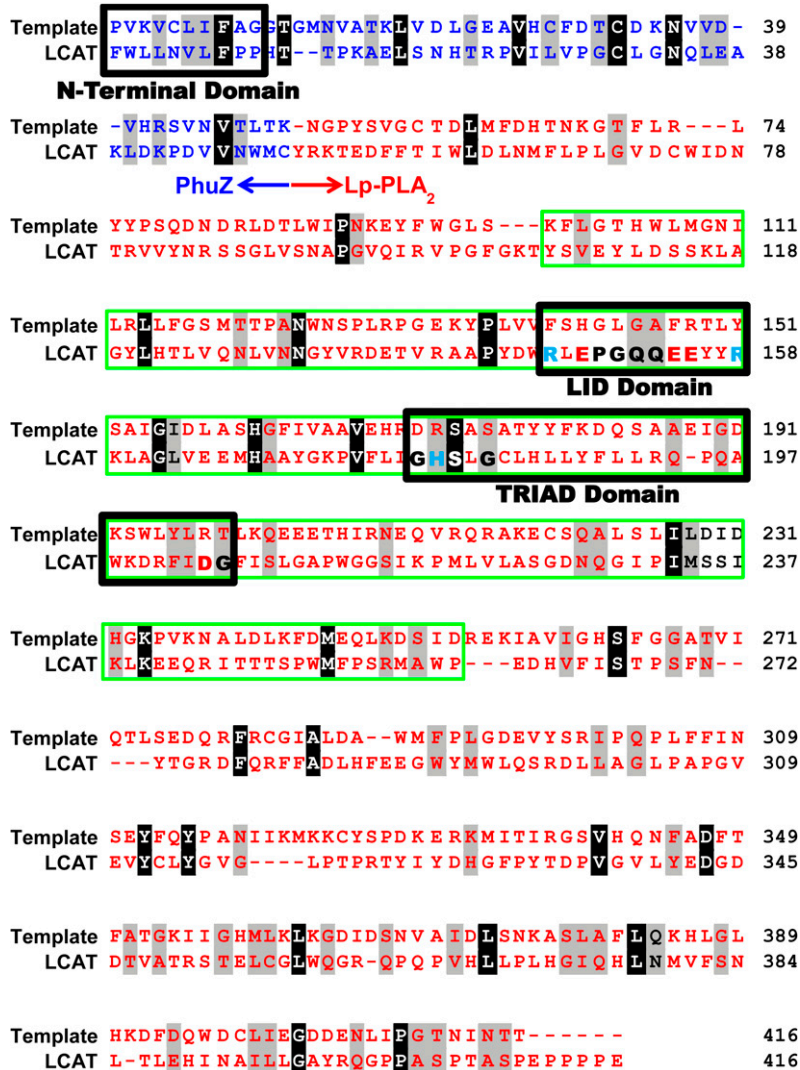**LCAT** `L-TLEHI NAILLGAYRQGPPASPTASPEPPPPE` 416

Fig. 2. Alignment of the dual template (PhuZ and Lp-PLA₂) with human LCAT for creation of homology model *m5*. The core domain, residues 107-260, is boxed in green. The three key domains, N terminus, lid, and triad are boxed in bold black. Key residues R147, E149-E155, R158, G179-S181, and D204-G205 are bolded. Identity between template and target is denoted by black, similarity by gray.

and the PhuZ structure as templates (Fig. 1A) to create homology models for human LCAT. After initial models looked promising, we proceeded to create a total of 100 models with MODELLER and chose the top 10 lowest energy models for further study. All 10 of the lowest energy models are quite similar, an alignment of the 10 structures is shown in **Fig. 3A**. Their alignments differ predominantly in the C-terminal domain, where eight prolines occur in the last fifteen residues of LCAT. The terminal six residues of this region are unmatched in the template (Fig. 2).

We examined each of these 10 models in detail using molecular graphics and, although each was similar, the fifth best scoring model (designated *m5**) was chosen for further modeling. It was immediately apparent that the model *m5** possessed a number of structural attributes that matched known LCAT features. We proceeded to the formation of disulfide bonds.

Disulfide bonds

Disulfide bonds were specified in *m5** between the correct cysteine disulfide pairs, C50-C74 and C313-C356 (3, 4), the cross-linked model was subjected to harmonic constraints and secondary-structure restraints, energy minimized, and

the resulting structure (designated *m5*) is shown in Fig. 3B. The cysteine residues were sufficiently close in *m5* that disulfide bond formation created no significant structural distortion following energy minimization; the relative motions of *m5* during energy minimization after disulfide bond formation are illustrated in supplementary Fig. 1. All four cysteines are completely conserved in all LCAT sequences, except for C356 that is exceptional only because it falls in a deleted region of one mammalian species (**Fig. 4**, red arrowheads).

We also created disulfide bonds using the alternate approach of specifying the disulfide bonds within MODELLER via a routine instead of within NAMD minimization. We created 100 disulfide-containing models and selected the 10 with the lowest energy for further examination. Alignment of the top 10 models is shown in supplementary Fig. 2a. Alignment of the models using the disulfide-forming routine was good, but not quite as good as models created with no disulfide bonds; in particular, the N-terminal domain was more mobile. We will compare the principle features of model *m5*, all top 10 disulfide-free models, and all top 10 models created using the disulfide-forming routine as we move through the Results.
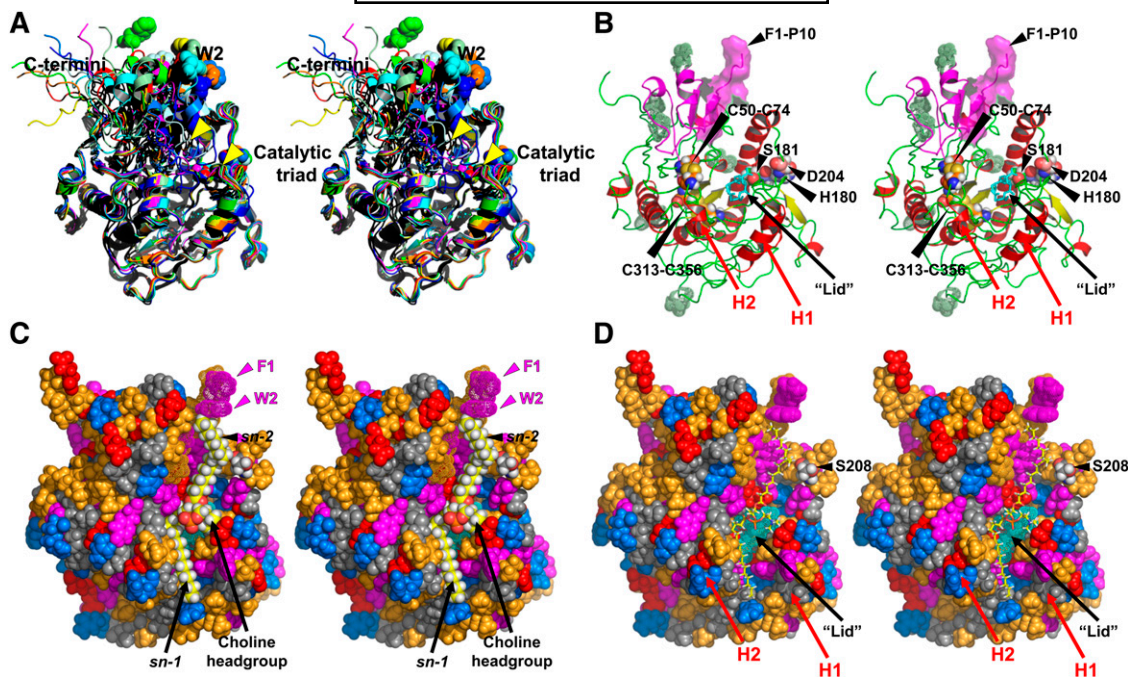
Fig. 3. Relaxed-eyed stereo images of LCAT homology model highlighting noteworthy features. A: Alignment of Cα carbons of the 10 lowest energy models (in 10 different colors) of 100 produced by the multi-alignment module of MODELLER. The 100% conserved W2 residue and the three residues corresponding to a catalytic triad are shown in spacefilling mode and the poorly aligned C-terminal domain is indicated. B: Ribbons image of model *m5* showing: active site triad as spacefilling, S181 (green), H180 (pale blue), and D204 (pink); disulfide bonds, C50-C74 and C313-C356 (gold spacefilling); lipase lid (cyan stick); PhuZ homologous region (residues 1–50) (magenta); residues 1–10 (magenta surface); known glycosylation sites (dotted pale green spacefilling). Helical domains, H1 and H2, located at the end of the *sn-1* trough, red arrows (see Fig. 8). C: Space filling image of model *m5* with same view as (B) containing a POPC molecule inserted into the hydrophobic crevice and energy minimized (see Materials and Methods). Color code: aromatic residues (magenta); hydrophobic residues (L, V, M, I, A, P) (gold); basic residues (blue); acidic residues (red); conserved residues F1 and W2 (magenta arrowheads); other LCAT residues (gray); POPC, (yellow spacefilling). D: Same image as (C) with docked POPC in stick format to show location of the lid (dotted spacefilling cyan). Helical domains, H1 and H2, located at the end of the *sn-1* trough, red arrows (see Fig. 8).

Secondary and tertiary structure. Model *m5* has a helical content of 25%, identical to the experimentally determined value of 24–25% (40) for human LCAT. Further, the model has features similar to other members of the α/β-hydrolase fold family, with 12 helical segments, 9 at least four residues in length, that surround a core of eight β-strands (Fig. 3B). Values for the 10 lowest energy models were similar. The models created without disulfide bonds have an average helical content of 27 ± 1.2% with an average of 15 helical segments surrounding a core of 13 β-strands. The models created with the disulfide-forming routine have an average helical content of 28 ± 1.5% with an average of 15 helical segments surrounding a core of 11 β-strands.

Hydrophobic PC binding trough

The five N-terminal residues, FWLLN, known to be required for LCAT activation by apoA-I (4), form in *m5* one end of an L-shaped trough lined by numerous aromatic and hydrophobic residues (Fig. 3B–D). The first 10 residues of LCAT, FWLLNVLFPP, are highly conserved and six (underlined) are completely conserved in all 45 LCAT orthologs that have been sequenced, from mammals to fish (see box in sequence alignment, Fig. 4). These 10 residues, shown as a transparent surface representation in Fig. 3B extend the 7 Å aromatic-rich binding pocket for the short *sn-2* oxidized acyl chain in Lp-PLA₂ (Fig. 1D) by

15–20 Å in *m5*. Examination of the alignment of the ten lowest energy states for the disulfide-free models, finds that in five models the hydrophobic trough is completely open, in three models the trough is partially closed by contact of the loops marked by yellow arrowheads (Fig. 3A), and in two models the trough closure is more extensive. In the top ten models created using the disulfide bond-forming routine (supplementary Fig. 2a), the hydrophobic trough is completely open in four models (supplementary Fig. 3), but is in varying states of closure in the remaining six models primarily due to contact of the loops marked by yellow arrowheads in supplementary Fig. 2a.

Remarkably, POPC, a good substrate for LCAT, can be docked into the L-shaped hydrophobic trough of *m5* in an orientation similar to that of platelet-activating factor or oxidized PC in the Lp-PLA₂ crystal structure (Fig. 1E). For docking, a POPC molecule was manually inserted into the trough in *m5* with the terminal methyl of the *sn-2* chain in Van der Waals contact with the completely conserved W2 of the N-terminal aromatic cluster and the *sn-1* chain and the polar head group bent at approximately 90° to the *sn-2* chain to fit snugly into the shorter end of the L-shaped hydrophobic trough. Then the structure was energy minimized with parts of the protein subjected to harmonic constraints, but the POPC allowed to move freely, and the results (designated *m5p*) are shown in Fig. 3C, D.
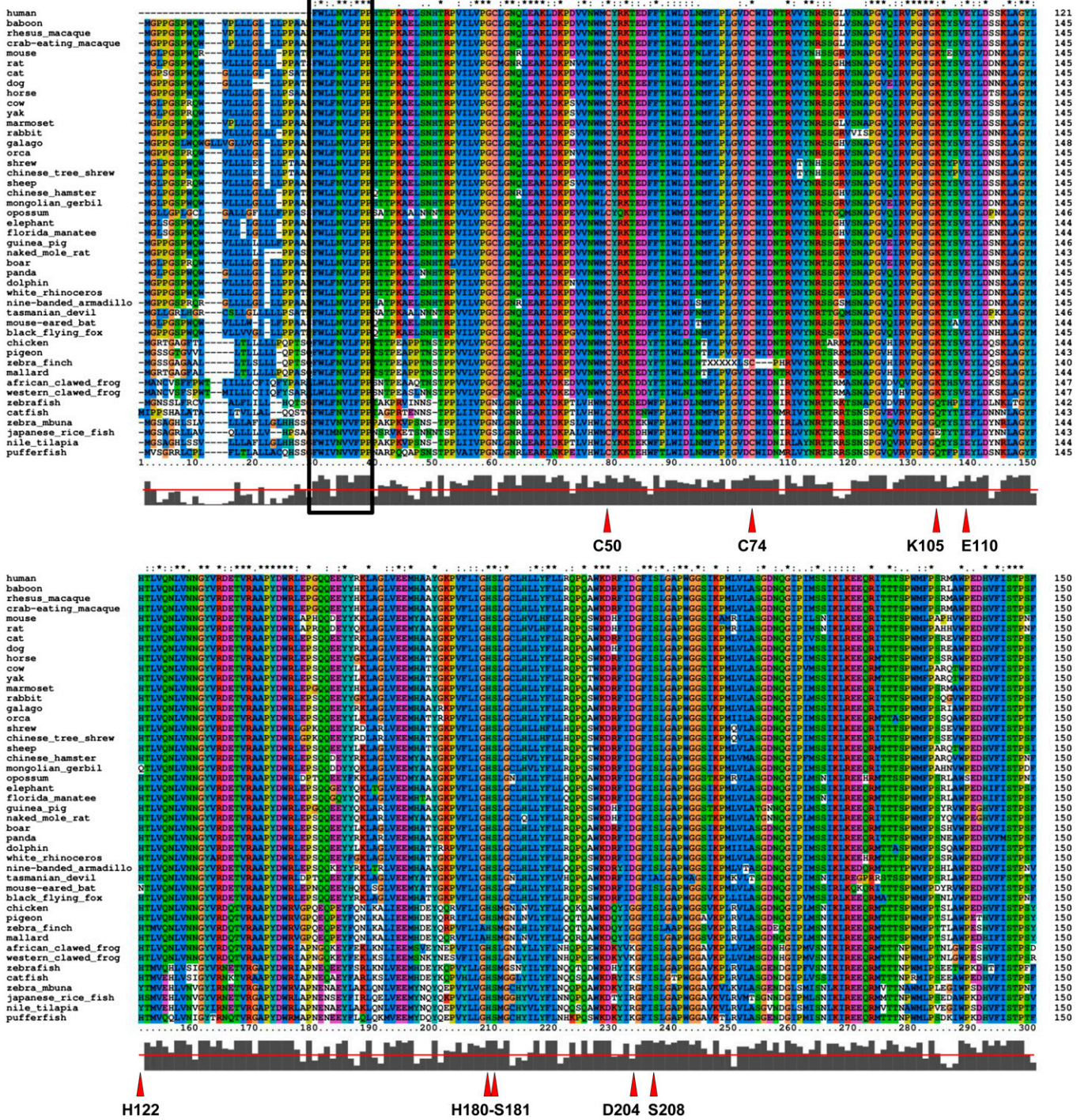
Fig. 4. Alignment of sequences of 45 LCAT orthologs from mammals to fish. Sequences were aligned using ClustalX, packaged through MultiCLUSTAL script, performing 98 cycles of alignments and using six different parameters, as follows: different multiple protein substitution matrices; costs for gap openings between 0 and 20 in steps of 4; pairwise gap extensions between 0 and 0.5 with 0.1 intervals; and multiple gap extensions between 0.02 and 0.1 with an interval of 0.02. The highest scoring alignment, according to the MultiCLUSTAL estimate, was chosen for further analysis. The first 10 residues that are highly conserved are indicated by a box. The symbols at the top are: (*) 100% conserved; (:) very well conserved; (.) conservation of polar or nonpolar properties. The histogram at the bottom plots the overall degree of conservation of each sequence position; the red line denotes approximate 50% conservation.

There are several detailed features of the complete POPC/LCAT model, *m5p*, that are worthy of note: *i*) The *sn-2* chain binding trough is capped by a pair of π-stacked Trp/Phe residues in both the model and the Lp-PLA₂ template, W2/F1 in the LCAT model (Fig. 3C) and W298/ F300 in the Lp-PLA₂ template (Fig. 1D). *ii*) The POPC molecule has settled deeply into the trough (Fig. 3C, D). *iii*) The terminal methyl of the *sn-2* chain remains in Van der Waals contact with the completely conserved residue W2 (Fig. 3C). *iv*) The *sn-2* chain has kinked at the double
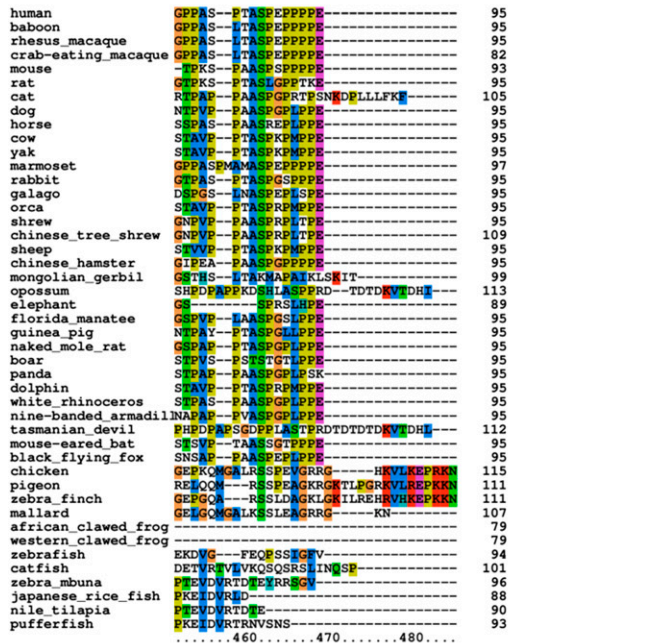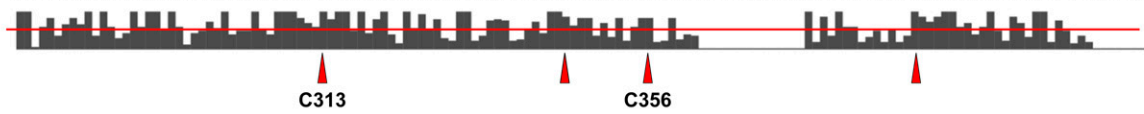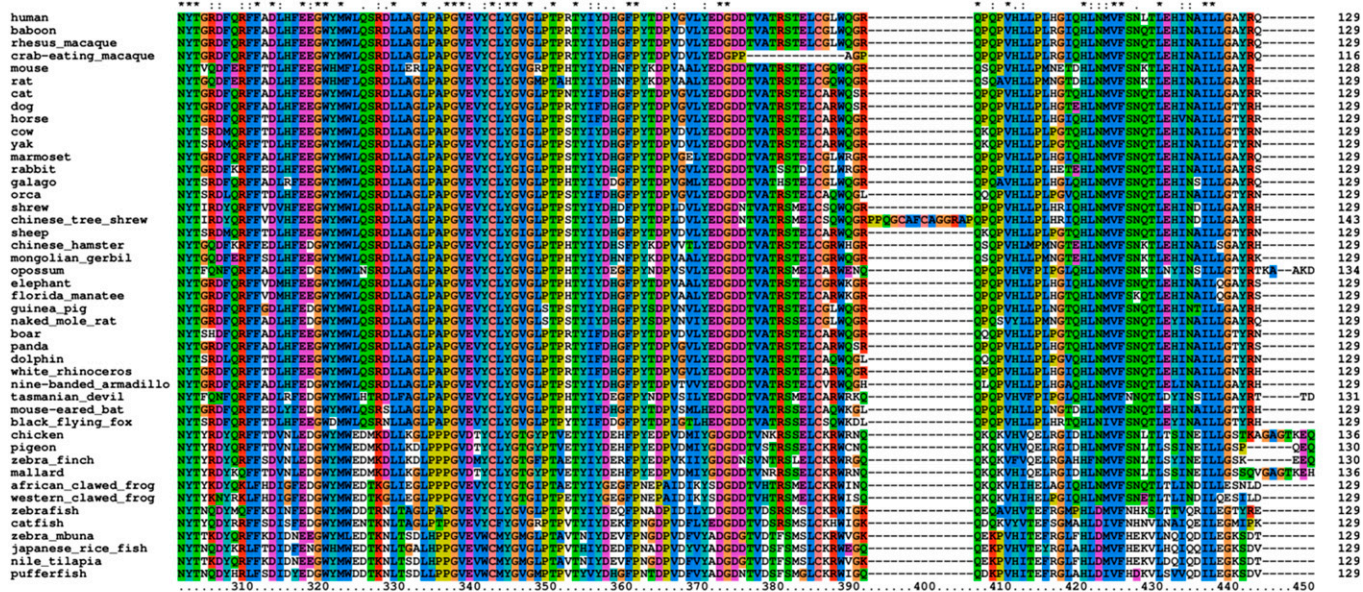
Fig. 4. Continued.

bond (Fig. 3C, D) adjacent to the almost 100% conserved residue S208 (Fig. 4).

Catalytic triad

A catalytic triad refers to three amino acid residues that function together at the active site of certain hydrolases, such as proteases and lipases (41). LCAT is unusual because it is the sole example of an enzyme that combines lipase and transferase activities. A common method for generating a nucleophilic residue for covalent hydrolysis is an acid-base-nucleophile triad. The residues form a charge-relay network to polarize and activate the nucleophile, which attacks the substrate, forming a covalent intermediate that in the hydrolase is hydrolyzed to regenerate the free enzyme. One key feature of the charge-relay network is that the acid, most often Asp, forms a hydrogen bond with the base, most often His, to drive formation of a salt bridge during activation of the nucleophile, most often Ser (41).

Three residues, S181, H180, and D204, are juxtaposed in Van der Waals contact in *m5* to create a putative catalytic triad in the correct linear order, S↔H↔D (Fig. 3B).

When the 10 lowest energy MODELLER structures of the models formed without disulfide bonds were aligned by their α carbons, the side chains of S181, H180, and D204 were remarkably well aligned (**Fig. 5A**); none of the non-hydrogen atoms of the key nucleophile, S181, known experimentally to be the nucleophile in the catalytic triad of LCAT (12), deviated from their position relative to the rest of the alignment by more than 1 Å. The same is true for the top 10 models created with the disulfide bond-forming routine (supplementary Fig. 2b).

The linear order S181↔H180↔D204 is identical to that in Lp-PLA₂, except reversed in direction relative to the hydrophobic trough (Fig. 5B). It is especially noteworthy that, as required by the catalytic triad charge-relay network concept, D204 forms a hydrogen bond with H180 (Fig. 5B), a geometry that is maintained in both sets of 10 lowest energy models (Fig. 5A; supplementary Fig. 2b).

S181, experimentally determined to be the nucleophile in the catalytic triad of LCAT (42), is at the elbow of the L-shaped hydrophobic trough in *m5* at the opposite end of the long arm from the conserved N-terminal cluster, F1-P10 (Fig. 3B–D), in a comparable position to the key catalytic residue, S273, of Lp-PLA₂ (compare Fig. 1F and Fig. 3B). H180 and S181 are completely conserved in all 45 LCAT ortholog sequences, while D204 is conserved in all 31 mammalian species (Fig. 4, red arrowheads).

The fact that S181 and H180 are adjacent in the linear sequence of *m5* is a unique feature of the proposed catalytic triad in our model that requires discussion. In other catalytic triads determined from crystal structures (41), the nucleophile and base are in separate segments of the protein. The juxtaposition of the nucleophile and the base in our model is not a problem in principle because: *i)* many catalytic triads with numerous combinations of nucleophile and base have been created by convergent evolution (41); and *ii)* LCAT is unique in that it possesses both lipase and transferase activities (4).

Not only is LCAT only distantly related to lipases (11), but the enzyme itself acts as the donor in the acyl transferase step (4). The other enzymes that catalyze formation of CE, acyl-CoA (CoA):cholesterol acyltransferases, ACAT1, and ACAT2, and indeed most acyl transferases, require an activated donor (43). Because there is no evidence for more than one active site in LCAT, to serve as the donor, the acyl ester enzyme intermediate created by the catalytic triad must have a high chemical potential. Thus the evolutionary origin of the active site of LCAT is likely to be distinct from the evolutionary origin of the active site of pure lipases. Indeed, Das, Davis, and Rudel (44) have suggested that ACAT1 may involve a Ser/His/Asp catalytic triad of S456, H460, and D400, placing Ser and His near one another in the primary sequence of ACAT1. However, the possibility exists that another histidine residue acts as the base. Given its interaction with the lid, H186 is a possible candidate.

For adjacent Ser-His residues to function as a base-nucleophile charge relay system, one of the nitrogens of the His imidazole ring must be able to hydrogen-bond to the
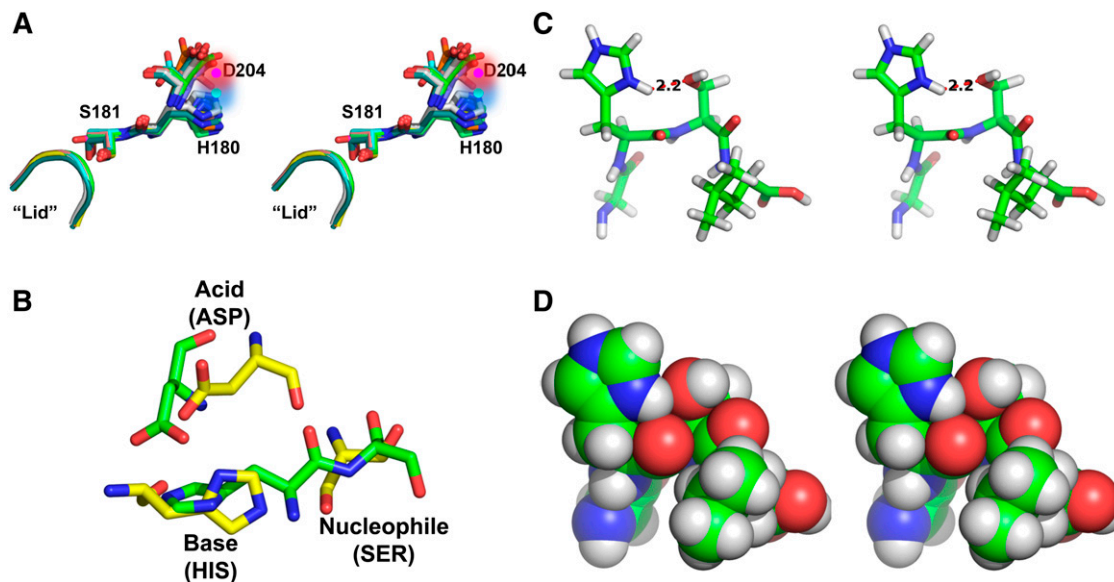

Fig. 5. Several noteworthy features of the lid, the "catalytic triad," and their interactions. A: Relaxed-eyed stereo image of the alignment of the lipase lid (residues 150–153) and the catalytic triad (S181, H180, D204) from the full alignment of the 10 lowest energy models created without disulfide bonds shown in Fig. 2A to illustrate persistent formation of a hydrogen bond between the acid, D204, and the base, H181, of the catalytic triad. The lid is in ribbons and the catalytic triad is in stick. The persistent hydrogen bond is indicated by a blue-pink dotted line. B: Overlay of the catalytic triad of our LCAT homology model 5 (green) with the known catalytic triad of Lp-PLA₂ (yellow). C: Relaxed-eyed stereo image in stick representation of the tetra-peptide sequence, G179-H180-S181-L182, in which the hydroxyl of S181 and the nitrogen adjacent to the Cγ atom of the protonated H180 imidazole ring were brought within hydrogen bond distance by rotation of [ϕ, φ], the peptide was energy minimized and then was subjected to 1 ns MD simulation at 310 K in vacuo. D: Relaxed-eyed stereo image of (C) in spacefilling representation.

hydroxyl of the Ser. To determine whether such a conformation was possible in our model, we created a peptide from the tetrapeptide sequence, G179-H180-S181-L182, rotated the [ϕ, φ] of H180 and S181 to move the hydroxyl moiety of S181 within hydrogen bond distance of the H180 δ and ε imidazole nitrogens and energy minimized the two peptides. Only the association of the S181 hydroxyl with the Nδ of the H180 imidazole ring maintained hydrogen bond distance. When we created a hydrogen bond by placing a charge on the His and subjected that peptide to MD simulations at 310 K in vacuo for 3 ns to mimic the low dielectric environment of a catalytic triad, the hydrogen bond between the hydroxyl of S181 and the Nδ of the H180 imidazole ring remained intact, creating the structure shown in Fig. 5C, D. The [ϕ, φ] for the H180-S181 pair, [−82, −97] and [60, −137], although appearing somewhat strained, fall into allowable regions of the Ramachandran diagram (supplementary Fig. 4).

β-Turn lipase lid

Many lipases have a lid that separates the active site from solvent and maintains a low dielectric environment. For example, PLA$_2$ from cobra venom that hydrolyzes the *sn-2* chain of PC contains a lipase lid (31-GGSGTP-36) structured as a β-turn (45). In model *m5*, a short β-turn (Fig. 3B) formed by four residues (150-PGQQ-153) covers the key catalytic residue, S181, in the hydrophobic trough (cyan, Fig. 3B–D), thus separating the active site from solvent approximately halfway between S181 and the docked

sn-2 ester bond (Fig. 3B–D). The location and conformation of this β-turn is quite invariant in both sets of 10 lowest energy models (Fig. 5A; supplementary Fig. 2b).

To mimic opening of the lid, we removed the lid residues 150–153 from *m5p*, created a peptide bond between residues E149 and E154, minimized with harmonic constraints and secondary structure restraints, and performed a 5.5 ns MD simulation at 310 K to create model *m5d*. In this simulation, we harmonically constrained the terminal methyls of the two POPC acyl chains and all residues greater than 10 Å from the N and P atoms of the POPC molecule. The results are shown in **Fig. 6A**. Noteworthy features of *m5d* include: *i*) significant disordering of the POPC acyl chains induced in part by reorientation of several aromatic residues allowing greater insertion into the hydrophobic trough (white arrowheads, Fig. 6A); and *ii*) deeper insertion and reorientation of the POPC headgroup to allow the *sn-2* carbonyl moiety to come within ∼4 Å of the hydroxyl of the nucleophilic residue of the catalytic triad, S181 (Fig. 6B).

The nucleophile S181 is restricted by direct or indirect interactions to four functionally critical residues

It is noteworthy that S181 in model *m5* has direct interactions with one residue, E149, that determines the activity of LCAT with PC substrates containing different *sn-2* fatty acyl compositions (46) and a second residue whose mutation, K218N, is known to cause LCAT deficiency (47). Further, S181 has indirect interactions with a third
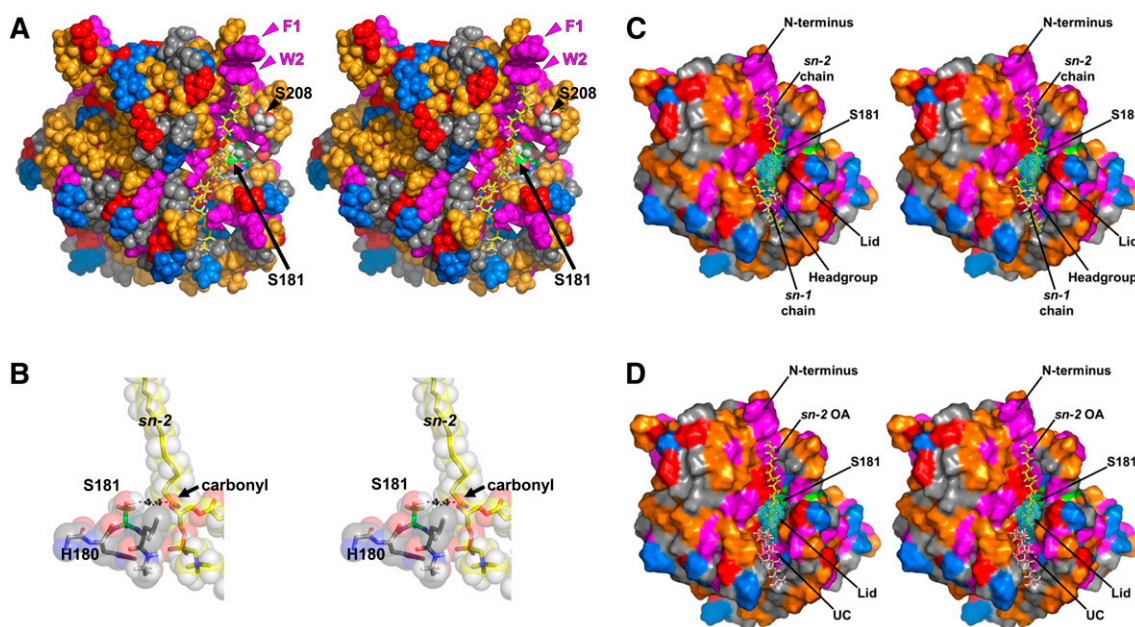

Fig. 6. Details of the lipid-binding trough. A: Space filling relaxed-eyed stereo image of the structure in Fig. 2C in which the lid has been removed, harmonic restraints applied to the terminal methyl groups of the acyl chains and to all residues >10 Å from the phosphorus or nitrogen atoms of the POPC and the model subjected to a 5 ns MD simulation (see Materials and Methods). Color code: same as Fig. 2C except the catalytic triad is spacefilling green. B: Space filling relaxed-eyed stereo image of the juncture of the POPC molecule and the catalytic triad from the structure shown in (A) illustrating the close juxtaposition (4.4 Å) of the *sn-2* carbonyl moiety and the –OH group of S181 of the catalytic triad. C: Relaxed-eyed surface image of model *m5* containing a docked POPC. Color code: aromatic (magenta); hydrophobic (gold); basic (blue) acidic (red); other residues (gray). Color code POPC: carbons (yellow); hydrogens (white). D: Relaxed-eyed surface image with same view as (C) showing proposed binding of *sn-2* OA (yellow and white) after hydrolysis and UC (peach and white) in the site that previous held the *sn-1* acyl chain and headgroup.

residue whose mutation, R147W, also causes LCAT deficiency (47) (**Fig. 7**).

E149 in human LCAT, the residue immediately preceding the lid in *m5*, is critical in determining the activity of LCAT with PC substrates containing different *sn-2* fatty acyl compositions (46). It is probably no coincidence that E149 forms a hydrogen bond to the hydroxyl moiety of S181 (Fig. 7A) and forms salt bridges to two basic residues on the opposite side of the lid that are nearby in the primary sequence, R147 and R158 (Fig. 7B). The salt bridge fluctuates between R147 and R158 in different low energy models (data not shown). It is equally significant that a mutation of one of the two basic residues salt bridged to E149, R147W, is associated with LCAT deficiency (47). Finally, K218, also associated with LCAT deficiency via the mutation, K218N, forms a hydrogen bond to the backbone carbonyl of H180 (Fig. 7A, B), restricting the motion of both the nucleophile S181 and the base H180.

One additional residue of note, Q152 in the lid, is directly associated with S181 via a hydrogen bond of its backbone

NH to the backbone carbonyl of S181 (Fig. 7A, B); this hydrogen bond is maintained in both sets of lowest energy models, the top 10 created with the disulfide bond-forming routine (supplementary Fig. 2b) and the top 10 created without disulfide bonds (supplementary Fig. 5).

It is intuitively tempting to speculate that the breaking of the interactions of E149, K128 and Q152 to S181, and E154 and E155 to H186 and R99, respectively (Fig. 7B), would trigger the opening of the lid away from the nucleophile, S181, exposing the nucleophile to the substrate and freeing it to form a hydrogen bond with the base, H180 or perhaps H186.

Glycosylation sites

LCAT is a glycoprotein. Attached carbohydrate residues are expected to be on the protein surface exposed to the solvent. The known *N*-glycosylated residues, N20, N84, N272, and N384, and the *O*-glycosylated residues, T407 and S409 (3), are all located on loops on the surface of
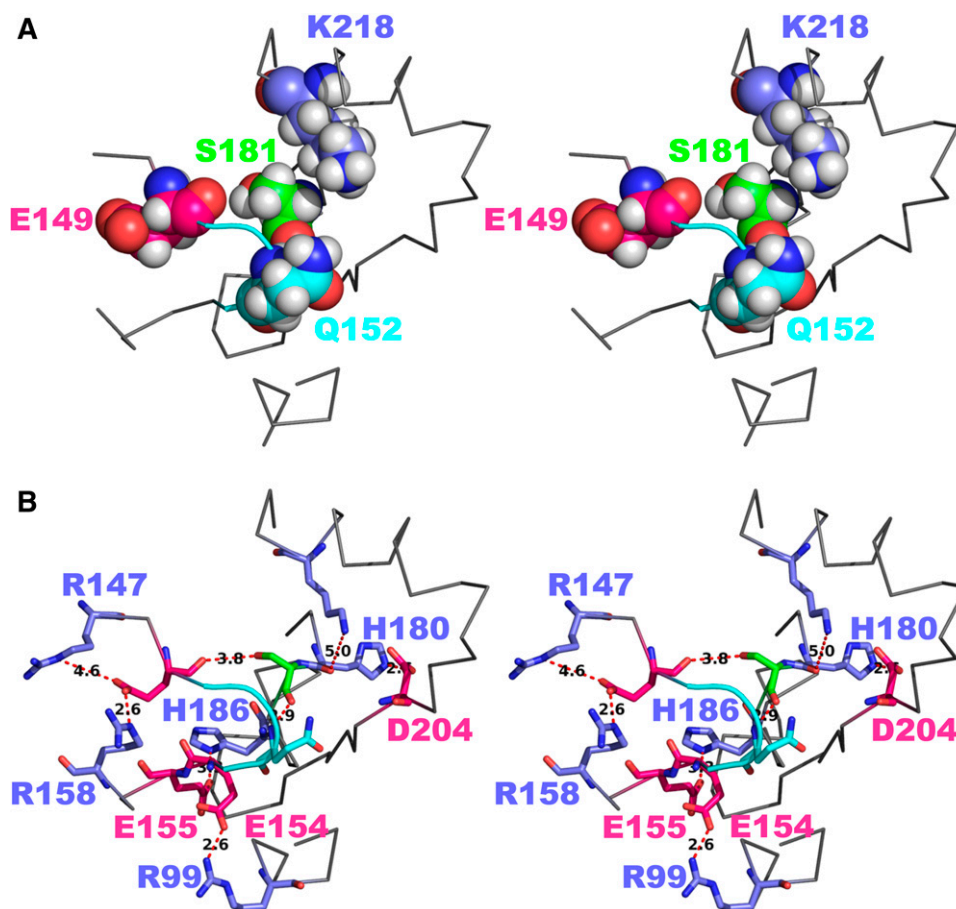

Fig. 7. Relaxed-eyed images of salt bridges and hydrogen bonds between key molecular moieties in the vicinity of the lipase lid. A: Relaxed-eyed stereo view of the lid-triad region showing interactions of S181 with Q152, E149, and K218. The lid loop region, R147-R158, the triad loop-helix region, I178-L221, and the R99 helix region, V96-G102 are in gray backbone representation; the lid (P150-Q153) is in cyan ribbons; E149, Q152, K218, and S181 are spacefilling. B: Relaxed-eyed stereo view of the same region as (A) showing the extensive salt bridges and hydrogen bonds between the lid and the triad loop on the right side and the lid and other residues on the left side (the opposite side of the lid from the triad). Interacting residues are stick. All the interactions on the left side are salt bridges: E149↔R158; E149↔R147; E155↔R99. Interactions on the right side are mixed: hydrogen bonds, Q152···S181 and K218···H180 and a salt bridge, E154↔H186.

the model on the side of the LCAT model opposite the hydrophobic trough (dotted pale green spacefilling representation in Fig. 3B).

DISCUSSION

Goodness of fit of model *m5* to known LCAT attributes

The human LCAT homology model *m5* presented here contains multiple structural attributes that match known LCAT features previously determined by biochemical and biophysical studies (3, 4). These include: *i*) a putative catalytic triad (13) in which an aspartic acid is hydrogen bonded to a histidine; *ii*) a β-turn lid (45) covering and hydrogen bonding to the backbone carbonyl of the known nucleophile of the catalytic triad, S181; *iii*) a unique location for E149, a residue that determines the specificity of LCAT toward the *sn-2* chain of the PC substrate (48), allowing its backbone carbonyl to hydrogen bond to the hydroxyl of S181 and salt bridge to a pair of arginine residues, one associated with LCAT deficiency, on the opposite side of the lid from S181; *iv*) a direct interaction of K218, a residue associated with LCAT deficiency, with the nucleophile, S181; *v*) an L-shaped hydrophobic trough containing the nucleophile (S181), the lid, E149, and K218 (associated with LCAT deficiency) at its elbow; *vi*) hydrophobic trough dimensions that snugly accommodate the *sn-2* chain of a PC molecule in its longer end and the *sn-1* chain and the choline head group (or UC) in its shorter end; *vii*) disulfide bonds (3, 4) formed without structural distortion; *viii*) location of the six known glycosylation sites (3) on surface loops; *ix*) a structure belonging to the α/β-hydrolase fold family; and *x*) an α-helical content identical to experimentally determined values (40).

The close juxtaposition and proper ordering of a correct catalytic triad of residues, S181↔H180↔D204, seems unlikely to be coincidental. S181 has been shown to be the primary catalytic residue in human LCAT (42), so the presence of H180 immediately adjacent to it in the primary sequence is significant; no other His is within 10 Å of S181. Even more unlikely to have occurred randomly is the hydrogen bond formation between H180 and D204 (Fig. 5), a close juxtaposition seen in both sets of 10 lowest energy LCAT homology models (Fig. 5A; supplementary Fig. 2b) with Cα distances varying between 5 and 6 Å; no other Asp is within 10 Å of H180.

However, the juxtaposition of a nucleophile next to a base in the primary sequence is unprecedented for a catalytic triad. We have shown that it is possible for H180 and S181 to rearrange their [φ, φ] angles to allow hydrogen bonding between the hydroxyl moiety of S181 and one of the nitrogen moieties of the imidazole ring of H180. The resulting conformation might be a bit strained, but it is known that there is usually unfavorable geometry between the serine and histidine of a catalytic triad, a strain that reduces the energy of the hydrogen bond to facilitate the relaying of the proton to the substrate (41). In an alternative scenario, after the lid opens, H186, salt bridged to the

lid residue E155, might move into position to act as the base.

The occurrence precisely at the elbow of the L-shaped hydrophobic trough of S181, E149, K218, and a β-turn structure with a sequence similar to previously described lipase-like lids (45), moieties that interact with each other, is also unlikely to be coincidental. Further, the dimensions of the L-shaped hydrophobic trough that fit the size and shape of a docked PC molecule especially snugly after energy minimization and a short MD simulation (*m5p* and *m5d*) suggest that the general features of the proposed PC binding site are plausible and likely correct. Finally, the spatial closeness allowing creation of the correct disulfide bonds without significant structural distortion and the presence of all potentially glycosylated residues on the solvent-exposed surface of the LCAT structure directly opposite the lipid-binding trough are also unlikely to be coincidental.

Despite relatively low sequence conservation and identity between human LCAT and the template structures, the structural attributes of our model *m5* that match known LCAT features (3, 4) support the general correctness of the human LCAT homology model presented here. Although the general consensus is that sequences falling below 20% sequence identity can have very different structures (49), detectable levels of sequence similarity usually imply significant structural similarity (17). The most unique features of our model, the interactions between S181, E149, and the lid, and between H181-A204, E149-R147, E149-R158, E154-R99, and E155-H186 are located in the core region of the alignment (Fig. 2). The likely reason that we were able to derive a plausible model for LCAT under conditions of low similarity between target and template is because the core of our alignment has only a single residue gap; core regions tend to have conserved structural folds and are more amenable to homology modeling than peripheral loop regions (39).

While the α-helical content of the model is 25%, spot on to the experimentally determined value of 24–25% (40), we are somewhat less confident about the β-sheet content of 14–17% calculated by the SS program used by RasMol. A previous report using circular dichroism spectroscopy suggested a β-sheet content of 40% (40) for human LCAT. While circular dichroism spectroscopy is not considered a particularly accurate way to calculate β-sheet content (50), the β-sheet content calculated for our LCAT model may be somewhat low.

Examination of the ribbons representation of our model in Fig. 3B shows a number of mostly parallel extended chain segments, many meeting the structural criteria of β-strands, which thread back and forth through the center of the structure between the surface α-helixes. The parallel extended chains not recognized as β-sheet by the PyMol program used for the illustration or by the SS program used by RasMol for percent calculation may be, in several instances, β-sheets in the actual LCAT structure. Because the spatial distribution of α carbons are similar for extended chains and individual strands in β-sheets, the

global packing of residues will also be similar between the homology model and actual human LCAT.

LCAT is known to esterify cholesterol with POPC *sn-2* chains of varying lengths. To accommodate itself to a given *sn-2* chain, the length of the trough model *m5* can be altered by movement of residues F1 and W2 (Fig. 3C). Measurements of the 10 lowest energy LCAT homology models (formed without disulfide bonds) give an average distance between W2 and S181 of 26 Å and a range of 22–31 Å (not shown); the extended acyl chain of oleic acid (OA) (Fig. 6D) is approximately 24 Å in length.

Stability of the hydrophobic trough of the homology model of LCAT

The exposure to solvent of the hydrophobic trough of our model *m5* poses a thermodynamic conundrum. The hydrophobic trough for the *sn-1* chain of oxidized PC in the published structure for a truncated form of Lp-PLA₂ is presented as solvent exposed (13). In model *m5d* (Fig. 6), although both the *sn-1* and *sn-2* chains of POPC are deeply buried in the long L-shaped acyl trough, the hydrophobic trough is also solvent exposed. The importance of the hydrophobic effect in biological systems (51) suggests that additional conformational changes or macromolecular interactions may be required before either Lp-PLA₂ or LCAT achieve their enzymatic functions.

Possible changes or interactions to stabilize the hydrophobic trough in *m5* during LCAT function include: *i*) dimerization (supplementary Fig. 6a); *ii*) scissor-like articulation about the catalytic triad along the long axis of the hydrophobic trough resulting in juxtaposition of the *sn-1* and *sn-2* acyl chains (supplementary Fig. 6b); *iii*) zipper-like association of the opposite edges of the long-axis of the hydrophobic trough (supplementary Fig. 6c); *iv*) association of the complete hydrophobic trough with apoA-I on HDL particles (supplementary Fig. 6c). To evaluate possible mechanisms for trough closure in the monomeric state, we used NMA.

NMA is a computational method that can reveal potential changes in the conformation of large proteins (52). The biologically relevant modes for proteins are low-frequency modes that represent the most likely large-scale motions of the protein (53). When we applied all atom NMA to model *m5* (54), the motions of the lowest frequency modes, 7 and 8, were located predominantly in the L-shaped hydrophobic trough.

In mode 7 (see supplementary Fig. 7a; supplementary Video 7a), the diagonal quadrants of the hydrophobic trough move toward and away and orthogonal to one another to create a sliding motion of each edge along the long axis of the trough. This motion could be involved in transferring PC and UC molecules from the apoA-I presentation tunnel (55, 56) to the LCAT active site. In mode 8 (see supplementary Fig. 7b; supplementary Video 7b), the horizontal quadrants of the hydrophobic trough move toward and away from one another to open and close the trough, a motion that could be helpful in opening and closing the trough to shield the acyl tails from solvent. It is worthy of note that, as might be expected, the *sn-2* end of

the hydrophobic trough is significantly more conserved that the *sn-1*/UC end (supplementary Fig. 7b).

Mechanism of interaction of LCAT with lipid-associated apoA-I

MD simulations of dHDL containing both POPC and UC consistently result in a gap forming between the pairwise helix 5 repeats (55, 56) of the antiparallel molecular double belt structure (57) of apoA-I in dHDL particles. Examination of an ensemble of 16 MD simulations (55, 58, 59) showed that, in 14 of the 16, the methyl ends of POPC acyl chains (POPC-Me) from the bilayer center are inserted into the gap and exposed to solvent (one example is shown in **Fig. 8A**). In the other two examples of the ensemble, the hydroxyl moiety of UC is inserted into the gap and exposed to solvent (one example is shown in Fig. 8C).

We proposed several years ago that, after attachment of LCAT to dHDL, the gap between pairwise helix 5 repeats in apoA-I creates an amphipathic "presentation" tunnel for migration of hydrophobic acyl chains and amphipathic
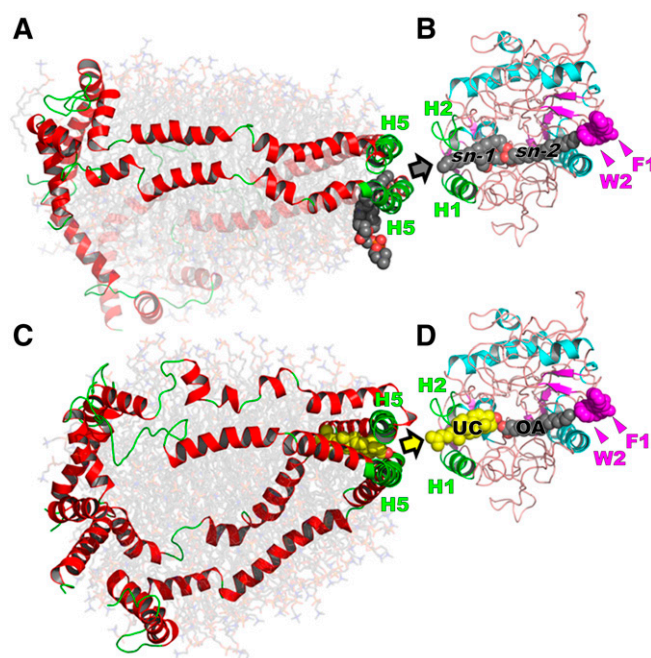

Fig. 8. Proposed interactions of LCAT homology model *m5* with dHDL. A: Ribbons image of an MD simulation of a dHDL particle (55) showing the pairwise H5 helices of apoA-I forming the acyl chain/free cholesterol "presentation" tunnel (green). The methyl end of the *sn-2* chain of one POPC inserts with 80–90% fidelity during simulation into the presentation tunnel, likely due to an increased mobility relative to the saturated *sn-1* chain. B: LCAT model containing docked POPC indicating the *sn-2* and *sn-1* binding troughs and the helical pair, H1 and H2, located at the end of the *sn-1* trough. C: A second simulated dHDL particle (59) showing the –OH of an UC inserted, in 5–10% of the MD simulations, into the 5/5 presentation tunnel. D: LCAT model containing docked OA and UC in the *sn-2* and *sn-1* binding end, respectively, and the helical pair, H1 and H2, located at the end of the *sn-1* trough. Gray and yellow arrows show proposed transfer of POPC and UC, respectively, from dHDL. All molecular model figures were created with the PyMOL molecular graphics system (version 1.3, Schrödinger).

UC from the bilayer to the active site of LCAT (55, 56). It is especially noteworthy that the gap between pairwise helix 5 domains has also been observed by X-ray crystallography (60). We considered the novel and consistent conformation of the residues forming the pairwise helix 5 gap after MD simulations to be suggestive of a role of the gap in activation of LCAT by lipid-associated apoA-I. In essentially every simulation of dHDL, either one or two terminal methyl groups or UC are inserted into the tunnel gap (55). For a movie of the insertion of two acyl methyl groups in the "presentation" tunnel during MD simulation see supplementary Video 8.

It has been suggested that the *sn-2* chains of PC and UC have separate binding sites within LCAT (61). We note that the cavity on the short side of the L-shaped trough (opposite the *sn-2* acyl chain) that initially binds the *sn-1* chain in model *m5* (Fig. 6C) is also of the correct size and shape for the binding of UC (Fig. 6D). Once the *sn-2* chain of PC passes into the LCAT hydrophobic trough and is converted to an acyl ester enzyme intermediate, UC could displace the bound lysolecithin molecule as illustrated in Fig. 6D. It seems likely that the displaced lysolecithin would partition back into the lipid bilayer of the LCAT-associated dHDL before being taken up by albumin.

Although LCAT may initially bind to parts of apoA-I in addition to the pairwise helix 5 domain, e.g., arginine residues in tandem helical repeat 6 (62), the presentation tunnel hypothesis implies that during CE formation the pairwise helix 5 repeats interact directly with LCAT. While a scissor-like articulation about the catalytic triad along the long axis of the hydrophobic trough resulting in juxtaposition of the *sn-1* and *sn-2* acyl chains might be an intermediate in apoA-I/ LCAT interactions (supplementary Fig. 6b), it seems unlikely this conformation alone would allow PC transfer from dHDL. Further, although it may be the state of LCAT in solution, it seems improbable that a dimer of LCAT (supplementary Fig. 6a) is directly involved in PC transfer from dHDL. This leaves two more likely conformations for lipid transfer: a zipper-like association of the opposite edges of the long-axis of the hydrophobic trough (supplementary Fig. 6c) and/or the association of the complete hydrophobic trough with apoA-I on HDL particles (supplementary Fig. 6d).

The gap between helix 5 repeats of apoA-I in HDL particles is lined on both sides by four charged residues, E125, K133, E136, and K140 (55, 56). A potential candidate for LCAT docking to the pairwise helix 5 gap is the methyl end of the *sn-1* chain binding trough. Figure 8B shows that the methyl end of the *sn-1* chain is held in place by antiparallel helical domains H1 (residues 113–125) and H2 (residues 345–354) of the LCAT model (also see Fig. 3B, D). Both H1 and H2 contain acidic and basic residues (D133, K116, H122, and D345, D346, R351, E354, respectively) that could salt bridge to the pairwise helix 5 gap in lipid-associated apoA-I. In this way, the H1-H2 end of the hydrophobic trough of LCAT would bind to the presentation tunnel of phospholipid-rich HDL particles to open a passageway from the bilayer interior to the active sites of LCAT for movement of amphipathic PC acyl chains and

UC to the active site of LCAT (Fig. 8). At some point during the interaction, the association of LCAT with HDL might shift to a full trough orientation (supplementary Fig. 6d).

In this model, apoA-I becomes part of the LCAT molecule when substrate is in the tunnel and the tunnel connects to the active site of LCAT. Studies by Parks and Gebre (7) suggest that the poor LCAT reactivity for long-chain *sn-2* acyl groups is multifactorial, involving turnover of substrate molecules at the active site of the enzyme and activation of the enzyme by its cofactor apoA-I. Our model is compatible with these observations because it suggests that turnover at the active site should be indistinguishable from activation, i.e., both would depend upon apoA-I structure.

CONCLUSIONS AND EXPERIMENTAL PREDICTIONS

The juxtaposition of multiple complicated structural motifs in a single model of LCAT, each motif matching extensive experimental biological and structural information (3, 4), is compelling evidence for the general correctness of many features of the model. We consider it unlikely, however, that the hydrophobic lipid-binding trough is open to the solvent and we suggest as a working hypothesis that LCAT is dimeric in solution and monomeric when performing its enzymatic functions while associated with HDL. Because the model provides a detailed molecular solution to the complicated thermodynamic problem of how hydrophobic acyl chains and amphipathic UC can be transferred from HDL to the catalytic triad of LCAT, it has important implications for understanding the antiatherogenic role of HDL in reverse cholesterol transport.

In lieu of a full crystal structure for human LCAT, our model predicts the detailed location and identity of residues in four major functional motifs: *i*) the catalytic triad (13); *ii*) the β-turn lipase lid (45) covering the key catalytic triad residue, S181; *iii*) the L-shaped hydrophobic trough for uptake and binding of a PC molecule and its product after hydrolysis, an *sn-2* fatty acyl chain and the UC substrate in the last step of esterification; and *iv*) activation of the lipase-lid by a trigger mechanism that involves E149, K218 and K147.

Powerful tests of the model would be to assay the effects on biological activity of LCAT of rationally designed site-directed mutants of key residues of these four functional motifs: *i*) the catalytic triad residues H180 and D204; *ii*) lid residues G151 and Q152; *iii*) aromatic residues in the L-shaped hydrophobic trough, e.g., W291, Y292, and Y157; and *iv*) residues in the putative lid trigger mechanism H186, E149, and K147. An additional test would be to perform chemical cross-linking of specific residues in solution phase human LCAT and fitting of observed crosslinks to contact map/cross-link matrix plot analyses (63).▉

Mechanical Engineering for use of the commodity cluster Cheaha that they jointly maintain.

REFERENCES

1. Toth, P. P., P. J. Barter, R. S. Rosenson, W. E. Boden, M. J. Chapman, M. Cuchel, R. B. D'Agostino, Sr., M. H. Davidson, W. S. Davidson, J. W. Heinecke, et al. 2013. High-density lipoproteins: a consensus statement from the National Lipid Association. *J. Clin. Lipidol.* **7:** 484–525.

2. Glomset, J. A. 1968. The plasma lecithins:cholesterol acyltransferase reaction. *J. Lipid Res.* **9:** 155–167.

3. Jonas, A. 2000. Lecithin cholesterol acyltransferase. *Biochim. Biophys. Acta.* **1529:** 245–256.

4. Sorci-Thomas, M. G., S. Bhat, and M. J. Thomas. 2009. Activation of lecithin:cholesterol acyltransferase by HDL ApoA-I central helices. *Clin. Lipidol.* **4:** 113–124.

5. Subbaiah, P. V., M. Liu, and F. Paltauf. 1994. Role of sn-2 acyl group of phosphatidylcholine in determining the positional specificy of lecithin-cholesterol acyltransferase. *Biochemistry.* **33:** 13259–13266.

6. Liu, M., J. D. Bagdade, and P. V. Subbaiah. 1995. Specificity of lecithin:cholesterol acyltransferase and atherogenic risk: comparative studies on the plasma composition and in vitro synthesis of cholesteryl esters in 14 vertebrate species. *J. Lipid Res.* **36:** 1813–1824.

7. Parks, J. S., and A. K. Gebre. 1997. Long-chain polyunsaturated fatty acids in the sn-2 position of phosphatidylcholine decrease the stability of recombinant high density lipoprotein apolipoprotein A-I and the activation energy of the lecithin:cholesterol acyltransferase reaction. *J. Lipid Res.* **38:** 266–275.

8. Forte, T., K. R. Norum, J. A. Glomset, and A. V. Nichols. 1971. Plasma lipoproteins in familial lecithin: cholesterol acyltransferase deficiency: structure of low and high density lipoproteins as revealed by elctron microscopy. *J. Clin. Invest.* **50:** 1141–1148.

9. Norum, K. R., J. A. Glomset, A. V. Nichols, and T. Forte. 1971. Plasma lipoproteins in familial lecithin: cholesterol acyltransferase deficiency: physical and chemical studies of low and high density lipoproteins. *J. Clin. Invest.* **50:** 1131–1140.

10. Peelman, F., N. Vinaimont, A. Verhee, B. Vanloo, J. L. Verschelde, C. Labeur, S. Seguret-Mace, N. Duverger, G. Hutchinson, J. Vandekerckhove, et al. 1998. A proposed architecture for lecithin cholesterol acyl transferase (LCAT): identification of the catalytic triad and molecular modeling. *Protein Sci.* **7:** 587–599.

11. Sensi, C., S. Simonelli, I. Zanotti, G. Tedeschi, G. Lusardi, G. Franceschini, L. Calabresi, and I. Eberini. 2014. Distant homology modeling of LCAT and its validation through in silico targeting and in vitro and in vivo assays. *PLoS ONE.* **9:** e95044.

12. Francone, O. L., and C. J. Fielding. 1991. Structure-function relationships in human lecithin:cholesterol acyltransferase. Site-directed mutagenesis at serine residues 181 and 216. *Biochemistry.* **30:** 10074–10077.

13. Samanta, U., and B. J. Bahnson. 2008. Crystal structure of human plasma platelet-activating factor acetylhydrolase: structural implication to lipoprotein binding and catalysis. *J. Biol. Chem.* **283:** 31617–31624.

14. Kraemer, J. A., M. L. Erb, C. A. Waddling, E. A. Montabana, E. A. Zehr, H. Wang, K. Nguyen, D. S. Pham, D. A. Agard, and J. Pogliano. 2012. A phage tubulin assembles dynamic filaments by an atypical mechanism to center viral DNA within the host cell. *Cell.* **149:** 1488–1499.

15. Eswar, N., B. Webb, M. A. Marti-Renom, M. S. Madhusudhan, D. Eramian, M. Y. Shen, U. Pieper, and A. Sali. 2006. Comparative protein structure modeling using MODELLER. *Curr. Protoc. Protein Sci.* **15:** 5.6.1–5.6.30.

16. Fiser, A., R. K. Do, and A. Sali. 2000. Modeling of loops in protein structures. *Protein Sci.* **9:** 1753–1773.

17. Martí-Renom, M. A., A. C. Stuart, A. Fiser, R. Sanchez, F. Melo, and A. Sali. 2000. Comparative protein structure modeling of genes and genomes. *Annu. Rev. Biophys. Biomol. Struct.* **29:** 291–325.

18. Sali, A., and T. L. Blundell. 1993. Comparative protein modelling by satisfaction of spatial restraints. *J. Mol. Biol.* **234:** 779–815.

19. Pettersen, E. F., T. D. Goddard, C. C. Huang, G. S. Couch, D. M. Greenblatt, E. C. Meng, and T. E. Ferrin. 2004. UCSF Chimera–a visualization system for exploratory research and analysis. *J. Comput. Chem.* **25:** 1605–1612.

20. Karlin, S., and S. F. Altschul. 1990. Methods for assessing the statistical significance of molecular sequence features by using general scoring schemes. *Proc. Natl. Acad. Sci. USA.* **87:** 2264–2268.

21. Henikoff, S., and J. G. Henikoff. 1992. Amino acid substitution matrices from protein blocks. *Proc. Natl. Acad. Sci. USA.* **89:** 10915–10919.

22. Chenna, R., H. Sugawara, T. Koike, R. Lopez, T. J. Gibson, D. G. Higgins, and J. D. Thompson. 2003. Multiple sequence alignment with the Clustal series of programs. *Nucleic Acids Res.* **31:** 3497–3500.

23. Larkin, M. A., G. Blackshields, N. P. Brown, R. Chenna, P. A. McGettigan, H. McWilliam, F. Valentin, I. M. Wallace, A. Wilm, R. Lopez, et al. 2007. Clustal W and Clustal X version 2.0. *Bioinformatics.* **23:** 2947–2948.

24. Higgins, D. G., J. D. Thompson, and T. J. Gibson. 1996. Using CLUSTAL for multiple sequence alignments. *Methods Enzymol.* **266:** 383–402.

25. Jeanmougin, F., J. D. Thompson, M. Gouy, D. G. Higgins, and T. J. Gibson. 1998. Multiple sequence alignment with Clustal X. *Trends Biochem. Sci.* **23:** 403–405.

26. Thompson, J. D., T. J. Gibson, F. Plewniak, F. Jeanmougin, and D. G. Higgins. 1997. The CLUSTAL_X windows interface: flexible strategies for multiple sequence alignment aided by quality analysis tools. *Nucleic Acids Res.* **25:** 4876–4882.

27. Brooks, B., and M. Karplus. 1983. Harmonic dynamics of proteins: normal modes and fluctuations in bovine pancreatic trypsin inhibitor. *Proc. Natl. Acad. Sci. USA.* **80:** 6571–6575.

28. Go, N., T. Noguti, and T. Nishikawa. 1983. Dynamics of a small globular protein in terms of low-frequency vibrational modes. *Proc. Natl. Acad. Sci. USA.* **80:** 3696–3700.

29. Levitt, M., C. Sander, and P. S. Stern. 1985. Protein normal-mode dynamics: trypsin inhibitor, crambin, ribonuclease and lysozyme. *J. Mol. Biol.* **181:** 423–447.

30. Perahia, D., and L. Mouawad. 1995. Computation of low-frequency normal modes in macromolecules: improvements to the method of diagonalization in a mixed basis and application to hemoglobin. *Comput. Chem.* **19:** 241–246.

31. Lindahl, E., C. Azuara, P. Koehl, and M. Delarue. 2006. NOMAD-Ref: visualization, deformation and refinement of macromolecular structures based on all-atom normal mode analysis. *Nucleic Acids Res.* **34:** W52–W56.

32. Humphrey, W., A. Dalke, and K. Schulten. 1996. VMD: visual molecular dynamics. *J. Mol. Graph.* **14:** 33–38.

33. Kalé, L. V., R. Skeel, M. Bhandarkar, R. Brunner, A. Gursoy, N. Krawetz, J. C. Phillips, A. Shinozaki, K. Varadarajan, and K. Schulten. 1999. NAMD2: greater scalability for parallel molecular dynamics. *J. Comput. Phys.* **151:** 283–312.

34. Catte, A., J. C. Patterson, M. K. Jones, W. G. Jerome, D. Bashtovyy, Z. Su, F. Gu, J. Chen, M. P. Aliste, S. C. Harvey, et al. 2006. Novel changes in discoidal high density lipoprotein morphology: a molecular dynamics study. *Biophys. J.* **90:** 4345–4360.

35. Jones, M. K., A. Catte, J. C. Patterson, F. Gu, J. Chen, L. Li, and J. P. Segrest. 2009. Thermal stability of apolipoprotein A-I in high-density lipoproteins by molecular dynamics. *Biophys. J.* **96:** 354–371.

36. Renault, L., V. Negre, T. Hotelier, X. Cousin, P. Marchot, and A. Chatonnet. 2005. New friendly tools for users of ESTHER, the database of the alpha/beta-hydrolase fold superfamily of proteins. *Chem. Biol. Interact.* **157–158:** 339–343.

37. Wallner, B., and A. Elofsson. 2005. All are not equal: a benchmark of different homology modeling programs. *Protein Sci.* **14:** 1315–1327.

38. Venclovas, C., and M. Margelevicius. 2005. Comparative modeling in CASP6 using consensus approach to template selection, sequence-structure alignment, and structure assessment. *Proteins.* **61(Suppl 7):** 99–105.

39. Greer, J. 1981. Comparative model-building of the mammalian serine proteases. *J. Mol. Biol.* **153:** 1027–1042.

40. Jin, L., Y. P. Lee, and A. Jonas. 1997. Biochemical and biophysical characterization of human recombinant lecithin: cholesterol acyltransferase. *J. Lipid Res.* **38:** 1085–1093.

41. Dodson, G., and A. Wlodawer. 1998. Catalytic triads and their relatives. *Trends Biochem. Sci.* **23:** 347–352.

42. Francone, O. L., and C. J. Fielding. 1991. Effects of site-directed mutagenesis at residues cysteine-31 and cysteine-184 on lecithin-cholesterol acyltransferase activity. *Proc. Natl. Acad. Sci. USA.* **88:** 1716–1720.

43. Chang, T. Y., B. L. Li, C. C. Chang, and Y. Urano. 2009. Acyl-coenzyme A:cholesterol acyltransferases. *Am. J. Physiol. Endocrinol. Metab.* **297:** E1–E9.

A robust homology model for LCAT   633

44. Das, A., M. A. Davis, and L. L. Rudel. 2008. Identification of putative active site residues of ACAT enzymes. *J. Lipid Res.* **49:** 1770–1781.

45. Segelke, B. W., D. Nguyen, R. Chee, N. H. Xuong, and E. A. Dennis. 1998. Structures of two novel crystal forms of Naja naja naja phospholipase A2 lacking Ca2+ reveal trimeric packing. *J. Mol. Biol.* **279:** 223–232.

46. Zhao, Y., J. Wang, A. K. Gebre, J. W. Chisholm, and J. S. Parks. 2003. Negative charge at amino acid 149 is the molecular determinant for substrate specificity of lecithin: cholesterol acyltransferase for phosphatidylcholine containing 20-carbon sn-2 fatty acyl chains. *Biochemistry.* **42:** 13941–13949.

47. Calabresi, L., L. Pisciotta, A. Costantin, I. Frigerio, I. Eberini, P. Alessandrini, M. Arca, G. B. Bon, G. Boscutti, G. Busnach, et al. 2005. The molecular basis of lecithin:cholesterol acyltransferase deficiency syndromes: a comprehensive study of molecular and biochemical findings in 13 unrelated Italian families. *Arterioscler. Thromb. Vasc. Biol.* **25:** 1972–1978.

48. Wang, J., A. K. Gebre, R. A. Anderson, and J. S. Parks. 1997. Amino acid residue 149 of lecithin:cholesterol acyltransferase determines phospholipase A2 and transacylase fatty acyl specificity. *J. Biol. Chem.* **272:** 280–286.

49. Chothia, C., and A. M. Lesk. 1986. The relation between the divergence of sequence and structure in proteins. *EMBO J.* **5:** 823–826.

50. Whitmore, L., and B. A. Wallace. 2008. Protein secondary structure analyses from circular dichroism spectroscopy: methods and reference databases. *Biopolymers.* **89:** 392–400.

51. Tanford, C. 1978. The hydrophobic effect and the organization of living matter. *Science.* **200:** 1012–1018.

52. Hinsen, K. 1998. Analysis of domain motions by approximate normal mode calculations. *Proteins.* **33:** 417–429.

53. Bray, J. K., D. R. Weiss, and M. Levitt. 2011. Optimized torsion-angle normal modes reproduce conformational changes more accurately than Cartesian modes. *Biophys. J.* **101:** 2966–2969.

54. Mouawad, L., and D. Perahia. 1993. Diagonalization in a mixed basis: a method to compute low-frequency normal modes for large macromolecules. *Biopolymers.* **33:** 599–611.

55. Jones, M. K., A. Catte, L. Li, and J. P. Segrest. 2009. Dynamics of activation of lecithin:cholesterol acyltransferase by apolipoprotein A-I. *Biochemistry.* **48:** 11196–11210.

56. Li, L., S. Li, M. K. Jones, and J. P. Segrest. 2012. Rotational and hinge dynamics of discoidal high density lipoproteins probed by interchain disulfide bond formation. *Biochim. Biophys. Acta.* **1821:** 481–489.

57. Segrest, J. P., M. K. Jones, A. E. Klon, C. J. Sheldahl, M. Hellinger, H. De Loof, and S. C. Harvey. 1999. A detailed molecular belt model for apolipoprotein A-I in discoidal high density lipoprotein. *J. Biol. Chem.* **274:** 31755–31758.

58. Catte, A., J. C. Patterson, D. Bashtovyy, M. K. Jones, F. Gu, L. Li, A. Rampioni, D. Sengupta, T. Vuorela, P. Niemela, et al. 2008. Structure of spheroidal HDL particles revealed by combined atomistic and coarse-grained simulations. *Biophys. J.* **94:** 2306–2319.

59. Jones, M. K., F. Gu, A. Catte, L. Li, and J. P. Segrest. 2011. "Sticky" and "promiscuous", the yin and yang of apolipoprotein A-I termini in discoidal high-density lipoproteins: a combined computational-experimental approach. *Biochemistry.* **50:** 2249–2263.

60. Mei, X., and D. Atkinson. 2011. Crystal structure of C-terminal truncated apolipoprotein A-I reveals the assembly of high density lipoprotein (HDL) by dimerization. *J. Biol. Chem.* **286:** 38570–38582.

61. Szedlacsek, S. E., E. Wasowicz, S. A. Hulea, H. I. Nishida, F. A. Kummerow, and T. Nishida. 1995. Esterification of oxysterols by human plasma lecithin-cholesterol acyltransferase. *J. Biol. Chem.* **270:** 11812–11819.

62. Roosbeek, S., B. Vanloo, N. Duverger, H. Caster, J. Breyne, I. De Beun, H. Patel, J. Vandekerckhove, C. Shoulders, M. Rosseneu, et al. 2001. Three arginine residues in apolipoprotein A-I are critical for activation of lecithin:cholesterol acyltransferase. *J. Lipid Res.* **42:** 31–40.

63. Walker, R. G., X. Deng, J. T. Melchior, J. Morris, P. Tso, M. K. Jones, J. P. Segrest, T. B. Thompson, and W. S. Davidson. 2014. The structure of human apolipoprotein A-IV as revealed by stable isotope-assisted cross-linking, molecular dynamics, and small angle x-ray scattering. *J. Biol. Chem.* **289:** 5596–5608.