





Peptide-encoding mRNA barcodes for the high-throughput in vivo screening of libraries of lipid nanoparticles for mRNA delivery

Received: 19 August 2022

Accepted: 26 March 2023

Published online: 1 May 2023

 Check for updates

Luke H. Rhym^{1,2,3}, Rajith S. Manan ^{1,2,3}, Antonius Koller¹, Georgina Stephanie¹ & Daniel G. Anderson ^{1,2} 


Developing safe and effective nanoparticles for the delivery of messenger RNA (mRNA) is slow and expensive, partly due to the lack of predictive power of in vitro screening methods and the low-throughput nature of in vivo screening. While DNA barcoding and batch analysis present methods for increasing in vivo screening throughput, they can also result in incomplete or misleading measures of efficacy. Here, we describe a high-throughput and accurate method for the screening of pooled nanoparticle formulations within the same animal. The method uses liquid chromatography with tandem mass spectrometry to detect peptide barcodes translated from mRNAs in nanoparticle-transfected cells. We show the method's applicability by evaluating a library of over 400 nanoparticle formulations with 384 unique ionizable lipids using only nine mice to optimize the formulation of a biodegradable lipid nanoparticle for mRNA delivery to the liver. Barcoding lipid nanoparticles with peptide-encoding mRNAs may facilitate the rapid development of nanoparticles for mRNA delivery to specific cells and tissues.

RNA therapeutics hold immense potential for the treatment of a wide variety of diseases by directly controlling protein production within specific cells in the body^{1–5}. In particular, non-viral delivery of mRNA using lipid nanoparticles (LNPs) is an attractive method for the delivery of therapeutic proteins due to the transient nature of expression from mRNA and the potential for repeat dosing of synthetic nanoparticles (NPs)^{2–4}. Furthermore, mRNA LNPs have already been applied in pre-clinical models for protein replacement therapy⁶, vaccination^{7,8} and gene editing^{9–12}, as well as for the prevention of COVID-19 (refs. 13,14).

Despite this progress, there continues to be a need for safe, specific and efficacious delivery vehicles. One challenge to delivery lipid development is the poor capability of in vitro studies to predict in vivo

biodistribution and efficacy, as well as the low-throughput nature of traditional in vivo studies^{15,16}. While these issues have been partly addressed by the application of methods such as batch analysis⁸ and DNA and/or RNA barcoding^{17,18}, these methods each have their own shortcomings that limit their use as screening approaches in pre-clinical studies. For example, the extent to which batch analysis can reduce the number of animals required for library screening depends on the composition of the library itself; in cases where a large fraction of a library consists of similarly performing LNPs, batch analysis can require a comparable number of animals to single LNP analysis. Furthermore, while DNA and/or RNA barcoding can provide a wealth of information regarding the separate biodistributions of dozens of NPs within the

¹David H. Koch Institute for Integrative Cancer Research, Massachusetts Institute of Technology, Cambridge, MA, USA. ²Department of Chemical Engineering, Massachusetts Institute of Technology, Cambridge, MA, USA. ³These authors contributed equally: Luke H. Rhym, Rajith S. Manan.

 e-mail: dgander@mit.edu

same animal, previous studies have shown that trafficking of LNPs to a specific tissue does not necessarily lead to expression within that tissue and that, often, tissues that have very low accumulation of a certain LNP will have functional delivery at a much higher rate than tissues with substantially higher accumulation^{19,20}.

Recently, a method for *in vivo* screening of NPs for functional mRNA delivery using LNPs coformulated with Cre mRNA and DNA barcodes in conjunction with the transgenic Ai14 mouse, termed 'FIND', has been reported²¹. Briefly, in this system, successfully transfected cells will constitutively express tdTomato, and DNA barcodes that are colocalized with tdTomato signal can be quantified as an indirect measure of the extent of functional delivery^{21,22}. While this presents a significant improvement over the original DNA barcoding assay, it still has several important limitations. For example, NPs that are internalized or retained on the surface of a target cell will result in positive hits in a FIND screen as long as the target cell is successfully transfected by a single NP type (Fig. 1a). This may lead to misleading results, especially given that studies have shown uptake of multiple NPs per cell across multiple endosomes^{23,24}. To avoid this 'bystander' effect, an ideal multiplexed assay for functional delivery would instead directly measure the amount of protein production effected by each NP. In addition, LNPs containing mRNA barcodes and DNA barcodes have been shown to result in significantly different biodistributions, which may lead to additional bias¹⁸. Finally, FIND is dependent on the use of the transgenic Ai14 mouse model, precluding its use as a screening system in larger animal models and in specific disease models for which Cre-Lox reporter systems are not available.

Here, we have developed a peptide barcode-based screening system that allows for the simultaneous evaluation of multiple LNPs within the same animal by packaging a different peptide barcode encoding mRNA into each distinct type of LNP. Barcoded LNPs can then be pooled and administered in a single dose to the same animal and, once translation of successfully delivered mRNA has occurred in a target tissue or cell type, expression of each peptide barcode sequence can be quantified using liquid chromatography followed by tandem mass spectrometry (LC-MS/MS). Thus, each unique LNP formulation yields its own distinct measurement that is a direct quantification of protein production effected by it within the analysed tissue or cell type. Because this method does not rely on the presence of any reporter transgene, it is inherently model independent and theoretically may be used in any preclinical model. In this study we show that, by applying MS/MS, we are able to detect and quantify the expression of peptide barcodes from up to 65 distinct mRNA LNPs in both *in vitro* and *in vivo* models accurately and within a wide dynamic range. We then apply this system to screen a library of 384 ionizable lipids over more than 400 different LNPs for their ability to effect hepatic protein production and, as a result, develop a new biodegradable LNP, RM133-3-21, for potent functional mRNA delivery *in vivo*.

Results

We reasoned that an ideal screening system for functional mRNA delivery would be model independent so that it could be applied in any preclinical model of disease, would consist of multiple measures of protein production that are each orthogonal to any others such that multiple formulations could be tested within the same assay and would be highly sensitive and quantitative over a wide dynamic range. The simplest approach to having an orthogonal functional output for each NP within a pool would be to have each NP effect production of a unique protein. The target tissues or cell populations could then be collected and lysed, and each expressed protein could be detected and its expression level measured using LC-MS/MS, which is a selective and sensitive method for protein detection and quantification (Fig. 1b). Since the only prerequisite for this assay is that the targeted cell type be able to translate functional protein from exogenously delivered mRNA, the proposed method would inherently be model independent.

Thus, the assay as described would meet all three of our design criteria for an ideal screening system for functional delivery.

To that end, we developed an assay in which NPs are formulated with unique peptide barcode encoding mRNAs, which, when functionally delivered to the cytoplasm of target cells, express a protein fused with a peptide barcode sequence that can be detected and quantified through LC-MS/MS and a short epitope tag for quantification of expression using enzyme-linked immunosorbent assay. To reduce background noise from non-barcode proteins present within target cells, we used a monomeric variant of streptavidin (monomeric streptavidin, mSA)^{25,26}, which can easily be enriched from crude lysates using a biotinylated scaffold, as the carrier protein for each peptide barcode (Fig. 1c). After barcode-tagged mSA is immobilized on biotinylated beads, the C-terminal end of the carrier protein is cleaved by tobacco etch virus (TEV) protease, releasing the peptide barcode into solution. Finally, the resulting peptide barcode mixture is separated from the biotinylated beads through filtration and run on LC-MS/MS to quantify the relative amounts of each peptide present in solution.

As an initial proof of concept, we wanted to determine whether this approach could be used to accurately measure functional delivery to cells cultured *in vitro*. To that end, we prepared a set of 24 mRNAs that, upon functional delivery to the cytosol of target cells, would be translated into functional carrier proteins tagged with unique peptide barcode sequences (Supplementary Table 1). To determine whether this assay would enable simultaneous quantification of functional delivery of each mRNA, we created two pools (F2 and F3) that were each composed of the same 24 peptide barcode mRNAs but in different relative amounts and used them to transfect human embryonic kidney 293T (HEK293T) cells (Fig. 2a). F2 was constructed by adding each peptide barcode mRNA (except bc28 and bc32, which were each added at the median dose) at a different dose ranging from 13.6 ng to 1.95 μ g; F3 was constructed by forming six groups each consisting of four peptide barcode mRNAs, with the mRNAs of each group added at a relative dose four times lower than that of the previous group, such that the dynamic range covered (2.42 ng to 2.48 μ g) was roughly three orders of magnitude (Supplementary Table 1). To account for natural variations in measured peptide abundance due to different extents to which peptides ionize and fragment, as well as any differences in translation efficiencies between mRNAs, we transfected a separate set of cells with an additional pool composed of each mRNA in an equal amount (equal delivery reference). By normalizing the MS2 intensity of each peptide with respect to its intensity in the equal delivery reference, we are able to obtain a relative peptide abundance with respect to the other peptides within the pool. In general, we observed a very strong correlation ($R^2 = 0.951$ and $R^2 = 0.991$ for pools F2 and F3, respectively) between the normalized abundance of each peptide barcode measured via LC-MS/MS to the dose of the corresponding mRNA (Fig. 2b).

Encouraged by these results, we performed a similar proof of concept experiment *in vivo*, using previously described cKK-e12 LNPs²⁷ to deliver an mRNA pool composed of paired peptide barcode mRNAs at doses linearly spaced from 0.001 to 0.1 mg kg⁻¹ (Fig. 2a and Supplementary Table 2). We again found that, when normalized to an equal delivery reference, the measured peptide abundances align closely ($R^2 = 0.929$) with the administered doses of the corresponding peptide barcode mRNAs (Fig. 2b). Altogether, the results from these proof of concept experiments demonstrate that the peptide barcoding assay can accurately measure the amount of protein translated from multiple mRNAs within a single biological sample.

We then sought to optimize and expand the peptide barcode mRNA library to increase the screening throughput of the system. Our previous proof of concept experiments were performed using HA-tagged barcode sequences; since this tag comprises over half of the amino acid sequence of each barcode, we reasoned that varying the epitope tag could substantially affect sensitivity and quantification using the assay. Thus, we explored the effect of using four different

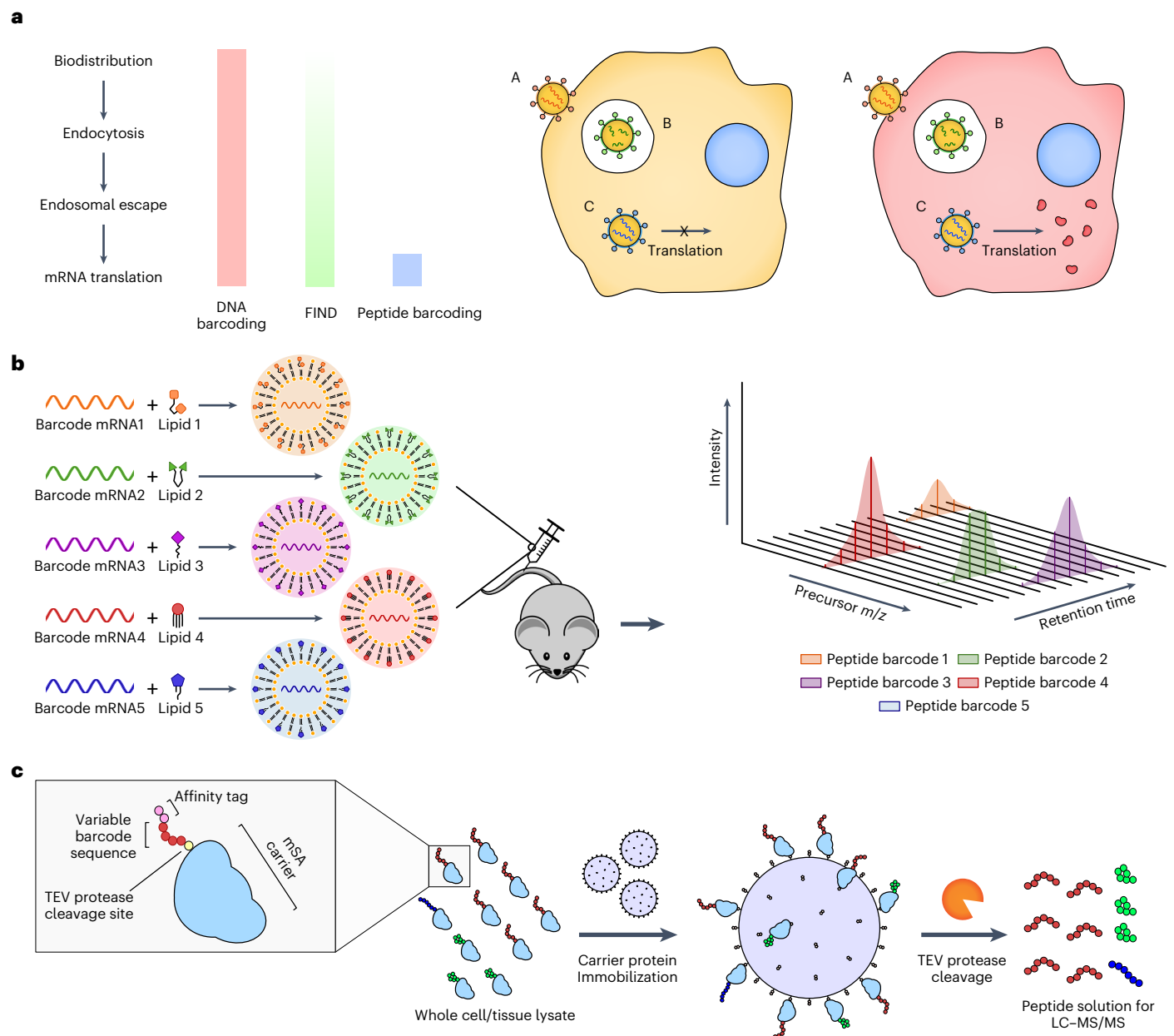


Fig. 1 | Overview of the peptide barcoding assay. a, Comparison of DNA barcoding, FIND and peptide barcoding. While DNA barcoding is unable to distinguish between functional and non-functional delivery, both FIND and peptide barcoding result in positive signal only when functional delivery has occurred. However, FIND will also detect any LNPs that are localized to the same cell, whereas peptide barcoding will only result in signal for LNPs that have resulted in protein production. For example, in the scenario to the right, FIND would identify LNPs A, B and C as hits, when only C has actually resulted

in functional delivery. **b**, Overview of the peptide barcoding assay, where NPs are each formulated with a unique peptide barcode mRNA, the products of which can then be detected and quantified using LC-MS/MS. **c**, Structure of the peptide barcode protein, which consists of a carrier protein (mSA, monomeric streptavidin) fused to a variable barcode sequence and an affinity tag through a TEV protease recognition site. The carrier protein and associated peptide barcode can be purified from whole lysate by immobilization onto biotinylated beads.

epitope tags (HA, c-myc, FLAG and 6xHis) on measured peptide abundance. As expected, the choice of epitope tag had a significant impact on peptide abundance, with myc-tagged barcodes having more than an order of magnitude increase in their measured precursor ion intensity relative to those that were HA-tagged; conversely, FLAG-tagged and His-tagged barcodes were either undetected or resulted in intensities that were several orders of magnitude lower than their HA-tagged and myc-tagged counterparts (Supplementary Fig. 1).

To increase throughput of the assay, we generated a library of transcription templates encoding 114 unique myc-tagged peptide barcode mRNAs (Supplementary Table 3). These were then combined in equal

amounts and subsequently transcribed to produce a single pool of peptide barcode mRNAs; this pool was then delivered *in vivo* to the mouse liver using cCK-e12 LNPs, and the resulting peptide barcode mixture was analysed using LC-MS/MS to identify a subset of suitable barcode sequences (Fig. 2c). Of the 114 unique barcode sequences, we identified a set of 65 peptides with high peptide dot products in liver lysates from treated mice, low dot products in PBS injected liver lysates, high signal to noise ratio (measured as the ratio of peptide abundance in treated liver lysates to abundance in untreated liver lysates) and some degree of separation in either precursor mass or retention time from other peptides in the library (Fig. 2d and Supplementary Table 3). We then

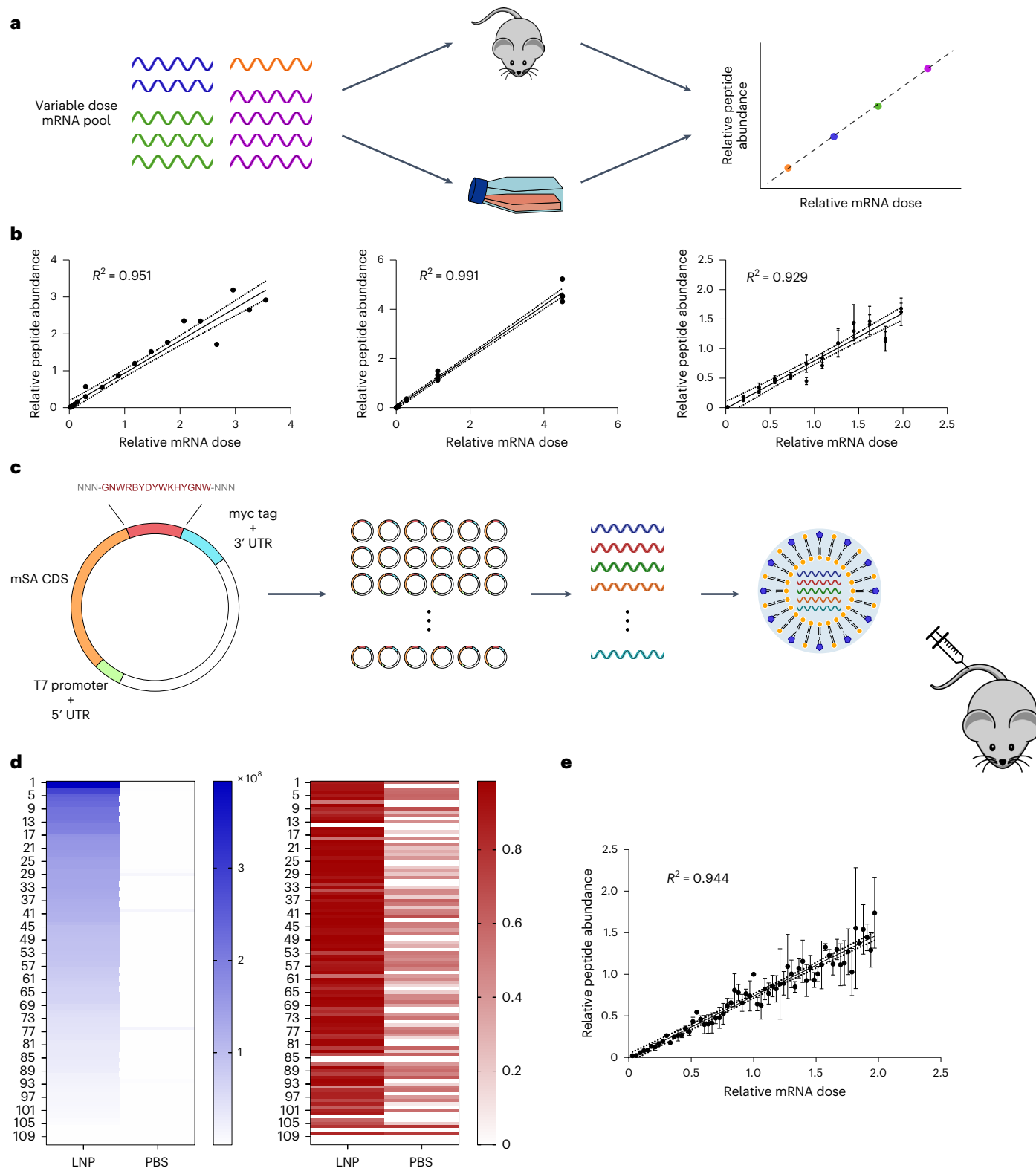


Fig. 2 | Validation and expansion of the peptide barcoding assay. a, An mRNA pool was constructed by combining different doses of each peptide barcode mRNA; the mRNA pool was then packaged into LNPs or lipoplexes and administered to mice or HEK293T cells, respectively; peptide barcode solutions purified from cell lysates were then run on LC-MS/MS, and the relative abundance of each peptide was compared to its dose in the mRNA pool. Relative mRNA dose is calculated as the amount of each mRNA normalized to the average amount of each mRNA in the pool, while relative peptide abundance is calculated as the MS2 intensity measured for each peptide normalized to the MS2 intensity of that peptide in a sample generated using equal amounts of each mRNA (equal delivery reference). **b**, Parity charts comparing relative peptide abundance to relative mRNA dose using 24 peptide barcode mRNAs

in HEK293T cells (left two panels) and in mouse liver (rightmost panel). **c**, A library of IVT templates were generated, pooled together in equal amounts and transcribed to produce a pool of 114 peptide barcode mRNAs, which was then formulated into LNPs and administered to mice. **d**, MSI intensity (left) and peptide dot product (right) for each of the 114 peptide barcode sequences that were screened. **e**, Parity chart comparing relative peptide abundance and relative mRNA dose for a subset of 65 peptide barcodes chosen from the original pool of 114 on the basis of their MSI intensities and dot products. For in vivo experiments in **b**, **d** and **e**, values shown are the mean of $n = 3$ mice per group, with error bars representing the s.d. For the in vitro experiments of **b**, values shown are the result from a single experiment. UTR, untranslated region.

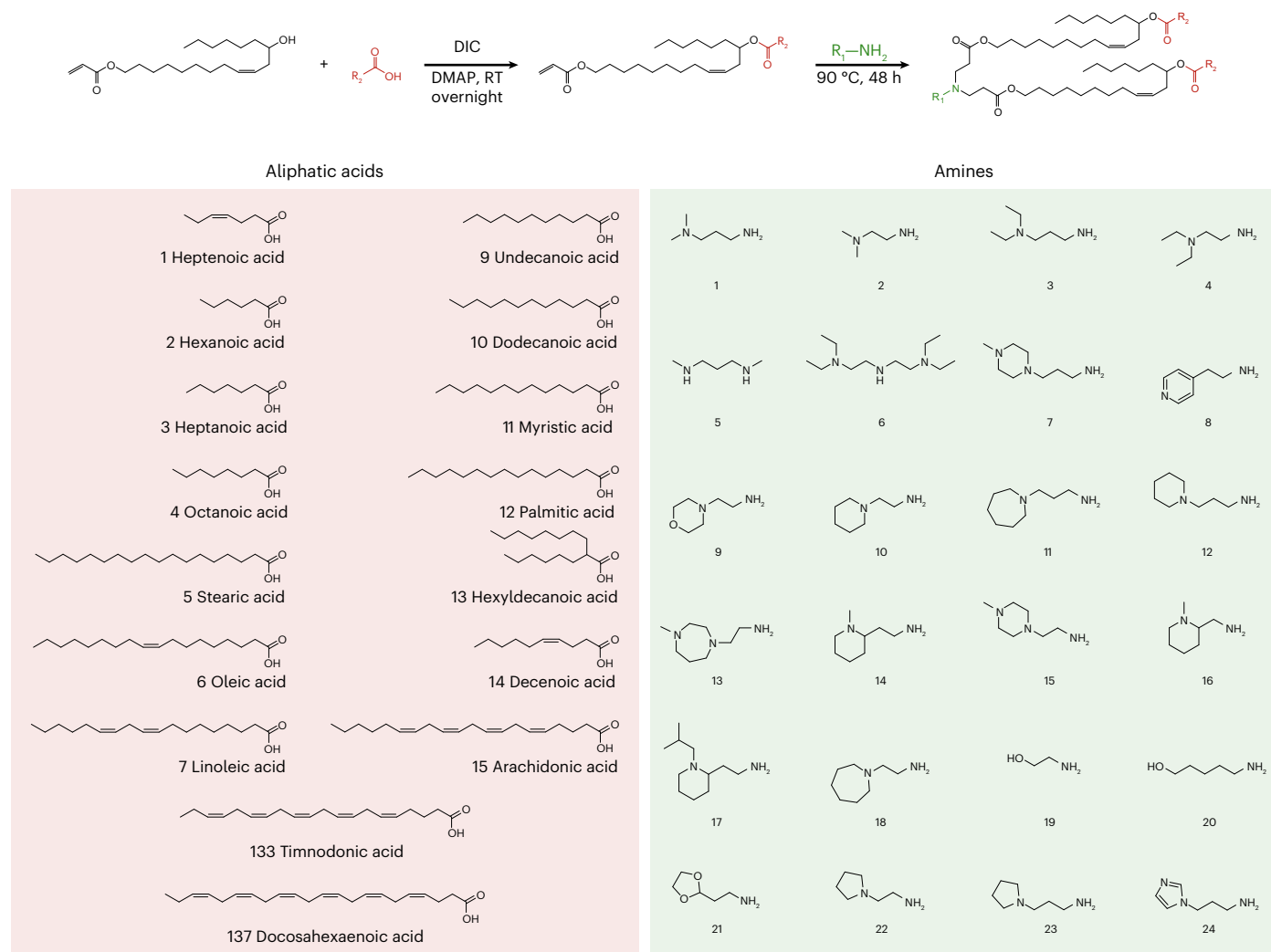


Fig. 3 | Synthesis of 384 ionizable lipids with biodegradable ester linkages and extensive tail branching. Full list of the esters (left) and amines (right) used to generate the library are shown. RT, room temperature. DIC, N,N'-diisopropylcarbodiimide.

performed an *in vivo* validation experiment in the mouse liver using either multiple reaction monitoring (MRM) or data-independent acquisition (DIA) for quantification of peptide abundance and observed excellent correlations between mRNA dose and measured precursor ion intensity, with slightly better quantification using DIA ($R^2 = 0.944$) than MRM ($R^2 = 0.825$) (Fig. 2e and Supplementary Fig. 2). *In vitro* validation experiments in HEK293T cells yielded similar results, indicating that the expanded library of 65 peptide barcodes can be used to accurately and simultaneously measure functional delivery rates both *in vitro* and *in vivo* (Supplementary Fig. 3). We note that only a weak correlation was observed between MS1 intensities of peptides expressed from the equal delivery controls in cultured HEK293T cells and mouse liver lysates ($R^2 = 0.412$), which may be due to differences in translation efficiency in the two contexts, as well as differences in peptide extraction efficiencies from the two matrices (Supplementary Fig. 4).

We then sought to demonstrate that peptide barcoding can be used to develop and optimize LNP formulations. To that end, we synthesized a library of 384 unique ionizable lipids, all featuring the same general branched tail structure containing several biodegradable ester linkages, using the combinatorial reaction and components shown in Fig. 3. We then screened this library for mRNA delivery efficacy by formulating each ionizable lipid into a separate LNP with 1,2-dioleoyl-*sn*-glycero-3-phosphoethanolamine (DOPE), cholesterol and C14-PEG2000, and assaying for hepatic protein production effected by each LNP. Using

the peptide barcoding assay, this entire library was screened using only eight mice (48 unique LNPs per mouse, not including replicates). To avoid potential false positives from background peptides that may be mistaken for barcodes, we set a minimum peptide dot product cut-off of 0.8. Of the 384 LNPs evaluated, 43 resulted in peptide dot products greater than the cut-off, with most (37 out of 43) being derived from either timnodonic acid (133-) or docosahexaenoic acid (137-) (Fig. 4a). We then validated the top four hits identified in the peptide barcoding analysis using the firefly luciferase (FLuc) assay and found that all four LNPs resulted in high levels of FLuc protein production in the liver when evaluated individually (Fig. 4b).

Next, we explored whether we could further improve performance of the lead compound, RM133-3, by optimizing its formulation. Several previous ionizable lipid screens have been performed using either a constant 1,2-distearoyl-*sn*-glycero-3-phosphocholine (DSPC)^{28,29} or DOPE^{20,30} formulation; in addition, a systematic study of the factors important for LNP efficacy found that phospholipid identity and ionizable lipid to weight ratio were the most significant predictors of LNP performance²⁷. Thus, we chose to vary the phospholipid identity (DOPE versus DSPC) and the ionizable lipid to mRNA weight ratio (10, 15 and 20), as well as the molar ratios of lipid components in the formulation, which we varied linearly between the constant DSPC and DOPE formulations previously mentioned (Fig. 5a and Supplementary Table 4). Using a full factorial design of these variables, we generated a library of

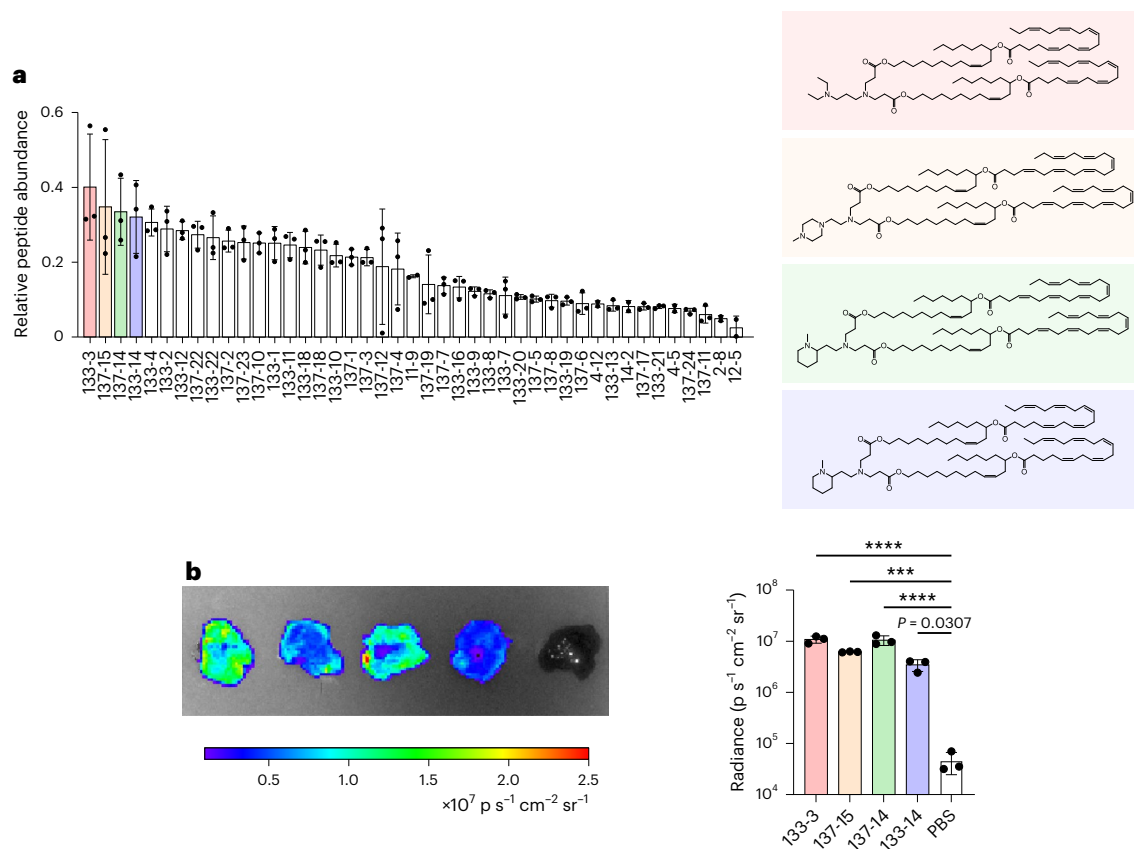


Fig. 4 | Evaluation of an ionizable lipid series using peptide barcoding.

a, Results from the peptide barcoding analysis of the combinatorial library (left) and structures of the top four hits (right). Only lipids that resulted in a peptide dot product more than 0.8 are shown. As in Fig. 2, relative peptide abundance is calculated as the MS1 intensity measured for each peptide normalized to the MS1 intensity of that peptide in a sample generated using equal amounts of each mRNA delivered using a reference lipid, cKK-e12 (equal delivery reference).

b, Representative luminescence image of the top four hits 6 h after intravenous (i.v.) administration of 0.5 mg kg^{-1} FLuc mRNA (left; from left to right, RM133-3, RM137-15, RM137-14, RM133-14) and quantification of average radiance (right). Values shown are the mean of $n = 3$ mice, with error bars representing the s.d. Statistical significance in **b** was determined using a one-way ANOVA with Dunnett's correction.

24 unique peptide-barcoded LNPs with RM133-3 as the ionizable lipid and administered these to the same mouse. From this library, we identified three formulations that performed at a significantly higher level than the original RM133-3 formulation (Fig. 5b). In accordance with previous findings^{27,31}, we found that formulations containing DOPE significantly outperformed those that contained DSPC (Extended Data Fig. 1). To verify that the barcoding analysis can be used to quantitatively assess individual LNP performance in a pooled setting, we compared the top three performing LNPs (RM133-3-20, RM133-3-14 and RM133-3-21) as well as the worst performing LNP (RM133-3-04) to the original formulation (RM133-3-01) using the human erythropoietin (hEPO) assay (Fig. 5c). While the ranking of the top three LNPs identified in the peptide barcoding analysis was not preserved in the results of the hEPO assay, we found that all three top performers did, in fact, outperform the original formulation; in addition, the worst performing LNP in the barcoding analysis resulted in lower protein production than any other formulation tested. Finally, the top-performing LNP, RM133-3-21, was found using the FLuc assay to be roughly 4.5 times more potent than DLin-MC3-DMA, which is a benchmark ionizable lipid currently used in the clinic (Fig. 5d).

Discussion

Over the years, it has become increasingly clear that in vitro screening methods are, in many cases, ill-suited for predicting the in vivo behaviour of nucleic acid delivery vehicles^{15,16}, highlighting the need for increasing the throughput of in vivo testing. To address this, we

have developed an LC-MS/MS based assay for directly measuring the functional delivery of mRNA in vivo by multiple NP carriers within the same animal. By barcoding each NP with an mRNA that encodes a distinct peptide tag, we are able to directly measure the protein production from up to 65 different NPs simultaneously within the same model, making it an ideal assay for applications such as protein replacement therapy and mRNA vaccination, where maximal protein production is desirable. By contrast, other barcoding methods rely on the measurement of DNA or RNA barcodes that have been delivered to the desired cells or tissues^{17,18}; while these methods can be significantly improved by selecting only cells that have successfully been transfected²¹, they all rely on the assumption that the level of distribution of a NP to target cells is directly related to the level of protein production effected by that NP.

Although peptide barcoding of NPs presents a promising method for the development and optimization of mRNA delivery vehicles, it does have a few important limitations. First, while peptide barcoding will be useful for screening NPs for their capacity to deliver mRNA and can easily be adapted for plasmid delivery, it cannot be used to screen for the functional delivery of payloads that result in knockdown of a target protein. For these applications, DNA barcoding methods for indirectly measuring the level of protein knockdown effected by pooled NPs may be useful^{32,33}. Second, multiple studies have now demonstrated that there can be non-linear effects associated with mixing different ionizable lipid species^{34,35}. Indeed, in a pilot screen of the combinatorial library explored in this study, we observed some degree of non-linearity

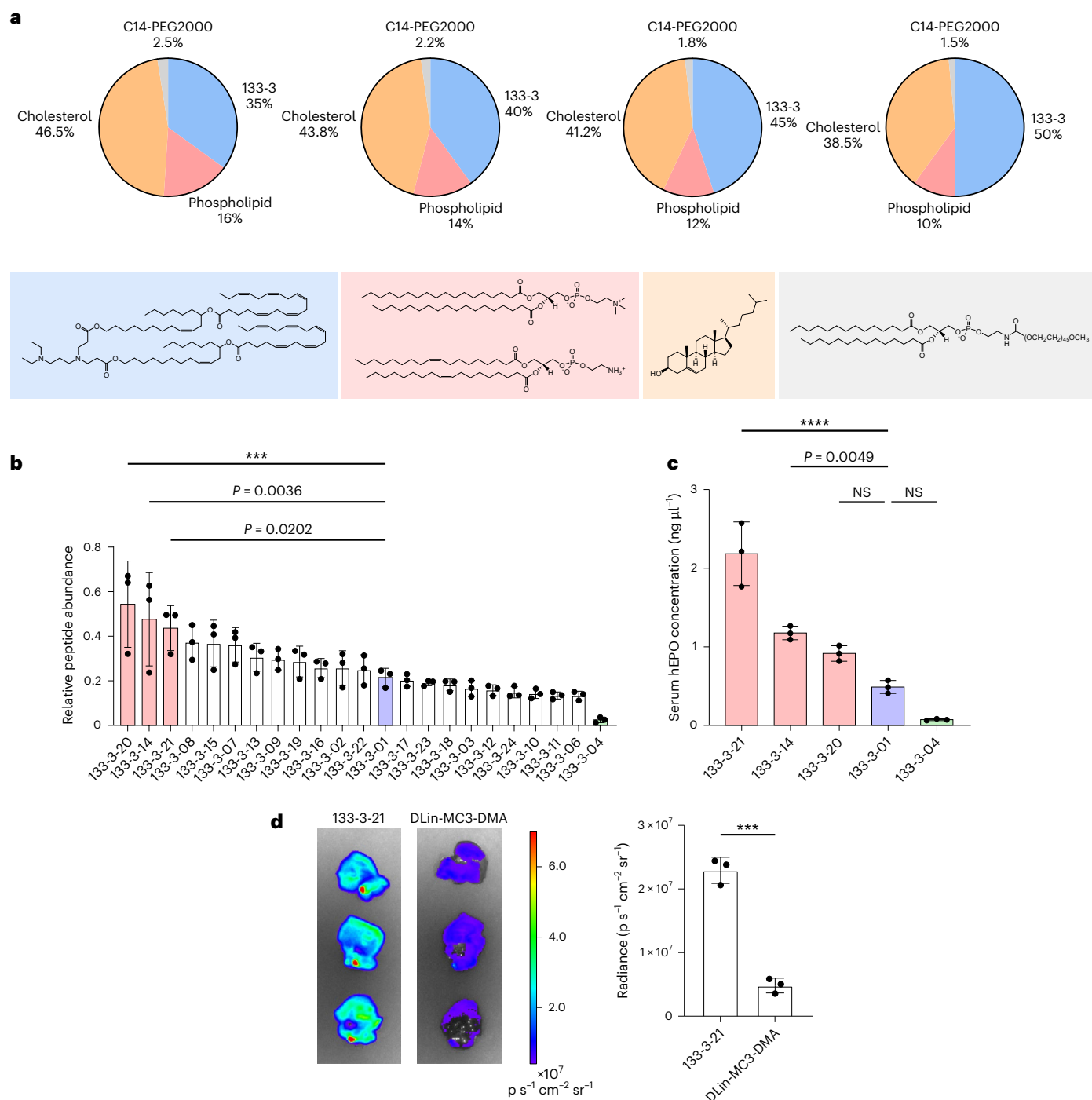


Fig. 5 | Optimization and characterization of lead lipid, RM133-3. a, Overview of the formulation parameters used in the optimization of RM133-3 LNPs.

b, Peptide barcoding analysis of the LNP library. Several of the formulations resulted in significantly higher protein production than the original formulation, RM133-3-01. As in earlier figures, relative peptide abundance is calculated as the MSI intensity measured for each peptide normalized to the MSI intensity of that peptide in a sample generated using equal amounts of each mRNA delivered using a reference lipid, cKK-e12 (equal delivery reference). **c**, Validation of the

barcode analysis using the hEPO assay. The top three performers identified in the barcode analysis (red), as well as the worst performer (green), were compared to the original formulation (blue) using the hEPO assay. **d**, FLuc expression following i.v. administration of RM133-3-21 (left) and MC3 (right) LNPs. Values shown are the mean of $n = 3$ mice, with error bars representing the s.d. Statistical significances in **b** and **c** were determined using a one-way ANOVA with Dunnett's correction. Statistical significance in **d** was determined using an unpaired, two-tailed Student's t -test. NS, not significant.

that resulted in a significantly higher measurement of protein production using our pooled barcoding analysis than with a traditional single NP analysis (Extended Data Fig. 2a–c). We further demonstrated that this non-linearity is inherent to NP pooling in general and is not specific to peptide barcoding (Extended Data Fig. 2d). We speculate that this

may be the cause of discrepancies in Fig. 5, where the rank ordering of LNP efficacies in the hEPO and peptide barcoding assays differed, which further highlights the need for secondary validation assays following any pooled NP analysis. While further studies are required to determine the exact cause of this non-linearity and how to predict its occurrence,

we speculate that it may be related to the level of chemical diversity, particularly within the lipid tails, present within a given NP pool.

A major advantage of DNA and RNA barcoding is the presence of an inherent signal amplification step, which affords a high level of sensitivity. While we have shown that peptide barcoding is also sensitive and can be used to detect and quantify hepatic protein expression at doses as low as 0.001 mg kg⁻¹, further studies will need to be performed to determine whether this method can be used for NP development in applications where protein production may be several orders of magnitude lower. For example, we observed in initial experiments measuring splenic expression of peptide barcodes that, while expressed peptide barcodes can be detected in whole spleen lysates, quantification is much less reliable than in liver lysates (Supplementary Fig. 5). This suggests that further improvements to our assay should be explored to expand peptide barcoding to more challenging applications, particularly those in which the desired levels of protein expression are significantly lower than in the liver.

In this study, we have demonstrated a proof of concept for using peptide barcoding as a screening tool for mRNA delivery vehicles. However, there are several areas for further exploration that could potentially improve the sensitivity, accuracy and throughput of this assay. Although several affinity tag sequences were screened as part of this study, we note that due to its length relative to the entire peptide barcode, it is likely that further engineering of this affinity tag sequence, as well as its length, will result in better sensitivity and more accurate quantification. In addition, use of more recent proteomics methods may further improve the sensitivity and accuracy of peptide barcode quantification. While we have generated a library of 65 peptide barcodes, we anticipate that the throughput of this assay can be improved significantly by systematically screening larger peptide barcode libraries as we have done in Fig. 2. We note that incorporation of additional steps in this screening approach, such as cell-free translation to address potential variability in translation efficiency between different peptide sequences and evaluation of peptide barcodes in different tissue lysates, would likely result in a more streamlined process and ultimately a more robust assay. Furthermore, RNA sequencing can be performed in parallel to peptide barcoding to evaluate both protein expression in and biodistribution to target tissues. Last, the translation efficiency of exogenously delivered mRNA is expected to vary between different cell types. While this study has primarily been focused on the delivery of mRNA to the liver, which is predominantly composed of hepatocytes, tissues with more heterogeneous cell populations may require fractionation before analysis to ensure measurement of delivery to the correct cell type.

As a proof of concept, we have demonstrated that peptide barcoding can be used to quickly and efficiently screen a large library of NPs for their ability to effect protein production in vivo. To summarize, we have screened 384 new ionizable lipids spanning more than 400 distinct NP formulations using only a small fraction of the animals that would have been required using traditional FLuc or hEPO assays. Through this process, we have demonstrated that peptide barcoding can be used for both discovery of ionizable lipids for mRNA delivery and formulation screening for a more fine-tuned optimization of existing ionizable lipids. As a result, we have developed RM133-3-21, which is an LNP that performs at a level 4.5 times that of the benchmark lipid, DLin-MC3-DMA, while incorporating multiple ester linkages that are likely to improve its pharmacokinetics. Whereas here we have demonstrated the use of peptide barcoding for hepatic protein production, we expect that this method can readily be applied to screening for functional delivery to other tissues that have historically been more difficult to target using LNPs. As such, we anticipate that peptide barcoding of LNPs will greatly facilitate the rapid development of improved NPs for mRNA delivery to the liver, as well as to non-hepatic targets.

Methods

Plasmid and mRNA synthesis

Plasmids containing a T7 promoter upstream of a mSA coding sequence (CDS) with a C-terminal peptide barcode were used as templates for mRNA in vitro transcription (IVT). In short, a cassette containing the mSA CDS separated from an affinity tag by two Bsal cut sites was cloned into the pUC19 vector to generate an IVT template cloning vector. Individual templates were then generated by cloning in annealed DNA oligonucleotides corresponding to the desired variable barcode sequences. For the peptide barcode screen of Fig. 2c, a DNA oligonucleotide containing several degenerate bases flanked by two Bsal sites was amplified, and the resulting degenerate double-stranded DNA fragment was cloned into the generic IVT cloning vector using Golden Gate assembly. The degenerate oligonucleotide was designed to exclude certain amino acids, such as cysteine and methionine, to prevent any by-products that could reduce assay sensitivity. IVT templates were linearized using EcoRI, and mRNA was transcribed using the HiScribe T7 High Yield RNA Synthesis Kit (New England Biolabs). Capping was performed cotranscriptionally using CleanCap Reagent AG (TriLink Biotechnologies). Tailing was done enzymatically posttranscription using *Escherichia coli* Poly(A) Polymerase (New England Biolabs). The resulting capped and tailed mRNA was then purified using the Monarch RNA Cleanup Kit (New England Biolabs).

Ionizable lipid synthesis

Details on the synthesis and characterization of the ionizable lipid library of Fig. 3 are provided in the Supplementary Information.

LNP formulation

LNPs were formulated as previously described²⁷. In short, a lipid solution consisting of an ionizable lipid, a phospholipid (either DOPE (Avanti Polar Lipids) or DSPC (Avanti Polar Lipids)), cholesterol (Sigma) and 14:0 PEG2000 PE (Avanti Polar Lipids) was prepared in ethanol. For benchmarking of RM133-3-21, DLin-MC3-DMA was purchased from Organix. The molar composition used was 35:16:46.5:2.5 unless otherwise noted. The lipid solution was mixed by flowing through a microfluidic channel³⁶ with an aqueous mRNA solution containing 10 mM citrate buffer (pH 3) at a 1:3 organic to aqueous volume ratio. The resulting LNPs were then dialysed against PBS (pH 7.4) for 2 h at 4 °C with a buffer to sample volume ratio greater than 1,000. For all pooled analyses, LNPs were dialysed separately against PBS and pooled before dosing.

LNP characterization

The z-average diameter and polydispersity index (PDI) of LNPs were measured using dynamic light scattering (Zetasizer Nano ZS (Malvern Panalytical), DynaPro Plate Reader (Wyatt Technology)). Total mRNA content and encapsulation efficiency were determined by performing a modified Ribogreen (ThermoFisher) assay, as previously described³⁷.

Cell culture and transfections

HEK293T cells were cultured at 37 °C and 5% CO₂ in high glucose Dulbecco's Modified Eagle Medium with GlutaMAX supplement (ThermoFisher) containing 10% fetal bovine serum and 1% penicillin-streptomycin. For in vitro transfections, 3 × 10⁶ cells were plated without antibiotics in a T75 flask and allowed to adhere and grow for a period of 24 h, after which Lipofectamine RNAiMAX (ThermoFisher) lipoplexes containing 13.2 µg of a pool of peptide barcode mRNAs were added to each flask. Whole cell lysate was collected 24 h posttransfection and processed as described below.

Animal experiments

Animal experiments for this study were approved by the Massachusetts Institute of Technology Institutional Animal Care and Use Committee

and were consistent with local, state and federal regulations as applicable. LNPs were administered to female 6–8-week-old C57BL/6 mice (Charles River laboratories) intravenously via the tail vein. For barcoding experiments, livers were collected and snap-frozen in liquid nitrogen 6 h postinjection and processed as described below. For hEPO experiments, blood was collected via tail-vein nicking into a clotting activator lined Microvette 100 Z tube (Sarstedt) 6 h postinjection. For FLuc experiments, whole organs were resected from mice at the indicated time point postinjection and imaged using an IVIS imaging system (PerkinElmer). Mice were euthanized through carbon dioxide asphyxiation.

Extraction of peptide barcodes from lysate

Cultured cells were collected 24 h posttransfection and lysed using RIPA lysis and extraction buffer (ThermoFisher) to obtain whole cell lysates. Snap-frozen livers were thawed on ice, and 150 mg of each were weighed out into 2 ml of soft tissue homogenizing CK14 tubes (Bertin). Then, 1.5 ml of Triton X-100 lysis buffer (150 mM NaCl, 1.0% Triton X-100, 50 mM Tris) were added to each tube, and livers were homogenized using a Precellys 24 tissue homogenizer (Bertin) to obtain liver lysates. All lysis buffers were supplemented with protease inhibitors by diluting Protease Inhibitor Cocktail (100×) (Cell Signaling Technology) to 1× in the appropriate lysis buffer. Cell and liver lysates were then precleared with 100 µl of G-Sep Agarose CL-6B beads (G-Biosciences) for 1 h and subsequently incubated with 100 µl of Immobilized Biotin Resin (G-Biosciences) for 3 h at room temperature. Beads containing the immobilized protein were then washed and resuspended in a 50 µl solution containing TEV protease (New England Biolabs) and incubated overnight at 4 °C. The resulting bead solution was then filtered using 0.65 µm Ultrafree-MC centrifugal filters (Millipore Sigma) and the filtrate was run on LC–MS/MS.

LC–MS/MS for peptide barcode analysis

Here, 1 µl of the peptide sample was loaded with an autosampler directly onto a 50 cm EASY-Spray C18 column (Thermo Scientific). Peptides were eluted from the column using a Dionex Ultimate 3000 Nano LC system with a 5 min gradient from 1% buffer B to 7% buffer B (100% acetonitrile, 0.1% formic acid), followed by a 29.8 min gradient to 25% and a 10.2 min gradient to 36% B, followed by a 0.5 min gradient to 80% B and held constant for 4.5 min. Finally, the gradient was changed from 80% buffer B to 99% buffer A (100% water, 0.1% formic acid) over 0.1 min and then held constant at 99% buffer A for another 19.9 min. The application of a 2.2 kV distal voltage electrospayed the eluting peptides directly into the Thermo Exploris480 mass spectrometer equipped with an EASY-Spray source (Thermo Scientific). Mass spectrometer-scanning functions and high-performance liquid chromatography gradients were controlled by the Xcalibur data system (Thermo Scientific). MS1 scan parameters were 60,000 resolution, scan range m/z 375–1,600, AGC at 300%, injection time (IT) at 50 ms. DIA scans were 33 2-Da windows from m/z 605–672 with 0.5 Da overlap and parameters of 15,000 resolution, HCD collision energy at 28%, AGC target at 200% and IT set to auto.

Statistics

All statistical analyses were performed using an unpaired, two-tailed Student's *t*-test (when comparing only two groups) or a one-way analysis of variance (ANOVA) with Dunnett's correction (when comparing three or more groups) in Graphpad Prism 9. In all figures, statistical significance is denoted by asterisks when $P < 0.001$, where *, **, *** and **** denote $P < 0.05$, $P < 0.01$, $P < 0.001$ and $P < 0.0001$, respectively.

Reporting summary

Further information on research design is available in the Nature Portfolio Reporting Summary linked to this article.

Data availability

The main data supporting the results in this study are available within the paper and its Supplementary Information. All data generated or analysed during the study are available from the corresponding author on reasonable request.

References

- Yin, H. et al. Non-viral vectors for gene-based therapy. *Nat. Rev. Genet.* **15**, 541–555 (2014).
- Kauffman, K. J., Webber, M. J. & Anderson, D. G. Materials for non-viral intracellular delivery of messenger RNA therapeutics. *J. Control. Release* **240**, 227–234 (2016).
- Li, B., Zhang, X. & Dong, Y. Nanoscale platforms for messenger RNA delivery. *Wiley Interdiscip. Rev. Nanomed. Nanobiotechnol.* **11**, e1530 (2019).
- Kowalski, P. S., Rudra, A., Miao, L. & Anderson, D. G. Delivering the messenger: advances in technologies for therapeutic mRNA delivery. *Mol. Ther.* **27**, 710–728 (2019).
- Rhym, L. H. & Anderson, D. G. Nanoscale delivery platforms for RNA therapeutics: challenges and the current state of the art. *Med* **3**, 167–187 (2022).
- DeRosa, F. et al. Therapeutic efficacy in a hemophilia B model using a biosynthetic mRNA liver depot system. *Gene Ther.* **23**, 699–707 (2016).
- Pardi, N. et al. Zika virus protection by a single low-dose nucleoside-modified mRNA vaccination. *Nature* **543**, 248–251 (2017).
- Miao, L. et al. Delivery of mRNA vaccines with heterocyclic lipids increases anti-tumor efficacy by STING-mediated immune cell activation. *Nat. Biotechnol.* **37**, 1174–1185 (2019).
- Yin, H. et al. Therapeutic genome editing by combined viral and non-viral delivery of CRISPR system components in vivo. *Nat. Biotechnol.* **34**, 328–333 (2016).
- Miller, J. B. et al. Non-Viral CRISPR/Cas gene editing in vitro and in vivo enabled by synthetic nanoparticle co-delivery of Cas9 mRNA and sgRNA. *Angew. Chem. Int. Ed.* <https://doi.org/10.1002/anie.201610209> (2017).
- Yin, H. et al. Structure-guided chemical modification of guide RNA enables potent non-viral in vivo genome editing. *Nat. Biotechnol.* **35**, 1179–1187 (2017).
- Finn, J. D. et al. A single administration of CRISPR/Cas9 lipid nanoparticles achieves robust and persistent in vivo genome editing. *Cell Rep.* **22**, 2227–2235 (2018).
- Polack, F. P. et al. Safety and efficacy of the BNT162b2 mRNA Covid-19 vaccine. *N. Engl. J. Med.* **383**, 2603–2615 (2020).
- Baden, L. R. et al. Efficacy and safety of the mRNA-1273 SARS-CoV-2 vaccine. *N. Engl. J. Med.* **384**, 403–416 (2021).
- Whitehead, K. A. et al. In vitro - in vivo translation of lipid nanoparticles for hepatocellular siRNA delivery. *ACS Nano* **6**, 6922–6929 (2012).
- Paunovska, K. et al. A direct comparison of in vitro and in vivo nucleic acid delivery mediated by hundreds of nanoparticles reveals a weak correlation. *Nano Lett.* **18**, 2148–2157 (2018).
- Dahlman, J. E. et al. Barcoded nanoparticles for high throughput in vivo discovery of targeted therapeutics. *Proc. Natl Acad. Sci. USA* **114**, 2060–2065 (2017).
- Guimaraes, P. P. et al. Ionizable lipid nanoparticles encapsulating barcoded mRNA for accelerated in vivo delivery screening. *J. Control. Release* **316**, 404–417 (2019).
- Fenton, O. S. et al. Synthesis and biological evaluation of ionizable lipid materials for the in vivo delivery of messenger RNA to B lymphocytes. *Adv. Mater.* **29**, 1606944 (2017).
- Fenton, O. S. et al. Customizable lipid nanoparticle materials for the delivery of siRNAs and mRNAs. *Angew. Chem. Int. Ed.* **57**, 13582–13586 (2018).

21. Sago, C. D. et al. High-throughput in vivo screen of functional mRNA delivery identifies nanoparticles for endothelial cell gene editing. *Proc. Natl Acad. Sci. USA* **115**, E9944–E9952 (2018).
22. Kauffman, K. J. et al. Rapid, single-cell analysis and discovery of vectored mRNA transfection in vivo with a loxP-flanked tdTomato reporter mouse. *Mol. Ther. Nucl. Acids* **10**, 55–63 (2018).
23. Gilleron, J. et al. Image-based analysis of lipid nanoparticle-mediated siRNA delivery, intracellular trafficking and endosomal escape. *Nat. Biotechnol.* **31**, 638–646 (2013).
24. Vermeulen, L. M. P. et al. Endosomal size and membrane leakiness influence proton sponge-based rupture of endosomal vesicles. *ACS Nano* **12**, 2332–2345 (2018).
25. Demonte, D., Drake, E. J., Lim, K. H., Gulick, A. M. & Park, S. Structure-based engineering of streptavidin monomer with a reduced biotin dissociation rate. *Proteins* **81**, 1621–1633 (2013).
26. Lim, K. H., Huang, H., Pralle, A. & Park, S. Stable, high-affinity streptavidin monomer for protein labeling and monovalent biotin detection. *Biotechnol. Bioeng.* **110**, 57–67 (2013).
27. Kauffman, K. J. et al. Optimization of lipid nanoparticle formulations for mRNA delivery in vivo with fractional factorial and definitive screening designs. *Nano Lett.* **15**, 7300–7306 (2015).
28. Hassett, K. J. et al. Optimization of lipid nanoparticles for Intramuscular administration of mRNA vaccines. *Mol. Ther. Nucleic Acids* **15**, 1–11 (2019).
29. Sabnis, S. et al. A novel amino lipid series for mRNA delivery: improved endosomal escape and sustained pharmacology and safety in non-human primates. *Mol. Ther.* **26**, 1509–1519 (2018).
30. Hajj, K. A. et al. Branched-tail lipid nanoparticles potently deliver mRNA in vivo due to enhanced ionization at endosomal pH. *Small* **15**, e1805097 (2019).
31. Rui, Z. et al. Helper lipid structure influences protein adsorption and delivery of lipid nanoparticles to spleen and liver. *Biomater. Sci.* **9**, 1449–1463 (2021).
32. Lokugamage, M. P., Sago, C. D., Gan, Z., Krupczak, B. R. & Dahlman, J. E. Constrained nanoparticles deliver siRNA and sgRNA to T cells in vivo without targeting ligands. *Adv. Mater.* **31**, 1902251 (2019).
33. Sago, C. D. et al. Nanoparticles that deliver RNA to bone marrow identified by in vivo directed evolution. *J. Am. Chem. Soc.* **140**, 17095–17105 (2018).
34. Miao, L. et al. Synergistic lipid compositions for albumin receptor mediated delivery of mRNA to the liver. *Nat. Commun.* **11**, 2424 (2020).
35. Cheng, Q. et al. Selective organ targeting (SORT) nanoparticles for tissue-specific mRNA delivery and CRISPR-Cas gene editing. *Nat. Nanotechnol.* **15**, 313–320 (2020).
36. Chen, D. et al. Rapid discovery of potent siRNA-containing lipid nanoparticles enabled by controlled microfluidic formulation. *J. Am. Chem. Soc.* **134**, 6948–6951 (2012).
37. Heyes, J., Palmer, L., Bremner, K. & MacLachlan, I. Cationic lipid saturation influences intracellular delivery of encapsulated nucleic acids. *J. Control. Release* **107**, 276–287 (2005).

Acknowledgements

We acknowledge project funding from Translate Bio (Lexington, MA, USA) and the Marble Center for Cancer Nanomedicine, as well as support from the Cancer Center Support (core) (grant no. P30-CA14051) from the National Cancer Institute. We thank the Koch Institute Swanson Biotechnology Center for technical support, specifically the Biopolymers & Proteomics core and Preclinical Imaging and Testing core.

Author contributions

L.H.R. developed the peptide barcoding concept. L.H.R., R.S.M. and D.G.A. conceived the study and designed experiments. L.H.R. and G.S. performed all in vitro and in vivo barcoding, hEPO and FLuc assays, including LNP formulation and animal work. A.K. developed and performed the LC–MS/MS method for all samples. R.S.M. designed and synthesized the combinatorial ionizable lipid library screened in this study. L.H.R., R.S.M. and D.G.A. contributed to the analysis and interpretation of the results and to the writing of the paper.

Competing interests

Three of the authors (L.H.R., R.S.M. and D.G.A.) have filed a patent (US provisional application 63/289,343) on the technology described in this manuscript. The remaining authors declare no competing interests.

Additional information

Extended data is available for this paper at <https://doi.org/10.1038/s41551-023-01030-4>.

Supplementary information The online version contains supplementary material available at <https://doi.org/10.1038/s41551-023-01030-4>.

Correspondence and requests for materials should be addressed to Daniel G. Anderson.

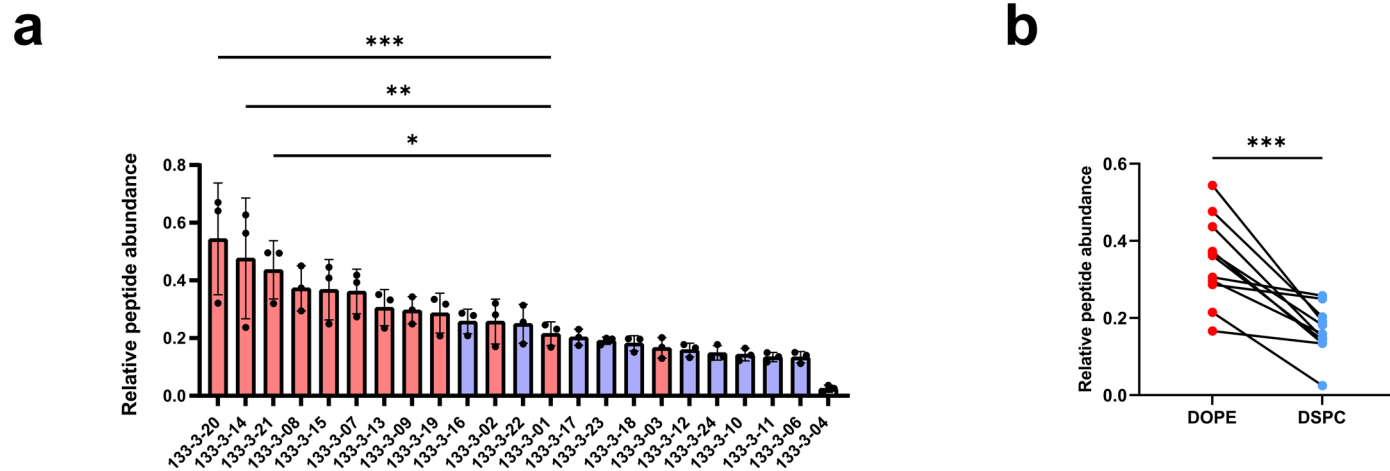
Peer review information *Nature Biomedical Engineering* thanks Xiangrong Song and the other, anonymous, reviewer(s) for their contribution to the peer review of this work.

Reprints and permissions information is available at www.nature.com/reprints.

Publisher's note Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.

Springer Nature or its licensor (e.g. a society or other partner) holds exclusive rights to this article under a publishing agreement with the author(s) or other rightsholder(s); author self-archiving of the accepted manuscript version of this article is solely governed by the terms of such publishing agreement and applicable law.

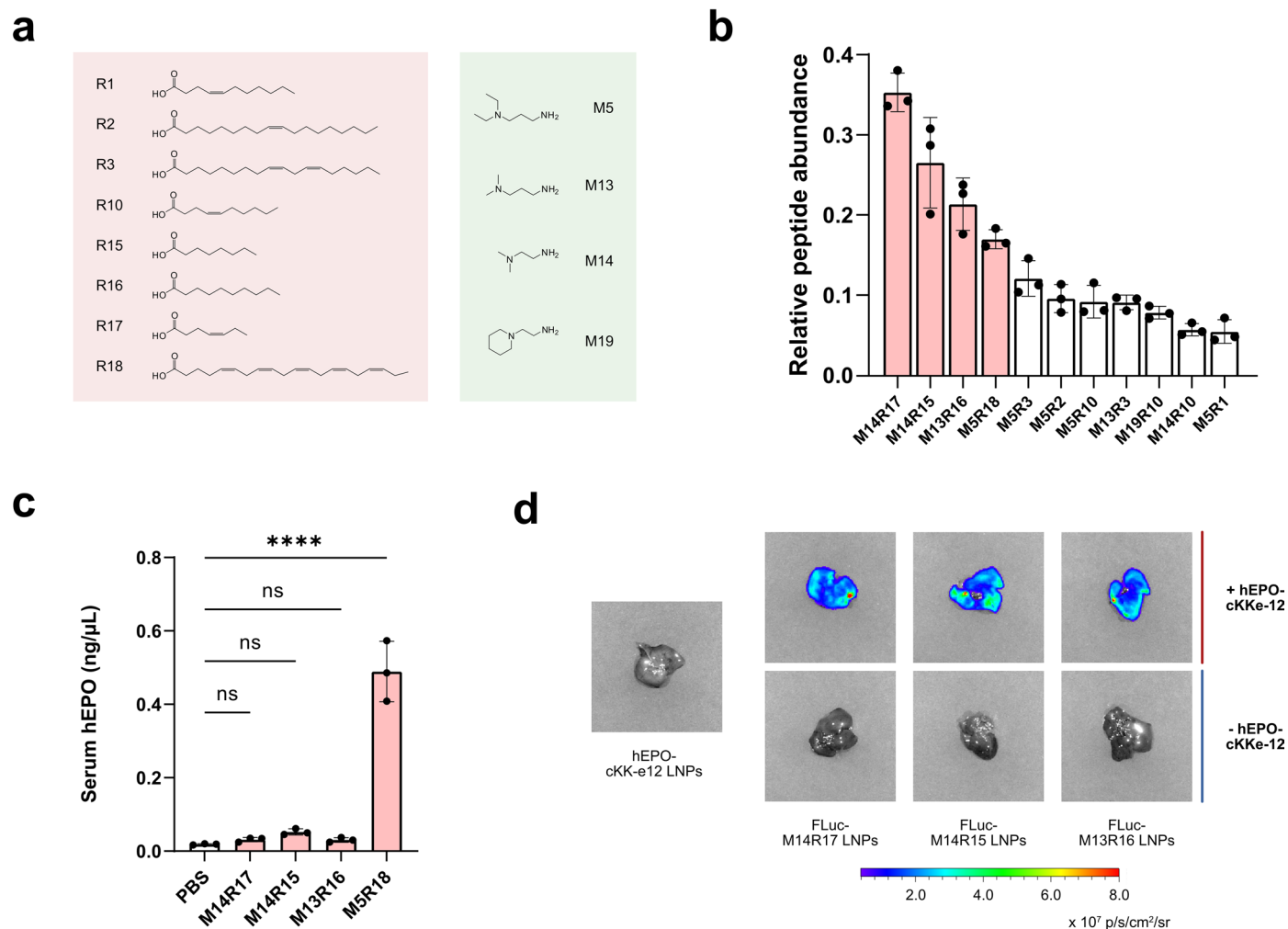
© The Author(s), under exclusive licence to Springer Nature Limited 2023



Extended Data Fig. 1 | Comparison of formulations containing DOPE or DSPC.

a Data of Fig. 4 replotted with DOPE formulations highlighted in red and DSPC formulations highlighted in blue. **b** Comparison of LNPs with identical molar

compositions, differing only in phospholipid identity. Values shown are mean of $n = 3$ mice/group, with error bars representing the SD. * $p < 0.05$, ** $p < 0.005$, *** $p < 0.0005$.



Extended Data Fig. 2 | Pilot screen using a small subset of the combinatorial library. **a** A list of esters (left) and amines (right) used to generate lipids for the pilot screen. **b** Peptide barcoding analysis of the pilot library. Only lipids that resulted in a peptide dot product > 0.8 are shown. **c** Individual particle analysis using the top four hits, showing that the top three lipids from the peptide barcoding analysis do not result in significant protein production. **d** FLuc assay

confirming that FLuc mRNA-containing LNPs composed of any of the three ‘false positives’ alone do not result in detectable protein expression, whereas a 1:1 mixture of hEPO mRNA-containing cKKe-12 LNPs and FLuc mRNA-containing ‘false positive’ LNPs results in robust protein expression. Values shown are the mean of $n = 3$ mice/group, with error bars representing the SD. The FLuc assay of panel **d** was performed twice with similar results.

Reporting Summary

Nature Portfolio wishes to improve the reproducibility of the work that we publish. This form provides structure for consistency and transparency in reporting. For further information on Nature Portfolio policies, see our [Editorial Policies](#) and the [Editorial Policy Checklist](#).

Statistics

For all statistical analyses, confirm that the following items are present in the figure legend, table legend, main text, or Methods section.

- | n/a | Confirmed |
|-------------------------------------|--|
| <input type="checkbox"/> | <input checked="" type="checkbox"/> The exact sample size (n) for each experimental group/condition, given as a discrete number and unit of measurement |
| <input type="checkbox"/> | <input checked="" type="checkbox"/> A statement on whether measurements were taken from distinct samples or whether the same sample was measured repeatedly |
| <input type="checkbox"/> | <input checked="" type="checkbox"/> The statistical test(s) used AND whether they are one- or two-sided
<i>Only common tests should be described solely by name; describe more complex techniques in the Methods section.</i> |
| <input checked="" type="checkbox"/> | <input type="checkbox"/> A description of all covariates tested |
| <input type="checkbox"/> | <input checked="" type="checkbox"/> A description of any assumptions or corrections, such as tests of normality and adjustment for multiple comparisons |
| <input type="checkbox"/> | <input checked="" type="checkbox"/> A full description of the statistical parameters including central tendency (e.g. means) or other basic estimates (e.g. regression coefficient) AND variation (e.g. standard deviation) or associated estimates of uncertainty (e.g. confidence intervals) |
| <input type="checkbox"/> | <input checked="" type="checkbox"/> For null hypothesis testing, the test statistic (e.g. F , t , r) with confidence intervals, effect sizes, degrees of freedom and P value noted
<i>Give P values as exact values whenever suitable.</i> |
| <input checked="" type="checkbox"/> | <input type="checkbox"/> For Bayesian analysis, information on the choice of priors and Markov chain Monte Carlo settings |
| <input checked="" type="checkbox"/> | <input type="checkbox"/> For hierarchical and complex designs, identification of the appropriate level for tests and full reporting of outcomes |
| <input checked="" type="checkbox"/> | <input type="checkbox"/> Estimates of effect sizes (e.g. Cohen's d , Pearson's r), indicating how they were calculated |

Our web collection on [statistics for biologists](#) contains articles on many of the points above.

Software and code

Policy information about [availability of computer code](#)

Data collection Absorbance and fluorescence readouts on a microplate reader (such as for Ribogreen) were collected from a Tecan infinite 2000 PRO by using i-Control v1.8. Ex vivo imaging of mouse tissues was performed using an IVIS Spectrum In Vivo Imaging System from PerkinElmer.

Data analysis LC-MS/MS data were analysed using Skyline Version 21. IVIS image analysis was done using Living Image Software from PerkinElmer. All data plotting and statistical analyses were performed in Graphpad Prism Version 9.

For manuscripts utilizing custom algorithms or software that are central to the research but not yet described in published literature, software must be made available to editors and reviewers. We strongly encourage code deposition in a community repository (e.g. GitHub). See the Nature Portfolio [guidelines for submitting code & software](#) for further information.

Data

Policy information about [availability of data](#)

All manuscripts must include a [data availability statement](#). This statement should provide the following information, where applicable:

- Accession codes, unique identifiers, or web links for publicly available datasets
- A description of any restrictions on data availability
- For clinical datasets or third party data, please ensure that the statement adheres to our [policy](#)

The main data supporting the results in this study are available within the paper and its Supplementary Information. All data generated or analysed during the study are available from the corresponding author on reasonable request.

Human research participants

Policy information about [studies involving human research participants and Sex and Gender in Research](#).

Reporting on sex and gender	<input type="text" value="The study did not use human research participants."/>
Population characteristics	<input type="text" value="—"/>
Recruitment	<input type="text" value="—"/>
Ethics oversight	<input type="text" value="—"/>

Note that full information on the approval of the study protocol must also be provided in the manuscript.

Field-specific reporting

Please select the one below that is the best fit for your research. If you are not sure, read the appropriate sections before making your selection.

Life sciences Behavioural & social sciences Ecological, evolutionary & environmental sciences

For a reference copy of the document with all sections, see [nature.com/documents/nr-reporting-summary-flat.pdf](https://www.nature.com/documents/nr-reporting-summary-flat.pdf)

Life sciences study design

All studies must disclose on these points even when the disclosure is negative.

Sample size	<input type="text" value="Sample sizes were calculated via a power analysis designed to reach p-values lower than 0.05 on the basis of predicted differences between groups and of estimates of error derived from previous experience."/>
Data exclusions	<input type="text" value="No data points were excluded."/>
Replication	<input type="text" value="Experiments were, in general, repeated once more to confirm the results."/>
Randomization	<input type="text" value="Mouse cages were randomly allocated to the experimental groups."/>
Blinding	<input type="text" value="The experiments were not blinded because all data were quantitatively measured with standard equipment."/>

Reporting for specific materials, systems and methods

We require information from authors about some types of materials, experimental systems and methods used in many studies. Here, indicate whether each material, system or method listed is relevant to your study. If you are not sure if a list item applies to your research, read the appropriate section before selecting a response.

Materials & experimental systems

n/a	Involvement in the study
<input checked="" type="checkbox"/>	<input type="checkbox"/> Antibodies
<input type="checkbox"/>	<input checked="" type="checkbox"/> Eukaryotic cell lines
<input checked="" type="checkbox"/>	<input type="checkbox"/> Palaeontology and archaeology
<input type="checkbox"/>	<input checked="" type="checkbox"/> Animals and other organisms
<input checked="" type="checkbox"/>	<input type="checkbox"/> Clinical data
<input checked="" type="checkbox"/>	<input type="checkbox"/> Dual use research of concern

Methods

n/a	Involvement in the study
<input checked="" type="checkbox"/>	<input type="checkbox"/> ChIP-seq
<input checked="" type="checkbox"/>	<input type="checkbox"/> Flow cytometry
<input checked="" type="checkbox"/>	<input type="checkbox"/> MRI-based neuroimaging

Eukaryotic cell lines

Policy information about [cell lines and Sex and Gender in Research](#)

Cell line source(s)	<input type="text" value="HEK293T cells (CRL-3216) were purchased from ATCC."/>
Authentication	<input type="text" value="Authentication was provided by the manufacturer. Cells were further authenticated in the laboratory by analysing their morphology and growth rate."/>

Mycoplasma contamination

Mycoplasma-free cells were purchased.

Commonly misidentified lines
(See [ICLAC](#) register)

No commonly misidentified cell lines were used.

Animals and other research organisms

Policy information about [studies involving animals](#); [ARRIVE guidelines](#) recommended for reporting animal research, and [Sex and Gender in Research](#)

Laboratory animals

6–8-week old female C57BL/6 (Stock No: 027) mice were obtained from Charles River Laboratory.

Wild animals

The study did not involve wild animals.

Reporting on sex

Only female mice were used.

Field-collected samples

The study did not involve samples collected from the field.

Ethics oversight

All animal procedures were approved by the MIT Committee on Animal Care, and supervised by veterinary staff from the MIT Division of Comparative Medicine.

Note that full information on the approval of the study protocol must also be provided in the manuscript.