

Structural insights into human TM6SF2-mediated cholesterol coordination in apoB-100 interactions

Xiaoxiao Jia^{1,5}, Qiaoyue Chen^{1,5}, Zhihua Wang², Linlin Zheng¹, Chunping Sun³, Suowen Xu⁴,

Mengchan Liu^{1,6}, Hongwu Qian^{1,6}

¹MOE Key Laboratory for Membraneless Organelles and Cellular Dynamics, Hefei National Research Center for Interdisciplinary Sciences at the Microscale, Division of Life Sciences and Medicine, University of Science and Technology of China, Hefei 230027, China

²Anhui Provincial Laboratory of Inflammatory and Immunity Disease, Anhui Institute of Innovative Drugs, School of Pharmaceutical Sciences, Anhui Medical University, Hefei, 230032, China.

³Department of Endocrinology, The Third Affiliated Hospital of Anhui Medical University (The First People's Hospital of Hefei), Anhui Medical University, Hefei 230001, China

⁴Department of Endocrinology and Metabolism, the First Affiliated Hospital of USTC, Division of Life Sciences and Medicine, University of Science and Technology of China, 230026, Hefei, China.

⁵These authors contributed equally to this work.

⁶To whom correspondence should be addressed: liumengchan@ustc.edu.cn, hongwuqian@ustc.edu.cn.

Transmembrane 6 superfamily member 2 (TM6SF2) is a liver- and intestine-enriched ER membrane protein that promotes VLDL lipidation/secretion and protects against hepatic lipid accumulation. The common E167K mutation of TM6SF2 impairs protein stability and VLDL secretion while showing a paradoxical protective lipid/cardiovascular profile. However, it remains elusive how TM6SF2 contributes to lipoprotein maturation. Here, we first determined the cryo-EM structure of human TM6SF2 using HEK293F-expressed proteins. TM6SF2 forms a dimer, with each protomer comprising 10 transmembrane helices (TMs) that can be divided into an N-terminal domain (NTD, TM1-2) and two EXPERA domains—EXPERA-N (TM3-6) and EXPERA-C (TM7-10). Further analysis revealed that an endogenous cholesterol molecule was coordinated by the NTD and EXPERA-N domains. Disruption of cholesterol coordination compromised the TM6SF2-apoB-100 interaction in pull-down assays. When the structure was determined using proteins expressed under cholesterol-depleted conditions, the cholesterol densities disappeared, accompanied by more flexible conformations of the NTD and EXPERA-N domains. Molecular dynamics simulations suggested that the luminal loop between TM5 and TM6 of the EXPERA-N domain (TM5-6 linker) becomes more flexible under cholesterol-free conditions. The TM5-6 linker is essential for apoB-100 interaction, as supported by the AlphaFold3-predicted structure and pull-down assays. The common E167K mutation is located on the TM5-6 linker and weakens its direct interaction with apoB-100. Together, these observations suggest that cholesterol contributes to the TM6SF2-apoB-100 interaction by maintaining a stable conformation of the TM5-6 linker.

Introduction

Metabolic-associated steatotic liver disease (MASLD) is the most prevalent chronic liver disease worldwide, with an estimated overall prevalence of 32.4% in adults¹. Multiple genetic studies have established a strong association between transmembrane 6 superfamily member 2 (TM6SF2) and MASLD, particularly for the well-characterized disease-causing variant E167K of TM6SF2²⁻⁴. Moreover, TM6SF2 has been reported to be associated with alcohol-related cirrhosis⁵, type 2 diabetes mellitus⁶ and cardiovascular diseases⁷.

TM6SF2 is a membrane protein that is localized mainly to the endoplasmic reticulum (ER) and the ER-Golgi intermediate compartment (ERGIC), with the highest expression in the liver and small intestines^{8,9}. The E167K variant has been reported to reduce protein stability, thereby impairing protein function^{2,10}. Inactivation of TM6SF2 impairs the secretion of triglyceride-rich lipoproteins^{8,9}. Further investigations suggest that TM6SF2 may exert its function during the bulk lipidation stage of VLDL assembly^{9,11}, potentially via direct interaction with apoB100 proteins^{11,12}. Unsecreted triglycerides are stored as lipid droplets in hepatocytes, thereby causing abnormal lipid accumulation in the liver¹³. In addition to its role in hepatocytes, TM6SF2 plays essential roles in lipid processing in enterocytes^{14,15}. Recently, intestinal TM6SF2 deficiency has been suggested to exacerbate MASLD via the gut–liver axis¹⁶. Despite these advances, the molecular basis of TM6SF2 during VLDL assembly is still unclear.

TM6SF2 is predicted to contain 10 transmembrane helices (TMs)¹⁷. TM3-6 and TM7-10 are organized into two four-TM regions that share sequence similarity with TM2-5 of the emopamil binding protein (EBP) and TM1-4 of meningioma-associated protein 30 (MAC30, TM1-4)¹⁷ and the σ 2 receptor (TM1-4)¹⁸. On the basis of this similarity, this four-TM domain has been designated the EXPanded EBP (EXPERA) domain¹⁷. Structural studies of EBP and the σ 2 receptor have suggested that the EXPERA domain can bind cholesterol¹⁹ or its analogs²⁰. Nevertheless, it is unclear whether the EXPERA domains in TM6SF2 retain the capacity to bind cholesterol or its derivatives.

Here, we determined the cryo-EM structures of TM6SF2 purified from HEK293F cells, revealing a dimeric organization with C2 symmetry. Each protomer comprises 10 TMs and can be divided into an N-terminal domain with two TM segments (NTD, TM1-2) and two EXPERA domains—EXPERA-N (TM3-6) and EXPERA-C (TM7-10). An endogenous cholesterol molecule is coordinated in each protomer by the NTD and EXPERA-N domains and plays an important role in promoting the TM6SF2-apoB-100 interaction. Further structural determination under cholesterol-depleted conditions indicated that cholesterol is essential for the stability of the NTD and EXPERA-N domains. Molecular dynamics simulations also suggested that the TM5-6 linker of the EXPERA-N domain becomes more flexible under cholesterol-free conditions. This linker was further confirmed to be essential for the TM6SF2-apoB-100 interaction, as supported by the AlphaFold3-predicted structure and *in vitro* pull-down assays. Together, these findings suggest that cholesterol contributes to the TM6SF2-apoB-100

interaction by maintaining a stable conformation of the TM5-6 linker.

Results

Structural determination of human TM6SF2

The purification of human TM6SF2 (hTM6SF2) protein for structural determination is detailed in the Methods section. Briefly, the Bril protein was fused to the C-terminus of hTM6SF2 to increase expression. The overexpressed hTM6SF2 protein was then reconstituted into nanodiscs prior to purification following previously established protocols^{21,22}. After affinity purification, the size exclusion chromatography (SEC) results revealed that the TM6SF2 nanodisc sample was monodisperse and suitable for cryo-EM structure determination (Fig. 1a).

Preliminary cryo-EM analysis revealed that the TM6SF2 nanodisc sample exhibited a preferred particle orientation, with the majority of the particles adopting a top view. To address this problem, we added 0.0004% (w/v) laurylmaltose neopentylglycol (LMNG) prior to cryo-EM sample preparation, which resulted in the observation of side-view particles (Extended Data Fig. 1a). After large-scale data collection, we eventually determined the cryo-EM structure of TM6SF2 at an overall resolution of 3.21 Å with C2 symmetry using 103,799 high-quality particles (Fig. 1b, Extended Data Fig. 1b-c, Supplementary Table 1).

Based on the cryo-EM map, a dimeric model was built for hTM6SF2, with 362 residues

assigned to each protomer (Fig. 1c, Supplementary Table 1). Consistent with previous predictions¹⁷, each protomer comprises ten TM segments (TM1-10), which can be further divided into an N-terminal domain with two TM segments (NTD, TM1-2) and two EXPERA domains, designated EXPERA-N (TM3-6) and EXPERA-C (TM7-10), respectively (Fig. 1d-e, Extended Data Fig. 1d).

In the dimeric interface, several hydrophobic residues of TM5/7/8 in one protomer contact the hydrophobic residues from another protomer on the luminal side. In the middle of the transmembrane domain, a hydrogen bond is formed between Tyr275 of one protomer and Phe274 of another protomer. Additionally, the aromatic rings of two Phe274 adopt a stacked conformation, engaging in π - π interactions (Extended Data Fig. 2a). On the cytosolic side, no direct protein-protein contacts are observed between both protomers. Instead, some lipid-like densities are observed between the two protomers, suggesting that lipids may help stabilize the dimeric assembly at the cytosolic face (Extended Data Fig. 2a-b). In each protomer, the two EXPERA domains exhibit C2 pseudosymmetry, and their overall arrangement is reminiscent of the dimeric structure of EBP and the σ 2 receptor (Extended Data Fig. 2c). This observation implies that the TM6SF2 protomer may have originated from a dimeric ancestor.

Endogenous cholesterol is coordinated in the EXPERA-N domain

The EXPERA domain has been suggested to bind cholesterol or its analogs in other EXPERA-containing proteins^{19,20}. After close examination of the TM6SF2 structure,

extra and sterol-like densities were observed in both EXPERA domains. In the EXPERA-N domain, the densities occupy a deep pocket formed by the EXPERA-N and NTD domains, adopting a horizontal orientation (Fig. 2a-b). However, the sterol-like density in the EXPERA-C domain is accommodated in a membrane-facing cavity oriented perpendicular to the membrane plane (Extended Data Fig. 3a-b).

Given that the cholesterol analog cholesteryl hemisuccinate (CHS) was added during membrane protein extraction, we attempted to model both cholesterol and CHS at these densities. The sterol-like densities in the EXPERA-N domain could accommodate a cholesterol molecule well, but not CHS, owing to the lack of densities corresponding to the hemisuccinate moiety (Fig. 2c, Extended Data Fig. 3c). In contrast, the densities in the EXPERA-C domain permit unambiguous placement of a CHS molecule (Extended Data Fig. 3b-c). We therefore assigned cholesterol to the extra densities in the EXPERA-N domain and CHS to the extra densities in the EXPERA-C domain (Fig. 2c, Extended Data Fig. 3b-c).

Since CHS is an exogenous sterol molecule, we focused our subsequent discussion on the endogenous cholesterol bound in the EXPERA-N domain. Cholesterol binds within the luminal leaflet, reminiscent of the binding modes observed in the EBP and $\sigma 2$ receptor structures^{19,20} (Extended Data Fig. 3d). The steroid ring is surrounded by polar and hydrophobic residues from the four TMs of the EXPERA-N domains, including TM3-Thr74/Asp78/Ile81, TM4-Ile112/Asp116, TM5-Phe148/Ile152/Phe155, TM6-

Tyr178, and Tyr96/Glu101/Leu104 on the TM3-TM4 (TM3-4) linker. The isooctyl tail is coordinated by TM3-Phe73/Val77 and residues from the NTD (TM1-Ser21 and TM2-Ser40/Ile43/Leu44). Additionally, the 3-hydroxyl group forms a hydrogen bond with Ser151 on TM5 (Fig. 2d, Extended Data Fig. 3e).

To explore the functional role of this cholesterol, we established an *in vitro* pull-down assay using purified TM6SF2 and apoB-100. hTM6SF2 was purified using LMNG as the solubilizing detergent, whereas apoB-100 was obtained from human serum as VLDL according to published protocols²³. Purified hTM6SF2 and VLDL were then mixed, followed by affinity purification using the FLAG tag on TM6SF2. Copurified apoB-100 was detected by SDS-PAGE. Indeed, apoB-100 could be pulled down by WT TM6SF2 (Fig. 2e, Lane 3), supporting a direct interaction between these two proteins¹². We then introduced phenylalanine substitutions to disrupt cholesterol coordination, generating two mutants, T74F/D78F and D116F. Both mutants behaved similarly to the WT protein in SEC and were confirmed with a reduction in bound cholesterol, as measured by liquid chromatography-mass spectrometry (LC-MS) (Fig. 2f; Extended Data Fig. 3f). These two mutants exhibited compromised interactions with apoB-100, suggesting that cholesterol plays an essential role in maintaining the TM6SF2-apoB-100 interaction (Fig. 2e, Lanes 4-5).

To further confirm the role of cholesterol in the TM6SF2-apoB-100 interaction, we expressed and purified TM6SF2 from the cholesterol-auxotrophic insect cell line

Spodoptera frugiperda (Sf9) (Extended Data Fig. 3g). The interaction between this cholesterol-depleted TM6SF2 and apoB-100 was then examined using the established *in vitro* pull-down assay. Compared with HEK293F-derived TM6SF2, the Sf9-derived proteins pulled down less amount of apoB-100, indicating a compromised interaction (Fig. 2g).

Cholesterol maintains EXPERA-N domain stability

To explore how cholesterol affects the interaction between hTM6SF2 and apoB-100, we sought to determine the structure of cholesterol-depleted hTM6SF2 derived from the Sf9 expression system. The expressed protein was reconstituted into nanodiscs using the same protocol as that used for the HEK293F-derived sample (Extended Data Fig. 4a). Consistent with the detergent-solubilized proteins, the Sf9 nanodisc sample also exhibited reduced interaction with apoB-100 compared with that of the HEK293F-derived nanodiscs (Extended Data Fig. 4b). Cryo-EM analysis of the Sf9-expressed protein yielded a reconstruction at an overall resolution of 3.32 Å (Extended Data Fig. 5a-c). An atomic model was built based on this map and is hereafter designated TM6SF2-Sf9 (Extended Data Fig. 5d, Supplementary Table 1). The original TM6SF2 structure, derived from the HEK293F-expressed protein, is correspondingly renamed TM6SF2-HEK293F.

Although the overall structures of TM6SF2-HEK293F and TM6SF2-Sf9 are almost identical (Extended Data Fig. 6a-b), the sterol-like densities in the EXPERA-N domain

were replaced by weak, disconnected densities in the TM6SF2-Sf9 map (Figs. 2**a-b**, 3**a-b**), confirming successful cholesterol depletion upon switching to the Sf9 expression system. In contrast, the sterol-like densities in the EXPERA-C domain remained intact (Extended Data Fig. 6**c**), suggesting that they did not arise from endogenous cholesterol. Accordingly, we assigned CHS molecules to these densities in the TM6SF2-Sf9 structure (Extended Data Fig. 6**d-e**) because CHS was also added during membrane protein extraction for this sample.

Along with the disappearance of cholesterol densities, the map quality of the NTD and EXPERA-N domains became much noisier than that of the TM6SF2-HEK293F structure (Fig. 3**c**). In particular, TM1 and TM2 in the NTD were nearly invisible at the 6σ threshold (Fig. 3**c**). These observations indicate that the NTD and EXPERA-N domains become more flexible under cholesterol-depleted conditions. In contrast to the NTD and EXPERA-N domains, the EXPERA-C domain is almost identical to the TM6SF2-HEK293F structure (Fig. 3**c**, Extended Data Fig. 6**a-b**).

We also performed molecular dynamics (MD) simulations to examine the flexibility of TM6SF2 in the cholesterol-bound (TM6SF2-CLR) and cholesterol-free (TM6SF2-Apo) states. These simulations indicate that the TM6SF2-Apo structure is more flexible than the TM6SF2-CLR structure, particularly in two regions, residues 31-60 (Region 1) and residues 163-191 (Region 2) (Fig. 3**d**). Region 1 corresponds to TM2 of the NTD, and Region 2 comprises the luminal TM5-TM6 linker (designated the TM5-6 linker

hereafter) and TM6 of the EXPERA-N domain (Fig. 3e), echoing the increased flexibility of the NTD and EXPERA-N domains in the TM6SF2-Sf9 structure.

Predicted interaction between TM6SF2 and apoB-48

A detailed understanding of the interaction between TM6SF2 and apoB-100 should be essential to explain the role of cholesterol in this interaction. However, our efforts to determine this complex structure were unsuccessful owing to the heterogeneous size of VLDL particles. We therefore sought to predict the complex structure using AlphaFold3²⁴. Since apoB-100 exceeds the size limit of the Alphafold3 online server, we predicted the complex structure using a shorter isoform of apoB, apoB-48, which has been suggested to interact with TM6SF2 similarly to the full-length protein¹². Although apoB-48 did not adopt a spherical conformation, likely because of the lack of lipids during prediction, we identified a region of apoB-48 (residues 1594-1804) with high predicted confidence for interaction with the luminal side of the TM6SF2 dimer, as indicated by the expected position error plot (Extended data Fig. 7a-b). We also predicted the complex structure using the C-terminal region of apoB-100 (residues 2001-4563); however, no specific region was predicted to interact with TM6SF2 with high confidence (Extended data Fig. 7c-d). We therefore employed the predicted TM6SF2-apoB-48 structure as a representative model for analyzing the TM6SF2-apoB-100 interaction.

On the basis of the predicted structure, we observed extensive polar interactions

between the luminal region of the TM6SF2 dimer and residues 1594-1804 of apoB-100. Within TM6SF2, one protomer contributes a greater number of interacting residues, as exemplified by the TM6SF2-B protomer in Fig. 4a. In this protomer, the luminal loops of the EXPERA-C domain, comprising TM7-8 and TM9-10 linkers, constitute the major binding sites for apoB-100 (Fig. 4b). Specifically, the charged residues Asp245 and Arg259 in the TM7-8 linker, together with Glu330 and Arg327 in the TM9-10 linker, form salt bridges with oppositely charged residues on apoB-100. Moreover, several hydrogen bonds are observed between the EXPERA-C domain and apoB-100, including Tyr326-Glu1629, Thr325-Ser1655, and Tyr326-Glu1670 (Fig. 4b). In addition to the EXPERA-C domain, direct interactions are also observed between the EXPERA-N domain and apoB-100, including two salt bridges (Arg105-Asp1650 and Lys163-Glu1674) and two hydrogen bonds (Ser166-Lys1702 and Glu167-Ser1678) (Fig. 4c). The other protomer, TM6SF2-A, interacts with apoB-100 via two salt bridges (Asp245-Lys1730 and Asp331-Lys1769) and two hydrogen bonds (Thr325-Asp1722 and between the main-chain amine of Phe324 and the side chain of Ser1745) (Fig. 4d).

We then validated the predicted interfaces using the *in vitro* pull-down assay. Mutations, including R105A, K163A/R259A and Y326A/R327A, were introduced into TM6SF2 to disrupt the predicted interactions. These variants display behavior comparable to that of the WT protein in SEC (Extended data Fig. 7e), but exhibit reduced capacities to interact with apoB-100, thereby validating the predicted interfaces (Fig. 4e).

The disease-causing mutation E167K has previously been reported to reduce protein stability, thereby impairing TM6SF2 function^{2,10}. In the predicted structure, Glu167 is located on the TM5-6 linker and forms a hydrogen bond with Ser1678 of apoB-100 (Fig. 4c). These observations suggest that the E167K mutation may additionally weaken the direct interaction between TM6SF2 and apoB-100. In support of this speculation, our *in vitro* pull-down assay demonstrated that E167K actually compromised the interaction between TM6SF2 and apoB-100 (Fig. 4f, Lane 6; Extended data Fig. 7f). In addition to E167K, two other mutations in the TM5-6 linker, K163D and K163A/S166A, also impaired the TM6SF2-apoB-100 interaction (Fig. 4f, Lanes 4-5; Extended data Fig. 7f), indicating that the TM5-6 linker plays an essential role in mediating the TM6SF2-apoB-100 interaction.

These biochemical observations remind us of our previous MD simulation data, which revealed that residues 163-191 (Region 2) become more flexible under cholesterol-depleted conditions (TM6SF2-Apo) (Fig. 3d). This region comprises the TM5-6 linker (residues 163-168) and TM6 (residues 169-191) (Fig. 3e). Among all the residues in Region 2, two apoB-100-interacting residues (Ser166 and Glu167) exhibit the most significant increase in flexibility (Fig. 3d). Our previous structural observations suggested that the EXPTRA-N domain is more flexible under cholesterol-depleted conditions. Upon closer inspection of the EM map, the map quality of the TM5-6 linker is even poorer than that of other regions of the EXPTRA-N domain, such as TM5, in the TM6SF2-Sf9 structure (Fig. 3c). These observations suggest that the coordinated

cholesterol in the EXPERA-N domain may be essential for stabilizing the conformation of the TM5-6 linker, thereby contributing to the direct interaction between TM6SF2 and apoB-100.

Discussion

TM6SF2 plays essential roles during the bulk lipidation of VLDL and is closely associated with MASLD. Here, we determined the cryo-EM structures of TM6SF2 using proteins overexpressed in cholesterol-rich HEK293F cells (TM6SF2-HEK293F) and cholesterol-auxotrophic insect cells (TM6SF2-Sf9), revealing a cholesterol binding pocket formed by the NTD and EXPERA-N domain. By combining MD simulations, an *in vitro* TM6SF2-apoB-100 pull-down assay, and AlphaFold3-based structure prediction, we revealed that cholesterol contributes to the TM6SF2-apoB-100 interaction by stabilizing the conformation of the TM5-6 linker in the EXPERA-N domain.

We proposed a model for the function of TM6SF2 in VLDL maturation based on our structural and biochemical analyses. Under cholesterol-depleted conditions, TM6SF2 adopts a more dynamic conformation, particularly the NTD and EXPERA (EXPERA-N) domains. Cholesterol binding then stabilizes these two domains, as evidenced by the TM6SF2-HEK293F and TM6SF2-Sf9 structures (Fig. 3c). Consequently, the TM5-6 linker adopts a more stable conformation that interacts with premature VLDL, as suggested by MD simulations, AlphaFold3 prediction, and *in vitro* pull-down assays

(Figs. 3d, 4). Lipids can be transported from the ER membrane to premature VLDL. After maturation, VLDL is released via the secretory pathway (Extended data Fig. 8a). This model is admittedly preliminary, and many molecular details remain unclear. For example, how are lipids transported from the ER membrane to VLDL, and whether TM6SF2 plays additional roles during lipid transport, such as in the recruitment of some lipid transport proteins? Moreover, how does the TM6SF2-VLDL interaction is regulated after maturation? Further investigations integrated with structural, biochemical, and cellular methods are needed to address these problems.

Based on our model, the pathogenic effects of several disease mutations, including E167K and L156P²⁵, may be partially attributed to weakened interactions between TM6SF2 and apoB-100. Glu167, located in the TM5-6 linker, directly interacts with apoB-100 in VLDL (Fig. 4c). E167K substitution may weaken the TM6SF2-apoB-100 interaction, thereby impairing lipidation. L156P substitution may disrupt the TM5 helix and induce conformational changes in the TM5-6 linker, thereby impairing the TM6SF2-apoB interaction (Extended data Fig. 8b).

Cholesterol has long been known to be required for VLDL assembly and secretion²⁶⁻³², and this requirement has been attributed to cholesterol serving as an essential surface component for VLDL assembly³⁰. Here, our structural studies on TM6SF2 provide an additional possible explanation: reduced cholesterol levels may weaken the interaction between VLDL and TM6SF2, thereby impairing the lipidation stage of VLDL assembly.

To validate this hypothesis, further *in vivo* studies are needed, particularly in animal models.

Despite these remaining questions, our structural studies advance the understanding of the molecular basis of TM6SF2 function during the bulk lipidation of VLDL and provide a structural basis for future functional studies and potential therapeutic development for MASLD as well as cardiovascular diseases.

Acknowledgments

We are grateful to the Cryo-EM Center of the University of Science and Technology of China (USTC) and the Thermo Scientific Krios G4 system of Anhui Medical University for EM facility support and to the staff members for their technical support in cryo-EM data collection. This work was supported by the National Key R&D Program of China (2024YFA1307900 & 2024YFA1803104 to H.Q.), the National Natural Science Foundation of China (32271241 to H.Q.), and the start-up funding from the University of Science and Technology of China (KY9100000034 & KJ2070000082 to H.Q.). This study was also supported by “USTC Research Funds of the Double First-Class Initiative” (YD9105202629 to M.L. and YD9110002089 to S.X.) and the Research Funds of Centre for Leading Medicine and Advanced Technologies of IHM (2025IHM01040 to S.X.).

Author Contributions

H.Q. conceived the project. H.Q., M.L. and S.X. designed the experiments. X.J., Q.C. and L.Z. performed cloning and protein purification. C.S. and X.J. isolated VLDL from the plasma. X.J. and Q.C. performed the *in vitro* pull-down assays. Z.W. performed the MD simulation analysis. X.J. prepared the cryo-EM samples and collected the data. M.L. and H.Q. determined the structures. All the authors contributed to the data analysis and preparation of the manuscript. H.Q. and M.L. prepared the figures, and H.Q. wrote the manuscript.

Competing interests

The authors declare that they have no competing interests.

Methods

Cell culture

HEK293F cells (Invitrogen) were cultured at 37°C with 5% CO₂ in SMM 293T-TII medium (Sino Biological Inc.) with shaking. Sf9 cells (Invitrogen) were cultured at 27°C in SIM SF Expression Medium (Sino Biological Inc.) with shaking. *E. coli* BL21 (DE3) cells (Alpalifebio) were grown in Luria-Bertani (LB) medium (1% tryptone, 1% NaCl, and 0.5% yeast extract) at 37°C with shaking.

Molecular cloning

The gene encoding full-length human *TM6SF2* (UniProt accession code: Q9BZW4) was subcloned from a cDNA library and inserted into the pCAG and pFastBac vectors for expression in HEK293F and Sf9 cells, respectively. A Bril protein was fused to the C-terminus to increase expression, after which FLAG and Strep tags were fused for affinity purification. All the mutants were generated with a standard two-step PCR-based strategy.

Expression and purification of MSP1E3D1 from BL21 cells

E. coli BL21 (DE3) transformed with the MSP1E3D1 plasmid was cultured in LB medium supplemented with kanamycin (50 µg/mL) at 37°C, and 0.2 mM isopropyl β-D-thiogalactoside (IPTG) was added when the OD₆₀₀ reached 1.0. After further culture at 37°C for 4 h, the cells were collected and resuspended in lysis buffer (25 mM Tris pH 8.0, 150 mM NaCl) supplemented with 1 mM phenylmethanesulfonylfluoride

(PMSF) and 1% Triton X-100. After sonication, the cell debris was pelleted via centrifugation at $20,000\times g$ at 4°C for 1 h. The supernatant was collected and applied to Ni-NTA agarose. After rinsing with lysis buffer plus 10 mM imidazole, the proteins were eluted with lysis buffer plus 250 mM imidazole. Eluates were subjected to ion exchange chromatography (Source™ 15Q, Cytiva) for further purification.

Expression and purification of hTM6SF2 from HEK293F cells

HEK293F cells were transfected with *TM6SF2* plasmids at a density of 2.5×10^6 cells per mL using polyethylenimine (PEI; Yeasen Biotechnology, Shanghai). Briefly, 1 mg of total plasmid DNA was premixed with 3 mg of PEI in 50 mL of fresh medium and incubated for 15-30 min at room temperature. The transfection mixture was then added to 1 L of cell culture. Forty-eight hours after transfection, the cells were collected and resuspended in buffer containing 25 mM Tris-HCl pH 8.0, 150 mM NaCl and protease inhibitor cocktail (Sigma-Aldrich).

[Purification for biochemical assays]

The resuspended cells were solubilized with 1% (w/v) LMNG (Anatrace) at 4°C for 2 h. After centrifugation at $27,000 \times g$ for 1 h, the supernatant was collected and incubated with Strep-Tactin®XT 4Flow® resin (IBA). After rinsed with buffer A containing 25 mM Tris (pH 8.0), 150 mM NaCl, 0.005% LMNG, the target protein was eluted with buffer A plus 50 mM biotin (Harveybio). The eluate was then loaded onto anti-FLAG M2 affinity resin (Sigma-Aldrich), followed by washing with buffer A and eluting with

buffer A plus 0.2 mg/mL FLAG peptide (Bankpeptide Biological Technology Co., Ltd.).

The purified protein was then concentrated and subjected to size-exclusion chromatography (SEC) (Superose 6 Increase 10/300 GL, Cytiva) in buffer A. The peak fractions were flash-frozen and stored at -80°C.

[Nanodisc reconstitution and purification for cryo-EM studies]

The resuspended cells were extracted with 1% (w/v) LMNG plus 0.1% CHS (w/v) (Bluepus) at 4°C for 2 h. After centrifugation at $27,000 \times g$ for 0.5 h, the supernatant was collected and mixed with the nanodisc scaffold protein MSP1E3D1 (100 mg per liter HEK293F cells). The mixture was incubated at 4°C for 1 h before detergent-absorbing beads (Bio-Beads SM-2, Bio-Rad) were added. After overnight incubation, the solution was subjected to affinity chromatography and SEC as described above, using detergent-free buffer (25 mM Tris-HCl 8.0, 150 mM NaCl). The peak fractions were concentrated to approximately 10 mg/ml and used immediately for cryo-EM sample preparation.

Expression and purification of hTM6SF2 from Sf9 cells

Recombinant baculovirus was generated using the Bac-to-Bac baculovirus expression system (Invitrogen). Briefly, the TM6SF2 pFastBac plasmids were transformed into DH10Bac competent cells to produce recombinant bacmid via Tn7 transposition, which was screened by blue-white selection. Positive recombinant bacmid was purified and transfected into Sf9 cells seeded in 12-well plates at a density of 0.3×10^6 cells/ml using

Transfection Reagent (ABclone). The cell culture supernatant was harvested at 96 h post-transfection to yield a primary P1 virus stock. P1 virus stock was added into Sf9 cells at a density of 2.0×10^6 cells/ml at a volume ratio of 1:100 (virus:culture medium). After incubation at 27°C for 72 h, the supernatant was collected as a P2 virus stock. Sf9 cells were infected with the P2 virus at a density of 3.0×10^6 cells/ml using a 1:25 volume ratio (virus:culture medium). Forty-eight hours after transfection, the cells were collected and resuspended in buffer containing 25 mM Tris-HCl (pH 8.0), 150 mM NaCl and protease inhibitor cocktail (Sigma-Aldrich). Protein purification for biochemical assays and nanodisc reconstitution for cryo-EM studies were then performed following the same procedure as that used for protein expressed in HEK293F cells.

***In vitro* pull-down between TM6SF2 and ApoB100**

[Separation of VLDL]

300 μ L of Solution A ($\rho = 1.006$ g/mL, containing 11.40 g NaCl, 0.1 g EDTA-2Na and 1 mL of 1 N NaOH in 1 L 18 M Ω ultrapure water) was layered onto 600 μ L serum, and ultracentrifugation was performed at 150,000 rpm ($899,744 \times g$) and 16°C for 60 min (Acceleration 5, Deceleration 7). The top 300 μ L of the VLDL- containing fraction was collected and combined with an equal volume of buffer containing 50 mM Tris-HCl pH 8.0, 105 mM NaCl. The collection of human serum was approved by the Institutional Research Ethics Committee of the Third Affiliated Hospital of Anhui Medical University (Approval No. 2026-299-01).

[Pull-down assay]

Pull-down assays were performed using FLAG-tagged TM6SF2 (WT and its mutant proteins) and VLDL isolated from human serum. TM6SF2 and VLDL were mixed at a 16:1 molar ratio and then incubated on ice for 1 h. Subsequently, the mixtures were incubated with anti-FLAG M2 affinity resin (Sigma-Aldrich) for an additional 1 hour at 4 °C. After extensive washing with 25 mM Tris pH 8.0, 150 mM NaCl, 0.005% LMNG (the detergent-free buffer was used for nanodisc sample), bound proteins were eluted with wash buffer plus 0.4 mg/mL FLAG peptide (Bankpeptide Biological Technology Co., Ltd.). The inputs and elutes were subjected to SDS-PAGE.

Quantification of cholesterol in purified proteins

Cholesterol was extracted from purified proteins and analyzed by liquid chromatography-mass spectrometry (LC-MS) as described previously³³. Briefly, 120 µL of purified protein (0.84 mg/mL) were mixed with 400 µL PBS buffer, 500 µL methanol and 1 mL methyl tert-butyl ether (MTBE). Then, 20 µg D₇-cholesterol (MCE, HY-N0322S) was added as an internal standard. The mixture was vortexed and centrifuged at 3,500 rpm for 5 min to induce phase separation. The upper organic phase (about 800 µL) was transferred to a new tube. Extraction was repeated by adding 800 µL MTBE to the original tube, and the organic phases were combined. Samples were dried under a stream of nitrogen. Dried extracts were resuspended in 200 µL of 90% methanol and incubated at room temperature for 1 h, followed by centrifuging at

10,000×g for 10 min. Then, 100 μ L supernatant were transferred to a 2 mL autosampler vial containing 250 μ L inserts and analyzed by LC-MS. Cholesterol were quantified using internal standard normalization.

Cryo-EM sample preparation and data collection

The cryo-samples were prepared using Thermo Fisher Vitrobot Mark IV. Quantifoil R1.2/1.3 Cu grids were glow-discharged with air in a PDC-32G-2 plasma cleaner (Harrick) under vacuum for 2 min and mid-force for 50 s. Aliquots of 3.5 μ L of purified TM6SF2-HEK293F and TM6SF2-Sf9 at concentrations of approximately 10 mg/mL were applied to the glow-discharged grids. After blotting with filter paper for 3.5 s (100% humidity and 8°C), the grids were plunged into liquid ethane cooled with liquid nitrogen. For the LMNG-containing sample, 0.0004% (w/v) LMNG was added immediately prior to cryo-EM sample preparation.

TM6SF2-HEK293F grids were loaded onto a Titan Krios electron microscope operating at 300 kV and equipped with an energy filter and a Falcon 4i DED detector (0.92 Å/pixel). For the LMNG-containing sample, data were collected using a K3 direct electron detector under the super-resolution mode (0.41 Å/pixel), which was binned to 0.82 Å/pixel by twofold after motion correction. TM6SF2-Sf9 grids were loaded onto another Titan Krios electron microscope operating under the same conditions and equipped with a K3 direct electron detector. For all datasets, movie stacks were automatically collected with EPU using a total dose of 50 e⁻ Å⁻² and defocus values

ranging from $-1.7\ \mu\text{m}$ to $-2.1\ \mu\text{m}$. Each movie stack was fractionated into 32 frames and motion corrected using Patch Motion Correction in cryoSPARC³⁴. The defocus values were estimated using Patch CTF estimation in cryoSPARC³⁴. The pixel size of the TM6SF2-Sf9 dataset was calibrated to $0.832\ \text{\AA}$ based on map comparison with the TM6SF2-HEK293F map.

Cryo-EM data processing

In total, 502 and 2,051 movies were collected for LMNG-free and LMNG-containing TM6SF2-HEK293F samples, respectively. For the TM6SF2-Sf9 samples, 3,128 movies were collected.

[TM6SF2-HEK293F]

In total, two datasets were collected for TM6SF2-HEK293F: Dataset 1 ($0.92\ \text{\AA}/\text{pixel}$, LMNG-free sample) and Dataset 2 ($0.82\ \text{\AA}/\text{pixel}$, LMNG-containing sample). After motion correction and blob picking, 191,832 particles and 807,005 particles were extracted from Datasets 1 and 2, respectively. A box size of 246 pixels was used for Dataset 1, whereas particles from Dataset 2 were extracted with a box size of 276 pixels and then rescaled to 246 pixels to match the pixel size of Dataset 1. After 2D classification, 170,660 good particles (9,786 from Dataset 1 and 160,874 from Dataset 2) were selected and subjected to multiple rounds of *ab initio* reconstructions and heterogeneous refinements. Finally, we obtained a 3D reconstruction at $4.0\ \text{\AA}$, which served as a good reference for subsequent data processing.

Using the good 2D averages as templates, more particles were selected from Dataset 2, yielding 1,112,782 particles. These particles were combined with the previously selected Dataset 1 particles and subjected to heterogeneous refinements, resulting in a subset of 206,762 good particles. After *ab initio* reconstruction, 127,650 particles were selected as seeds for seed-facilitated classification, yielding 234,254 good particles. After 3D classification, 80,288 particles were selected to generate a 3D reconstruction at 3.6 Å. After another round of seed-facilitated classification, 103,799 particles were selected to generate a 3.48 Å map. Finally, the map quality was improved to 3.21 Å after applying C2 symmetry and a local mask on the transmembrane region.

[TM6SF2-Sf9]

A total of 2,575,148 particles were extracted from the TM6SF2-Sf9 dataset. After heterogeneous refinement and *ab initio* reconstruction, 242,139 good particles were selected. These particles were selected as seeds for seed-facilitated classification to recover additional good particles from the initial particle pool, yielding 655,786 particles. After the bin1 particles were re-extracted, another round of seed-facilitated classification was applied, resulting in 335,862 good particles. These particles yielded a 3D reconstruction at an overall resolution of 3.32 Å.

Resolutions were estimated on the basis of the gold-standard Fourier shell correlation 0.143 criterion³⁵ with high-resolution noise substitution³⁶. Local resolution maps were

calculated using the “Local Resolution Estimation” in cryoSPARC³⁴.

Model building and refinement

The TM6SF2 structure predicted by AlphaFold3³⁷ was used as the initial model for docking into the maps via Chimera³⁸. After manual adjustment in Coot³⁹, the structures were further refined in real space via PHENIX⁴⁰, applying secondary structure and geometry restraints. All structural figures were generated with ChimeraX⁴¹.

Molecular Dynamics Simulations

All-atom molecular dynamics (MD) simulations were performed using the academic release of the Desmond molecular dynamics engine (D. E. Shaw Research, New York, NY)⁴², with system construction and trajectory analysis carried out through the Maestro interface (Schrödinger, LLC, New York, NY). The OPLS4 all-atom force field was employed for all simulations⁴³. The analysis followed the previous research methods^{44,45}. Briefly, separate simulation systems were constructed for dimeric TM6SF2 in the cholesterol-free (TM6SF2-Apo) and cholesterol-bound (TM6SF2-CLR) states. Each protein system was embedded in a preequilibrated 1-palmitoyl-2-oleoyl-sn-glycero-3-phosphocholine (POPC) lipid bilayer, with the membrane oriented such that the transmembrane domain was aligned along the bilayer normal. The systems were subsequently solvated with TIP4P explicit water molecules and neutralized by the addition of 0.15 M NaCl. The resulting systems were subjected to a default multi-step relaxation protocol, which included restrained minimization and short equilibration

stages designed to relieve unfavorable contacts while progressively relaxing restraints on the protein backbone and side chain atoms. Following relaxation, 150 ns of production MD was carried out in the NPT ensemble at 300 K and 1 bar, using a Nosé–Hoover thermostat and a Martyna–Tobias–Klein barostat, with a time step of 2 fs. Trajectory coordinates were saved every 500 ps for subsequent analysis.

Post-simulation analyses were conducted using the Simulation Analysis module within the Desmond/Maestro environment. Specifically, C α root-mean-square deviation (R.M.S.D) and root-mean-square fluctuation (R.M.S.F) were computed to assess global and residue-level structural stability, respectively. Time-dependent pocket-volume analysis was carried out to monitor fluctuations in the ligand-binding cavity. Hierarchical trajectory clustering was applied to the production ensemble to extract the most populated representative conformer for each system. Structural superimpositions were performed on C α atoms using the same module. All structural figures were exported at 300 DPI in TIFF format.

Statistics and Reproducibility

No statistical methods were used to predetermine the sample size. The experiments were not randomized, and the investigators were not blinded to allocation during the experiments or outcome assessment. Each experiment was conducted independently at least three times with similar results.

Data Availability

The atomic coordinates of TM6SF2-HEK293F and TM6SF2-Sf9 have been deposited in the PDB (<http://www.rcsb.org>) under the accession codes 45XK and 45XL. The corresponding electron microscopy map has been deposited in the Electron Microscopy Data Bank (<https://www.ebi.ac.uk/emdb/>) under the accession codes EMD-83602 and EMD-83603.

Code availability

No custom code or mathematical algorithm was used in this study.

References

- 1 Riazi, K. *et al.* The prevalence and incidence of NAFLD worldwide: a systematic review and meta-analysis. *Lancet Gastroenterol* **7**, 851-861, doi:10.1016/S2468-1253(22)00165-0 (2022).
- 2 Kozlitina, J. *et al.* Exome-wide association study identifies a variant that confers susceptibility to nonalcoholic fatty liver disease. *Nat Genet* **46**, 352-+, doi:10.1038/ng.2901 (2014).
- 3 Liu, Y. L. *et al.* rs58542926 influences hepatic fibrosis progression in patients with non-alcoholic fatty liver disease. *Nat Commun* **5**, doi:ARTN 4309 10.1038/ncomms5309 (2014).
- 4 Sookoian, S. *et al.* Genetic Variation in Transmembrane 6 Superfamily Member 2 and the Risk of Nonalcoholic Fatty Liver Disease and Histological Disease Severity. *Hepatology* **61**, 515-525, doi:10.1002/hep.27556 (2015).
- 5 Buch, S. *et al.* A genome-wide association study confirms PNPLA3 and identifies TM6SF2 and MBOAT3 as risk loci for alcohol-related cirrhosis. *Nat Genet* **47**, 1443-+, doi:10.1038/ng.3417 (2015).
- 6 Liu, D. J. *et al.* Exome-wide association study of plasma lipids in >300,000 individuals. *Nat Genet* **49**, 1758-+, doi:10.1038/ng.3977 (2017).
- 7 Holmen, O. L. *et al.* Systematic evaluation of coding variation identifies a candidate causal variant in TM6SF2 influencing total cholesterol and myocardial infarction risk. *Nat Genet* **46**, 345-+, doi:10.1038/ng.2926 (2014).
- 8 Mahdessian, H. *et al.* TM6SF2 is a regulator of liver fat metabolism influencing triglyceride secretion and hepatic lipid droplet content. *P Natl Acad Sci USA* **111**, 8913-8918, doi:10.1073/pnas.1323785111 (2014).
- 9 Smagris, E., Gilyard, S., BasuRay, S., Cohen, J. C. & Hobbs, H. H. Inactivation of Tm6sf2, a Gene Defective in Fatty Liver Disease, Impairs Lipidation but Not Secretion of Very Low Density Lipoproteins. *J Biol Chem* **291**, 10659-10676, doi:10.1074/jbc.M116.719955 (2016).
- 10 Ehrhardt, N. *et al.* Hepatic Tm6sf2 overexpression affects cellular ApoB-trafficking, plasma lipid levels, hepatic steatosis and atherosclerosis. *Hum Mol Genet* **26**, 2719-2731, doi:10.1093/hmg/ddx159 (2017).
- 11 Luo, F. *et al.* Hepatic TM6SF2 Is Required for Lipidation of VLDL in a Pre-Golgi Compartment in Mice and Rats. *Cell Mol Gastroenter* **13**, 879-899,

- doi:10.1016/j.jcmgh.2021.12.008 (2022).
- 12 Li, B. T. *et al.* Disruption of the ERLIN-TM6SF2-APOB complex destabilizes APOB and contributes to non-alcoholic fatty liver disease. *Plos Genet* **16**, doi:ARTN e1008955 10.1371/journal.pgen.1008955 (2020).
- 13 Newberry, E. P. *et al.* Liver-Specific Deletion of Mouse Tm6sf2 Promotes Steatosis, Fibrosis, and Hepatocellular Cancer. *Hepatology* **74**, 1203-1219, doi:10.1002/hep.31771 (2021).
- 14 O'Hare, E. A. *et al.* TM6SF2 rs58542926 Impacts Lipid Processing in Liver and Small Intestine. *Hepatology* **65**, 1526-1542, doi:10.1002/hep.29021 (2017).
- 15 Zhang, X. *et al.* Intestinal TM6SF2 protects against metabolic dysfunction-associated steatohepatitis through the gut-liver axis. *Nat Metab* **7**, doi:10.1038/s42255-024-01177-7 (2025).
- 16 Chen, L. Z. *et al.* Intestinal Depletion of TM6SF2 Exacerbates High-fat Diet-induced Metabolic Dysfunction-associated Steatotic Liver Disease through the Gut-liver Axis. *J Clin Transl Hepato* **13**, 443-455, doi:10.14218/Jcth.2024.00407 (2025).
- 17 Sanchez-Pulido, L. & Ponting, C. P. TM6SF2 and MAC30, new enzyme homologs in sterol metabolism and common metabolic disease. *Front Genet* **5**, doi:ARTN 439 10.3389/fgene.2014.00439 (2014).
- 18 Moebius, F. F., Striessnig, J. & Glossmann, H. The mysteries of sigma receptors: New family members reveal a role in cholesterol synthesis. *Trends Pharmacol Sci* **18**, 67-70, doi:10.1016/S0165-6147(96)01037-1 (1997).
- 19 Alon, A. *et al.* Structures of the σ_2 receptor enable docking for bioactive ligand discovery. *Nature* **600**, 759-+, doi:10.1038/s41586-021-04175-x (2021).
- 20 Long, T. *et al.* Structural basis for human sterol isomerase in cholesterol biosynthesis and multidrug recognition. *Nat Commun* **10**, doi:ARTN 2452 10.1038/s41467-019-10279-w (2019).
- 21 Shirzad-Wasei, N. *et al.* Rapid transfer of overexpressed integral membrane protein from the host membrane into soluble lipid nanodiscs without previous purification. *Biol Chem* **396**, 903-915, doi:10.1515/hsz-2015-0100 (2015).
- 22 Qian, H. W. *et al.* Structural Basis of Low-pH-Dependent Lysosomal Cholesterol Egress by NPC1 and NPC2. *Cell* **182**, 98-+, doi:10.1016/j.cell.2020.05.020 (2020).
- 23 Bronzert, T. J. & Brewer, H. B. New Micromethod for Measuring Cholesterol in Plasma Lipoprotein Fractions. *Clin Chem* **23**, 2089-2098 (1977).

- 24 Abramson, J. *et al.* Accurate structure prediction of biomolecular interactions with AlphaFold 3. *Nature* **630**, doi:10.1038/s41586-024-07487-w (2024).
- 25 Huang, H. Y. R. *et al.* Deep metabolic phenotyping of humans with protein-altering variants in TM6SF2 using a genome-first approach. *Jhep Rep* **7**, doi:ARTN 101243 10.1016/j.jhepr.2024.101243 (2025).
- 26 Isusi, E. *et al.* Short- and long-term effects of atorvastatin, lovastatin and simvastatin on the cellular metabolism of cholesteryl esters and VLDL secretion in rat hepatocytes. *Atherosclerosis* **153**, 283-294, doi:Doi 10.1016/S0021-9150(00)00407-X (2000).
- 27 Watts, G. F. *et al.* Inhibition of cholesterologenesis decreases hepatic secretion of apoB-100 in normolipidemic subjects. *Am J Physiol-Endoc M* **273**, E462-E470, doi:DOI 10.1152/ajpendo.1997.273.3.E462 (1997).
- 28 Watts, G. F. *et al.* Direct Correlation between Cholesterol-Synthesis and Hepatic Secretion of Apolipoprotein B-100 in Normolipidemic Subjects. *Metabolism* **44**, 1052-1057, doi:Doi 10.1016/0026-0495(95)90104-3 (1995).
- 29 Prinsen, B. H. C. M. T. *et al.* Endogenous cholesterol synthesis is associated with VLDL-2 apoB-100 production in healthy humans. *J Lipid Res* **44**, 1341-1348, doi:10.1194/jlr.M300023-JLR200 (2003).
- 30 Khan, B., Wilcox, H. G. & Heimberg, M. Cholesterol Is Required for Secretion of Very-Low-Density Lipoprotein by Rat-Liver. *Biochem J* **258**, 807-816, doi:DOI 10.1042/bj2580807 (1989).
- 31 Khan, B. V., Fungwe, T. V., Wilcox, H. G. & Heimberg, M. Cholesterol Is Required for the Secretion of the Very-Low-Density Lipoprotein - Invivo Studies. *Biochim Biophys Acta* **1044**, 297-304, doi:Doi 10.1016/0005-2760(90)90073-7 (1990).
- 32 Fungwe, T. V., Cagen, L., Wilcox, H. G. & Heimberg, M. Regulation of Hepatic Secretion of Very Low-Density-Lipoprotein by Dietary-Cholesterol. *J Lipid Res* **33**, 179-191 (1992).
- 33 McDonald, J. G., Smith, D. D., Stiles, A. R. & Russell, D. W. A comprehensive method for extraction and quantitative analysis of sterols and secosteroids from human plasma. *J Lipid Res* **53**, 1399-1409, doi:10.1194/jlr.D022285 (2012).
- 34 Punjani, A., Rubinstein, J. L., Fleet, D. J. & Brubaker, M. A. cryoSPARC: algorithms for rapid unsupervised cryo-EM structure determination. *Nat Methods* **14**, 290-296, doi:10.1038/nmeth.4169 (2017).
- 35 Rosenthal, P. B. & Henderson, R. Optimal determination of particle orientation, absolute hand, and contrast loss in single-particle electron cryomicroscopy. *J Mol Biol* **333**, 721-

- 745 (2003).
- 36 Chen, S. *et al.* High-resolution noise substitution to measure overfitting and validate resolution in 3D structure determination by single particle electron cryomicroscopy. *Ultramicroscopy* **135**, 24–35, doi:10.1016/j.ultramic.2013.06.004 (2013).
- 37 Tunyasuvunakool, K. *et al.* Highly accurate protein structure prediction for the human proteome. *Nature* **596**, 590–596, doi:10.1038/s41586-021-03828-1 (2021).
- 38 Pettersen, E. F. *et al.* UCSF Chimera--a visualization system for exploratory research and analysis. *J Comput Chem* **25**, 1605–1612, doi:10.1002/jcc.20084 (2004).
- 39 Emsley, P., Lohkamp, B., Scott, W. G. & Cowtan, K. Features and development of Coot. *Acta Crystallogr D Biol Crystallogr* **66**, 486–501, doi:10.1107/S0907444910007493 (2010).
- 40 Adams, P. D. *et al.* PHENIX: a comprehensive Python-based system for macromolecular structure solution. *Acta Crystallogr D Biol Crystallogr* **66**, 213–221, doi:10.1107/S0907444909052925 (2010).
- 41 Meng, E. C. *et al.* UCSF ChimeraX: Tools for structure building and analysis. *Protein Sci* **32**, e4792, doi:10.1002/pro.4792 (2023).
- 42 Bowers, K. J. *et al.* in *SC '06: Proceedings of the 2006 ACM/IEEE Conference on Supercomputing*. 43–43.
- 43 Lu, C. *et al.* OPLS4: Improving Force Field Accuracy on Challenging Regimes of Chemical Space. *J Chem Theory Comput* **17**, 4291–4300, doi:10.1021/acs.jctc.1c00302 (2021).
- 44 Manzano, J. A. H. *et al.* Dimeric Anthraquinone Rugulosin A Induces Apoptosis in Lung Adenocarcinoma and Targets PI3K/AKT/MAPK Pathways. *Acs Omega* **11**, 18744–18759, doi:10.1021/acsomega.5c09585 (2026).
- 45 Luchini, A., Lenz, J., Larsen, A. H. & Keller, S. Effect of Cholesterol on the Structure and Composition of Glyco- DIBMA Lipid Particles. *Langmuir* **39**, 3569–3579, doi:10.1021/acs.langmuir.2c03019 (2023).

Figure legends

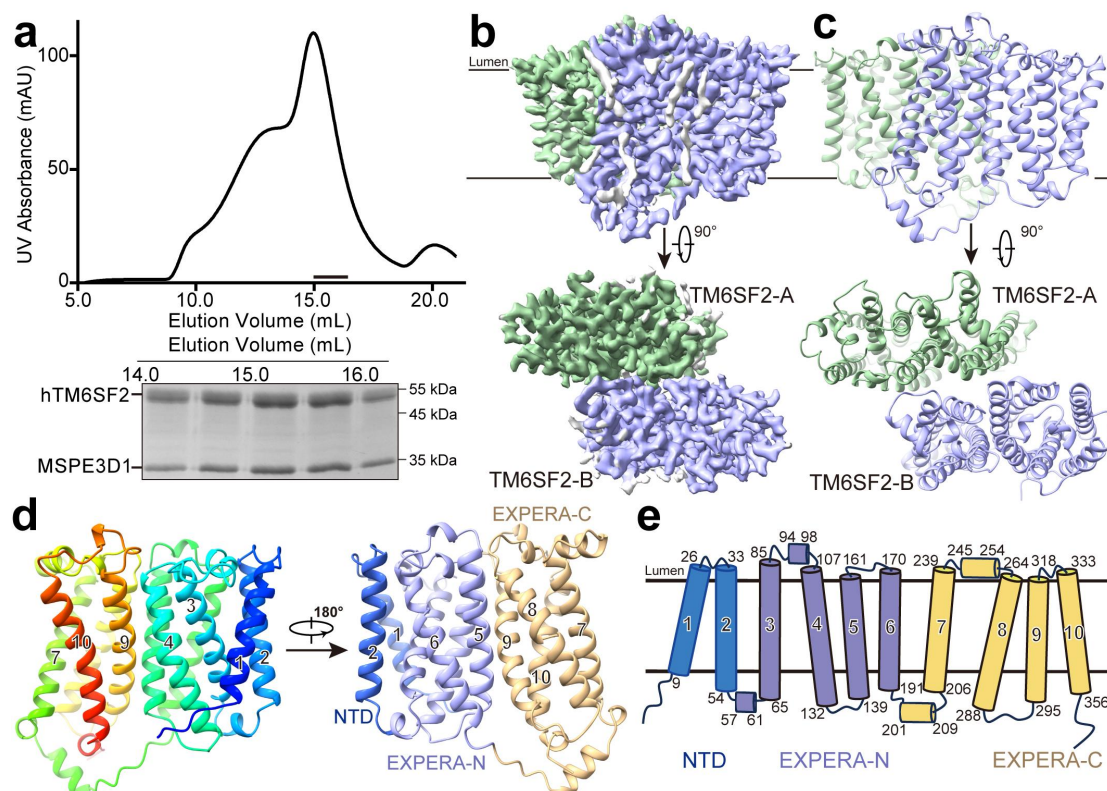


Figure 1 | Cryo-EM structure of human TM6SF2. **a**, Representative size-exclusion chromatography (SEC) profile of human TM6SF2 (hTM6SF2) reconstituted into nanodiscs with MSPE3D1 as the scaffold protein. After SDS-PAGE analysis, the peak fractions (indicated by the black line) were concentrated for cryo-EM analysis. **b**, **c**, Membrane (*top*) and luminal (*bottom*) views of the final cryo-EM reconstruction (**b**) and atomic model (**c**). EM maps are contoured at 5σ . The two hTM6SF2 protomers are colored green and light blue, and labeled TM6SF2-A and TM6SF2-B, respectively. **d**, Structure of a TM6SF2 protomer. *Left*: The structure is rainbow colored from the amino (N-, blue) terminus to the carboxyl (C-, red) terminus. *Right*: The structure is domain colored, with the N-terminal domain (NTD, TM1-2) in blue, the N-terminal EXPERA domain (EXPERA-N, TM3-6) in light blue, and the C-terminal EXPERA domain

(EXPERA-C, TM7-10) in wheat. **e**, Topological structure of hTM6SF2 based on the cryo-EM structure. The structural elements are colored according to the same scheme in panel **d**.

purified TM6SF2 variants were measured using LC-MS analysis. *Top*, bar graphs showing cholesterol levels in the indicated variants. *Bottom*, SDS-PAGE results showing comparable protein levels of the TM6SF2 variants used for LC-MS. Data are presented as the means $\pm$ S.D. of three independent experiments. Significant differences were determined with one-way ANOVA with Tukey's multiple comparisons tests. **g**, Comparison of the interaction between apoB-100 and TM6SF2 purified from the cholesterol-rich HEK293F expression system (TM6SF2-HEK293F) and the cholesterol-depleted Sf9 (TM6SF2-Sf9) expression system.

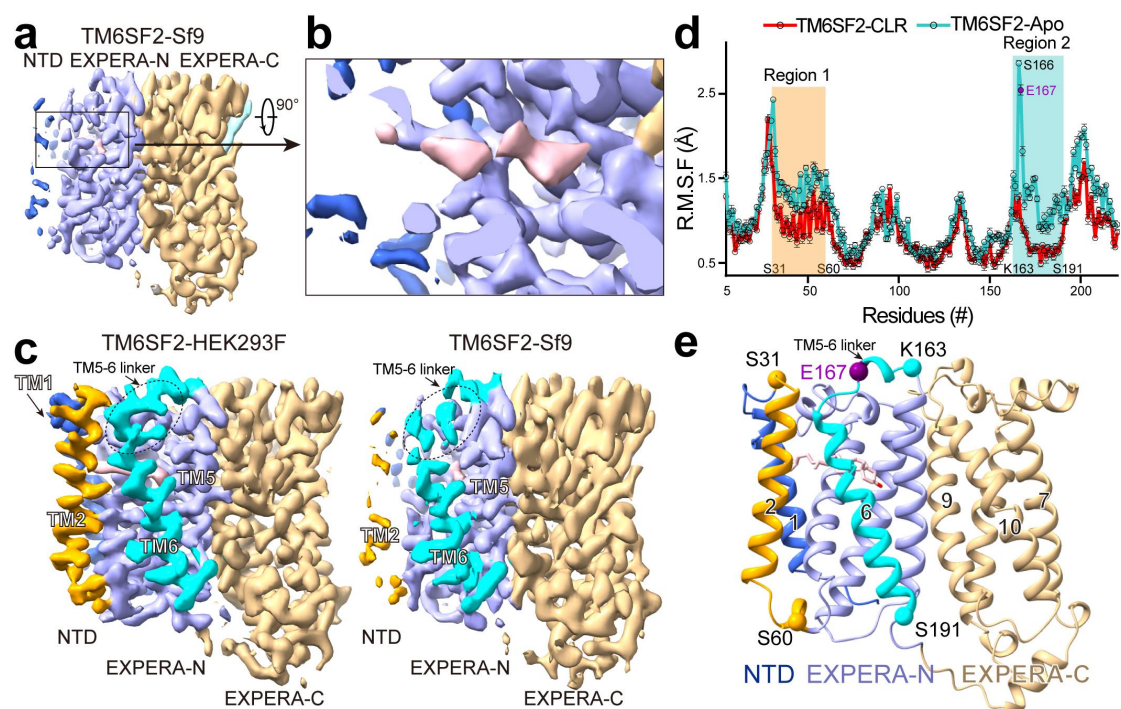


Figure 3 | Cholesterol maintains the stability of the NTD and EXPERA-N domains.

a, A membrane view of the TM6SF2-Sf9 map showing that the sterol-like densities in the EXPERA-N domain are absent and that the EM densities for the NTD and EXPERA-N domains become weaker and noisier. **b**, A zoomed-in view to show that the extra densities in the cholesterol-binding pocket of the EXPERA-N domain become weaker and disconnected. See Fig. 2b for comparison with the sterol-like densities in the TM6SF2-HEK293F structure. The EM maps are contoured at 6σ , which is identical to the contour level used in Fig. 2b. **c**, TM6SF2-HEK293F (*left*) and -Sf9 (*right*) maps are presented with identical views and contoured levels (6σ) for comparison. **d**, R.M.S.F. analysis of TM6SF2 residues during a 150-ns MD simulation under cholesterol-bound

(TM6SF2-CLR) and cholesterol-free (TM6SF2-Apo) conditions. Residues 31-60 (Region 1) and 163-191 (Region 2) exhibit obvious differences in R.M.S.F. values between these two conditions and are highlighted with orange and cyan boxes, respectively. R.M.S.F., Root mean square fluctuation. Data are presented as means $\pm$ S.E.M of three independent simulations. **e**, Mapping of the Region 1 and Region 2 residues onto our TM6SF2-HEK293F structure. In the structure, the residues in Regions 1 and 2 are colored orange and cyan, respectively. In panel **c**, the EM maps corresponding to these two regions are also colored orange and cyan.

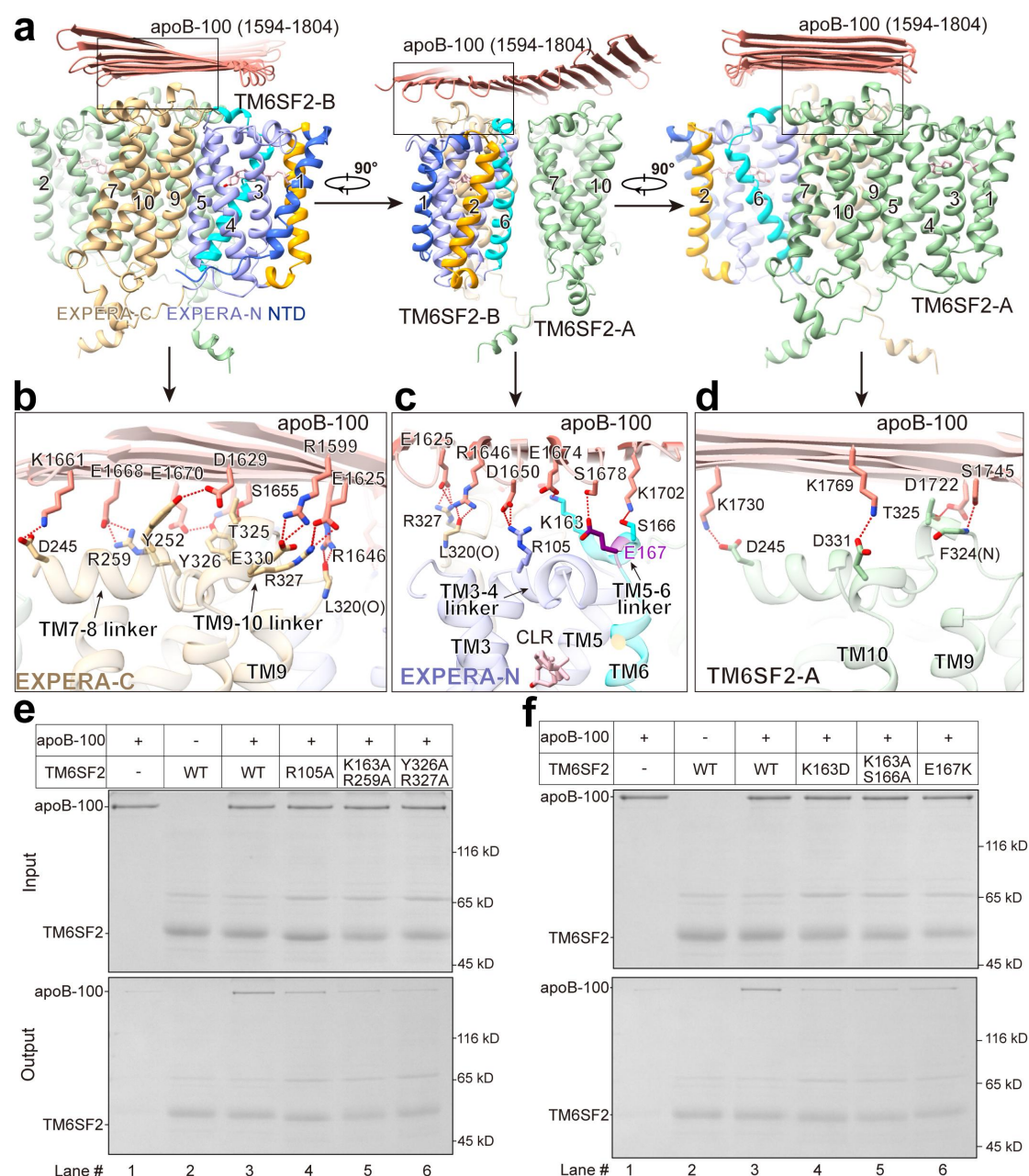


Figure 4 | Predicted interaction between hTM6SF2 and apoB-100. **a**, Three membrane views showing the interaction between dimeric TM6SF2 and residues 1594-1804 of apoB-100. Among the two protomers in the TM6SF2 dimer, TM6SF2-B constitutes the major interaction site for apoB-100. **b**, **c**, Two zoomed-in views showing the predicted interactions of apoB-100 with the EXPERA-C (**b**) and EXPERA-N (**c**)

domains in the TM6SF2-B protomer. **d**, A zoomed-in view showing the predicted interaction between apoB-100 and the other protomer, TM6SF2-A. Salt bridges and hydrogen-bonds are indicated with red dotted lines in panels **b-d**. **e**, Biochemical validation of the predicted TM6SF2-apoB-100 interfaces using the *in vitro* pull-down assay. **f**, Biochemical validation of the residues in the TM5-6 linker involved in the TM6SF2-apoB-100 interaction by the *in vitro* pull-down assay.