The ghrelin O-acyltransferase structure reveals a catalytic channel for transmembrane hormone acylation

Received for publication, June 10, 2019, and in revised form, August 12, 2019 Published, Papers in Press, August 14, 2019, DOI 10.1074/jbc.AC119.009749

Maria B. Campaña, Flaviyan Jerome Irudayanathan, Tasha R. Davis, Kayleigh R. McGovern-Gooch, Rosemary Loftus, Mohammad Ashkar, Najae Escoffery, Melissa Navarro, Michelle A. Sieburg, Shikha Nangia, and James L. Hougland

This work was supported by Syracuse University, American Diabetes Association Grants 1-16-JDF-042 and 7-18-MUI-001 (to J. L. H.), and National Science Foundation Grant CAREER CBET-1453312 (to S. N.). The authors declare that they have no conflicts of interest with the contents of this article.

This article contains Tables S1 and S2, Figs. S1–S10, and Files S1–S4.

Integral membrane proteins represent a large and diverse portion of the proteome and are often recalcitrant to purification, impeding studies essential for understanding protein structure and function. By combining co-evolutionary constraints and computational modeling with biochemical validation through site-directed mutagenesis and enzyme activity assays, we demonstrate here a synergistic approach to structurally model purification-resistant topologically complex integral membrane proteins. We report the first structural model of a eukaryotic membrane-bound O-acyltransferase (MBOAT), ghrelin O-acyltransferase (GOAT), which modifies the metabolism-regulating hormone ghrelin. Our structure, generated in the absence of any experimental structural data, revealed an unanticipated strategy for transmembrane protein acylation with catalysis occurring in an internal channel connecting the endoplasmic reticulum lumen and cytoplasm. This finding validated the power of our approach to generate predictive structural models for other experimentally challenging integral membrane proteins. Our results illuminate novel aspects of membrane protein function and represent key steps for advancing structure-guided inhibitor design to target therapeutically important but experimentally intractable membrane proteins.

Integral membrane proteins represent a large and essential portion of the proteome, including a growing number of enzymes, receptors, and transporters that serve as desirable drug targets (1, 2). However, these proteins often prove recalcitrant to purification and structural analysis due to their hydrophobic nature and reliance on interactions with lipid bilayers for both stability and activity (3, 4). We report here a synergistic approach to develop a structural model of a topologically complex integral membrane protein by combining co-evolutionary contact constraints and computational modeling with biochemical validation. Building solely upon the protein’s primary sequence and a biochemical assay for its function, the approach provides an accessible and efficient route to build structural models of intractable membrane protein targets. We demonstrate this approach by developing a structural model for ghrelin O-acyltransferase (GOAT), a member of the membrane-bound O-acyltransferase (MBOAT) enzyme family responsible for octanoylation of the peptide hormone ghrelin (Fig. 1) (5, 6). One of three protein-modifying MBOAT family members alongside Hedgehog acyltransferase (Hhat) and Porcupine (Porcn) (7–9), GOAT plays a central role in regulating energy homeostasis and metabolism through octanoylated ghrelin-dependent signaling pathways (10). Whereas the unique chemistry and biology of ghrelin and GOAT have inspired continued efforts to target this system for therapeutic benefit, the inability to purify active GOAT and determine its structure has hampered progress toward this goal (11–13). In this work, we report the first structural model for an eukaryotic MBOAT family member. Our human GOAT (hGOAT) structure is highly consistent with a recently reported crystal structure for the bacterial MBOAT homolog D-alanyl transferase DltB (14). Our structure suggests a novel strategy for solving the topological challenge presented by transmembrane protein acylation, where protein targets and co-substrates are separated by a cellular membrane. In an unanticipated mechanism, ghrelin octanoylation occurs in an internal channel within hGOAT without the octanoyl-CoA donor being transported into the endoplasmic reticulum (ER) lumen. The availability of this therapeutically interesting enzyme’s structure opens the door to the structure-guided design of inhibitors targeting GOAT and other MBOAT family members. Looking beyond the MBOATs, our success in modeling GOAT and predicting specific protein–ligand interactions validates the power of our approach for creating molecular models for other experimentally challenging integral membrane proteins.

The abbreviations used are: GOAT, ghrelin O-acyltransferase; hGOAT, human ghrelin O-acyltransferase; MBOAT, membrane-bound O-acyltransferase; ER, endoplasmic reticulum; TM, transmembrane; IM, intramembrane; MSA, multiple-sequence alignment; hmm, hidden Markov model.

© 2019 Campaña et al. Published under exclusive license by The American Society for Biochemistry and Molecular Biology, Inc.
Results and discussion

Computational model for human GOAT structure

In generating our hGOAT structural model, we utilized state-of-the-art co-evolutionary contact predictions with computational protein folding and structure optimization methods (Fig. S1) (15, 16). Coevolutionary contact analysis exploits the tendency of residues interacting with each other within folded proteins to co-evolve to maintain energetically beneficial interactions (17–19). Analysis of many protein sequences employing a multiple-sequence alignment identifies pairs of co-evolving residues, from which it is inferred that these residues lie in proximity to each other. Assigning pairs of residues as co-evolving supports assignment of spatial interactions between them, providing constraints that can define major features of protein structure. Using metagenomics protein databases, we generated a multiple-sequence alignment to predict residues that are potentially in contact (defined as $\Delta_{C_{\beta}} < 8$ Å) with each other in the folded structure of hGOAT (File S1) (15, 20). This set of contacts (File S2), represented by the contact map (Fig. 1C and Fig. S2), guided our hGOAT structural modeling (17, 20, 21). Experimental information on the membrane topology of mouse GOAT and co-evolutionary contact constraints were iteratively

Figure 1. Structural model of hGOAT generated by computational methods. A, schematic of ghrelin octanoylation by hGOAT showing the predicted transmembrane topology of hGOAT containing 11 transmembrane helix domains (TM1–11), two intramembrane domains (IM1–2), and loop regions generated using Protter (43). B, octanoylation of a ghrelin-mimetic fluorescent peptide by recombinant hGOAT. C, contact maps for hGOAT showing the probability for a co-evolutionary contact from RaptorX analysis (i) and amino acid contacts in the final optimized hGOAT structure (ii). D, structure of hGOAT in an ER-mimetic lipid membrane, correlated to color-coded membrane topology in A. E, illustration of the internal channel within hGOAT (green) transiting from the ER lumen to the cytoplasm, with the channel determined by the CAVR 3.0 plugin in PyMOL (33). F, structural overlay of hGOAT and DltB showing the absolutely conserved histidine residues (hGOAT His-338 (teal) and DltB His-336 (purple) (Protein Data Bank code 6BUG, chain C) within these acyltransferases.
combined in protein-folding simulations to generate ~30,000 potential hGOAT structures (11, 17, 22). The generated structures were clustered, and the lowest-energy structures that satisfied the contact map were isolated (Fig. S3) (22). Representative structures from the top five clusters were then subjected to further structural refinement to yield the optimal hGOAT model (23). The optimal model was embedded in a lipid membrane and subjected to structural relaxation in explicit solvent using all-atom molecular dynamics simulations (24–26). This simulation used an ER-mimetic lipid bilayer to ensure optimization of hydrophobic protein–lipid interactions (27).

**Features of the human GOAT structure**

Our computationally derived structure for hGOAT is consistent with the previously reported topological model of the mouse GOAT ortholog containing a total of 11 transmembrane helices with slightly altered helix boundaries (Fig. 1A) (11), indicating the two sets of constraints from our coevolutionary contact analysis and previous topological studies support a common hGOAT structural model. To determine how strongly our hGOAT structure depends on the experimental topological constraints from mouse GOAT (11), we excluded these constraints and repeated our analysis, which generated an identical hGOAT membrane topology. This indicates that co-evolutionary contact constraints alone are sufficient to predict the membrane topology of hGOAT, suggesting this approach for topology modeling of integral membrane proteins to complement established algorithms for predicting membrane protein topology.

Ramachandran analysis indicates that 92.4% (400 of 433) of hGOAT residues lie in favored (98%) regions, 98.2% (425 of 433) lie in allowed (>99.8%) regions, and 1.9% (8 of 433) are outliers (Fig. S4) (28). The enzyme forms an ellipsoidal cone composed of transmembrane helices, with the narrow end facing the ER lumen (Fig. 1D). The exposed ends of five transmembrane helices (TM1, TM4, TM5, TM7, and TM11) converge to form a pore through which the interior of hGOAT is connected to the ER lumen. At the cytoplasmic membrane interface, the predicted cytoplasmic loops fold up to form a core region bounded by the lipid-contacting perimeter helices. As a result, there is minimal cytoplasmic exposure of hGOAT residues beyond the plane of the membrane.

The hGOAT structure contains a contiguous internal channel through the enzyme core that transits from the ER lumen space to the cytoplasm (Fig. 1E). The channel is bent within hGOAT, with the restriction formed by the C-terminal end of helix TM8 and the N-terminal end of TM9. This positions an absolutely conserved histidine residue (His-338) in direct contact with the internal channel (7), consistent with proposals for this histidine to serve as a general base for catalyzing ghrelin acylation. Following completion of our hGOAT structure and during subsequent biochemical validation experiments (described below), the release of a crystal structure for bacterial MBOAT alanyl transferase DltB provided an independent basis for comparison and validation of our hGOAT structure (14). The His-338 residue in hGOAT closely matches the location of the analogous histidine residue (His-336) in the DltB structure (Fig. 1F) (5, 14). Further comparison of the hGOAT model and the DltB structure reveals remarkable similarities in overall topology and structure, with a TM-score of 0.6 and root mean square deviation of 2.23 Å for ~100 aligned conserved residues between the structural models for these distantly related MBOAT family members (12.3% sequence identity, 26.8% sequence similarity, E-value 2.7 × 10^{-8} and bit score 48.7; Figs. S5 and S6 and Table S1) (29). However, the low overall homology between DltB and hGOAT leads to very poor structure prediction for nonhomologous sequence positions, as would be expected for this type of comparison. The demonstrated ability of our hGOAT modeling based on coevolutionary contact restraints to arrive at the same protein fold as DltB, in the absence of any experimental structural information, underscores the power of this approach to accurately predict protein structures.

**Mutagenesis analysis of hGOAT structural model**

To validate our computational hGOAT structural model biochemically, we mutated ~10% of the residues within hGOAT to alanine and determined the impact of these mutations on hGOAT octanoylation activity in a peptide-based acylation assay (Fig. 2 and Fig. S7) (30, 31). These 42 alanine mutations were spread across a range of amino acids and degrees of conservation, with the majority of sites chosen conserved at >75% among GOAT orthologs (File S3). In narrowing the pool of mutations to ~40 positions, residues with surface-exposed side chains were deemphasized compared with residues predicted to lie within the enzyme interior. Approximately half of the mutation sites were selected based on the residue’s side chain contacting the internal void, as we propose this channel will likely contain the substrate-binding sites and catalytic residues within hGOAT.

In this pool of alanine mutants, we observed a range of activities from near/above WT ghrelin octanoylation activity to complete loss of detectable activity (Table S2 and Fig. S8). When mapped onto the hGOAT structural model, mutations leading to a marked decline (>3-fold; purple) or loss of enzyme activity (red) appear clustered within the core of hGOAT (Fig. 2). For quantitative analysis of the impact of these mutations, we determined whether alanine mutagenesis of residues contacting the internal void is more likely to yield reduced enzyme activity compared with non-void-contacting mutations. Within the pool of mutations, the void-contacting alanine mutations were significantly more likely to result in loss of enzyme activity (p < 0.03; Fig. 2D). This mutation activity mapping defines a functionally essential core within hGOAT and expands the number of residues within hGOAT known to be required for enzyme activity (5, 6, 11).

**The octanoyl-CoA–binding site within hGOAT**

We expect the octanoyl-CoA acyl donor to enter the hGOAT active site through interaction with the cytoplasmic face of the enzyme, based on the availability of acyl-CoAs within the cell. When docked into our hGOAT model, octanoyl-CoA binds to hGOAT through interactions of both its CoA and octanoyl chain regions with residues in TM6, the TM7–TM8 connecting loop, TM8, and TM9 (Fig. 3). In the docked complex, the CoA portion forms both polar and nonpolar interactions with mul-
multiple hGOAT residues while remaining exposed to the cytoplasm (Fig. 3, A and B). The phosphoadenosine group binds into a discrete pocket while the phosphopantetheine chain is in contact with multiple polar amino acid side chains (Fig. 3, C–E). Among these CoA-contacting amino acids, all alanine mutations examined except one lead to a loss of hGOAT activity.

In the docked hGOAT:octanoyl-CoA structure, the acyl chain of octanoyl-CoA makes a sharp turn and penetrates upward into the interior of hGOAT following a channel that terminates at Trp-351 (Fig. 3, C–E). Given the unique preference of hGOAT for an octanoyl acyl donor (6, 13, 32), we examined alanine mutagenesis of predicted contacts within this acyl-binding pocket to determine the impact of those mutations on hGOAT acyl donor selectivity. As alanine mutagenesis would provide additional space within the acyl-binding site, we determined the ability of hGOAT alanine variants to accept 12-carbon (lauryl-CoA) and 14-carbon (myristoyl-CoA) acyl donors in place of octanoyl CoA. The WT enzyme and the majority of hGOAT alanine variants exhibited the expected preference for an eight-carbon acyl donor, but alanine mutagenesis of Trp-351 and Phe-331 resulted in loss of appreciable reactivity with octanoyl-CoA but engendered new activity with the longer acyl donors (Fig. 3F and Fig. S9). The F331A variant gained activity with the C12 donor, whereas W351A hGOAT could acylate a ghrelin-derived peptide with both C12 and C14 acyl chains. This altered selectivity supports the modeled positions of Trp-351 and Phe-331 as forming the end of the acyl-binding pocket. The altered preference for longer acyl donors by the F331A and W351A variants was also observed in a direct competition assay where hGOAT variants were provided acyl donors ranging from six to 12 carbons (Fig. 3G). This altered selectivity was not observed for any other alanine variants with detectable ghrelin acylation activity (Fig. S10). Acyl donor reengineering upon targeted alanine mutagenesis localizes Phe-331 and Trp-351 to the distal end of the acyl donor–binding site within hGOAT and provides further biochemical validation of our hGOAT structural model.

Although structural studies play a central role in developing our understanding of protein function, the limited availability of integral membrane proteins within structural databases creates a particularly acute challenge for structurally modeling these proteins (see http://blanco.biomol.uci.edu/mpstruc/). In this work, we demonstrate the development and validation of a structural model for an integral membrane protein that leverages bioinformatics constraints from coevolutionary contact analysis and model evaluation by biochemical analysis while circumventing the requirement of protein purification.

Our model provides indispensable and novel insights into several long-standing questions regarding the mechanism for

---

6 Please note that the JBC is not responsible for the long-term archiving and maintenance of this site or any other third party hosted site.
MBOAT-catalyzed transmembrane protein acylation. The topological separation of two essential conserved residues, His-338 and Asn-307, is explained by these two residues playing roles in distinct aspects of GOAT activity. The location of His-338 within the central channel of GOAT, identical to the position observed for the analogous histidine in DltB, (14) is consistent with this residue acting as a general base to activate the ghrelin serine hydroxyl side chain for octanoyl transfer. In contrast, our hGOAT:octanoyl-CoA model implicates Asn-307 in the binding site for the acyl donor. Based on our models of hGOAT and the hGOAT:octanoyl-CoA complex, we propose that hGOAT catalyzes transmembrane acylation of ghrelin by binding both substrates within the hGOAT internal channel and “handing off” the octanoyl group from CoA to ghrelin within this channel.
 Whereas many aspects of this proposed pathway—such as the ghrelin-binding site and location of catalytic residues—remain to be functionally validated by ongoing studies, the established ability of our hGOAT model to efficiently guide biochemical studies demonstrates a novel approach to advance investigations of similar membrane proteins that are intractable to current structural approaches.

**Experimental procedures**

**Co-evolutionary contact analysis of hGOAT**

A multiple-sequence alignment (MSA) was performed with the hGOAT sequence against the UNIREF90 database utilizing the JackHMMER tool (34). The MSA parameters were set to eight iterative searches ($n = 8$) with an $e$-value threshold of $1 \times 10^{-40}$. The resulting alignment was filtered to exclude highly similar sequences using the HHfilter tool with 90% identity and 75% sequence coverage cut-offs. This MSA was used as the input for the hmmbuild tool to construct a hidden Markov model (hmm) curated specifically for the MSA (35), which would represent the consensus sequence of hGOAT and its closest homologs. This hmm was then utilized to search against a master database that included uniref100 and metagenome database (metaclust_2018_01) using the hmmsearch tool with a bit score cut-off of 27 (15, 21). The resulting MSA was filtered again using the HHfilter tool with 90% identity and 75% sequence coverage against hGOAT. Furthermore, sequences with unidentified amino acids (X; this is to accommodate for RaptorX) and sequence positions with ≥50% gaps were also filtered from the MSA using trimAL (36). The resulting MSA had an $M_{eff} - 0.8/N$ of 551.7, which is greater than the recommended value of 64 for reliable model prediction using co-evolutionary contacts (15, 17). These contacts were used to guide the hGOAT folding. The MSA analysis and curations were performed using in-house Python scripts and the ConKit Python library (38).

**Folding simulations**

The folding simulations were performed in two stages. In both stages, contact restraints were used, and the models were iteratively clustered, refined, and scored based on their overall backbone energy. Full details of the folding simulation protocols and software are provided in the supporting information.

**Refinement and relaxation using molecular dynamics**

The optimized hGOAT model from stage 2 was oriented with respect to a membrane bilayer using the PPM server (25). The calculated hydrophobic thickness of the hGOAT structural model is $25.2 \pm 2.4 \, \text{Å}$, with a tilt angle of 3° relative to the membrane normal vector. The oriented protein was then embedded in an ER-mimetic lipid bilayer (1:1 dipalmitoylphosphatidylcholine/dioleoyl phosphatidylcholine) using the CHARMM-GUI web server and subject to an all-atom equilibration at 310.15 K in explicit solvent and 150 mM NaCl counterions (24, 39). The simulation was carried out for 500 ns using GROMACS 2016.4, and the structural deviations were monitored (40). The equilibrated structure was isolated and utilized for prediction of internal channels and docking studies.

---

**Figure 4. Proposed pathway for transmembrane ghrelin octanoylation by GOAT.** Ghrelin (GSSFL-ghrelin) and octanoyl-CoA enter the GOAT internal channel from the ER luminal pore and cytoplasmic acyl donor–binding sites, respectively, followed by acyl transfer to the ghrelin serine side chain hydroxyl. Octanoylated ghrelin dissociates to the ER lumen, resulting in the octanoyl chain transiting through the GOAT interior, and CoA is released back to the cytoplasm. The red and blue rectangles represent perimeter helices, the green rectangle represents intramembrane domains forming the cytoplasmic surface of hGOAT, and dotted lines represent binding interactions between the octanoyl-CoA acyl donor and its binding site within hGOAT. (Fig. 4).

---

**ACCELERATED COMMUNICATION: Molecular structure of GOAT**

*J. Biol. Chem.* (2019) 294(39) 14166–14174 14171
ACCELERATED COMMUNICATION: Molecular structure of GOAT

Molecular docking and relaxation of hGOAT:octanoyl-CoA complex

To build a model of the hGOAT:octanoyl-CoA bound complex, we performed docking using Autodock Vina implemented in the YASARA software suite (41, 42); full details of the docking procedure are provided in the supporting information.

General experimental methods

Data plotting and curve fitting were carried out with Kaleidagraph (Synergy Software, Reading, PA). Membrane topology schematics were generated using Protter (http://wlab.ethz.ch/protter/start/) (43), and structural figures were generated using Chemdraw Prime 15.1 and PyMOL. Hexanoyl-CoA (hexanoyl-CoA, free acid) (Crystal Chem Inc.), octanoyl-CoA (octanoyl-CoA, free acid) (AdventBio), decanoyl-CoA (decanoyl-CoA, free acid) (Crystal Chem Inc.), lauroyl-CoA (lauroyl-CoA, free acid) (Crystal Chem Inc.), and myristoyl-CoA (myristoyl-CoA, tetracanoyl-CoA, free acid) (Crystal Chem Inc.) were solubilized to 5 mM in 10 mM Tris-HCl (pH 7.0), aliquoted into low-adhesion microcentrifuge tubes, and stored at −80 °C. Methoxy arachidonyl fluorophosphonate was purchased from Cayman Chemical (Ann Arbor, MI) and solubilized with DMSO. Unlabeled GSSFLC_{NH2} peptide was synthesized by Sigma-Genosys (The Woodlands, TX), solubilized in 1:1 acetonitrile/H$_2$O, and stored at −80 °C. Acrylodan (Anaspec) for peptide substrate labeling was solubilized in acetonitrile with the stock concentration determined by absorbance at 393 nm in methanol ($\varepsilon_{393} = 18,483$ M$^{-1}$ cm$^{-1}$, per the manufacturer’s data sheet). GSSFLC_{NH2} peptide concentrations were determined by reaction of the cysteine thiol with 5,5′-dithiobis(2-nitrobenzoic acid) and absorbance at 412 nm, using $\varepsilon_{412} = 14,150$ M$^{-1}$ cm$^{-1}$ (44).

Peptide substrate fluorescent labeling

The GSSFLC_{NH2} peptide substrate used in the hGOAT acylation assay was fluorescently labeled and purified using protocols reported previously (30, 45, 46). The concentration of acrylodan-labeled GSSFLC_{NH2} was calculated using absorbance of acrylodan at 360 nm ($\varepsilon = 13,300$ M$^{-1}$ cm$^{-1}$) (30, 47).

Construction of hGOAT mutants

Site-directed mutagenesis was performed on our previously reported hGOAT expression construct as described in the supporting information and File S4 (30). This construct was commercially synthesized by Integrated DNA Technologies (Corvally, IA) containing a C-terminal FLAG epitope tag, a polyhistidine (His$_6$) tag, and 3′X human influenza hemagglutinin tags appended downstream of a tobacco etch virus protease site (48).

Expression and enrichment of hGOAT in membrane protein fractions

hGOAT WT and mutants were expressed in insect (S9) cell membrane fractions using procedures published previously (13, 30, 46, 49).

hGOAT expression analysis by anti-FLAG Western blotting

Expression of hGOAT was determined by anti-FLAG Western blotting using published protocols (Fig. S7) (46). Each gel contained an empty vector microsomal protein as negative control and N-terminal FLAG-BAP fusion protein as a positive control (Millipore Sigma, P7582-100UG, 1:150 dilution, 30-μl total volume).

Following electrophoretic separation and transfer to a polyvinylidene difluoride membrane, the membrane was probed with a FLAG antibody (horseradish peroxidase–conjugated DYKDDDDK tag antibody, Invitrogen catalogue no. PA1-984B-HRP, 1:1000 dilution, 10-ml total volume) in 5% nonfat milk in TBST buffer (Tris-buffered saline, 0.1% Tween 20) overnight at 4 °C. The membrane was treated with West Pico Chemiluminescent substrate-imaging reagent (Thermo Scientific) followed by imaging on a ChemiDoc XRS+ gel documentation system (Bio-Rad).

hGOAT activity assay: Standard reaction conditions

hGOAT activity assays under standard conditions were performed with 50 μg of membrane protein, 1.5 μM fluorescent peptide substrate, 300 μM octanoyl-CoA, 1 μM methoxy arachidonyl fluorophosphonate, and 50 μM HEPES, pH 7.0, in a total volume of 50 μl as described previously (46). All components except for the peptide and acyl-CoA substrates were incubated at room temperature for 30 min prior to reaction initiation by the addition of peptide and acyl-CoA substrates. Reactions were incubated at room temperature for 2 h in the dark and then stopped by the addition of 50 μl of 20% acetic acid in isopropyl alcohol. Reaction solutions were clarified and analyzed by reverse-phase HPLC (46). Substrate and acylated products were detected by fluorescence ($\lambda_{ex}$ 360 nm, $\lambda_{em}$ 485 nm), with the substrate eluting with a retention time of 5–6 min and the octanoylated peptide eluting with a retention time of 11–12 min. Chromatogram analysis and peak integration was performed using Chemstation for LC (Agilent Technologies) (30). Product conversion was calculated by dividing the integrated fluorescence for the product peak by the total integrated peptide fluorescence (substrate and product) in each run. Percentage activity for each hGOAT mutant was calculated by normalizing the product conversion for the mutant to that of WT hGOAT in a reaction run in parallel on the same day using the same reagents.

Statistical analysis of hGOAT alanine variant reactivity

Full details of hGOAT alanine variant statistical testing using a Wilcoxon signed-rank test ($n = 42$, test statistic $W = 294.5$, $p = 0.02978$) are provided in the supporting information, including the R script utilized (50).

Single acyl donor reactivity assay

To determine the reactivity of hGOAT variants with different length acyl donors, hGOAT activity was measured in the presence of a 100 μM octanoyl-CoA, lauryl (dodecanoyl)-CoA, or myristoyl (tetradecanoyl)-CoA. Characteristic retention times for each acylated form of the peptide substrate provided confirmation of the nature of the attached acyl chain, with dodecanoyl-GSSFLC$_{AcDan}$ eluting at ~17 min and tetradecanoyl-GSSFLC$_{AcDan}$ eluting at ~19 min. For each acyl donor, relative activity was calculated normalized...
to the highest activity observed across the panel of WT hGOAT and hGOAT variants.

**Acyl donor competition assay**

To determine the relative preference of each hGOAT variant for acyl donors ranging from six to 12 carbons, hGOAT activity was measured in the presence of a 100 μM concentration each of four potential acyl donors (hexanoyl-CoA, octanoyl-CoA, decanoyl-CoA, and laurlyl (dodecanoyl)-CoA). Each potential product peak was assigned by retention time compared with a standard reaction containing only one acyl donor for each potential product. Competition experiments including myristoyl-CoA were unsuccessful, potentially due to low critical micelle concentration for this acyl donor lying near 100 μM (51).

**Author contributions**—M. B. C., F. J. I., K. R. M.-G., M. A. S., S. N., and J. L. H. conceptualization; M. B. C., F. J. I., T. R. D., K. R. M.-G., R. L., M. A., N. E., M. N., M. A. S., S. N., and J. L. H. investigation; M. B. C., F. J. I., T. R. D., K. R. M.-G., R. L., M. A., N. E., M. N., M. A. S., S. N., and J. L. H. methodology; M. B. C., F. J. I., S. N., and J. L. H. writing-original draft; M. B. C., F. J. I., T. R. D., K. R. M.-G., R. L., M. A. S., S. N., and J. L. H. writing-review and editing; S. N. and J. L. H. supervision; S. N. and J. L. H. funding acquisition.

**Acknowledgments**—We thank Profs. Jason Fridley and John Chisholm (Syracuse University) for assistance with statistical analysis and figure generation. We also thank Prof. Jinbo Xu (Toyota Technological Institute at Chicago), Prof. Sergey Ovchinnikov (Harvard University), and Prof. Badri Adhikari (University of Missouri, St. Louis) for advice regarding computational modeling. This material is based in part upon work supported by the National Science Foundation under Grant CHE-1659775. We also gratefully acknowledge computational resources provided by Syracuse University research computing and the Extreme Science and Engineering Discovery Environment (XSEDE), which is supported by National Science Foundation Grant NSF-ACI-1053575.

**References**

1. Rask-Andersen, M., Almén, M. S., and Schiotth, H. B. (2011) Trends in the exploitation of novel drug targets. *Nat. Rev. Drug Discov.* 10, 579–590 [CrossRef Medline]

2. Overington, J. P., Al-Lazikani, B., and Hopkins, A. L. (2006) How many drug targets are there? *Nat. Rev. Drug Discov.* 5, 993–996 [CrossRef Medline]

3. Stansfeld, P. J., Goose, J. E., Caffrey, M., Carpenter, E. P., Parker, J. L., Newstead, S., and Sansom, M. S. (2015) Memprotmd: automated insertion of membrane proteins into explicit lipid membranes. *Structure* 23, 1350–1361 [CrossRef Medline]

4. Bill, R. M., Henderson, P. J., Iwata, S., Kunji, E. R., Michel, H., Neutze, R., Newstead, S., Poolman, B., Tate, C. G., and Vogel, H. (2011) Overcoming barriers to membrane protein structure determination. *Nat. Biotechnol.* 29, 335–340 [CrossRef Medline]

5. Yang, J., Brown, M. S., Liang, G., Grishin, N. V., and Goldstein, J. L. (2008) Identification of the acyltransferase that octanoylates ghrelin, an appetites-timulating peptide hormone. *Cell* 132, 387–396 [CrossRef Medline]

6. Gutierrez, J. A., Solenberg, P. J., Perkins, D. R., Willency, J. A., Knierman, M. D., Jin, Z., Witcher, D. R., Luo, S., Onyia, J. E., and Hale, J. E. (2008) Ghrelin octanoylation mediated by an orphan lipid transferase. *Proc. Natl. Acad. Sci. U.S.A.* 105, 6320–6325 [CrossRef Medline]

7. Hofmann, K. (2000) A superfamily of membrane-bound O-acyltransferases with implications for wnt signaling. *Trends Biochem. Sci.* 25, 111–112 [CrossRef Medline]

8. Masumoto, N., Lanyon-Hogg, T., Rodgers, U. R., Konitsiotis, A. D., Magee, A. I., and Tate, E. W. (2015) Membrane bound O-acyltransferases and their inhibitors. *Biochem. Soc. Trans.* 43, 246–252 [CrossRef Medline]

9. Resh, M. D. (2016) Fatty acylation of proteins: the long and the short of it. *Prog. Lipid Res.* 63, 120–131 [CrossRef Medline]

10. Müller, T. D., Nogueiras, R., Andermann, M. L., Andrews, Z. B., Anker, S. D., Argente, J., Batterham, R. L., Benoi, S. C., Bowers, C. Y., Broglio, F., Casanueva, F. F., D’Alessio, D., Depoortere, I., Gebeler, A., Ghigo, E., et al. (2015) Ghrelin. *Mol. Metab.* 4, 437–460 [CrossRef Medline]

11. Taylor, M. S., Ruch, T. R., Hsiao, P. Y., Hwang, Y., Zhang, P., Dai, L., Huang, C. R., Berndsen, C. E., Kim, S. M., Pandey, A., Wolberger, C., Marmorstein, R., Machamer, C., Boeke, J. D., and Cole, P. A. (2013) Architectural organization of the metabolic regulatory enzyme ghrelin O-acyltransferase. *J. Biol. Chem.* 288, 32211–32228 [CrossRef Medline]

12. Taylor, M. S., Dempsey, D. R., Hwang, Y., Chen, Z., Chu, N., Boeke, J. D., and Cole, P. A. (2015) Mechanistic analysis of ghrelin-O-acyltransferase using substrate analogs. *Bioorg. Chem.* 62, 64–73 [CrossRef Medline]

13. Darling, I. E., Zhao, F., Loftus, R. J., Patton, L. M., Gibbs, R. A., and Hougland, J. L. (2015) Structure-activity analysis of human ghrelin O-acyltransferase reveals chemical determinants of ghrelin selectivity and acyl group recognition. *Biochemistry* 54, 1100–1110 [CrossRef Medline]

14. Ma, D., Wang, Z., Merrikh, C. N., Lang, K. S., Lu, P., Li, X., Merrikh, H., Rao, Z., and Xu, W. (2018) Crystal structure of a membrane-bound O-acyltransferase. *Nature* 562, 286–290 [CrossRef Medline]

15. Ovchinnikov, S., Park, H., Varghese, N., Huang, P. S., Pavlopoulos, G. A., Kim, D. E., Kamisetty, H., Kyprides, N. C., and Baker, D. (2017) Protein structure determination using metagenome sequence data. *Science* 355, 294–298 [CrossRef Medline]

16. Marks, D. S., Hopf, T. A., and Sander, C. (2012) Protein structure prediction from sequence variation. *Nat. Biotechnol.* 30, 1072–1080 [CrossRef Medline]

17. Ovchinnikov, S., Kinch, L., Park, H., Liao, Y., Pei, J., Kim, D. E., Kamisetty, H., Grishin, N. V., and Baker, D. (2015) Large-scale determination of previously unsolved protein structures using evolutionary information. *Elife* 4, e09248 [CrossRef Medline]

18. Hopf, T. A., Scharfe, C. P., Rodrigues, J. P., Green, A. G., Kohlbacher, O., Sander, C., Bonvin, A. M., and Marks, D. S. (2014) Sequence co-evolution gives 3d contacts and structures of protein complexes. *Elife* 3, e03430 [CrossRef Medline]

19. Nicoludis, J. M., and Gaudet, R. (2018) Applications of sequence coevolution in membrane protein biochemistry. *Biochim. Biophys. Acta Biomembr.* 1860, 895–908 [CrossRef Medline]

20. Wang, S., Sun, S., and Xu, J. (2018) Analysis of deep learning methods for blind protein contact prediction in casp12. *Proteins* 86, 67–77 [CrossRef Medline]

21. Steinegger, M., and Söding, J. (2018) Clustering huge protein sequence sets in linear time. *Nat. Commun.* 9, 2542 [CrossRef Medline]

22. Adhikari, B., and Cheng, J. (2018) Confold2: improved contact-driven ab initio protein structure modeling. *BMC Bioinformatics* 19, 22 [CrossRef Medline]

23. Yang, J., Yan, R., Roy, A., Xu, D., Poisson, J., and Zhang, Y. (2015) The i-tasser suite: protein structure and function prediction. *Nat. Methods* 12, 7–8 [CrossRef Medline]

24. Huang, J., Rauscher, S., Nawrocki, G., Ran, T., Feig, M., de Groot, R. L., Grubmüller, H., and Mackerrell, A. D., Jr. (2017) Charmm36m: an improved force field for folded and intrinsically disordered proteins. *Nat. Methods* 14, 71–73 [CrossRef Medline]

25. Lomize, M. A., Pogozheva, I. D., Joo, H., Mosberg, H. I., and Lomize, A. L. (2012) OPM database and ppm webserver: resources for positioning of proteins in membranes. *Nucleic Acids Res.* 40, D370–D376 [CrossRef Medline]

26. Irudayanathan, F. J., Trasatti, J. P., Karande, P., and Nangia, S. (2016) Molecular architecture of the blood brain barrier tight junction proteins—a synergistic computational and in vitro approach. *J. Phys. Chem. B* 120, 77–88 [CrossRef Medline]

27. Rajagopal, N., Irudayanathan, F. J., and Nangia, S. (2019) Palmitoylation of claudin-5 proteins influences their lipid domain affinity and tight junction
assembly at the blood-brain barrier interface. J. Phys. Chem. B 123, 983–993 CrossRef Medline
28. Chen, V. B., Arendall, W. B., 3rd, Head, J. J., Keedy, D. A., Immormino, R. M., Kapral, G. J., Murray, L. W., Richardson, J. S., and Richardson, D. C. (2010) Molprobity: all-atom structure validation for macromolecular crystallography. Acta Crystallogr. D Biol. Crystallogr. 66, 12–21 CrossRef Medline
29. Xu, J., and Zhang, Y. (2010) How significant is a protein structure similarity with tm-score = 0.5? Bioinformatics 26, 889–895 CrossRef Medline
30. Darling, J. E., Rodrigues, T., Sieburg, M. A., Abizaid, A., and Hougland, J. L. (2016) Ghrelin octanoylation is completely stabilized in biological samples by alkyl fluorophosphonates. Endocrinology 157, 4330–4338 CrossRef Medline
31. McGovern-Gooch, K. R., Rodrigues, T., Darling, J. E., Sieburg, M. A., Abizaid, A., and Hougland, J. L. (2016) Ghrelin octanoylation is completely stabilized in biological samples by alkyl fluorophosphonates. Endocrinology 157, 4330–4338 CrossRef Medline
32. Hougland, J. L. (2019) Ghrelin octanoylation by ghrelin O-acyltransferase: unique protein biochemistry underlying metabolic signaling. Biochem. Soc. Trans. 47, 169–178 CrossRef Medline
33. Chovancova, E., Pavelka, A., Benes, P., Strnad, O., Brezovsky, J., Kozlikova, B., Gora, A., Sustr, V., Klvana, M., Medek, P., Biedermannova, L., Sochor, J., and Damborsky, I. (2012) Caver 3.0: a tool for the analysis of transport pathways in dynamic protein structures. PLoS Comput. Biol. 8, e1002708 CrossRef Medline
34. Johnson, L. S., Eddy, S. R., and Portugaly, E. (2010) Hidden Markov model speed heuristic and iterative HMM search procedure. BMC Bioinformatics 11, 431 CrossRef Medline
35. Remmert, M., Biegert, A., Hauser, A., and Söding, J. (2011) HHblits: lightning-fast iterative protein sequence searching by HMM-HMM alignment. Nat. Methods 9, 173–175 Medline
36. Capella-Gutiérrez, S., Silla-Martínez, J. M., and Gabaldón, T. (2009) Triimal: a tool for automated alignment trimming in large-scale phylogenetic analyses. Bioinformatics 25, 1972–1973 CrossRef Medline
37. Wang, S., Li, Z., Yu, Y., and Xu, J. (2017) Folding membrane proteins by deep transfer learning. Cell Syst. 5, 202–211.e3 CrossRef Medline
38. Simkovic, F., Thomas, J. M. H., and Righo, D. J. (2017) ConKit: a Python interface to contact predictions. Bioinformatics 33, 2209–2211 CrossRef Medline
39. Lee, J., Cheng, X., Swals, J. M., Yeom, M. S., Eastman, P. K., Lemkul, J. A., Wei, S., Buckner, J., Jeong, J. C., Qi, Y., Jo, S., Pande, V. S., Case, D. A., Brooks, C. L., 3rd, MacKerell, A. D., Jr., Klauda, J. B., and Im, W. (2016) CHARMM-GUI input generator for NAMD, GROMACS, AMBER, OPENMM, and CHARMM/OPENMM simulations using the CHARMM36 additive force field. J. Chem. Theory Comput. 12, 405–413 CrossRef Medline
40. Abraham, M. J., Murtola, T., Schulz, R., Pall, S., Smith, J. C., Hess, B., and Lindahl, E. (2015) Gromacs: high performance molecular simulations through multi-level parallelism from laptops to supercomputers. SoftwareX 1, 19–25 CrossRef Medline
41. Seeliger, D., and de Groot, B. L. (2010) Ligand docking and binding site analysis with PyMOL and Autodock/Vina. J. Comput. Aided Mol. Des. 24, 417–426 CrossRef Medline
42. Trott, O., and Olson, A. J. (2010) AutoDock Vina: improving the speed and accuracy of docking with a new scoring function, efficient optimization, and multitreading. J. Comput. Chem. 31, 455–461 CrossRef Medline
43. Omasits, U., Ahrens, C. H., Müller, S., and Wollscheid, B. (2014) ProTrer: interactive protein feature visualization and integration with experimental proteomic data. Bioinformatics 30, 884–886 CrossRef Medline
44. Riddles, P. W., Blakeley, R. L., and Zerner, B. (1979) Ellman’s reagent: 5,5’-dithiobis(2-nitrobenzoic acid)—a reexamination. Anal. Biochem. 94, 75–81 CrossRef Medline
45. Tricerri, M. A., Behling Agree, A. K., Sanchez, S. A., and Jonas, A. (2000) Characterization of apolipoprotein A-I structure using a cysteine-specific fluorescence probe. Biochemistry 39, 14682–14691 CrossRef Medline
46. Sieburg, M. A., Cleverdon, E. R., and Hougland, J. L. (2019) Biochemical assays for ghrelin acylation and inhibition of ghrelin O-acyltransferase. Methods Mol. Biol. 2009, 227–241 CrossRef Medline
47. Post, P. L., Trybus, K. M., and Taylor, D. L. (1994) A genetically engineered, protein-based optical biosensor of myosin ii regulatory light chain phosphorylation. J. Biol. Chem. 269, 12880–12887 Medline
48. Roth, A. F., Feng, Y., Chen, L., and Davis, N. G. (2002) The yeast dhhc cysteine-rich domain protein akr1p is a palmitoyl transferase. J. Cell Biol. 159, 23–28 CrossRef Medline
49. Wellman, M. K., Patterson, Z. R., MacKay, H., Darling, J. E., Mani, B. K., Zigmans, J. M., Hougland, J. L., and Abizaid, A. (2015) Novel regulator of acylated ghrelin, cf801, reduces weight gain, rebound feeding after a fast, and adiposity in mice. Front. Endocrinol. (Lausanne) 6, 144 CrossRef Medline
50. R Core Team (2018) R: a language and environment for statistical computing. R Foundation for Statistical Computing, Vienna, Austria
51. Smith, R. H., and Powell, G. L. (1986) The critical micelle concentration of some physiologically important fatty acyl-coenzyme A’s as a function of chain length. Arch. Biochem. Biophys. 244, 357–360 CrossRef Medline