ACS NANO

www.acsnano.org

ARTICLE

A Machine Learning Assisted Tool and Numerical Model for Analyzing Lipid Nanoparticles

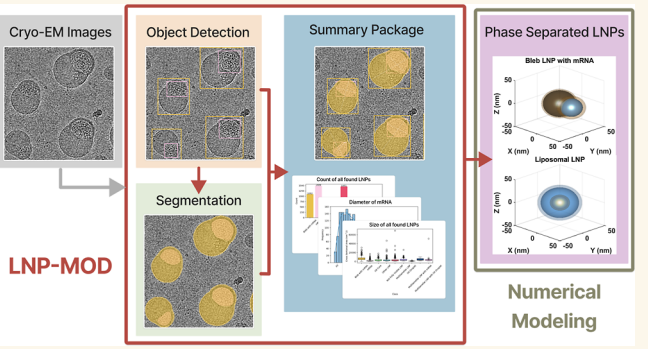
Owen Yuk Long Ip, Harrison D. E. Fan, Yao Zhang, Jerry Leung, Colton Strong, Janell Ko, Pieter R. Cullis, and Miffy Hok Yan Cheng*

Cite This: *ACS Nano* 2025, 19, 33387−33398

Read Online

ACCESS | Metrics & More | Article Recommendations | Supporting Information

ABSTRACT: **The transfection potency and biological fate of gene-loaded lipid nanoparticles (LNPs) are often determined by their morphological and physicochemical properties. Cryogenic-electron microscopy (cryo-EM) remains the most effective tool to analyze LNP morphology and internal structures in their native state, but analysis of cryo-EM micrographs is time-consuming and inefficient due to the diversity in size, shape, and structure of LNPs. In this study, we developed the Lipid Nanoparticle Morphology and Object Detector (LNP-MOD) pipeline. We adopted a modular design by using the You Only Look Once (YOLO) model for object detection and the Segmentation Anything model 2 (SAM 2) for LNP compartmental segmentation. We trained the model and demonstrated that LNP-MOD can effectively identify and segment different classes of LNPs and their corresponding internal structures with ∼80% accuracy. We further compared the image analysis data with mathematical modeling of LNPs containing water and mRNA (Liposomal LNPs and Bleb LNPs) according to the phase preferences of the lipids, and showed correspondence between LNP-MOD output, modeling results, and experimental data. Our approach of combining single-particle cryo-EM imaging with LNP-MOD is complementary to other analytical techniques. It allows for rapid identification and segmentation of a variety of LNP-nucleic acid morphologies and presents a powerful tool to inform the design of next-generation LNPs.**



KEYWORDS: *lipid nanoparticles, mRNA, LNP morphologies, numerical model, cryo EM, machine learning, computer vision*

INTRODUCTION

Lipid nanoparticles (LNPs) have become a promising delivery system for nucleic acid therapeutics as demonstrated by two clinically approved nanomedicines, namely Onpattro for the treatment of transthyretin induced amyloidosis and the COVID-19 LNP mRNA vaccines.[1,2] While FDA guidelines require characterization of LNP particle size distribution, charge, and mRNA encapsulation,[3,4] many other properties such as loading efficiency, morphology, and empty LNP populations are often not characterized. However, these properties can significantly impact the biological activity and therapeutic outcome.[5] Most recently, our research group and others have shown that the delivery and transfection potency of gene-loaded LNPs is dependent on their morphological and related structural properties.[6−9] To study LNP structure and morphology, transmission cryogenic-electron microscopy (cryo-EM) has become the most popular tool to analyze their most native state since the vitrification process prevents artifacts that would be induced by chemical staining and dehydration. This analytical method also enables measurement of size, shape, internal structure, and lamellarity of the LNPs. Moreover, cryo-EM provides high-resolution, single particle structure analysis, and allows visualization of the inner structure of encapsulated drugs, nucleic acid cargo, and even membrane proteins.[9−11] Cryo-EM structural analysis of LNPs[12] can benefit the formulation development of LNPs in the pharmaceutical industry by enabling the design of homogeneous next-generation LNP-based vaccines and therapeutics.

Received: June 13, 2025
Revised: August 25, 2025
Accepted: August 26, 2025
Published: September 11, 2025

https://doi.org/10.1021/acsnano.5c09956
ACS Nano 2025, 19, 33387−33398

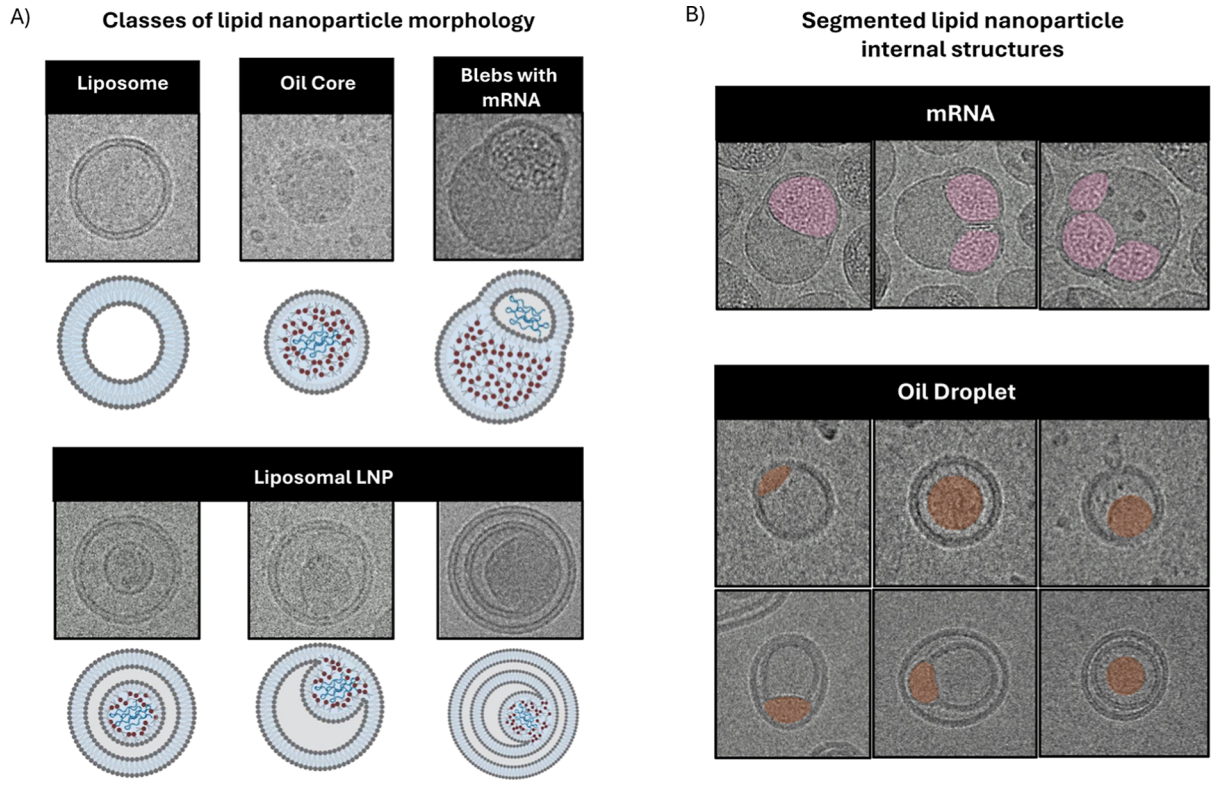


Figure 1. LNP-MOD training parameters. (A) Classes of lipid nanoparticle morphology and internal structure with example cryo-EM images and illustrations. Classes include Liposome, Oil Core, Blebs with mRNA, and Liposomal LNP. (B) Examples of different internal structures in highlighted regions: mRNA (pink) and oil droplet (brown).

However, high-throughput cryo-EM morphological analysis has typically been limited by the diversity in shape, size, phase, and orientation of LNPs. The low signal-to-noise ratio (SNR) and low contrast in LNP cryo-EM images also limits the use of masking and thresholding image analysis techniques for confident classification. These conditions make the process of particle annotation and measurement time-consuming and labor intensive. Despite different semi- or fully automated approaches[13−15] that have been proposed to leverage recent advancements using machine learning for cryo-EM particle picking, those techniques have been optimized for protein particle picking of a single purified protein, rather than a self-assembled LNP system with lower uniformity. Thus, there is an increasing need to develop new image analysis tools that can enable rapid and high-throughput analysis of LNP cryo-EM images.

In this study, we have established an AI model pipeline called Lipid Nanoparticle Morphology and Object Detector (LNP-MOD) with a modular design that consist of 3 parts: Object Detection, Segmentation, and Post-Processing. The Object Detection module was first used to locate the LNPs and their internal structures from cryo-EM images. Since the size, shape, and internal features of an LNP can vary drastically between different formulations, an object detection model with a strong baseline performance was trained to locate target instances while only needing a small data set. We explored the You Only Look Once (YOLO) model for its strong foundation model performance, variety of model sizes, and lower computation demand to train and deploy. We fine-tuned and evaluated different versions and sizes of the YOLO models with our custom data set and implemented a refined YOLOv12-s (YOLO version 12—size s) model for LNP-MOD to strike a balance between performance and computational complexity.

For the Instance Segmentation Module, we then utilized the Segment Anything model 2 (SAM 2) that demonstrated promising results in segmenting specific objects for both medical and cryo-EM imaging.[16−21] Leveraging SAM 2's promptable framework, outputs from the Object Detection Module served as prompts for specific targets of interest to generate precise outline masks at the pixel level. Our fine-tuned SAM 2 showed significant improvements in segmenting LNP internal structures and showed strong zero-shot generalization for unseen LNPs, allowing LNP-MOD to study new LNP morphologies by building only a small data set to fine-tune the Object Detection Module.

The Post-Processing Module then combined the outputs from the previous two modules and generated a summary package to provide an overview of the LNPs in terms of their size, distribution, and segmented internal features. Lastly, we demonstrated that the particle features characterized from LNP-MOD agreed well with numerical modeling of two different classes of LNP morphologies (Liposomal LNPs and Bleb LNPs), which revealed some important biophysical behaviors of these LNP systems.

RESULTS AND DISCUSSION

A Modular Approach to Building a Lipid Nanoparticle Analysis Pipeline. LNPs can adopt various morphologies and with multiple compartments due to phase separation and cargo encapsulation.[22−25] In order to train and evaluate LNP-MOD, a data set was curated from 228 cryo-EM images containing LNPs from the Nanomedicines Research Group at the University of British Columbia. The classification

https://doi.org/10.1021/acsnano.5c09956
ACS Nano 2025, 19, 33387−33398

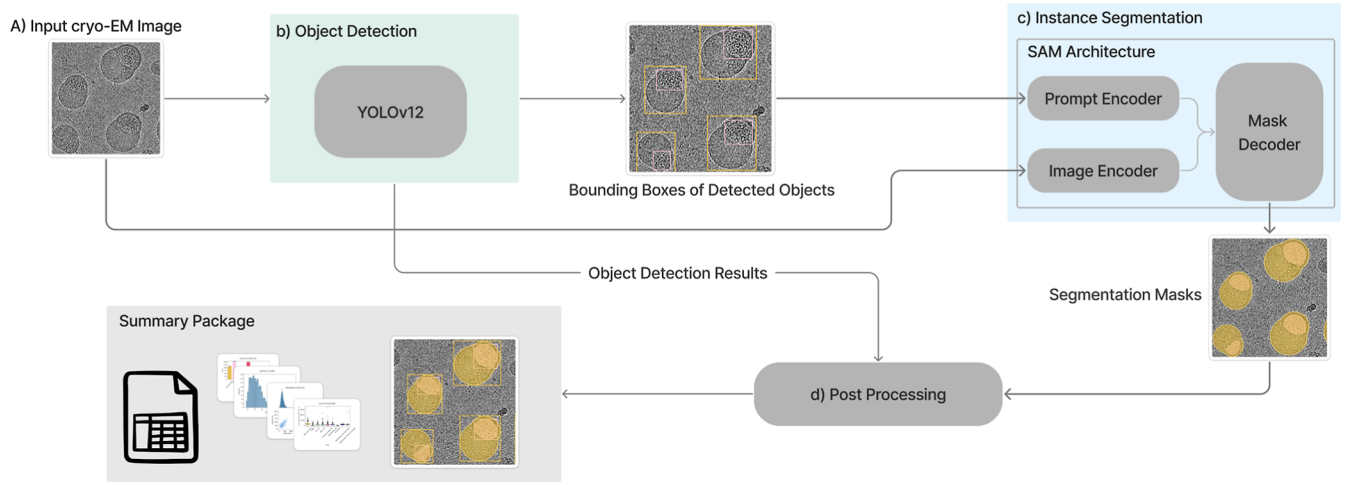


Figure 2. Pipeline design of LNP-MOD. (A) A cryo-EM image is uploaded to LNP-MOD and forwarded to the Object Detection Module. (B) The Object Detection Module processes the input image and generates inference results consisting of object classes and pixel coordinates of bounding boxes (B.1), and confidence of detection of all detected LNPs and internal structures in the image. (C) Input cryo-EM images from A are fed into the Image Encoder of SAM architecture, while the bounding boxes output from previous module (B.1) are fed into the Prompt Encoder. The Mask Decoder then processes the output from both the Prompt Encoder and Image Encoder to generate a segmentation mask (C.1) for each detected object from Object Detection. (D) The outputs from the Object Detection Module and Instance Segmentation Module are further processed to generate a summary package that provides an overview of the detected LNPs, such as their diameter, area, type, and distribution.

of LNPs and internal structures from cryo-EM images was based on previous work on LNP formulation development.[9,26−29] LNPs from the training data set fell under 4 major identifiable classes: Liposome, Oil Core, Blebs with mRNA, and Liposomal LNP (Figure 1A), and 2 classes for internal structures: mRNA and oil droplet (Figure 1B). LNPs that did not belong to the 4 major identifiable classes were classified as Other LNP. Any LNPs not fully visible in the image were classified as Not Fully Visible LNP. The annotations were then reviewed by three scientists to independently correct annotations and to minimize bias. The data set was then split in an 80:20 ratio for model training and validation, respectively (refer to Methods).

The pipeline of LNP-MOD consisted of 3 modules: Object Detection, Instance Segmentation, and Post-Processing (Figure 2). For Object Detection, the model You Only Look Once version 12 (YOLOv12)[30] architecture was utilized to classify different LNP structural morphologies (object classes) and their corresponding internal compartments. It also identified and located LNPs (pixel coordinates) and confidence of detection from cryo-EM images with better accuracy using fewer parameters, as compared to previous YOLO models.

As shown in Figure S1, the architecture of YOLOv12 consists of 3 main modules: Backbone, Neck, and Head. The Backbone is a feature extractor that produces feature maps at different resolutions, while the neck combines these features, and the head outputs final detections (bounding boxes and class predictions). This enables the network to capture broader contextual information, thus improving its ability to identify LNPs and internal structures of various sizes and shapes.

Subsequently, the coordinates of the generated bounding boxes, along with the input cryo-EM images, were forwarded from the Object Detection Module to the Instance Segmentation Module. A key criterion of the segmentation model is the ability to handle unfamiliar objects, as LNP morphologies and internal structures can vary from different formulation processes and lipid compositions, and can readily adopt a different compartment at various locations within the LNP. For this step, Segmentation Anything model 2 (SAM 2)[20] was employed as it is a state-of-the-art promptable segmentation model that takes points, bounding boxes, or masks as inputs to specify segmentation targets, and demonstrates strong zero shot capability with unseen objects and images.[16,19,21,31]

The architecture of SAM 2 consists of 3 major components for image segmentation: Image Encoder, Prompt Encoder, and Mask Decoder (Figure S2). The Image Encoder extracts high level feature embeddings from the input image. The Prompt Encoder converts given prompts into embeddings that guides the segmentation process. Finally, the Mask Decoder combines both the image and prompt embeds to produce the final segmentation masks. By utilizing the bounding boxes outputted from the Object Detection Module as a prompt, this allows SAM 2 to segment the target LNPs or internal structures. This is especially important in the scenario when segmenting different internal structures of LNPs, or even overlapping LNPs, to provide a better segmentation result.

Finally, outputs from the Object Detection and Instance Segmentation Modules were then processed through the Post-Processing Module, producing data sheets and visualizations to provide an overview of the processed cryo-EM images, enabling further quantitative analysis. Bounding boxes and class labels identified by the Object Detection Module were integrated with segmentation masks from the Instance Segmentation Module to characterize each identified LNP and its internal structures. The segmentation masks were then converted into polygons to enable geometric measurement including calculation of the longest diameter and cross-sectional area in pixel units. These pixel-based measurements were then converted into nanometers through using the number of pixels of the cryo-EM scale bar as a reference. Based on the pixel coordinates of each LNP and its internal structures, the structures of each LNP object were identified

https://doi.org/10.1021/acsnano.5c09956
ACS Nano 2025, 19, 33387−33398

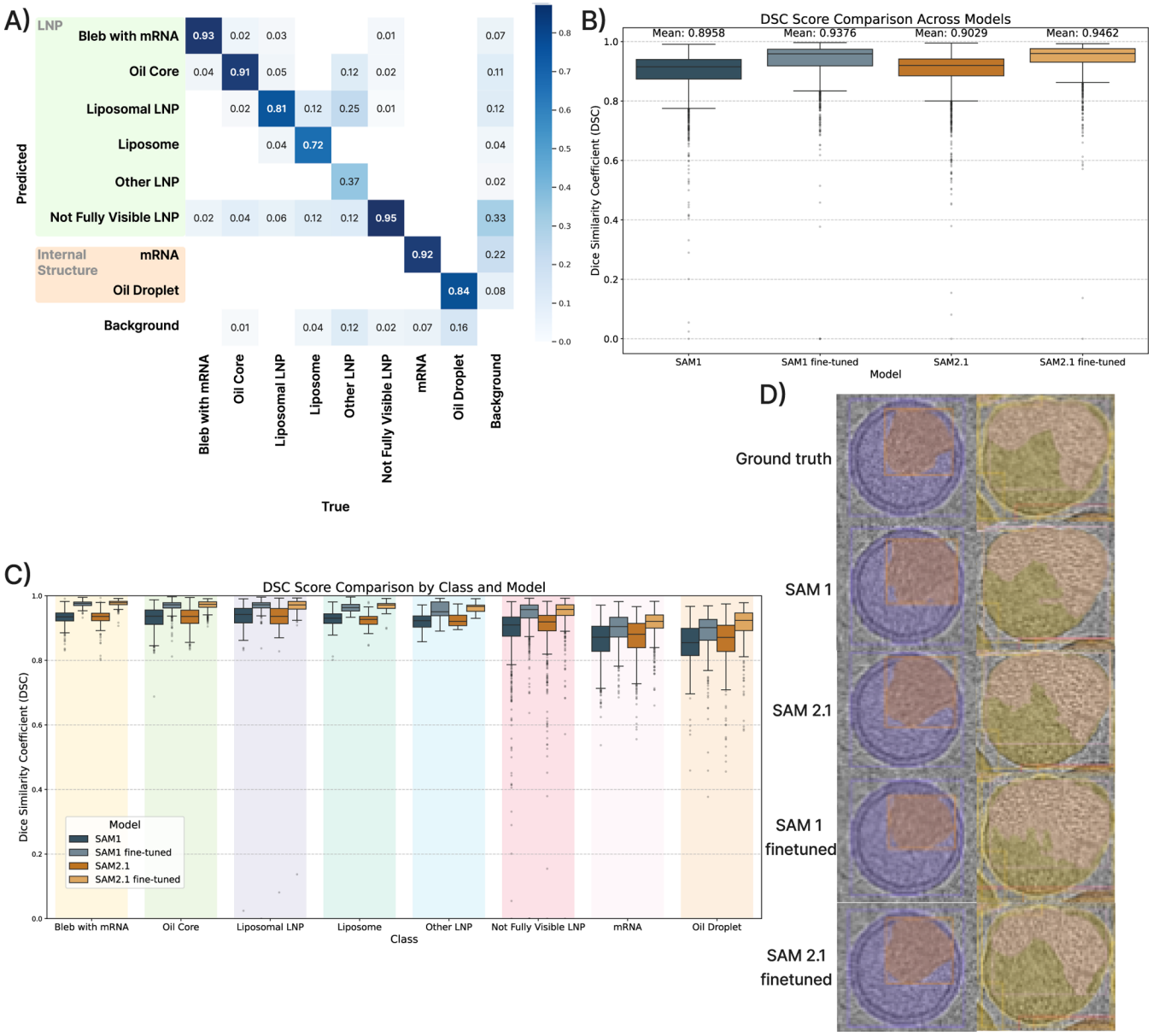


Figure 3. LNP-MOD evaluation on validation data set. (A) normalized confusion matrix of LNP-MOD object detection model with YOLOv12-s architecture on validation data set. The model achieved an average identification accuracy of 74% for LNPs and 82% for internal structures. Notably, the matrix indicates a clear differentiation between the two groups of classes, with no misclassifications between them. (B) Overall Dice Similarity Coefficient (DSC) Score comparison across different segmentation models. The box plot compares the overall performance between foundation and fine-tuned models, demonstrating consistently high segmentation accuracy with mean DSC scores above 0.9. (C) Per-class DSC score comparison. This plot provides a detailed breakdown of the DSC scores for each specific LNP and internal structure class, comparing the performance of the four models from Figure 3B. (D) Visual examples of segmentation performance on validation data set. The panel displays a comparison between the ground truth annotation and the segmentation mask generated by the foundation models and fine-tuned models. The left column images show segmentation of Liposomal LNP (purple region) containing an oil droplet internal structure (orange region). The right column images show segmentation of a Blebs with mRNA LNP (yellow region) containing an internal mRNA structure (pink region).

and characterized. The measurement metrics (diameter, cross-sectional area, and volume) and data visualizations (LNP size distribution, class population, and segmented properties) were compiled to provide a comprehensive data package (Figure S3).

For both ease of use and modularity, LNP-MOD was developed as a Python package and embedded in a Jupyter Notebook with a graphical user interface (GUI) suitable for both experienced programming personnel and scientists.

LNP-MOD Performance. The performance of LNP-MOD was evaluated using a validation data set comprised of cryo-EM images that the models had not seen during training. The validation data set contained 48 cryo-EM images with ~2400 various LNPs and internal structures. To evaluate the performance between different generations of YOLO architecture models (YOLOv8-s, YOLOv9-s, YOLOv10-s, YOLOv11-s and YOLOv12-s), each model was trained with the same training data set and compared for the Object Detection Module. The YOLOv12-s model achieved the best

https://doi.org/10.1021/acsnano.5c09956
ACS Nano 2025, 19, 33387−33398

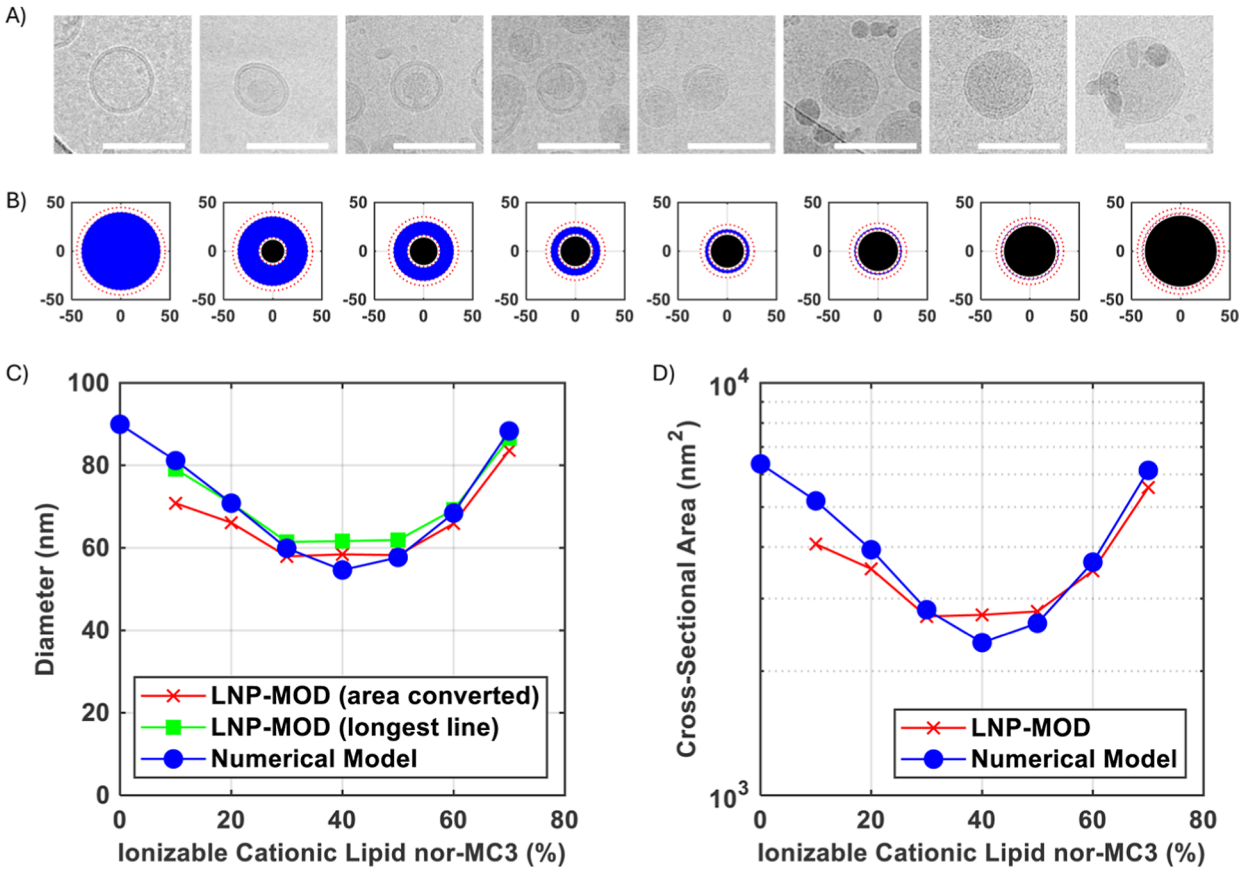


Figure 4. Liposomal LNP cryo-EM analysis and numerical modeling. (A) Cryo-EM images of LNP-mRNA systems containing NanoLuciferase-encoded mRNA (808 nt; $N/P = 6$) with lipid compositions nor-MC3/ESM:cholesterol/PEG-DMG at ratios of cationic lipid from 0% to 70%. Scale bar = 100 nm. (B) Cross-section of numerical model at ratios of cationic lipid from 0% to 70% and no PEG-lipid, showing outline of membrane lipid heads (red dashed), oil droplet (black), and 4 copies NanoLuc mRNA + water (blue). Axes units in nm. (C) Average diameter of LNP systems determined by LNP-MOD image analysis by area converted (red x's) and longest line (green ■'s) methods compared with numerical model (blue ●'s). (D) Average cross-sectional area of LNP systems determined by LNP-MOD (red x's) compared with numerical model (blue ●'s).

performance with a Mean Average Precision averaged (mAP) of 0.796 under an intersection over union (IoU) threshold from 50% to 95%. (summarized in Table S2 and Figure S4). YOLOv12 integrates attention mechanisms from vision transformers, and its Area Attention Module enlarges the receptive field compared to earlier models. Therefore, the YOLOv12-s model was employed as the object detection model to maintain a balance between performance and computation efficiency (Table S1).

As shown in Figure 3A, the classification performance of the LNP-MOD with the normalized confusion matrix was further studied. The model correctly identified LNP morphologies with an average accuracy of 78% and internal structures with 83% accuracy. The significantly lower performance for Other LNP was likely due to the lack of samples from the training data set. Additionally, the reduced accuracy in identifying Liposomal LNP and Liposome likely stemmed from the inherent difficulty in distinguishing these classes, as shown in Figure S5. Overall, the matrix shows a clear separation between the LNP morphologies and internal structures, with no misclassifications between the two groups of classes, demonstrating the model's robust discriminative capability.

The segmentation performance of LNP-MOD was evaluated with the validation data set and compared against the SAM 1 and SAM 2.1 foundation models. As shown in Figure 3B,C, LNP-MOD outperformed both foundation models, achieving a higher overall and per-class Dice Similarity Coefficient (DSC). The mean DSC increased from 0.8958 (SAM 1) and 0.9029 (SAM 2) to 0.9462 (LNP-MOD), while the standard deviation decreased from 0.0749 to 0.0498, indicating improved accuracy and consistency. Figure 3D shows two qualitative segmentation examples from the validation task that further demonstrated the stronger performance of LNP-MOD on segmenting weak boundaries of LNP and their internal structures. While both foundation models were able to locate LNPs in cryo-EM images, they failed to accurately segment the LNPs and their internal structures. This is likely due to the low contrast and fuzzy background of cryo-EM images, as well as the absence of cryo-EM data images in the training data of the foundation models.[20,32] Overall, fine-tuning SAM 2.1 on cryo-EM images allowed LNP-MOD to yield significantly more accurate and robust segmentations.

Lastly, the Post-Processing Module was evaluated by comparing the generated summary packages with manually measured data for both the Liposomal LNP and Blebs with mRNA systems. Each package included information on particle diameter, cross-sectional area, and total count, and was broadly consistent with the data measured manually (Figures S6 and S7).

https://doi.org/10.1021/acsnano.5c09956
ACS Nano 2025, 19, 33387−33398

Table 1. Mathematical Formulations (eqs 4.1−4.16) Employed to Model Liposomal LNPs with Varying Ionizable Lipid Compositions

0% nor-MC3	10%–70% nor-MC3	Equation #
$MR_{\mathrm{nMC3}} = \frac{n_{\mathrm{nMC3}}}{n_{\mathrm{ESM}} + n_{\mathrm{chol}} + n_{\mathrm{nMC3}}}$		(4.1)
$n_{\mathrm{ESM}} = n_{\mathrm{chol}}$		(4.2)
$HA_{\mathrm{eff}} = HA_{\mathrm{ESM}} + HA_{\mathrm{chol}}$		(4.3)
$V_{\mathrm{LNP}} = \frac{4}{3}\pi r^3 = V_{\mathrm{in}} + V_{\mathrm{mid}} + V_{\mathrm{out}}$		(4.4)
$V_{\mathrm{in}} = V_{\mathrm{oil}} + V_{\mathrm{ml}} + V_{\mathrm{w,ml}}$		(4.5)
$V_{\mathrm{mid}} = V_{\mathrm{mRNA}} + V_{\mathrm{w}}$		(4.6)
$V_{\mathrm{out}} = V_{\mathrm{bl}} + V_{\mathrm{w,bl}}$		(4.7)
$V_{\mathrm{oil}} = \frac{4}{3}\pi r_{\mathrm{oil}}^3$		(4.8)
$V_{\mathrm{oil}} = 0$	$V_{\mathrm{oil}} = \frac{n_{\mathrm{nMC3}} MW_{\mathrm{nMC3}}}{\rho_{\mathrm{nMC3}} N_{\mathrm{A}}}$	(4.9)
$V_{\mathrm{ml}} = 0$	$V_{\mathrm{ml}} = \frac{4}{3}\pi((r_{\mathrm{oil}} + t_{\mathrm{ml}})^3 - r_{\mathrm{oil}}^3)$	(4.10)
$V_{\mathrm{w,ml}} = 0$	$V_{\mathrm{w,ml}} = \left(\frac{4\pi(r_{\mathrm{oil}} + t_{\mathrm{ml}})^2}{HA_{\mathrm{eff}}}\right)\left(\frac{MW_{\mathrm{w}}}{\rho_{\mathrm{w}} N_{\mathrm{A}}}\right) n_{\mathrm{w,a}}$	(4.11)
$V_{\mathrm{mRNA}} = 0$	$V_{\mathrm{mRNA}} = \pi\left(\frac{d_{\mathrm{mRNA}}}{2}\right)^2 l_{\mathrm{nt}} n_{\mathrm{b}} n_{\mathrm{mRNA}}$	(4.12)
$r = 45$	$V_{\mathrm{w}} = V_{\mathrm{w,0\% nMC3}}\left(\frac{1}{1 + e^{k(MR_{\mathrm{nMC3}} - MR_{\mathrm{x}})}}\right)$	(4.13)
$V_{\mathrm{bl}} = \frac{4}{3}\pi(r^3 - (r - t_{\mathrm{bl}})^3)$		(4.14)
$V_{\mathrm{w,bl}} = \left(\frac{4\pi(r - t_{\mathrm{bl}})^2}{HA_{\mathrm{eff}}}\right)\left(\frac{MW_{\mathrm{w}}}{\rho_{\mathrm{w}} N_{\mathrm{A}}}\right) n_{\mathrm{w,a}}$		(4.15)
$n_{\mathrm{ESM}} + n_{\mathrm{chol}} = \frac{4\pi(r - t_{\mathrm{bl}})^2 + 4\pi r^2}{HA_{\mathrm{eff}}}$	$n_{\mathrm{ESM}} + n_{\mathrm{chol}} = \frac{4\pi(r_{\mathrm{oil}} + t_{\mathrm{ml}})^2 + 4\pi(r - t_{\mathrm{bl}})^2 + 4\pi r^2}{HA_{\mathrm{eff}}}$	(4.16)

Mathematical Modeling of Phase Separated LNPs that Contain Water, mRNA, and Ionizable Lipids. To further evaluate LNP-MOD on LNPs with distinct morphologies and membrane biophysical features, cryo-EM images of two different classes of LNPs analyzed by LNP-MOD were compared against mathematical models of theoretical LNPs. The first class of LNP analyzed was the Liposomal LNP system, which comprised of a bilayer membrane of equimolar ratios of egg sphingomyelin (ESM) to cholesterol, surrounding an oil droplet comprised of the ionizable cationic lipid nor-MC3. As shown in Figure 4A, cryo-EM images demonstrated a transition from a bilayer system to a monolayer system with increasing ratios of the ionizable lipid from 10−70%. We have recently reported that Liposomal LNP formation is determined by the nucleation of the ionizable lipid with mRNA cargo at pH 4, followed by self-assembly of the bilayer (comprised of helper lipid, cholesterol, and the positively charged form of the ionizable lipid), followed by phase separation of the ionizable lipid from the bilayer into an oil core as the pH is raised to pH 7.4. The RNA cargo is partitioned into a polar environment surrounded by a monolayer or bilayer of the helper lipid and cholesterol.[29]

To study the size and morphological transition between a bilayer liposomal membrane system to an amorphous oil core structure, the ionizable cationic lipid content of a theoretical Liposomal LNP was modeled across the same range of nor-MC3 molar ratios. Polyethylene glycol (PEG)-lipids were excluded from theoretical calculations. Surrounding the internal oil droplet of neutral nor-MC3 ionizable lipids was a set of concentric spherical shells, corresponding to monolayer of ESM and cholesterol, encapsulated water, mRNA encoding NanoLuciferase (NanoLuc; 808 nucleotides), and finally an outer bilayer of ESM and cholesterol. A cross-sectional schematic of this structure is provided in Figure S8.

The following parameters were used to model the membrane properties of the equimolar ratio of ESM to cholesterol[33] in the monolayer and bilayer structures at pH 7.4: monolayer thickness of 2.5 nm, bilayer thickness[34] of 5 nm, and headgroup areas of 0.5 nm^2 and 0.37 nm^2 per ESM and cholesterol molecule.[34] Additional constraints in the model included adsorption of 10 water molecules per lipid headgroup onto the inner bilayer leaflet, based on molecular dynamics simulations of phosphatidylcholine (PC) lipid hydration shells.[35] A similar hydration level was assumed for the monolayer surrounding the oil droplet. The aqueous compartment also contained 4 copies of NanoLuc mRNA, based on an average mRNA loading from theoretical calculation based on wt/wt ratio of lipids and mRNA, respectively (Table 1).

The bulk volume of encapsulated water followed an empirically derived, inverse decaying sigmoidal function (refer to Methods), reflecting that water loss was attributed to effects such as spontaneous curvature and possibly increased membrane hydrophobicity from partial incorporation of ionizable lipid into the membrane. The model initially constrained the LNP diameter to 90 nm in the absence of nor-MC3 (0%), where size was governed by favorable bilayer curvature. As nor-MC3 content increased, ESM helper lipid and cholesterol amounts decreased. The incorporation of ionizable lipid was predicted to occupy the internal volume, resulting in predominantly an oil-saturated structure (Figure 4B).

https://doi.org/10.1021/acsnano.5c09956
ACS Nano 2025, 19, 33387−33398

Table 2. Mathematical Formulations (eqs 5.1−5.20) Employed to Model Bleb LNP Containing a Water and mRNA Compartment Protruded from KC2-Based Oil Droplet

$f_{\text{mRNA+water}} < 50\%$	$f_{\text{mRNA+water}} = 50\%$	$f_{\text{mRNA+water}} > 50\%$	Equation #
$n_{\text{KC2}} = n_{\text{b}} n_{\text{mRNA}} \frac{N}{P}$			(5.1)
$n_{\text{DSPC}} = n_{\text{KC2}} \frac{MR_{\text{DSPC}}}{MR_{\text{KC2}}}$			(5.2)
$n_{\text{chol}} = n_{\text{KC2}} \frac{MR_{\text{chol}}}{MR_{\text{KC2}}}$			(5.3)
$HA_{\text{eff}} = \frac{n_{\text{DSPC}} HA_{\text{DSPC}} + n_{\text{chol}} HA_{\text{chol}}}{n_{\text{DSPC}} + n_{\text{chol}}}$			(5.4)
$V_{\text{KC2}} = \frac{n_{\text{KC2}} MW_{\text{KC2}}}{\rho_{\text{KC2}} N_A}$			(5.5)
$V_{\text{aq}} = \frac{4}{3}\pi r_{\text{aq}}^3$			(5.6)
$V_{\text{aq}} = V_{\text{mRNA}} + V_{\text{w,ml}} + V_{\text{w}}$			(5.7)
$V_{\text{mRNA}} = \pi \left(\frac{d_{\text{mRNA}}}{2}\right)^2 l_{\text{nt}} n_{\text{b}} n_{\text{mRNA}}$			(5.8)
$V_{\text{w,ml}} = \left(\frac{4\pi r_{\text{aq}}^2}{HA_{\text{eff}}}\right)\left(\frac{MW_{\text{w}}}{\rho_{\text{w}} N_{\text{A}}}\right) n_{\text{w,a}}$			(5.9)
$f_{\text{mRNA+water}} = \frac{CSA_{\text{mRNA+water}}}{CSA_{\text{LNP}}}$			(5.10)
$r_{\text{oil,o}} = r_{\text{oil}} + t_{\text{ml}}$			(5.11)
$r_{\text{aq,o}} = r_{\text{aq}} + t_{\text{bl}}$			(5.12)
$d = r_{\text{oil}}$	$d = \frac{r_{\text{oil}} + r_{\text{aq}}}{2}$	$d = r_{\text{aq}}$	(5.13)
$CSA_{\text{mRNA+water}} = \pi r_{\text{aq}}^2$			(5.14)
$CSA_{\text{LNP}} = \pi r_{\text{oil,o}}^2 + \pi r_{\text{aq,o}}^2 - CSA_{\text{intersection}}$			(5.15)
$CSA_{\text{intersection}} = r_{\text{oil,o}}^2 \cos^{-1}\left(\frac{d^2 + r_{\text{oil,o}}^2 - r_{\text{aq,o}}^2}{2dr_{\text{oil,o}}}\right) + r_{\text{aq,o}}^2 \cos^{-1}\left(\frac{d^2 + r_{\text{aq,o}}^2 - r_{\text{oil,o}}^2}{2dr_{\text{aq,o}}}\right) - \frac{1}{2}\sqrt{(d + r_{\text{oil,o}} + r_{\text{aq,o}})(d + r_{\text{oil,o}} - r_{\text{aq,o}})(d - r_{\text{oil,o}} + r_{\text{aq,o}})(-d + r_{\text{oil,o}} + r_{\text{aq,o}})}$			(5.16)
$V_{\text{intersection}} = \frac{\pi(r_{\text{oil,o}} + r_{\text{aq,o}} - d)^2 \left(d^2 + 2d(r_{\text{oil,o}} + r_{\text{aq,o}}) - 3(r_{\text{oil,o}} - r_{\text{aq,o}})^2\right)}{12d}$			(5.17)
$r_{oil} = \sqrt[3]{\frac{3(V_{\text{KC2}} + V_{\text{intersection}})}{4\pi}}$			(5.18)
$n_{\text{DSPC}} + n_{\text{Chol}} = \frac{4\pi r_{\text{aq}}^2 + SA_{\text{LNP}}}{HA_{\text{eff}}}$			(5.19)
$d_{\text{LNP}} = t_{\text{ml}} + r_{\text{oil}} + d + r_{\text{aq}} + t_{\text{bl}}$			(5.20)

Experimentally, the diameter of NanoLuc-encapsulated Liposomal LNP systems was analyzed by LNP-MOD using two image-based sizing methods (Figure 4C). The first method (area converted, red x's) summed the pixels in the segmented cross-section, converted the area from units of pixels^2 to nm^2 using the cryo-EM scale bar as a reference, then computed the diameter using $A = \pi d^2/4$. The second method (longest line, green ■'s) measured the maximum end-to-end distance across the segmented region. Both approaches revealed similar diameter trends with increasing nor-MC3 content and were in good agreement with the numerical model (blue ●'s). Segmented cross-sectional areas (Figure 4D) also showed consistently similar agreement between LNP-MOD sizing measurements (red x's) and the numerical model (blue ●'s).

Interestingly, when larger Firefly Luciferase mRNA (FLuc; 1921 nucleotides) was encapsulated under the same lipid composition, a distinct size and morphological trend emerged. Cryo-EM analysis revealed that FLuc-encapsulated Liposomal LNPs exhibited larger sizes than NanoLuc-encapsulated Liposomal LNPs (Figure S9), with a corresponding larger oil droplet (~20 to 40%) population compared to the LNP-NanoLuc mRNA system, suggesting that when the charge density ratio is constant, the ionizable lipid content corresponding to the oil core would increase to accommodate for larger mRNA.

The second class of LNPs analyzed was the Blebs with mRNA system (50/10/40%; KC2/DSPC/cholesterol), which consisted of an ionizable cationic lipid (KC2) oil droplet encapsulated by a monolayer of distearoylphosphatidylcholine (DSPC) and cholesterol. A phase-separated aqueous blebbed compartment containing Firefly Luciferase encoded mRNA (FLuc; 1921 nucleotides) was located adjacent to the oil core and was surrounded by a bilayer formed from DSPC and cholesterol. Between the mRNA and the enclosing bilayer was a shell-like region of water (Figure S10).

The numerical model was developed to predict the expected size of a Bleb LNP containing an assumed 4 copies of FLuc mRNA. As with the previous system, PEG-lipids were not included in the theoretical calculations. Consistent with earlier assumptions, the model assumed that 10 water molecules were

https://doi.org/10.1021/acsnano.5c09956
ACS Nano 2025, 19, 33387−33398

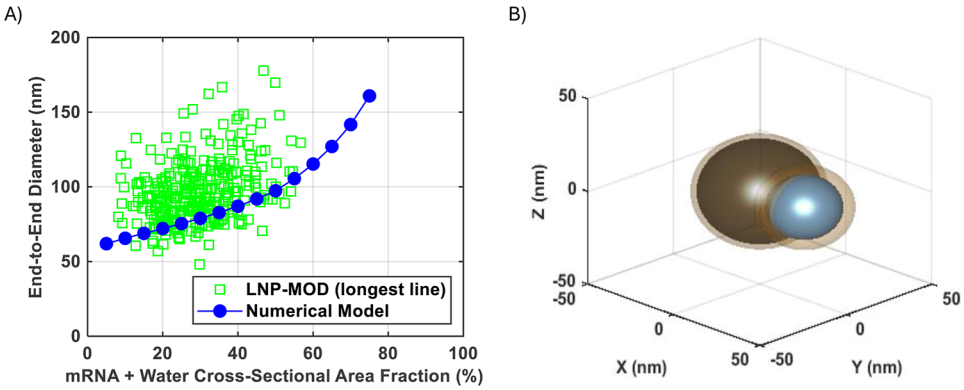


Figure 5. Blebs with mRNA LNP numerical modeling. (A) End-to-end diameter of LNPs determined by LNP-MOD image analysis by longest line method (green □'s) compared with numerical model (blue ●'s) at various mRNA + water cross-sectional areas. (B) 3D illustration of FLuc-encapsulated Bleb LNP numerical model at cross-sectional area fraction of mRNA + water = 25%, showing oil droplet (black), 4 copies FLuc mRNA + water aqueous compartment (blue), and monolayer and bilayer (brown translucent).

adsorbed onto each lipid head of the inner bilayer leaflet (Figure S10). The geometry was modeled as two overlapping spheres representing the oil and aqueous compartments. The model was defined by a separate system of equations (refer to Methods, Table 2) and solved for the overall cross-sectional area and volume, using the cross-sectional area fraction occupied by mRNA and water as an independent variable. This cross-sectional area fraction was varied from 5% to 75%.

As shown in Figure 5A, the end-to-end diameter of experimentally formulated FLuc-encapsulated Bleb LNP system was analyzed by LNP-MOD using the longest line method (green ■'s) over a large data set. Numerical model predictions (blue ●'s) generally followed the experimental trend and were in good agreement at the lower end of the size distribution. Figure 5B shows a 3D illustration of the predicted particle geometry at a mRNA + water cross-sectional area fraction of 25%. Although the model was solvable across the full range from 5%−75% (Figure S11 for cross-section views), two notable exceptions were observed.

First, at low fractions (5−15%), the solved value of internal water volume V_w (eq 5.7 in Methods) was negative, indicating that it may not be physically feasible to encapsulate 4 full copies of FLuc mRNA at such low internal area fractions. Second, at high fractions (65−75%), the solved outer surface area of the LNP SA_{LNP} (eq 5.19 in Methods) was negative. While not physically possible, this suggests that there were would not be enough DSPC and cholesterol lipids in the physical system to form the monolayer and bilayer membranes. This observation was reflected in the experimentally formulated particles, where Bleb LNPs did not exhibit a cross-sectional area fraction of mRNA and water greater than 60% cross-sectional area fraction.

CONCLUSION

This study introduced LNP-MOD, an AI model pipeline designed for high-throughput analysis of LNP cryo-EM images. LNP-MOD contains an Object Detection Module that leveraged YOLOv12's efficient architecture to generate bounding boxes that served as precise prompts for segmentation. These detections were then integrated into SAM 2's promptable segmentation framework, producing accurate and detailed segmentation masks. This combined approach delivered robust recognition and precise segmentation of LNPs and their internal structures, but also circumvented the need for extensive data sets typically required by traditional transformed-based models. Consequently, the pipeline demonstrated notable efficiency and practicality gains for cryo-EM image analysis. Comparing to the foundation models, LNP-MOD demonstrated higher accuracy at identifying and segmenting both LNP morphologies and their corresponding internal structures, as well as superior performance in machine learning evaluations.

To cross-validate the experimental result from LNP-MOD, numerical models were developed as an additional, biophysically grounded validation method, guided by our understanding and observations of various LNP morphologies. Two distinct LNP classes were modeled to challenge the platform's accuracy. LNP-MOD's analysis of the corresponding Liposomal LNP cryo-EM images revealed diameter and cross-sectional area measurements that were in line with the predicted trends. For the Blebs with mRNA system, LNP-MOD's measurement again closely matched the proposed geometrical model. We believe LNP-MOD can serve as a foundational tool enabling scientists to accelerate research in LNP morphology studies through an easily accessible cloud system. Additionally, each module within LNP-MOD can be utilized independently to facilitate the creation of customized training data sets and is highly tunable for targeted analyses of specific LNP types or internal structures. Lastly, we envision broader use of LNP-MOD as this tool has the potential to synergize with lipid chemistry AI-powered tools, extending into a pipeline of virtual genetic medicine discovery.

METHODS

Data Set. A data set of 228 cryo-EM images containing LNPs was curated from the Nanomedicines Research Group. All LNPs and internal structures in the images were annotated with segmentation masks and bounding boxes. The annotations were performed by human annotators following the classification method as shown in Figure 1. The annotations were then separately reviewed by 2 LNP experts to correct the annotations and to minimize bias. In total, 13,295 annotations were made consisting of 2039 Blebs with mRNA, 2253 oil core, 59 Other LNP, 3165 Not Fully Visible LNP, 1408 Liposomal LNP, 167 Liposome, 2757 mRNA, and 1447 Oil Droplet. Any LNPs that were partially covered, overlapping, or touching edges were classified as Not Fully Visible LNP. The Other LNP category included various new LNP structures yet to be explored, but not able to be classified as a class due to the small sample size. The data set was separated in an 80:20 ratio for model training and validation, respectively.

https://doi.org/10.1021/acsnano.5c09956
ACS Nano 2025, 19, 33387−33398

Model Training. The weight of LNP-MOD's object detection model was initialized with the foundation YOLO-v12s model, and trained for 200 epochs with the AdamW[36] optimizer with a learning rate of 0.01 and a weight decay of 0.0005. The training batch size was 2, which is the number of images the model trained with before updating parameters. Data augmentations were implemented to enrich the training data set to reduce overfitting and improve the generalization capability on unseen data. Mosaic augmentation, where 4 images are combined into a single mosaic image, was set to be always applied. Image translating, scaling, and horizontal flipping were set to be applied with a probability of 10%, 50%, and 50%, respectively.

For the Instance Segmentation Module, the model was initialized with the foundation SAM 2.1 checkpoint. Training was conducted with a batch size of 2 for 40 epochs using single-frame inputs and up to 10 objects per frame, employing a multicomponent loss function combing mask, Dice Similarity Coefficient (DSC), intersection over union (IoU), classification components with different weights. Data augmentation included random horizontal flipping, affine transformation and color jittering.

Evaluation Metrics. Following common practice and the recommendations in Metrics,[37] Mean Average Precision (mAP) and Dice Similarity Coefficient (DSC) were used to evaluate the performance of LNP-MOD Object Detection and Segmentation.

Mean Average Precision (mAP) is a widely used metric to evaluate the performance of object detection models. It combines precision and recall across multiple classes at various intersection over union (IoU) thresholds. A predicted bounding box is a true positive (TP) if the IoU with the ground truth bounding box is equal or greater than the threshold. Precision in this context refers to the ratio of correctly predicted positives to the total predicted positives. Recall is the ratio of true positives to all the actual positives. For each object category, the area under the Precision−Recall curve, known as Average Precision (AP), is defined as

$$\mathrm{AP} = \sum_{n=1}^{N-1} (R_n - R_{n-1}) P_n \tag{1}$$

where R_n and R_{n-1} are the recall values at threshold n and $n-1$, P_n is the precision at threshold n. The Mean Average Precision (mAP) is then determined by averaging the AP values for all categories, providing a comprehensive view of the model's accuracy and reliability, which is defined as

$$\mathrm{mAP} = \frac{1}{C} \sum_{c=1}^{C} \mathrm{AP}_c \tag{2}$$

where C is the total number of the classes and AP_c is the Average Precision for class c.

Dice Similarity Coefficient (DSC) is a statistical measure used to determine the similarity between 2 sets of data. It is commonly used in image segmentation, particularly in medical imaging, to compare the similarity between the predicted segmentation masks and the ground truth masks, which is defined as

$$\mathrm{DSC} = \frac{2|A \cap B|}{|A| + |B|} \tag{3}$$

where A is ground truth mask and B is prediction mask from fine-tuned SAM. $2|A \cap B|$ is the intersection area of A and B.

Numerical Modeling—Liposomal LNP. The Liposomal LNP numerical model was developed to simulate the structural and volumetric organization of these systems. The model is comprised of a system of equations (eqs 4.1−4.16) and is solved using MATLAB's vpasolve symbolic solver.

To anchor the simulation, the system is first evaluated at a nor-MC3 molar ratio of 0%, assuming a fixed outer radius r = 45 nm (eq 4.13) to match experimental observations. The total interior water volume at this baseline, denoted $V_{w,0\%\ nMC3}$, is stored as a reference. Subsequently, the model solves nor-MC3 molar ratios from 10% to 70%.

The nor-MC3 molar ratio MR_{nMC3}, the independent variable, is defined as the number of nor-MC3 molecules divided by the total number of molecules comprising the LNP (eq 4.1). In all formulations, ESM and cholesterol are held at equimolar concentrations (eq 4.2).

Cholesterol is assumed to contribute minimally to the lipid headgroup footprint due to its small polar hydroxyl group and its tendency to insert between the acyl chains of neighboring phospholipids. As such, its effective contribution to surface coverage is limited compared to that of phospholipids like ESM. The effective headgroup area, HA_{eff}, is therefore estimated as the sum of the headgroup areas, where HA_{ESM} = 0.5 nm^2 and HA_{chol} = 0.37 nm^2 (eq 4.3).

The total volume of the LNP V_{LNP} is composed of three concentric compartments—inner, middle, and outer—represented as spherical shells around a central nor-MC3 oil droplet (eqs 4.4−4.7):

1 **Inner volume** (V_{in}) includes:

Oil core volume V_{oil}, modeled as a solid sphere with radius r_{oil} (eq 4.8). This volume is derived from the number of nor-MC3 molecules n_{nMC3}, molecular weight MW_{nMC3} = 614.06 g/mol, oil density ρ_{nMC3} = 0.886 × 10^{-21} g/nm^3, and Avogadro's number N_A = 6.022 × 10^{23} mol^{-1} (eq 4.9).

Monolayer volume V_{ml}, a thin spherical shell around the oil droplet of fixed monolayer thickness t_{ml} = 2.5 nm (eq 4.10).

Adsorbed water volume on monolayer $V_{w,ml}$, which assumes 10 water molecules ($n_{w,a}$ = 10) are associated with each outward-facing monolayer lipid head. This is calculated by dividing the monolayer's surface area by HA_{eff}, and multiplying by the volume per water molecule derived from MW_w = 18.02 g/mol and ρ_w = 0.997 × 10^{-21} g/nm^3 (eq 4.11). At 0% nor-MC3, all components of V_{in} are zero.

2 **Middle volume** (V_{mid}) includes:

Encapsulated mRNA volume V_{mRNA}, modeled as a cylindrical nucleotide helix with diameter d_{mRNA} = 2 nm, nucleotide length l_{nt} = 0.33 nm, number of bases for NanoLuc n_b = 808, and number of encapsulated copies n_{mRNA} = 4 per LNP (eq 4.12). At 0% nor-MC3, no cationic lipid is present to complex mRNA, so V_{mRNA} = 0.

Bulk water volume V_w, which decreases as nor-MC3 is introduced. This depletion is modeled using an inverse decaying sigmoidal function governed by a steepness parameter k = 12%$^{-1}$ and a midpoint transition molar ratio MR_x = 15% (eq 4.13), corresponding to the molar ratio at the transition point between high and water content (i.e., approximately when the water content drops by half).

3 **Outer volume** (V_{out}) includes:

Bilayer volume (V_{bl}), a thin and spherical shell with of fixed bilayer thickness t_{bl} = 5 nm (eq 4.14).

Adsorbed water on bilayer $V_{w,bl}$, following the same rationale as the monolayer case, but accounting for both inward- and outward-facing bilayer leaflets (eq 4.15).

The total number of ESM and cholesterol molecules, n_{ESM} and n_{chol}, is calculated as the sum of surface areas of all relevant interfaces (oil monolayer (if present), inner bilayer leaflet, and outer bilayer leaflet) divided by HA_{eff} (eq 4.16). At 0% nor-MC3, there is no monolayer, so its surface area is excluded.

The outer cross-sectional area is defined geometrically as πr^2, with the radius r being a dependent variable solved from the system for each value of MR_{nMC3}.

Numerical Modeling−Bleb LNP. The Bleb LNP numerical model describes an aqueous mRNA-containing compartment blebbed out of a KC2-based oil droplet. The model solves for the aqueous and oil radii, r_{aq} and r_{oil}, using a defined system of equations (eqs 5.1−5.20).

https://doi.org/10.1021/acsnano.5c09956
ACS Nano 2025, 19, 33387−33398

The model begins by calculating the number of lipid molecules. The number of KC2 molecules, n_{KC2}, is determined from the number of nucleotide bases per FLuc mRNA. It is equal to the number of bases n_b = 1921, multiplied by the number of encapsulated mRNA copies per LNP n_{mRNA} = 4 and the N/P ratio $\frac{N}{P} = 6$ (eq 5.1). The number of DSPC and cholesterol molecules, n_{DSPC} and n_{chol}, are calculated based on the relative molar ratios of the three lipid species: MR_{DSPC} = 10%, MR_{chol} = 40%, MR_{KC2} = 50% (eqs 5.2 and 5.3).

Because DSPC and cholesterol are not equimolar in this system, the monolayer and bilayer they form will differ in composition. To account for this, the effective headgroup area, HA_{eff}, is calculated as a mole-fraction-weighted average of the individual lipid headgroup areas, where HA_{DSPC} = 0.65 nm^2 and HA_{chol} = 0.37 nm^2 (eq 5.4).

The **oil compartment volume**, V_{KC2}, is calculated using the number of KC2 molecules n_{KC2}, molecular weight MW_{KC2} = 642.11 g/mol, oil density $\rho_{KC2} = 0.886 \times 10^{-21}$ g/nm^3, and Avogadro's number $N_A = 6.022 \times 10^{23}$ mol^{-1} (eq 5.5).

The **aqueous compartment volume**, V_{aq}, is modeled as sphere of radius r_{aq}, excluding the surrounding bilayer (eq 5.6). Its total volume is the sum of the volume of encapsulated mRNA, the volume of water adsorbed to the bilayer, and the interior water volume (eq 5.7).

The **mRNA volume** V_{mRNA} is calculated just as before as a cylindrical nucleotide helix with diameter d_{mRNA} = 2 nm, nucleotide length l_{nt} = 0.33 nm, number of encapsulated copies n_{mRNA} = 4 per LNP, but with number of bases for FLuc n_b = 1921 (eq 5.8).

The **adsorbed water volume** $V_{w,ml}$ assumes 10 water molecules adsorbed onto each inner leaflet lipid head of the bilayer (eq 5.9).

The **remaining water volume** V_w is solved indirectly through a constraint on cross-sectional area.

The independent variable in the model is the mRNA and water internal structure fraction $f_{mRNA+water}$, defined as the ratio of the cross-sectional area $CSA_{mRNA+water}$ to the cross-sectional area of the entire LNP CSA_{LNP} (eq 5.10). The Bleb LNP is geometrically modeled as two overlapping spheres of outer radii $r_{oil,o}$ and $r_{aq,o}$, separated by a center-to-center distance d. These outer radii account for the monolayer (2.5 nm) and bilayer (5 nm) thicknesses, respectively (eqs 5.11 and 5.12). To simplify the geometry, the separation distance d is either set to r_{oil} if $f_{mRNA+water} < 50\%$, to r_{aq} if $f_{mRNA+water} > 50\%$, and to the average of the two radii if $f_{mRNA+water} = 50\%$ (eq 5.13).

The aqueous cross-sectional area $CSA_{mRNA+water}$ is calculated from r_{aq} (eq 5.14), and the total LNP cross-sectional area CSA_{LNP} is determined by summing the cross-sectional areas of both spheres and subtracting the overlapping intersection region (eq 5.15). The cross-sectional intersection area $CSA_{intersection}$ and the corresponding volume intersection $V_{intersection}$ are computed using derived geometric formulas for overlapping spheres (eqs 5.16 and 5.17).[38,39]

Since the aqueous compartment is treated as being "blebbed out" of the oil droplet, the oil radius r_{oil} is defined as the radius of a hypothetical sphere containing both the oil volume occupied by KC2 and the intersection volume (eq 5.18).

The number of DSPC and cholesterol molecules can then be summed together to find the outer surface area of the entire LNP SA_{LNP} (eq 5.19), which is the outermost monolayer lipid surface formed by DSPC and cholesterol encompassing both compartments.

Finally, the total end-to-end diameter of the LNP is simply the sum of the monolayer thickness, the radius of the oil compartment, the separation distance of the two spheres, the radius of the aqueous compartment, and the bilayer thickness (eq 5.20).

ASSOCIATED CONTENT

Data Availability Statement

All relevant data from this study are available within the article, source data and supplementary documents. The training and validation data set described in this paper are available upon reasonable request for research purposes. All code was implemented in Python and Jupyter Notebook. Several packages were used for training models, data analysis and visualization, including Pytorch (2.3.0), Ultralytics (8.2.2), numpy (1.26.4), pandas (2.2.2), matplotlib (3.8.4), plotly (5.22.0) and pycocotools (2.0.7). Model trainings were carried out at Google's Colab, which is a hosted Jupyter Notebook service with access to A100 GPUs. The code and interface for training and inference and trained model will be released at a GitHub Repository (https://github.com/owenip/LNP_MOD_PUBLIC).

Supporting Information

The Supporting Information is available free of charge at https://pubs.acs.org/doi/10.1021/acsnano.5c09956.

> High level overview of YOLOv12 Architecture, overview of LNP-MOD's Instance Segmentation Module, overview of summary package from LNP-MOD, comparison of 3 different object detection models, comparing Precision, Recall, and mean Average Precision (mAP) between YOLO object detection models, comparison of mean Average Precision (mAP) at IoU thresholds between YOLO object detection models, examples of tricky classification between Liposome and Liposomal LNP, an example of LNP-MOD's summary package with a set of cryo-EM images that contains the Blebs with mRNA LNP and the oil core LNP, an example of LNP-MOD's summary package with a set of cryo-EM images that contains Blebs with mRNA LNP, the oil core LNP, Liposomal LNP, Liposome and corresponding internal structures such as oil droplet and mRNA, Liposomal LNP structures and criterions for the numerical model, firefly luciferase encoded mRNA phase separated morphologies at lower ionizable lipid equimolar ESM/cholesterol Liposomal LNP system, Blebs with mRNA structures and criterions for the numerical model, cross-section of Blebs with mRNA LNP numerical model (PDF)

AUTHOR INFORMATION

Corresponding Author

Miffy Hok Yan Cheng − *Faculty of Pharmaceutical Sciences, Vancouver, British Columbia V6T 1Z3, Canada;* orcid.org/0000-0002-0261-4642; Email: miffy.cheng@ubc.ca

Authors

Owen Yuk Long Ip − *Polymorphic BioSciences, Vancouver, British Columbia V6T 1Z3, Canada*

Harrison D. E. Fan − *Department of Biochemistry and Molecular Biology, University of British Columbia, Vancouver, British Columbia V6T 1Z3, Canada; Department of Physics and Astronomy, University of British Columbia, Vancouver, British Columbia V6T 1Z1, Canada;* orcid.org/0000-0002-6672-1068

Yao Zhang − *Department of Biochemistry and Molecular Biology, University of British Columbia, Vancouver, British Columbia V6T 1Z3, Canada; School of Biomedical Engineering and Michael Smith Laboratories, University of British Columbia, Vancouver, British Columbia V6T 1Z4, Canada;* orcid.org/0000-0003-2783-1317

Jerry Leung − *Department of Biochemistry and Molecular Biology, University of British Columbia, Vancouver, British Columbia V6T 1Z3, Canada; Michael Smith Laboratories, University of British Columbia, Vancouver, British Columbia V6T 1Z4, Canada*

Colton Strong − *Department of Biochemistry and Molecular Biology, University of British Columbia, Vancouver, British Columbia V6T 1Z3, Canada; Michael Smith Laboratories, University of British Columbia, Vancouver, British Columbia V6T 1Z4, Canada*

Janell Ko − *Faculty of Pharmaceutical Sciences, Vancouver, British Columbia V6T 1Z3, Canada*

Pieter R. Cullis − *Polymorphic BioSciences, Vancouver, British Columbia V6T 1Z3, Canada; Department of Biochemistry and Molecular Biology, University of British Columbia, Vancouver, British Columbia V6T 1Z3, Canada;* orcid.org/0000-0001-9586-2508

Complete contact information is available at: https://pubs.acs.org/10.1021/acsnano.5c09956

Author Contributions

O.Y.L.I., H.D.E.F., and M.H.Y.C. drafted, wrote, and edited this manuscript. O.Y.L.I. built the LNP-MOD, designed the pipeline, and analyzed the LNP Cryo-TEM images. Y.Z., J.L., and M.H.Y.C. contributed to data collection for formulation, characterization and data analysis. O.Y.L.I., Y.Z, C.S., J.K., and M.H.Y.C. annotated the curated data set. P.R.C. and H.D.E.F. derived the numerical model equations, and H.D.E.F. programmed scripts to solve them. M.H.Y.C. conceptualized, designed, managed and supervised this project.

Notes

The authors declare the following competing financial interest(s): P.R.C. has a financial interest in Polymorphic BioSciences. O.Y.L.I. is an employee of Polymorphic BioSciences.

ACKNOWLEDGMENTS

We thank Meta AI and Ultralytics for making their source code of models and training framework publicly available. We acknowledge Google Colab for providing the necessary GPU support for our model training and experimentation processes. Cryo-TEM data was collected at the High-Resolution Macromolecular Electron Microscopy (HRMEM) facility at the University of British Columbia (https://cryoem.med.ubc.ca). We thank Claire Atkinson, Joeseph Felt, Liam Worrall and Natalie Strynadka. HRMEM is funded by the Canadian Foundation of Innovation and the British Columbia Knowledge Development Fund. M.H.Y.C was supported by CIHR Research Excellence, Diversity, and Independence (REDI) Early Career Transition Award (AWD-031113). Y.Z. (FBD 193487), C.S. and J.L. was supported by a Frederick Banting and Charles Best Canada Graduate Scholarships Doctoral Award.

REFERENCES

(1) Akinc, A.; Maier, M. A.; Manoharan, M.; Fitzgerald, K.; Jayaraman, M.; Barros, S.; Ansell, S.; Du, X.; Hope, M. J.; Madden, T. D.; Mui, B. L.; Semple, S. C.; Tam, Y. K.; Ciufolini, M.; Witzigmann, D.; Kulkarni, J. A.; van der Meel, R.; Cullis, P. R. The Onpattro Story and the Clinical Translation of Nanomedicines Containing Nucleic Acid-Based Drugs. *Nat. Nanotechnol.* **2019**, *14* (12), 1084−1087.

(2) Schoenmaker, L.; Witzigmann, D.; Kulkarni, J. A.; Verbeke, R.; Kersten, G.; Jiskoot, W.; Crommelin, D. J. A. mRNA-Lipid Nanoparticle COVID-19 Vaccines: Structure and Stability. *Int. J. Pharm.* **2021**, *601*, 120586.

(3) Peden, K. *Considerations for the Quality, Safety and Efficacy of Prophylactic Lipid Nanoparticle mRNA Vaccines*, 2022.

(4) Center for Drug Evaluation and Research. *Liposome Drug Products: Chemistry, Manufacturing, and Controls; Human Pharmacokinetics and Bioavailability; and Labeling Documentation.* (accessed 2024−09−19). https://www.fda.gov/regulatory-information/search-fda-guidance-documents/liposome-drug-products-chemistry-manufacturing-and-controls-human-pharmacokinetics-and.

(5) Simonsen, J. B. A Perspective on Bleb and Empty LNP Structures. *J. Controlled Release* **2024**, *373*, 952−961.

(6) Eygeris, Y.; Patel, S.; Jozic, A.; Sahay, G. Deconvoluting Lipid Nanoparticle Structure for Messenger RNA Delivery. *Nano Lett.* **2020**, *20*, 4543.

(7) Eygeris, Y.; Henderson, M. I.; Curtis, A. G.; Jozić, A.; Stoddard, J.; Reynaga, R.; Chirco, K. R.; Su, G. L.-N.; Neuringer, M.; Lauer, A. K.; Ryals, R. C.; Sahay, G. Preformed Vesicle Approach to LNP Manufacturing Enhances Retinal mRNA Delivery. *Small* **2024**, *20* (37), 2400815.

(8) Meulewaeter, S.; Nuytten, G.; Cheng, M. H. Y.; De Smedt, S. C.; Cullis, P. R.; De Beer, T.; Lentacker, I.; Verbeke, R. Continuous Freeze-Drying of Messenger RNA Lipid Nanoparticles Enables Storage at Higher Temperatures. *J. Controlled Release* **2023**, *357*, 149−160.

(9) Cheng, M. H. Y.; Leung, J.; Zhang, Y.; Strong, C.; Basha, G.; Momeni, A.; Chen, Y.; Jan, E.; Abdolahzadeh, A.; Wang, X.; Kulkarni, J. A.; Witzigmann, D.; Cullis, P. R. Induction of Bleb Structures in Lipid Nanoparticle Formulations of mRNA Leads to Improved Transfection Potency. *Adv. Mater.* **2023**, *35* (31), 2303370.

(10) Zhigaltsev, I. V.; Tam, Y. Y. C.; Kulkarni, J. A.; Cullis, P. R. Synthesis and Characterization of Hybrid Lipid Nanoparticles Containing Gold Nanoparticles and a Weak Base Drug. *Langmuir* **2022**, *38* (25), 7858−7866.

(11) Johnston, M. J. W.; Edwards, K.; Karlsson, G.; Cullis, P. R. Influence of Drug-to-Lipid Ratio on Drug Release Properties and Liposome Integrity in Liposomal Doxorubicin Formulations. *J. Liposome Res.* **2008**, *18* (2), 145−157.

(12) Yao, X.; Fan, X.; Yan, N. Cryo-EM Analysis of a Membrane Protein Embedded in the Liposome. *Proc. Natl. Acad. Sci. U.S.A.* **2020**, *117* (31), 18497−18503.

(13) Wagner, T.; Merino, F.; Stabrin, M.; Moriya, T.; Antoni, C.; Apelbaum, A.; Hagel, P.; Sitsel, O.; Raisch, T.; Prumbaum, D.; Quentin, D.; Roderer, D.; Tacke, S.; Siebolds, B.; Schubert, E.; Shaikh, T. R.; Lill, P.; Gatsogiannis, C.; Raunser, S. SPHIRE-crYOLO Is a Fast and Accurate Fully Automated Particle Picker for Cryo-EM. *Commun. Biol.* **2019**, *2* (1), 1−13.

(14) Bepler, T.; Kelley, K.; Noble, A. J.; Berger, B. Topaz-Denoise: General Deep Denoising Models for cryoEM and cryoET. *Nat. Commun.* **2020**, *11* (1), 5208.

(15) Dhakal, A.; Gyawali, R.; Wang, L.; Cheng, J. CryoTransformer: A Transformer Model for Picking Protein Particles from Cryo-EM Micrographs. *Bioinformatics* **2024**, *40* (3), btae109.

(16) Mazurowski, M. A.; Dong, H.; Gu, H.; Yang, J.; Konz, N.; Zhang, Y. Segment Anything Model for Medical Image Analysis: An Experimental Study. *Med. Image Anal.* **2023**, *89*, 102918.

(17) Huang, Y.; Yang, X.; Liu, L.; Zhou, H.; Chang, A.; Zhou, X.; Chen, R.; Yu, J.; Chen, J.; Chen, C.; Liu, S.; Chi, H.; Hu, X.; Yue, K.; Li, L.; Grau, V.; Fan, D.-P.; Dong, F.; Ni, D. Segment Anything Model for Medical Images? *Med. Image Anal.* **2024**, *92*, 103061.

(18) Gyawali, R.; Dhakal, A.; Wang, L.; Cheng, J. CryoSegNet: Accurate Cryo-EM Protein Particle Picking by Integrating the Foundational AI Image Segmentation Model and Attention-Gated U-Net. *Briefings Bioinf.* **2024**, *25* (4), bbae282.

(19) He, S.; Bao, R.; Li, J.; Stout, J.; Bjornerud, A.; Grant, P. E.; Ou, Y. Computer-Vision Benchmark Segment-Anything Model (SAM) in Medical Images: Accuracy in 12 Datasets. *arXiv* **2023**, 2304.09324.

(20) Ravi, N.; Gabeur, V.; Hu, Y.-T.; Hu, R.; Ryali, C.; Ma, T.; Khedr, H.; Rädle, R.; Rolland, C.; Gustafson, L.; Mintun, E.; Pan, J.; Alwala, K. V.; Carion, N.; Wu, C.-Y.; Girshick, R.; Dollár, P.; Feichtenhofer, C. SAM 2: Segment Anything in Images and Videos. *arXiv* **2024**, 2408.00714.

(21) Zhu, J.; Hamdi, A.; Qi, Y.; Jin, Y.; Wu, J. Medical SAM 2: Segment Medical Images as Video via Segment Anything Model 2. *arXiv* **2024**, 2408.00874.
(22) Maurer, N.; Wong, K. F.; Stark, H.; Louie, L.; McIntosh, D.; Wong, T.; Scherrer, P.; Semple, S. C.; Cullis, P. R. Spontaneous Entrapment of Polynucleotides upon Electrostatic Interaction with Ethanol-Destabilized Cationic Liposomes. *Biophys. J.* **2001**, *80* (5), 2310−2326.
(23) Semple, S. C.; Klimuk, S. K.; Harasym, T. O.; Santos, N. D.; Ansell, S. M.; Wong, K. F.; Maurer, N.; Stark, H.; Cullis, P. R.; Hope, M. J.; Scherrer, P. Efficient encapsulation of antisense oligonucleotides in lipid vesicles using ionizable aminolipids: formation of novel small multilamellar vesicle structures. *Biochim. Biophys. Acta* **2001**, *1510*, 152.
(24) MacLachlan, I.; Cullis, P. Diffusible-PEG-Lipid Stabilized Plasmid Lipid Particles. In *Advances in Genetics*; Elsevier, 2005; Vol. *53*, pp 157−188.
(25) Maurer, N.; Mori, A.; Palmer, L.; Monck, M. A.; Mok, K. W. C.; Mui, B.; Akhong, Q. F.; Cullis, P. R. Lipid-Based Systems for the Intracellular Delivery of Genetic Drugs. *Mol. Membr. Biol.* **1999**, *16* (1), 129−140.
(26) Kulkarni, J. A.; Darjuan, M. M.; Mercer, J. E.; Chen, S.; van der Meel, R.; Thewalt, J. L.; Tam, Y. Y. C.; Cullis, P. R. On the Formation and Morphology of Lipid Nanoparticles Containing Ionizable Cationic Lipids and siRNA. *ACS Nano* **2018**, *12* (5), 4787−4795.
(27) Zhigaltsev, I. V.; Cullis, P. R. Morphological Behavior of Liposomes and Lipid Nanoparticles. *Langmuir* **2023**, *39* (9), 3185−3193.
(28) Chander, N.; Basha, G.; Yan Cheng, M. H.; Witzigmann, D.; Cullis, P. R. Lipid Nanoparticle mRNA Systems Containing High Levels of Sphingomyelin Engender Higher Protein Expression in Hepatic and Extra-Hepatic Tissues. *Mol. Ther.—Methods Clin. Dev.* **2023**, *30*, 235−245.
(29) Cheng, M. H. Y.; Zhang, Y.; Fox, K.; Leung, J.; Strong, C.; Kang, E.; Chen, Y.; Tong, M.; Bommadevara, H.; Jan, E.; Ip, O. Y. L.; Rodríguez-Rodríguez, C.; Saatchi, K.; Häfeli, U. O.; Abdolahzadeh, A.; Witzigmann, D.; Cullis, P. R. Liposomal Lipid Nanoparticles for Extrahepatic Delivery of mRNA. *Nat. Commun.* **2025**, *16* (1), 4135.
(30) Tian, Y.; Ye, Q.; Doermann, D. YOLOv12: Attention-Centric Real-Time Object Detectors. *arXiv* **2025**, 2502.12524.
(31) Ma, J.; He, Y.; Li, F.; Han, L.; You, C.; Wang, B. Segment Anything in Medical Images. *Nat. Commun.* **2024**, *15* (1), 654.
(32) Kirillov, A.; Mintun, E.; Ravi, N.; Mao, H.; Rolland, C.; Gustafson, L.; Xiao, T.; Whitehead, S.; Berg, A. C.; Lo, W.-Y.; Dollar, P.; Girshick, R. Segment Anything. In *Proceedings of the IEEE/CVF International Conference on Computer Vision*, 2023, pp 4015−4026.
(33) Huang, J.; Buboltz, J. T.; Feigenson, G. W. Maximum Solubility of Cholesterol in Phosphatidylcholine and Phosphatidylethanolamine Bilayers. *Biochim. Biophys. Acta, Biomembr.* **1999**, *1417* (1), 89−100.
(34) Pandit, S. A.; Vasudevan, S.; Chiu, S. W.; Mashl, R. J.; Jakobsson, E.; Scott, H. L. Sphingomyelin-Cholesterol Domains in Phospholipid Membranes: Atomistic Simulation. *Biophys. J.* **2004**, *87* (2), 1092−1100.
(35) Chattopadhyay, M.; Krok, E.; Orlikowska, H.; Schwille, P.; Franquelim, H. G.; Piatkowski, L. Hydration Layer of Only a Few Molecules Controls Lipid Mobility in Biomimetic Membranes. *J. Am. Chem. Soc.* **2021**, *143* (36), 14551−14562.
(36) Loshchilov, I.; Hutter, F. Decoupled Weight Decay Regularization. *arXiv* **2018**, 1711.05101.
(37) Maier-Hein, L.; Reinke, A.; Godau, P.; Tizabi, M. D.; Buettner, F.; Christodoulou, E.; Glocker, B.; Isensee, F.; Kleesiek, J.; Kozubek, M.; Reyes, M.; Riegler, M. A.; Wiesenfarth, M.; Kavur, A. E.; Sudre, C. H.; Baumgartner, M.; Eisenmann, M.; Heckmann-Nötzel, D.; Rädsch, T.; Acion, L.; Antonelli, M.; Arbel, T.; Bakas, S.; Benis, A.; Blaschko, M. B.; Cardoso, M. J.; Cheplygina, V.; Cimini, B. A.; Collins, G. S.; Farahani, K.; Ferrer, L.; Galdran, A.; van Ginneken, B.; Haase, R.; Hashimoto, D. A.; Hoffman, M. M.; Huisman, M.; Jannin, P.; Kahn, C. E.; Kainmueller, D.; Kainz, B.; Karargyris, A.; Karthikesalingam, A.; Kofler, F.; Kopp-Schneider, A.; Kreshuk, A.; Kurc, T.; Landman, B. A.; Litjens, G.; Madani, A.; Maier-Hein, K.; Martel, A. L.; Mattson, P.; Meijering, E.; Menze, B.; Moons, K. G. M.; Müller, H.; Nichyporuk, B.; Nickel, F.; Petersen, J.; Rajpoot, N.; Rieke, N.; Saez-Rodriguez, J.; Sánchez, C. I.; Shetty, S.; van Smeden, M.; Summers, R. M.; Taha, A. A.; Tiulpin, A.; Tsaftaris, S. A.; Van Calster, B.; Varoquaux, G.; Jäger, P. F. Metrics Reloaded: Recommendations for Image Analysis Validation. *Nat. Methods* **2024**, *21* (2), 195−212.
(38) Weisstein, E. W. Circle-Circle Intersection. https://mathworld.wolfram.com/Circle-CircleIntersection.html (accessed, 2025 05 26).
(39) Weisstein, E. W. Sphere-Sphere Intersection. https://mathworld.wolfram.com/Sphere-SphereIntersection.html (accessed, 2025 05 26).

https://doi.org/10.1021/acsnano.5c09956
ACS Nano 2025, 19, 33387−33398