Journal of Controlled Release 394 (2026) 114911

Contents lists available at ScienceDirect

Journal of Controlled Release

journal homepage: www.elsevier.com/locate/jconrel

Review article

# A multiscale approach to lipid nanoparticle engineering from molecular structure to *in vivo* performance

Benjamin Winkeljann [a,b,c,d,e], Philipp Lapuhs [f,g,h], María J. Alonso [f,g,h], Ceren Kimna [e,i,*]

[a] Department of Pharmacy, Ludwig-Maximilians-Universität München, 81377 Munich, Germany
[b] Comprehensive Pneumology Center (CPC-M), Member of the German Center for Lung Research (DZL), 81377 Munich, Germany
[c] RNhale GmbH, 82152 Planegg, Germany
[d] Center for NanoScience (CeNS), Ludwig-Maximilians-Universität München, 80799 Munich, Germany
[e] Pharmaceutical Engineering and Technology Research Scientists (PETRS), Germany
[f] Center for Research in Molecular Medicine and Chronic Diseases (CiMUS), Campus Vida, Universidade de Santiago de Compostela, Santiago de Compostela 15782, Spain
[g] Department of Pharmacology, Pharmacy and Pharmaceutical Technology, Campus Vida, Universidade de Santiago de Compostela, Santiago de Compostela 15782, Spain
[h] Health Research Institute of Santiago de Compostela (IDIS), Santiago de Compostela 15706, Spain
[i] Institute for Intelligent Biotechnologies (iBIO), Helmholtz Center Munich, 85764 Neuherberg, Germany

ARTICLE INFO

*Keywords:*
Molecular dynamics simulations
Library screening
Molecular barcoding
Protein corona

ABSTRACT

The 2023 Nobel Prize in Physiology or Medicine recognized the discovery of base modifications that enabled effective mRNA vaccine development. Highlighting this, the former president of the Controlled Release Society, Twan Lammers, at the society's 2024 annual meeting, noted that it may soon be time for this recognition to extend to breakthroughs in drug delivery systems. Indeed, it was modern delivery technologies that unlocked the potential of RNA-based medicines, with lipid nanoparticles (LNPs) enabling their clinical translation less than a decade ago. While LNPs have proven essential for many nucleic acid delivery applications, their rational design remains challenged by complex lipid–RNA interactions, biological barriers, and the dynamic nano–bio interface. In this review, we highlight state-of-the-art strategies in LNP development, including computationally guided design, high-throughput screening, and insights into how protein corona formation shapes *in vivo* performance.

## 1. Introduction

The rise of nucleic acid (NA)-based therapies expanded the scope of modern medicine by enabling precise modulation of gene expression. However, the therapeutic potential of RNA and DNA is fundamentally limited by their instability, susceptibility to enzymatic degradation, and inability to cross cell membranes [1,2]. Effective delivery systems are therefore essential to protect NAs in physiological environments and enable their intracellular release. While current clinical use predominantly involves RNA, the fundamental principles of LNP design extend to other nucleic acid modalities, including DNA and gene editing components.

Lipid nanoparticles (LNPs) were developed to meet these requirements and have established themselves as a versatile platform for RNA delivery [3]. Viral vectors, although widely used, suffer from limitations such as restricted loading capacity, immunogenicity, and challenges with repeated dosing. Conjugation strategies offer an alternative but are mostly restricted to short RNAs such as siRNA or antisense oligonucleotides [4]. In contrast, LNPs can accommodate a broad range of cargo types, including mRNA, siRNA, and genomic DNA. Their modularity, scalability, and clinical performance have made them the most advanced non-viral delivery system [5].

To date, four LNP-based therapeutics have been approved by the U.S. Food and Drug Administration (FDA), including the siRNA formulation patisiran [6] (Onpattro, Alnylam, 2018), the mRNA COVID-19 vaccines BNT162b2 [7] (Comirnaty, BioNTech-Pfizer, 2020, including the Omicron BA.4/BA.5-adapted bivalent booster vaccine, 2022) and mRNA-1273 [8] (Spikevax, Moderna, 2020), as well as the mRNA RSV vaccine mRESVIA (Moderna, 2024). Additional LNP-formulated vaccines have received approval in the EU, China, Japan, and Indonesia [9].

* Corresponding author at: Institute for Intelligent Biotechnologies (iBIO), Helmholtz Center Munich, 85764 Neuherberg, Germany, Ingolstädter Landstr.
*E-mail address:* ceren.kimna@helmholtz-munich.de (C. Kimna).

https://doi.org/10.1016/j.jconrel.2026.114911
Received 31 July 2025; Received in revised form 14 December 2025; Accepted 6 April 2026
Available online 7 April 2026

Emerging candidates illustrate the expansion of LNP applications beyond vaccines for infectious diseases and into areas such as rare genetic disorders or oncology. Nexiguran ziclumeran (NTLA-2001), an *in vivo* CRISPR–Cas9 gene editing therapy developed by Intellia Therapeutics and Regeneron for transthyretin (TTR) amyloidosis, is currently recruiting for Phase 3 clinical trials (NCT06672237). ReCode Therapeutics' RCT2100, an inhaled mRNA–LNP candidate for cystic fibrosis, has entered Phase 1b (NCT06237335) and received orphan designation. ARCT-810 (Arcturus Therapeutics) targets ornithine transcarbamylase deficiency, received FDA Orphan Drug designation in July 2019, and has advanced into Phase 2 (NCT06488313). In oncology, two mRNA–LNP-based cancer vaccines targeting melanoma-associated antigens are in advanced clinical development. Intismeran autogene (mRNA-4157, Moderna and Merck) is being evaluated in a Phase 3 trial (NCT05933577) as adjuvant therapy following complete resection of high-risk stage III/IV melanoma while BNT111 (BioNTech) is currently in Phase 2 trials (NCT04526899) for unresectable stage III or IV melanoma. Overall, the RNA therapy pipeline is rapidly expanding, with mRNA-based therapeutics alone accounting for over 150 candidates in clinical development as of Q1 2025 [9], many of which utilize LNP technology. This momentum highlights both the translational success of LNP platforms and the opportunity to review their evolving applications in clinical settings.

Modern LNP formulation pipelines typically integrate both bottom-up and top-down strategies. In general, a bottom-up approach refers to the *de novo* design of LNPs by systematically varying lipid components, structures, or ratios, often enabled by high-throughput synthesis and screening. These explorations are guided by insights into lipid chemistry and self-assembly mechanisms, which help define an initial formulation landscape. In contrast, a top-down approach involves the iterative refinement of a lead formulation within a predefined design space, *e.g.*, by adjusting buffer conditions, PEG-lipid content, lipid ratios, or mixing parameters. Such strategies are typically employed during lead optimization and scale-up phases, where the focus shifts to improving pharmacokinetics, manufacturability, and stability, while preserving previously validated *in vivo* efficacy. Importantly, the fundamental design principles of LNPs, such as requirement for ionizable lipids with an appropriate pKa, the use of helper and PEG-lipids for colloidal stability are broadly applicable across different classes of nucleic acids. However, cargo specific factors such as nucleic acid length, desired intracellular location, intended therapeutic application may require distinct optimization within the same carrier design.

Despite many advances, the current landscape of LNP development remains fragmented between empirical high-throughput screening and mechanistic understanding. Both academia and industry have increasingly resorted to large-scale *in vivo* screenings, involving extensive animal studies to identify promising formulations. While this approach has accelerated candidate discovery, it raises fundamental questions about whether such volume-based strategies overlook deeper physicochemical and biological principles that govern LNP behavior. Beyond concerns about efficiency, the widespread use of animals in screening has raised ethical issues that have gained urgency as global initiatives push to reduce reliance on animal-based research. These concerns call for a critical reassessment of current practices and a shift toward uncovering the molecular mechanisms that enable truly rational design.

Fast screening methodologies typically trade resolution and mechanistic insight for speed. This often generates data sets that are difficult to interpret within the context of nanocarrier structure–function relationships. Many classical characterization techniques fall short of capturing the complexity of dynamic lipid assemblies especially when interacting within biological milieus. Moreover, the overreliance on functional endpoints such as gene silencing, protein expression and efficacy fosters overlooking critical physicochemical factors that govern nanoparticle (NP) behavior, leading to suboptimal or non-predictive design criteria based on correlation rather than causation. Nonetheless, recent advances have seen the adaptation of cutting-edge experimental tools from related fields to drug delivery research. These include highly multiplexed molecular barcoding strategies enabling simultaneous *in vivo* evaluation of extensive LNP libraries [10] as well as high-resolution imaging [11] and proteomic profiling [12] to dissect nano–bio interactions at unprecedented depth.

Complementing experimental workflows with computational approaches has emerged as a powerful strategy in overcoming some of the limitations of traditional trial-and-error development. Multiscale molecular dynamics simulations allow for exploration of lipid self-assembly and NP structure at both atomistic and mesoscale levels. Continuum-scale methods, such as computational fluid dynamics (CFD), can provide additional tools for linking molecular properties to macroscopic behavior relevant for manufacturing [13]. Together, such *in silico* approaches promise to bridge gaps between molecular design principles and biological performance and may facilitate more targeted and rational optimization strategies.

While previous work has addressed strategies for optimizing lipid nanocarriers and accelerating clinical translation of nucleic acid therapeutics [14], this review highlights recent advances in LNP design that move the field beyond empirical screening toward a more mechanistic understanding. We first discuss multiscale simulation approaches that link molecular composition to biological performance metrics and manufacturability. We move on to high-throughput *in vivo* evaluation techniques based on molecular barcoding, combined with omics tools that reveal detailed nano–bio interactions. Finally, we explore the protein corona as a challenge but also as an opportunity to be used as a design feature for modulating biodistribution and immunogenicity. Each of these domains contribute to a broader goal: to ensure that every experiment – whether computational or *in vivo* – generates insights that inform the others. This interconnected approach builds a foundation for rational design and accelerates the collective process in the development of NA therapeutics.

## 2. Rational and library-based design of lipid nanoparticles

The design space for a bottom-up approach in LNP formulation is fundamentally dictated by lipid chemistry. This involves the strategic modulation of key structural components, including ionizable lipids, phospholipids, cholesterol, and polyethylene glycol (PEG)-conjugated lipids, to achieve optimal drug encapsulation, stability, and delivery efficiency. The physicochemical properties of these lipids, such as ionization behavior, hydrophobicity, and membrane interactions, govern the NP size, charge, biodistribution, and endosomal escape efficiency, all of which are critical for therapeutic efficacy. This predefined lipid framework helps ensure that key requirements for LNP formulations, such as low toxicity, colloidal stability, or scalability of the manufacturing process are met. Furthermore, it facilitates a more rapid regulatory approval process, as the overall concept is already well understood by regulatory agencies. Consequently, the necessary analytical assessments and regulatory standards required for clearance are well established. On the flipside however, it must be noted that reliance on these four core lipid classes imposes significant constraints on the design space, thereby limiting the identification of formulations with characteristics dramatically different from those observed to date.

Modern bottom-up methodology increasingly integrates computational modeling, structure-function relationship (SFR) analysis, and lipid libraries to streamline the development process. Computational tools can predict the self-assembly behavior, and delivery efficiency of novel lipids, reducing time and resource demands [15]. High-throughput lipid screening platforms and ML algorithms further accelerate the identification of optimized formulations [16]. These advancements can substantially enhance the efficiency of bottom-up strategies and therefore might drive the rational design of next-generation LNPs.

### 2.1. Simulation-assisted design of lipid nanoparticles

To overcome the inherent constraints of lipid selection in LNP formulation, computational modeling has become an essential tool in guiding rational design. By predicting self-assembly behavior, stability, and molecular interactions, computational approaches complement experimental methods and may accelerate the optimization of LNP formulations. These methods span multiple scales, from molecular-level simulations that capture lipid-lipid and lipid-RNA interactions to continuum models that describe macroscopic properties such as phase behavior and mechanical stability. The most widely applied techniques in this context are molecular dynamics (MD) simulations and continuum methods. MD simulations solve Newton's equations of motion for each interaction site, typically ranging from individual atoms to coarse-grained beads depending on the desired resolution, to model structural organization and dynamic interactions over time. Continuum approaches, such as computational fluid dynamics (CFD), apply conservation laws of mass, momentum, and energy to describe multiphase flow and mixing phenomena. While MD simulations enable detailed modeling of molecular interactions at different resolutions, continuum methods offer a broader perspective on bulk properties and large-scale structural organization. Other computational techniques, such as density functional theory (DFT) and kinetic Monte Carlo simulations, exist but have seen limited application in LNP research. DFT, commonly used in quantum chemistry, could provide insights into lipid-electrolyte interactions at an atomic level and has been used to estimate the apparent pKa of LNPs [17], but is impractical for simulating large biological systems. Similarly, kinetic Monte Carlo methods, which excel at modeling reaction kinetics and stochastic processes, have not been widely explored for LNP formation due to the complex nature of lipid self-assembly.

#### 2.1.1. Meso- and molecular level simulations

MD simulations are among the most widely used computational techniques to study the interactions and behavior of molecules in LNP systems. These simulations are based on solving Newton's equations of motion, in which the force acting on each atom is obtained from the gradient of the system's potential energy:

$$m_i \frac{d\mathbf{r}_i}{dt} = \mathbf{F}_i = -\nabla_{\mathbf{r}_i} U(\mathbf{r})$$

The total potential energy $U(\mathrm{r})$ is commonly decomposed into bonded and nonbonded contributions:

$$U(\mathbf{r}) = U_{\mathrm{bonded}}(\mathbf{r}) + U_{\mathrm{non-bonded}}(\mathbf{r})$$

Bonded interactions include bond stretching, angle bending, and dihedral rotations, which are frequently described as:

$$U_{\mathrm{bonded}}(\mathbf{r}) = U_{\mathrm{bond}}(r_{ij}) + U_{\mathrm{angle}}(\theta_{ijk}) + U_{\mathrm{di}}(\phi_{ijkl})$$

$$= \frac{1}{2}k_{bond}(r - r_{bond})^2 + \frac{1}{2}k_{angle}(\theta - \theta_0)^2 + \frac{1}{2}k_{di}(1 + \cos[n\phi - \phi_0])^2$$

Where $k_{\mathrm{bond}}$ is the bond force constant, $k_{\mathrm{angle}}$ is the angle force constant, $k_{\mathrm{di}}$ is the dihedral force constant, $r_0$, $\theta_0$, and $\phi_0$ are equilibrium values, and $n$ is the dihedral multiplicity. Nonbonded interactions consist of van der Waals interactions, typically modeled using a Lennard-Jones (12–6) potential, and electrostatic interactions described by a Coulombic term:

$$U_{\mathrm{non-bonded}}(\mathbf{r}) = U_{\mathrm{vdW}}(r_{ij}) + U_{\mathrm{coulomb}}(r_{ij})$$

$$= \sum_{i,j} 4\varepsilon_{ij}\left[\left(\frac{\sigma_{ij}}{r_{ij}}\right)^{12} - \left(\frac{\sigma_{ij}}{r_{ij}}\right)^{6}\right] + \sum_{i,j}\frac{q_i q_j}{4\pi\epsilon_0\epsilon_r r_{ij}}$$

Here, $\varepsilon_{ij}$ describes the depth of the Lennard-Jones potential well, $\sigma_{ij}$ is the distance at which the LJ potential crosses zero, $q_i$ and $q_j$ are atomic partial charges, $\varepsilon_0$ and $\varepsilon_r$ denote the vacuum and relative permittivity, respectively. Together, these bonded and nonbonded terms constitute the force field, a parameter matrix that encodes the interatomic interactions governing the behavior of molecular systems (**Fig. 1 A**).

In LNP research, two primary levels of resolution in MD simulations are commonly employed: all-atom (AA) and coarse-grained (CG) simulations (Fig. 1B), with even coarser approaches such as dissipative particle dynamics (DPD) and Langevin dynamics (LD) being applied in specific cases [18]. AA MD simulations offer a high-resolution representation of molecular systems, captures detailed atomic interactions including hydrogen bonding and π-stacking effects, which are particularly relevant when studying RNA-lipid interactions [19]. AA simulations have also been applied to study how different ionizable lipid morphologies, such as conical *versus* cylindrical structures, influence RNA complexation and LNP stability [20]. Additionally, AA simulations have provided atomic-level insights into the structural organization and pH-driven phase transitions of LNP formulations. Simulations revealed an oil-like LNP core at physiological pH, where aminolipids shield mRNA in inverted micellar structures. Upon acidification, a phase transition leads to lipid bilayer formation. Furthermore, the protonation state of aminolipids was shown to shift significantly depending on their localization, driving lipid rearrangement and influencing LNP stability and function [21,22]. Recent computational work utilizing the Umbrella Sampling technique has further refined our understanding of ionizable lipid protonation states by calculating apparent pKa values in LNP environments [23]. However, due to computational constraints, AA simulations are typically limited to small system sizes (on the order of tens of nanometers) and short time scales, usually in the nano- to microsecond range. These limitations make AA simulations suitable for investigating local molecular interactions but less practical for modeling large-scale structures such as fully assembled LNPs.

CG MD simulations, in contrast, offer an efficient means of modeling larger systems over extended time scales by simplifying molecular representations, typically mapping 2–5 heavy atoms to a single interaction site (Fig. 1B). This reduction in resolution significantly extends the accessible length and time scales, enabling the simulation of entire LNPs or even larger biological structures such as cellular organelles [24] Among the most widely used CG force fields in LNP research is the Martini force field, initially introduced in 2007 [25] and continuously updated, with the current version, Martini 3 [26], that extends to a library of already parameterized building blocks for lipid carrier design [27,28]. Martini-based simulations have been extensively applied in LNP studies, including investigations of self-assembly, fusion, and lipid rearrangement dynamics. However, despite these advantages, CG models inherently lose atomic-level detail, limiting their ability to capture fine structural features such as hydrogen bonding networks or localized charge distributions, which are crucial for accurately describing RNA-lipid interactions.

Often, researchers aim at utilizing computational models to predict experimentally verifiable properties, such as diffusion coefficients, NP size, and lipid-binding affinities. Key parameters such as the radius of gyration, binding distances, the Solvent Accessible Surface Area (SASA) and binding free energies (Fig. 1C) can be directly compared to experimental techniques, including dynamic light scattering (DLS) or nanoparticle tracking analysis (NTA), small-angle X-ray scattering (SAXS), isothermal titration calorimetry (ITC), and microscale thermophoresis (MST). Although experimental validation of CG studies is challenging, recent studies on non-lipid RNA carrier systems have demonstrated both qualitative and quantitative agreement between MD-derived energetic and structural parameters and corresponding experimental data [29,30].

It must be noted that CG force fields are everything but a one-fits-all system. It has been demonstrated for proteins that recalibrating Martini parameters can significantly improve quantitative accuracy when predicting experimental observables such as diffusion coefficients and molecular binding constants [31,32]. CG simulations have also been employed to investigate the structural organization and complexation

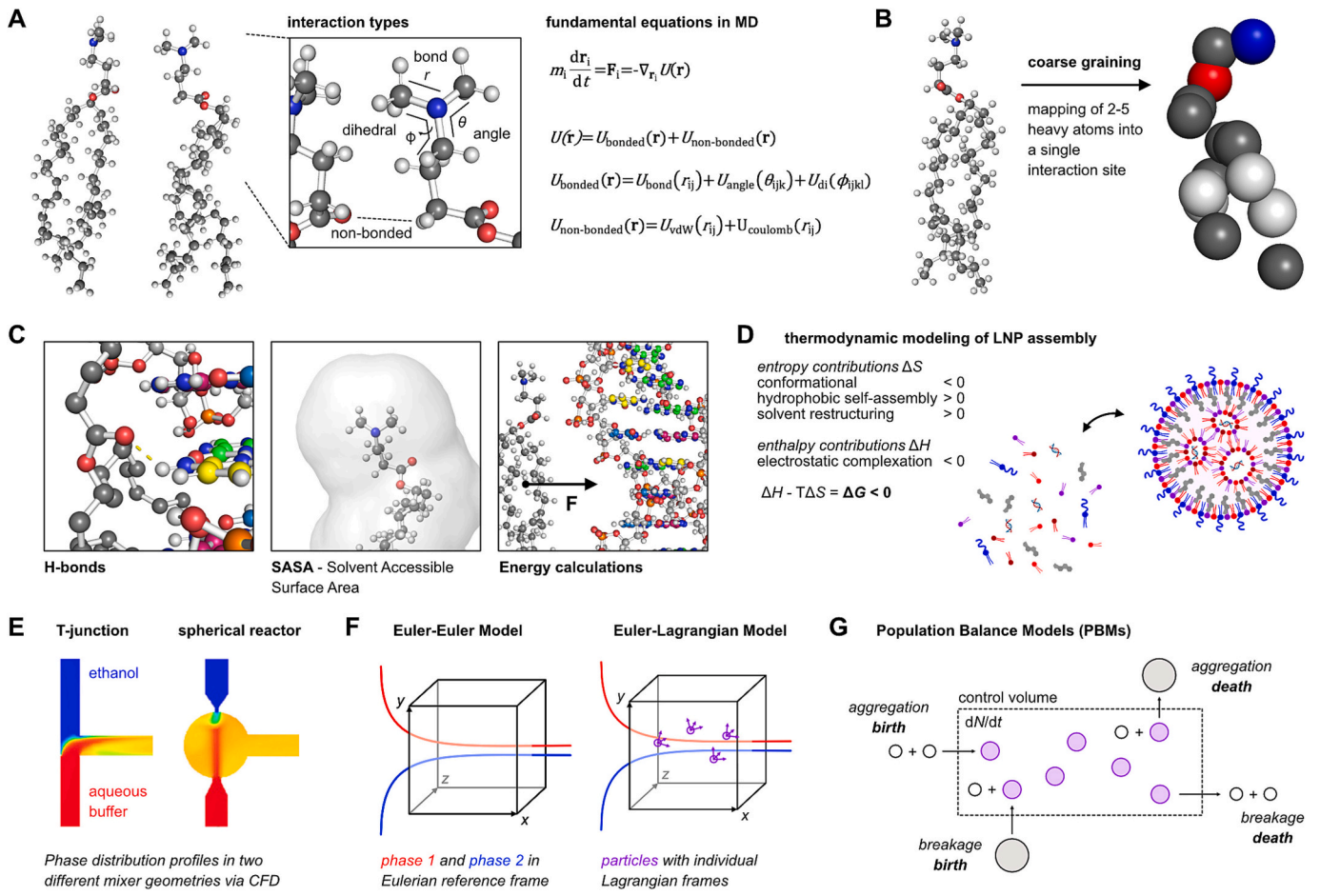


**Fig. 1.** Computational tools in LNP research. A) Fundamental equations used in molecular dynamics (MD) simulations, linking Newton's equations of motion with bonded and non-bonded interaction potentials. B) Coarse-graining of MD structures by mapping all-atom (AA) representations into reduced beads (typically 2–5 heavy atoms per bead) to enable longer timescale simulations. C) Representative readouts obtainable from MD simulations, including hydrogen-bond analysis, solvent-accessible surface area (SASA), and energy decomposition. D) Thermodynamic contributions relevant to LNP self-assembly. E) Computational fluid dynamics (CFD) simulations to determine phase distribution profiles in different reactor design. F) Modeling approaches in CFD, *i.e.*, Euler–Euler and Euler–Lagrange frameworks. G) Concept of population balance models (PBMs) for simulating particle nucleation, growth, aggregation, and breakage.

processes between ionizable lipids and mRNA, *e.g.* in the formation of structures such as blebs [18]. These structures were first observed experimentally and their importance on the QC and performance of LNPs have since been extensively discussed in the field [33,34]. Combined AA and CG simulations have further demonstrated the importance of the protonation state of ionizable lipid and the pH environment. At acidic pH, RNA is encapsulated within a lipid core in a hexagonal phase, shielded from ion interactions by charged ionizable lipids. In contrast, at neutral pH, LNPs expel most of the encapsulated RNA and water, which leads to the formation of distinct lipid-rich phases [35].

Given current computational capabilities, atomistic MD remains best suited for hypothesis-driven questions requiring atomic detail, such as protonation equilibria, hydrogen bonding, and local pKa estimation. In contrast, coarse-grained MD enables efficient exploration of lipid composition, PEG-lipid content, and pH-dependent self-assembly at mesoscale resolution. CG MD is well suited for routine high-throughput screening of formulation space, complemented by targeted AA MD for mechanistic validation of critical interactions. Recent work on non-lipid RNA carriers demonstrates that integrating CG MD with machine learning enables virtual screening of carrier structures [36]; similar strategies are likely to be adopted in LNP research, extending from carrier design to formulation screening. To ensure predictive value, MD outputs should be systematically cross validated with experimental observables. While MD will not replace wet-lab workflows, its judicious application can significantly reduce experimental cycles and uncover mechanistic principles that inform rational design.

### 2.1.2. *Continuum methods and complementary multiscale modeling approaches in lipid nanoparticle research*

While MD simulations offer detailed insights at the molecular level, continuum methods can provide a complementary approach by being capable of modeling the macroscopic behavior of LNPs and their surrounding fluid environments. These methods are particularly useful for studying transport properties and large-scale flow phenomena beyond the scope of atomistic or coarse-grained simulations [15]. By applying continuum-scale models, researchers can gain insights into viscosity, diffusion, and mixing behavior, which are critical for understanding both LNP formulation and large-scale production processes.

The foundation of continuum descriptions, such as computational fluid dynamics (CFD), reaction–diffusion models, or Population Balance Models (PBMs), lies in the conservation of mass, momentum, and energy, together with constitutive relations that describe how complex fluids respond to deformation. Although these models operate at the continuum scale, they depend on thermophysical parameters whose origins typically lie in molecular-level physics. Molecular statistical-mechanics approaches, including entropy-scaling correlations and thermodynamic modeling, are therefore often coupled to continuum frameworks to provide effective transport and thermodynamic parameters.

Entropy-scaling methods have been proposed as a potential approach

to estimate transport properties such as diffusion and viscosity in LNP systems. These methods are based on the premise that transport properties can be correlated with the residual entropy of the system as it has been successfully applied to alkanes, alkenes, and aromatics and their mixtures [37,38]. The appeal of this approach lies in its ability to establish generalizable relationships between thermodynamic and transport properties, which reduces the reliance on empirical measurements for model parameterization. To our knowledge, entropy-scaling methods have to date not been applied in LNP research, which might be attributed to challenges arising from complex interactions between ionizable lipids, RNA, and lipid self-assembly, which complicate universal scaling laws. Additionally, the lack of experimental data specific to intra-LNP transport properties limits the ability to validate entropy-scaling correlations.

Beyond entropy-based approaches, thermodynamic modeling can provide a framework for understanding LNPs at a continuum scale. It could be used to assess lipid self-assembly, phase transitions, and thermodynamic equilibria in different solvent environments. Additionally, thermodynamic models could help predict LNP stability, both in terms of colloidal interactions and nucleic acid encapsulation, which are critical for maintaining structural integrity and functional efficacy.

Free energy considerations form the basis of thermodynamic modeling. Depending on boundary conditions, either the Helmholtz free energy (F) or the Gibbs free energy (G) can be minimized to determine equilibrium properties. The total free energy of an LNP system comprises multiple contributions, including electrostatic interactions, hydrophobic association, hydrogen bonding, chain conformational entropy, and solvent entropy. While LNP formation leads to a reduction in configurational entropy due to molecular ordering, this can be offset by favorable enthalpic contributions from electrostatic complexation and hydrogen bonding, as well as entropic gains from hydrophobic assembly and solvent restructuring (Fig. 1D). A key thermodynamic parameter for LNP systems is the partition coefficient of NAs, which is governed by chemical potential balance and should ideally be maximized to favor NA retention within the LNP rather than its release into the external environment. Although, to the best of our knowledge, no dedicated thermodynamic model for LNPs currently exists, similar approaches have been applied to liposomal formulations, *e.g.*, for predicting their drug-loading capacity [39,40]. Potential modeling strategies include the Statistical Association Fluid Theory (SAFT) [41] and its modifications [42,43], as SAFT, unlike activity coefficient models or liquid-state theories, can explicitly capture dense-phase electrostatics and hydrophobic interactions.

A key challenge remains as model parameterization, as many necessary parameters rely on experimental data that are difficult to technically obtain. Importantly, the trade-off between predictive power and practical feasibility must be carefully considered. Excessive assumptions can lead to oversimplified models, reducing their utility for guiding LNP design and optimization.

CFD are traditionally used as a tool for studying the large-scale processes involved in LNP production. In contrast to simulations focusing on molecular-scale interactions, CFD models describe the macroscopic flow behavior of the multiphase fluids used in LNP synthesis. These methods are particularly valuable in understanding how mixing conditions influence NP formation, stability, and size distribution. Consequently, they can be employed to assess phase distribution profiles in reactors used for LNP manufacturing (Fig. 1E), including scalable systems such as T-junction mixers [44] and spheroidal reactor designs [45]. Two relevant multiphase flow models are the Eulerian – Eulerian and Eulerian–Lagrangian models (Fig. 1F), both of which can be coupled with PBMs to provide a more detailed understanding of LNP formation dynamics.

The Eulerian–Eulerian [46] model treats both the fluid phase (*e.g.*, water and ethanol) and the dispersed solid phase (*e.g.*, nanoparticles) as interpenetrating continua. It is based on the conservation equations for mass, momentum, and energy, with the Navier–Stokes equations governing momentum transport in both phases. In this framework, the velocity and pressure fields of each phase are described using a shared spatial (Eulerian) coordinate system. Interphase interactions, such as drag and lift forces, are incorporated to model NP dispersion and aggregation. This continuum approach is particularly useful for simulating large-scale mixing and phase interactions in LNP manufacturing, where different phases interact in a multiphase flow. However, since individual particles are not explicitly tracked, the model does not directly capture particle-scale processes.

In contrast, the Eulerian–Lagrangian model describes the fluid phase using a Eulerian framework while tracking individual NP in a Lagrangian reference frame. Each NP's motion follows Newton's second law, where forces such as drag, Brownian motion, electrostatic interactions, and steric effects influence its trajectory [47,48]. This method allows for a detailed representation of LNP transport and aggregation within microfluidic or chaotic-mixing systems. While computationally more expensive than the Euler–Euler approach, it provides deeper insight into how formulation parameters affect LNP properties at the particle level.

Notably, these CFD models treat LNPs as a static dispersed phase, meaning they do not account for particle formation and evolution over time. However, both Eulerian - Eulerian and Eulerian -Lagrangian approaches can be enhanced by coupling them with PBMs, which could be used to describe the evolution of LNP size distributions over time [49]. PBMs account for key processes such as (1) Nucleation and growth, *i.e.*, how lipid assemblies form and expand over time, (2) Aggregation, *i.e.*, how (precurser) LNPs collide and merge into larger clusters and eventually, (3) Breakage, *i.e.* breakdown of LNPs or LNP clusters under the shear forces within the mixer [50,51] (Fig. 1G). Up to now PBMs have not been used in LNP research but have shown to be useful in modeling dynamic multiphase systems such as bubble columns [52] or antibody solutions [53].

Despite their advantages, continuum methods have inherent limitations. While CFD models effectively describe large-scale mixing phenomena, they lack the molecular-level resolution needed to capture lipid self-assembly, RNA-lipid interactions, or nanoscopic structures. As a result, these models are complementary rather than standalone alternatives to MD simulations in LNP research. Moreover, many traditional continuum models struggle to describe non-equilibrium phenomena, which are crucial for LNP stability and drug release. For instance, the trapping of water within LNPs can significantly impact particle integrity and drug release kinetics, yet standard mass and energy balance equations fail to fully capture this process [54]. MD simulations, however, can quantify water confinement dynamics, lipid hydration states, and lipid phase transitions, information that could be incorporated into continuum models to enhance predictions of NP stability and drug release. Looking ahead, continued growth in computational power and expansion of experimental datasets may enable more rigorous integration of thermodynamic and entropy-scaling approaches into continuum frameworks, narrowing resolution gaps and strengthening predictive power for mRNA LNP design. At present, continuum approaches are most valuable for reactor and mixer design and scale-up, and their use in formulation workflows could help connect process conditions with particle size distribution and stability. Current models, however, treat LNPs as static dispersed phases and lack molecular resolution, underscoring their role as complementary tools for process optimization rather than direct formulation design.

### 2.2. Randomized design of LNP libraries: high-diversity formulation strategies

While bottom-up formulation strategies have become the dominant paradigm for LNP development, they are often implemented in a conventional sequential manner, *i.e.*, beginning with physicochemical characterization, followed by *in vitro* screening, and completed by selective *in vivo* validation. This linear workflow has proven increasingly

inefficient and restrictive. A key limitation lies in the widely recognized poor *in vitro–in vivo* correlation (IVIVC) [55] formulations that exhibit robust *in vitro* performance frequently fail *in vivo* due to the system-level complexity of bio–nano interactions. Conversely, promising *in vivo* candidates may be excluded early based on suboptimal *in vitro* metrics. These shortcomings have fostered a shift toward randomized and parallelized library designs that incorporate high formulation diversity and prioritize early *in vivo* or iterative screening cascades.

Over the past 15 years, numerous lipid libraries have been developed to address these challenges, with early efforts prior to 2015 focusing primarily on siRNA delivery. More recent work has shifted toward mRNA, reflecting evolving therapeutic priorities (Table 1).

#### 2.2.1. Experimental combinatorial libraries

Among the earliest large-scale studies in the field of RNA delivery, Akinc et al. [56], in 2008, generated a combinatorial library of over 1200 lipidoids *via* solvent-free Michael addition of alkyl-acrylates or -acrylamides to amines. The study identified structural trends, *i.e.,* multiple alkyl tails, amide linkages, and multivalent amine headgroups, that correlated with siRNA delivery efficacy *in vitro*. A refinement yielded 56 high-performing candidates with comparable potency to commercial transfection reagents. Selected lipidoids were formulated into LNPs and evaluated for systemic delivery in mice, rats, and non-human primates (NHPs). Lead candidates showed potent knockdown of hepatic Factor VII and ApoB, with effects persisting up to 4 weeks. Additionally, non-hepatic delivery was achieved *via* intranasal and intraperitoneal routes, with successful siRNA-mediated knockdown in the lung and macrophage populations. The same materials enabled efficient delivery of anti-miRs, broadening their functional applicability beyond siRNA.

Building on this, Whitehead et al. [61] synthesized a second-generation library of 1400 degradable lipidoids, incorporating ester linkages to enhance biocompatibility and facilitate intracellular degradation. *In vitro* screening identified 82 candidates achieving >50% knockdown, with top LNP formulations inducing potent silencing of Factor VII in mouse liver. Through SFR analysis, the authors identified four physicochemical predictors of *in vivo* efficacy: (i) ester degradability, (ii) alkyl tail structure, (iii) amine core type, and (iv) apparent pKa $\leq$ 5.4, with the apparent pKa emerging as the most influential factor. These criteria were validated using a second-round library, showing predictive power for delivery potency without requiring extensive *in vivo* testing. Lead LNPs showed minimal toxicity and extended applicability to monocyte and dendritic cell targeting, highlighting the translational potential of rational, degradable lipid design. However, the study also emphasized challenges in generalizing safety profiles across structurally related lipidoids, underscoring the complexity of predicting immunogenicity and toxicity.

Following many high-quality library studies on siRNA and ASO delivery using lipidoid-based nanoparticles, subsequent research, notably already before the SARS-CoV-2 pandemic, shifted toward systems optimized for mRNA delivery. This transition coincided with a refinement in materials chemistry, evolving from lipidoids, *i.e.*, structurally diverse, often non-degradable cationic lipids, to ionizable lipid (IL) platforms tailored for reduced cytotoxicity in combination with enhanced endosomal escape, and in some cases, immunostimulatory activity.

In this context, a major contribution in this area was made by Miao et al. [64], who introduced a library of heterocyclic ionizable lipids designed to enhance mRNA vaccine delivery and antitumor immunity. Using a one-pot, three-component reaction (3-CR), the authors synthesized a combinatorial library of 1080 ionizable lipids, systematically varying alkyl tail structure, linker chemistry, and cyclic amine head groups (Fig. 2A). Screening with luciferase mRNA revealed top-performing candidates with high transfection efficiency in both cell lines and antigen-presenting cells. Key structural motifs including unsaturated lipid tails, dihydroimidazole linkers, and cyclic amine head groups, were associated with both efficient mRNA delivery and STING-dependent activation of dendritic cells. To visualize transfection and antigen presentation *in vivo*, the authors employed a tdTomato reporter mouse model, in which successful delivery and expression of Cre mRNA triggers red fluorescence in target cells. This enabled spatial and temporal mapping of LNP performance in draining lymph nodes and injection sites. The study not only established SFR for potent, immunostimulatory LNPs, but also demonstrated their therapeutic applicability in multiple tumor models, validating that materials identified through *in vivo* screening can drive durable antitumor immunity.

While most early LNP library studies focused on delivery *via* intravenous (i.v.) or intramuscular (i.m.) injection, recent efforts have also expanded into alternative administration routes, such as pulmonary delivery. In this context, Li et al. [71] developed a combinatorial library of 720 biodegradable ionizable lipids (ILs) optimized for inhaled delivery of mRNA and CRISPR–Cas9 systems to lung epithelial cells. Using a three-component reaction scheme, the authors synthesized carbonate-derived lipids from ricinoleic acrylate cores, structurally validated *via* NMR and high-resolution mass spectrometry. LNPs were formulated using microfluidic mixing and screened *in vitro* for luciferase and Cas9/sgRNA delivery in A549 cells. A multi-tiered *in vivo* screen followed, employing C57BL/6 J mice for protein expression and Ai9 reporter mice for CRISPR-mediated tdTomato activation (Fig. 2B, C). The Ai9 model enables quantification of successful gene editing *via* Cre recombinase–mediated excision of a stop cassette upstream of the tdTomato reporter gene, leading to red fluorescence in edited cells. This screening cascade revealed several lead candidates for efficient pulmonary gene delivery, including previously AAV-exposed animals, indicating translational relevance for repeat dosing.

In 2025 the Mitchell lab introduced a modular “plug-and-play” platform for the synthesis of biodegradable ILs *via* a click-like Michael addition between dialkyl maleates and diverse amines or thiols [78]. A library of 500 ILs was generated and screened *in vitro* for luciferase delivery in HepG2 cells. Structural analysis revealed that IL efficacy was highly dependent on the amine architecture and the use of long, branched tails. The lead formulation showed enhanced performance *in vivo*, achieving >60% on-target gene editing of transthyretin (TTR) in mice. Beyond CRISPR/Cas9 delivery, the lead compound enabled adenine and cytosine base editing in hepatocytes and successfully corrected a mutation in the fumarylacetoacetate hydrolase (FAH) gene in a hereditary tyrosinemia type 1 (HT1) mouse model. Notably, to streamline discovery while reducing animal usage, the authors employed a pooled screening strategy in which mRNA-LNPs sharing the same amine structure were tested *in vivo* as batches. This enabled early identification of promising amines, which were subsequently deconvoluted and tested individually in purified form.

#### 2.2.2. Computationally supported library design

Despite remarkable advances in LNP formulation, the exploration of ionizable lipid chemical space remains constrained by synthetic complexity and high-throughput screening (HTS) limitations. Ionizable lipids are structurally diverse, with variation in headgroup ionizability, linker chemistry, and tail hydrophobicity creating a combinatorial explosion of potential structures. This vast design space poses a major bottleneck for traditional empirical screening. To further accelerate and refine LNP discovery, recent studies have therefore begun integrating computational tools, particularly ML, into lipid library design, enabling both virtual screening and guided synthesis. These models allow researchers to extract predictive SFRs from high-dimensional screening data, reduce the number of required synthesis iterations, and navigate complex design spaces that are impractical to explore empirically.

In one of the first applications of such hybrid approaches, Li et al. [74] developed a four-component reaction (4CR) platform to construct two sequential lipid libraries, one of 384 and another of 200 ILs, by permuting diverse head groups, linkers, and alkyl tails. A training dataset of 584 ILs was used to train an XGBoost ML model [79], an algorithm that learns patterns from data by combining many simple decision rules and gradually improving them to make accurate predictions.

**Table 1**
Summary of lipid and LNP library screening studies for nucleic acid delivery.

| Library specification | Library size | Nucleic acid cargo | Screening methodology | *In vivo* models (indiction/ target) | Main conclusion | Reference |
|---|---|---|---|---|---|---|
| lipid-like materials (lipidoids), different formulation ratios | 1200 lipidoids (combinatorial library) | siRNA (anti-luciferase), ASO (anti miRNAs 22 & 122) | *in vitro* (HeLa cells, HepG2 cells luciferase assay, primary macrophage cultures), selected for *in vivo* testing | Mice & rats (Factor VII & ApoB silencing, mir122 downregulation), NHPs (ApoB silencing | Synthesis strategy based on Micheal Addition for generating large libraries; Identification of molecular key features that enabled efficient delivery of small RNAs. | Akinc et al. - 2008 [56] |
| lipid-like materials (lipidoids), different formulation ratios | 126 lipidoids | siRNA (different target genes) | Screening *in vitro* (HeLa cells, luciferase assay), selected for *in vivo* testing. | Mice (Factor VII silencing), lead candidate further silencing of 5 hepatocellualr genes using siRNA pools; NHPs (transthyretin TTR silencing) | One step synthesis scheme for generating large libraries; Identification of C12–200 as a highly effective carrier material for siRNA delivery. | Love et al. - 2010 [57] |
| lipid-like materials (lipidoids), different core and tail structures | 17 lipidoids | siRNA (anti-luciferase, anti FVII) | Screening *in vitro* (HeLa cells, luciferase assay), selected for *in vivo* testing. | Mice (Factor VII silencing) | Variations in the core and tail structure and number influence the performance of the initial library screen (Akinc et al., 2008) | Mahon et al. - 2010 [57] |
| lipid-like materials (cationic lipoids) formulated into LNPs | 70 lipidoids | siRNA (anti-luciferase, anti FVII) | Full library tested *in vitro* (HeLa cells) and *in vivo* (mice) | Mice (Factor VII silencing) | Microfluidic mixing as a method for the rapid generation of LNP libraries. IVIVC was better for lipidoids formulated into LNPs than for lipoplexes. | Chen et al. - 2012 [58] |
| lipid-like materials (cationic lipoids) formulated into LNPs | 56 lipidoids | siRNA (anti-FVII, anti-TTR) | LNP formulation of all lipids and of selected lipid mixtures; ; characterization (pKa), *in vivo* (mice), lead formulation in NHPs | Mice (Factor VII silencing), NHPs (transthyretin TTR silencing) | Identification of a pKa range between 6.2 and 6.5 as optimal for efficient siRNA delivery. | Jayaraman et al. - 2012 [59] |
| lipid-like materials (cationic lipoids) formulated into LNPs | 32 lipidoids | siRNA (anti-FVII) | Full library characterized for pKa, stability, and tested *in vitro* (HeLa cells) for uptake, hemolysis, intracellular stability, silencing, and *in vivo* (mice) | Mice (Factor VII silencing) | Rapid, two-step solvent-free synthesis based on Michael addition and thiolyne 'click' photoaddition; identification of LNP pKa as a key determinant of function and activity, and strategies to minimize false negatives in IVIVC. | Alabi et al. - 2013 [60] |
| lipid-like materials (degradable cationic lipoids) | 1400 lipidoids | siRNA (anti-luciferase, anti FVII, anti CD45) | *in vitro* (HeLa cells), selected lipidoids for *in vivo* testing. | Mice (Factor VII silencing, CD45 silencing) | Key lipidoid features (O13 tail, $\geq$3 tails, $\geq$1 tertiary amine) predict high *in vivo* siRNA LNP efficiency; pKa $\geq$5.5 identified as indispensable for efficacy. | Whitehead et al. – 2014 [61] |
| C12–200 LNPs with variations in lipid:mRNA ratio, phospholipid, and molar composition | Library A: 14, Library B: 18, Library C: 6 | mRNA (Erythropoietin (EPO, luciferase) siRNA (anti FVII) | Design of Experiment (DoE) to create libraries, full *in vivo* screening (mice, serum EPO levels), | Mice (induction of EPO production luciferase expression, Factor VII silencing) | Adjusting formulation parameters significantly enhances mRNA LNP potency; mRNA-loaded LNPs show greater sensitivity to design space variations than siRNA-loaded LNPs. | Kauffman et al. - 2015 [62] |
| Lipid-like materials formulated into LNPs | 32 (2 rounds á 16 formulations) | Firefly luciferase and human factor IX mRNAs | *In vitro* (Hep3B cells, luciferase assay), *in vivo* validation | Mice (FIX-knockout) | Design of a synthetic route for a new class of lipid-like materials; systematic formulation adjustments in LNPs as key to maximise potency. | Li et al. – 2015 [63] |
| ionizable lipids formulated into LNPs | 1080 ILs (combinatorial library) | mRNA (luciferase, ovalbumin, Ai14/ Cre, Trp2) | *In vitro* (HeLa cells, BMDCs, BMDMs, luciferase assay, cell viability, immune marker analysis), *in vivo* validation (mice) | Mice (luciferase expression, TdTomato expression, Melanoma and human papillomavirus E7 tumor inhibition, antibody titers, T-cell responses) | One-step, three-component synthesis scheme for ionizable lipid library; ILs act as STING agonist in LNPs to enable targeted adjuvant stimulation; common structural features identified (unsaturated tail, dihydroimidazole linker, cyclic amine head groups). | Miao et al. - 2019 [64] |
| ionizable lipids formulated into LNPs | 24 ILs | mRNA (luciferase, CAR) | *in vitro* (Jurkat cells) & *ex vivo* (primary T cells), gene expression, cytotoxicity | n/a | C14–4 based LNPs identified as top performer for *ex vivo* T-cell transfection; LNPs exhibit lower cytotoxicity compared to electroporation | Billingsley et al. - 2020 [65] |

(*continued on next page*)

**Table 1** (*continued*)

| Library specification | Library size | Nucleic acid cargo | Screening methodology | *In vivo* models (indiction/ target) | Main conclusion | Reference |
|---|---|---|---|---|---|---|
| ionizable lipids formulated into LNPs | 80 ILs in 3 classes | siRNA (anti-Factor VII) | Lipid synthesis, formulation characterization (pKa, cLogD), *in vivo* (mice) | Mice (Factor VII silencing) | Modification of N-substitution (methyl → ethyl) decreases pKa; High *in vivo* FVII knockdown (>90%) achieved when cLogD ≥10 and measured pKa ≥5.8. | Rajappan et al. - 2020 [66] |
| C14–4 LNPs with varying lipid molar ratios | Library A: 16, Library B: 12 LNPs | mRNA (luciferase, CAR) | *in vitro* (Jurkat cells) & *ex vivo* (primary T cells), gene expression, cytotoxicity | n/a | Optimizing lipid ratios *via* DoE boosted mRNA delivery to T cells; achieved CAR expression comparable to electroporation with lower cytotoxicity. | Billingsley et al. - 2022 [67] |
| Dlin-MC3-DMA LNPs with varying compositions | 1080 LNPs | pDNA (Luciferase, mCherry, Cre), siRNA (STAT1, NF-κB2) | *In vitro*, top performers *in vivo* (1st intrahepatic, 2nd I. V.) | Mice (luciferase expression, TdTomato expression, STAT1 and NF-κB2 silencing) | Preferential liver transfection not primarily driven by biodistribution/ uptake but by intracellular trafficking; co-delivery of anti-inflammatory siRNA and pDNA extends therapeutic expression. | Zhu et al. - 2022 |
| linker-degradable ionizable lipids (LDILs) formulated into LNPs | 96 Ils | Luciferase mRNA | *In vitro* (IGROV1 cells) luciferase and endosomal escape assays, *in vivo* luciferase expression (mice) | Mice (luciferase expression, TdTomato expression, targeting of metastasis in breast cancer and melanoma models) | Synthesis router for linker-degradable ILs; GSH-responsive architecture drives improved endosomal escape and rapid mRNA release, improving *in vivo* mRNA delivery >100-fold *vs* DLin-MC3DMA. | Chen et al. - 2023 [68] |
| ionizable lipids formulated into LNPs | 161 ILs | Firefly luciferase mRNA | Combinatorial synthesis, LNP formulation, physicochemical characterization, *in vivo* screening | Mice (luciferase expression) | Single-step Ugi-4CR synthesis enabled systematic IL structure modulation for organ-specific mRNA delivery (liver to spleen) without helper lipids. Isomeric ILs showed marked differences in LNP performance. | He et al. – 2023 [69] |
| ionizable lipids formulated into LNPs | 16 ILs | siRNA (TTR), mRNA (EGFP, hEPO, Firefly luciferase, HA, OTC) | *In vitro* formulation and characterization, *in vivo* single and repeated dose studies with selected candidates | Mice (TTR silencing, EGFP expression, antibody levels in HA vaccine model), rats (EPO & OTC expression), NHPs (EPO expression) | Trialkyl ionizable lipid (Lipid 10) with three C10:1 chains and optimized pKa (6.3) enabled potent delivery across multiple RNA modalities (siRNA, mRNA, vaccines). Outperformed benchmark ILs; high-yield, scalable synthesis supports GMP manufacturing. | Lam et al. – 2023 [70] |
| ionizable lipids formulated into LNPs | 720 Ils (combinatorial library) | mRNA (Firefly luciferase, SpCas9, Cre) | High-throughput *in vitro* (A549 cells, luciferase), batch screening (by headgroup & tails), then 56 lead candidates individual *in vivo* | Mice (TdTomato expression, gene editing efficiency in Ai9 reporter model) | Library of biodegradable ILs identified RCB-4-8 with alkyne tails and carbonate linkers for efficient intratracheal mRNA delivery and repeat dosing; >100× potency *vs* MC3 and enabled CRISPR-Cas9 editing in airway epithelium. | Li et al. – 2023 [71] |
| ionizable lipids formulated into LNPs | 12 ILs | mRNA (Firefly luciferase, EGFP) ABE mRNA + sgRNA | *In vitro* characterization and optimization (formulation parameters), *in vivo* screening (neonatal mice brain, primary cells) and PoC (*in utero* delivery to NHP brains) | Mice fetuses/neonates (luciferase/ EGFP expression, gene editing efficiency in the brain) NHP fetus (EGFP expression gene editing efficiency in the brain) | Library of ionizable LNPs screened *via* ICV injection in fetal/neonatal mice; The lead IL enabled efficient perinatal brain mRNA delivery and base editing; partial biochemical rescue in MPS-IH model; validated in NHP fetus and human brain tissue. | Palanki et al. - 2023 [72] |
| ionizable lipids formulated into LNPs | 792 ILs (combinatorial library) | mRNA (Firefly luciferase, Cre) | Combinatorial synthesis, *in vivo* high-throughput screening (pooled LNPs) in several batches, followed by individual LNP testing and comparative analyses. | Mice (luciferase expression, TdTomato expression) | One-step Van Leusen 3CR enabled synthesis of imidazole-based ionizable lipids; Lead formulation achieved 98% spleen-selective mRNA delivery and dendritic cell transfection; >12× potency *vs* MC3 and 18× *vs* SORT without anionic lipids. | Dong et al. - 2024 [73] |
| ionizable lipids formulated into LNPs | Experimental: Library A: 384 ILs Library B: 200 | mRNA (Firefly luciferase, EPO) | *in vitro* (HeLa cells, luciferase assay), high throughput-batch testing *in* | Mice (luciferase/EPO expression) | ML-integrated with 4CR combinatorial synthesis accelerated IL discovery; ILs with branched tails and | Li et al. - 2024 [74] |

(*continued on next page*)

**Table 1** (*continued*)

| Library specification | Library size | Nucleic acid cargo | Screening methodology | *In vivo* models (indiction/ target) | Main conclusion | Reference |
|---|---|---|---|---|---|---|
| | ILs Virtual: 40,000 | | *vivo*, ML prediction (feature analysis) and validation | | tertiary amine headgroup outperformed MC3 and SM-102 for systemic and lung-targeted mRNA delivery; SORT formulation boosted lung potency 5× *vs* C12–200. | |
| ionizable lipids | Experimental: 1200, Virtual: ~60,000 | mRNA (Firefly luciferase, Cre) | Automated *in vitro* screening (HeLa & RAW 264.7 cells), computational platform (AGILE) integration, feature analysis) | Mice (luciferase/EGFP expression) | AGILE platform combines deep learning with combinatorial chemistry; identified structural motifs for cell-specific delivery (*e.g.*, tertiary amine headgroups and unsaturated tails); lead lipids enable muscle- and macrophage-targeted mRNA delivery. | Xu et al. – 2024 [75] |
| Dlin-MC3-DMA LNPs with varying helper lipids and compositions | 1080 LNPs | mRNA (Firefly luciferase, mCherry, OVA, Cre, Trp2, Gp100) | *In vitro* (DC2.4 cells, BMDCs) top performers *in vivo* (i.m. and i.v.) | Mice (TdTomato expression, OVA expression in tumor models) | Formulation screening reveals composition-driven immune polarization; Formulation with zwitterionic helper lipid induces dual Th1/Th2 responses, enhancing antitumor efficacy and synergy with checkpoint inhibitors. | Zhu et al. – 2024 [76] |
| ionizable lipids formulated into LNPs | 1700 ILs | mRNA (Firefly luciferase, Cas9mRNA + sgRNA | *In vitro* (HBE cells), 6 lead formulation *in vivo* (i.t.) | Mice (TdTomato expression) | Autonomous self-driving lab (LUMI-lab) integrates foundation models and robotic automation; discovery of brominated lipid tails as a novel structural motif improving endosomal escape; lead IL achieved record-high inhaled CRISPR-Cas9 editing efficiency in lung epithelium. | Cui et al. – 2025 [77] |
| ionizable lipids formulated into LNPs | 500 ILs | mRNA (Firefly luciferase, Cas9mRNA + sgTTR, sgPCSK9, sgHPD) siRNA (siTTR) | *In vitro* (HepG2 cells), formulations above performance threshold *in vivo* using a two-step pooled approach (i.v.) | Mice (luciferase expression, PCSK9 & HPD editing | Plug-and-play Michael addition strategy for rapid assembly of biodegradable ionizable lipids; lead lipid with multiple short tails and optimized headgroup enabled highly efficient systemic mRNA delivery and CRISPR gene editing across modalities; | Han et al. – 2025 [78] |

In this case, the model ranked candidate lipids based on expected transfection efficiency using *in vitro* luciferase expression data. This model was then applied to a virtual space of 40,000 lipids, from which 16 top candidates were synthesized and tested *in vivo*. Lead ILs enabled potent intramuscular transfection and were compatible with selective organ targeting (SORT [80]) formulations. This work represents a milestone in ML-guided LNP design, enabling performance-optimized lipids with broadened applicability across organs and use cases. Nevertheless, it must be noted that current HTS-compatible chemistries remain limited, and that ML models may exhibit design bias or fail to predict rare outliers, highlighting that the necessity for iterative validation remains.

Very recently, 'Large-scale Unsupervised Modeling followed by Iterative experiments' or short LUMI-lab presents a transformative step in automated LNP discovery, combining robotic synthesis with AI-driven design [77]. The platform integrates a molecular foundation model (LUMI-model), pretrained on 28 million chemical structures, with a closed-loop robotic workflow capable of autonomous synthesis, LNP formulation, and *in vitro* screening. The LUMI-model uses a modern deep learning approach called a transformer (a 3D transformer-based architecture), which learns general chemical patterns through unsupervised pretraining on millions of small molecules. This allows the model to build a broad understanding of molecular structures and then adapt quickly to new tasks with only a small amount of experimental data (few-shot learning). Across 10 iterative active learning cycles, over 1700 lipid candidates were evaluated in human bronchial epithelial cells. The system identified brominated lipid tails as a novel structural motif that enhances mRNA delivery. The best performing compound achieved >20% gene editing efficiency in lung epithelial cells after inhalation, surpassing previous benchmarks for pulmonary CRISPR-Cas9 delivery. This work exemplifies the potential of self-driving labs (SDLs) to accelerate molecular discovery in data-sparse therapeutic fields. Nonetheless, SDLs remain limited by sparse annotated datasets in emerging fields, as well as by the narrow range of assay modalities currently available. Integrating advanced phenotypic screening platforms, *e.g.*, organoids or multiomic profiling, may broaden discovery power. Ultimately, as richer datasets are accumulated, ML may enable reverse engineering of LNP compositions tailored to desired biodistribution profiles, therapeutic endpoints, and safety requirements.

In parallel with experimental and hybrid computational–empirical strategies, a growing number of studies have begun to explore *in silico* approaches for lead identification, that utilize ML models trained on preexisting experimental datasets to predict LNP performance without the need for new synthesis or screening [81–83]. These approaches include support vector machines (SVMs), which separate "high" and "low" performers by finding the best dividing line in the data; random forests, which combine many simple decision rules to make more robust predictions; and gradient boosting, which improves accuracy by

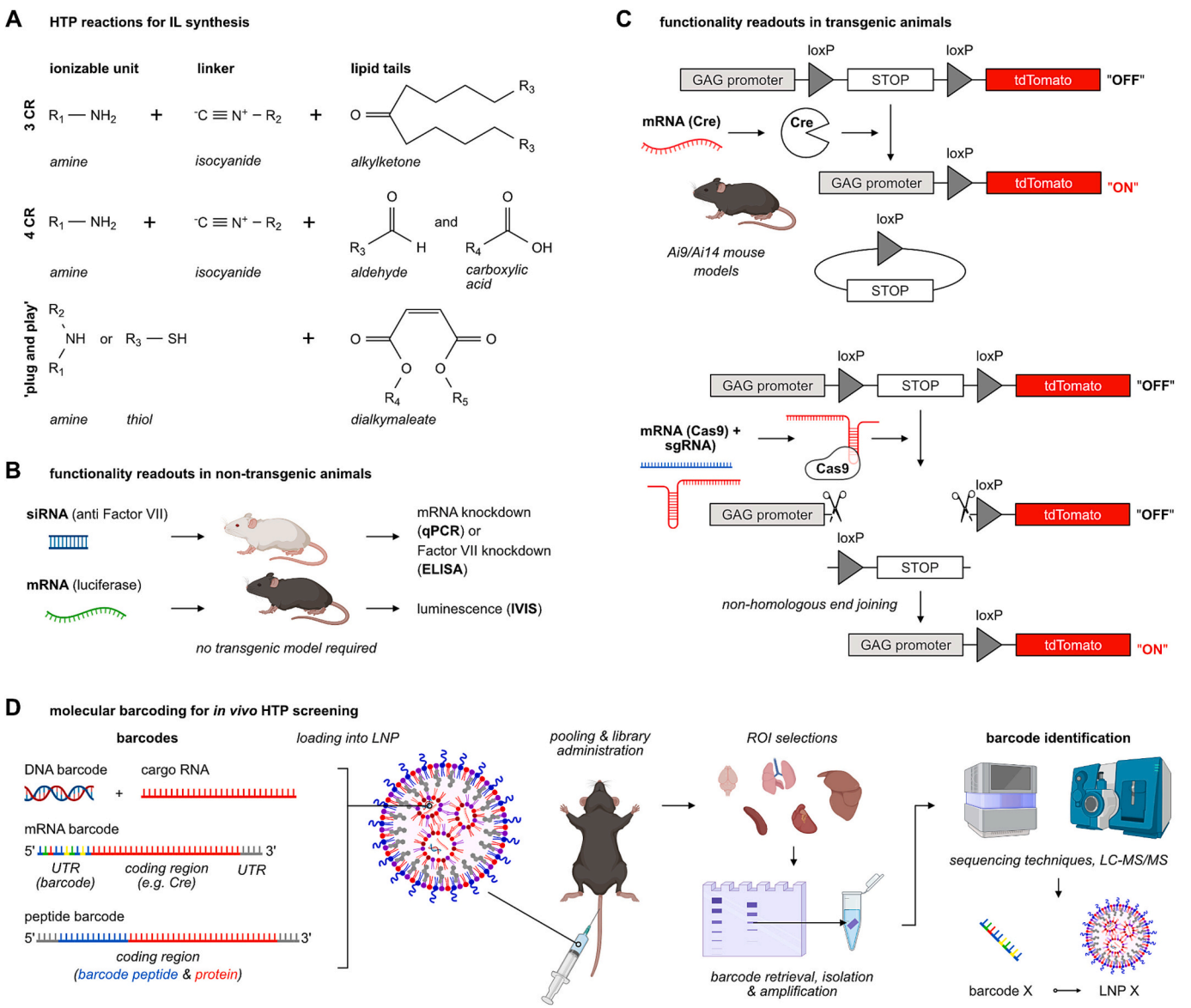


**Fig. 2.** Technologies for high throughput LNP libraries. A) One-step synthesis routes for creating combinatorial IL libraries, including three- (3CR) and four-component (4CR) reactions, as well as modular "plug-and-play" two-component schemes. B) Commonly employed reporter readouts for siRNA (Factor VII knockdown) and mRNA (luciferase expression) in non-transgenic mouse models. C) Reporter systems for mRNA and gene-editing payloads using Cre-lox recombination for quantitative tdTomato expression in Ai9/Ai14 reporter mouse models. D) Barcoding strategies, including DNA barcoding, mRNA barcoding, and peptide barcoding, and associated workflows for pooled LNP library screening (administration, region-of-interest selection, barcode retrieval, isolation and amplification, barcode–LNP association).

learning from previous errors step by step. More recently, graph-based neural networks have been used to represent molecules as networks of atoms and bonds, enabling property prediction, while transformer architectures, originally developed for language, can now generate new lipid structures by learning patterns from millions of chemical examples. While these studies use *in silico* tools during the predictive stage, they commonly comprise workflows where top-ranked candidates undergo experimental validation, underscoring that the predictive power of ML remains critically dependent on the quality, diversity, and consistency of the underlying data. As more high-quality LNP libraries, spanning diverse chemical scaffolds, administration routes, and biological endpoints, become available, ML-only models may gain in generalizability and exhibit reduced dataset bias.

For readers seeking further details and examples on ML-guided LNP design, we refer to the recent perspective by Hanna et al. [84]

## 3. High throughput *in vivo* performance evaluation using molecular barcoding

In July 2025, the National Institutes of Health (NIH) announced that it would no longer fund projects based solely on animal testing, highlighting its commitment to animal welfare. Similarly, the U.S. Food and Drug Administration (FDA) introduced new guidelines to encourage the replacement of animal testing with more efficient and predictive alternatives to reduce animal use and to lower research and development costs. These policy changes and the bottleneck in the classical approach underscores the growing need for more parallelized, multiplexed strategies. Multiplexed screenings enable broader exploration of the formulation design space, can detect synergistic effects, and provide a deeper insight into structure–function relationships, even in cases where no single formulation achieves highly specific tissue or cell targeting. Given the complex interplay of factors that influence LNP behavior *in*

*vivo*, molecular barcoding has emerged as a powerful tool and is increasingly recognized as a new standard for biodistribution and functionality assessment of pooled LNP libraries.

Through the assignment of unique molecular tags to each formulation, the co-administration of multiple formulations in a single experiment enables parallel assessment of biodistribution, cellular uptake, and functionality readouts while reducing inter-experimental variability. A rough estimation suggests that testing 100 LNP formulations *via* conventional methods would require over 300 mice, whereas a barcoded approach could achieve equivalent data using as few as 5–10 mice. Importantly, Zenhausern et al. demonstrated the implementation of barcoded LNP screening in end-of-life non-human primates (NHPs) already scheduled for euthanesia, which offers an enhanced translational relevance [85]. Interestingly, their findings highlighted that functional delivery results varied between species, with some formulations yielding to better mRNA translation efficiency in NHPs compared to mice.

Recently, placenta- [86], lung- [87], and gut-tropic [88] formulations were identified through molecular barcoding strategies, which otherwise would likely have been missed or significantly delayed using traditional sequential approaches. Among available molecular tags, nucleic acid based (DNA- and RNA-) barcoding and peptide barcoding have become the most widely used to the well-established pipelines for barcode decoding (Fig. 2D, Table 2). These methods are discussed in detail in the following section.

### 3.1. DNA barcoding

DNA barcoding is currently the most widely used high-throughput LNP screening strategy due to its vast diversity of unique sequences, scalability, and relatively low cost. By assigning a unique single stranded DNA (ssDNA) tag to each formulation, it is possible to assess the biodistribution of hundreds of LNPs in a single *in vivo* experiment [97]. These DNA tags are typically short oligonucleotides containing a barcode region, unique molecular identifiers (UMIs) and flanking primer binding sites that enable amplification and sequencing for decoding. The addition of specialized regions to the randomized barcode region is useful to minimize amplification errors and secondary structures, as well as the formation of G-quadruplexes.

Next generation sequencing (NGS) enables high-throughput, parallel sequencing, and provides relative quantification of barcode abundance within the samples. However, it is limited in cross-sample comparability. In contrast, ddPCR incorporates a probe sequence into the barcode for absolute quantification, enabling the total barcode concentration assessment with enhanced sensitivity and precision, which is useful especially for detecting low-abundance barcodes in

**Table 2**
Overview of HTP screening studies employing barcoding strategies.

| Library specification | *In vivo* screening size | Cargo | Barcode specifications | Barcode decoding | Identification | Reference |
|---|---|---|---|---|---|---|
| PEGylated lipid variants and different component ratios | 1st iteration: 112<br>2nd iteration: 105<br>3rd iteration: 64 | ssDNA barcode | 8 nt barcode region, 56 nt in total | NGS | Endothelial cells and macrophages in the spleen | Paunovska et al. - 2018 [55] |
| PEGylated lipid variants and different component ratios | 156 | ssDNA barcode | 8 nt barcode region, fluorescently labeled, QUANT barcodes | NGS | Cell type-specific effect of Cav1 in LNP delivery | Sago et al. - 2018 [89] |
| Lipid phase component ratios, N:P ratio | 16 | mRNA barcode | T7 promoter in the 5' UTR (luciferase), a barcode region at the 3' UTR | NGS | Barcoding strategy (mRNA *vs.* DNA) affects the LNP delivery | Guimaraes et al. - 2019 [90] |
| Piperazine-core ionizable lipids with varying chain length and lipid scaffolds | 65 | ssDNA barcode + Cre mRNA | 8 nt barcode region, 91 nt in total | NGS | Kupffer cells, spleen macrophages and dendritic cells | Ni et al. - 2022 [91] |
| Stereopure ionizable lipids, PEG-lipids and their varying component ratios | 128 | ssDNA barcode + siRNA | 8 nt barcode region, QUANT barcodes | NGS | ApoE-independent targeting of Kupffer cells, endothelial cells, and hepatocytes | Da Silva Sanchez et al. - 2022 [92] |
| PEG-lipids with varying molar amounts, alkyl lengths and molecular weights | 1st iteration: 41<br>2nd iteration: 53 | ssDNA barcode | 8 nt barcode region, 61 nt in total | NGS | Head and neck tumors | Huayamares et al. - 2023 [93] |
| Ionizable lipids with varying ester linkages and tail branching | 1st iteration: 48<br>2nd iteration: 24 | mRNA encoding peptide barcode | T7 promoter in the 5' UTR (peptide barcodes), a barcode region at the 3' UTR (carrier protein: streptavidin) | LC-MS/MS | Enhanced hepatic protein production | Rhym et al. - 2023 [94] |
| LNPs formulated with commercially available ionizable lipids, with two barcode mRNA loadings | 10 | mRNA encoding peptide barcode | T7 promoter in the 5' UTR (peptide barcodes), a barcode region at the 3' UTR (carrier protein: human erythropoietin) | LC-MS/MS | Workflow is not dependent on the carrier protein type | Odunze et al. - 2024 [95] |
| Cationic degradable (CAD) lipid-like materials with varying amine heads and aldehyde tails | 1st iteration: 96<br>2nd iteration: 4 | ssDNA barcode + FLuc mRNA | 8 nt barcode region, 66 nt in total | NGS | Lung tropism | Xue et al. - 2024 [87] |
| Ionizable lipids with varying cores and tails | 25 | ssDNA barcode + ribonucleoprotein (complexed Cas9 and sgRNA) | 10 nt barcode region, 61 nt in total | NGS | Endothelial and epithelial cells targeting in the lung | Haley et al. - 2025 [96] |
| Ionizable lipids with varying polyamine cores and alkyl chain length and the excipient composition | 98 | ssDNA barcode | 10 nt barcode region, 61 nt in total | NGS | Placenta tropism | Swingle et al. - 2022 [86] |

complex biological samples [98].

In a recent study, a library of 98 LNPs was generated by systematically varying the polyamine core and epoxide tails of the ionizable lipids, and the overall lipid molar ratios [86]. Each formulation was loaded with a unique 61-nucleotide DNA barcode. Following i.v. injection of the pooled library, barcode sequences were isolated from various tissues and analyzed *via* NGS. This approach led to the identification of "LNP55", a placenta tropic LNP that outperformed benchmark formulations in delivering luciferase encoding mRNA functionally to the placenta. Importantly, the barcoding strategy enabled comprehensive profiling across multiple organs that unlocked the correlation analyses between hepatic and extrahepatic delivery. These analyses revealed that formulations, where the majority of the signal is detected at the liver, exhibited weak correlation with extrahepatic delivery, whereas delivery efficiencies among extrahepatic tissues were more strongly interrelated.

Although DNA barcoding significantly reduces the number of animals required for *in vivo* screening, it does not directly report the functionality readout such as mRNA translation. To address this limitation, Sago et al. developed the FIND system, which co-delivers LNPs with Cre mRNA and a DNA barcode to a transgenic Ai14 mice [89]. In this model, Cre recombinase induces tdTomato expression by excising a stop cassette upstream of a reporter gene in successfully transfected cells. Colocalization of the DNA barcode with tdTomato expression therefore provides an indirect measure of functional delivery. While this approach enables simultaneous evaluation of biodistribution and functional delivery, its dependence on a transgenic model limits its broader applicability.

Dobrowolski et al. further focused on shifting the readout signal from cell surface marker based sorting to transcriptomic profiling [99]. The authors developed SENT-seq (single-cell nanoparticle targeting-sequencing), a method combining DNA-barcode quantification, functional mRNA readout, and whole-transcriptome analysis using single-cell RNA sequencing. In this method, DNA-conjugated beads capture a universal barcode sequence, cell hash oligonucleotide antibodies and endogenous mRNA that allows customizable sequencing read proportions and enables quantification of LNP delivery per mRNA and protein readout. The utility of SENT-seq was demonstrated with a 19-member LNP library varying in lipid molar ratios. They identified 17 transcriptionally distinct cell subtypes positive in cargo mRNA translated protein, some of which would not have been detected during standard FACS markers. Despite all formulations being tested in the liver, they exhibited distinct tropisms within the liver microenvironment, highlighting the importance of single-cell resolution in NP evaluation. Notably, SENT-seq proves to be more powerful than conventional approaches for defining on- and off-target cell types rather than organ-level tropism. By integrating LNP barcoding with transcriptomics, it can uncover cell-intrinsic features such as gene expression profiles that influence uptake and translation efficiency. This insight is particularly relevant for rare disease treatments or cell populations with transcriptionally defined phenotypes. However, SENT-seq has so far been demonstrated only in murine liver, and its broader applicability remains to be validated.

In a related study, Kim et al. used SENT-seq to screen a chemically diverse 128-member LNP library and aimed at targeting bone marrow in mice [100]. The authors quantified the functional delivery at the cell-type level and identified three lead LNPs that achieved functional mRNA delivery to murine hematopoietic stem cells (HSCs), with follow-up validation in primary human HSCs *ex vivo* and $CD34^+$ cells in rhesus macaques. Further evaluation across 24 cell types from bone marrow, liver, spleen, lung, heart, and kidney, revealed one formulation with a favorable bone marrow-to-liver targeting ratio. However, the study did not uncover statistically significant SFRs, highlighting a critical limitation: without rational or sufficiently diverse chemical design, even successful large-scale screens may fail to reveal the key mechanistic determinants of extrahepatic delivery.

While DNA barcoding is typically applied to compare multiple formulations within the same animal, its utility extends to comparative studies across species. Hatit et al. developed the SANDS (species-agnostic nanoparticle delivery screening) platform to assess cross-species predictability of LNPs [101]. Using a library of 89 LNPs, each carrying both a variable domain of heavy-chain-only antibody (VHH) reporter mRNA and the unique DNA barcode, they assessed functional delivery in mice with humanized and NHP hepatocytes.

Importantly, some correlations between LNP chemical structure and *in vivo* delivery were different between murinized, humanized and primatized hepatocytes. Furthermore, they found that humanized and murine cells underwent different transcriptional responses to LNPs. These findings demonstrate how barcoding strategy can be extended to probe interspecies delivery mechanisms to improve the predictive performance of preclinical experiments.

Although DNA barcoding is a powerful tool for LNP biodistribution assessment, it does not directly report on translation unless combined with orthogonal readouts. To address this, Wang et al. developed QuART (Quantitative Analysis of Reverse Transcribed Barcodes) which encodes a retron reverse transcriptase within the mRNA payload [102]. Upon cytosolic entry, the mRNA is reverse transcribed into a unique ssDNA barcode, which can then be sequenced for quantitative evaluation of functional mRNA delivery. The method was validated *in vivo*, demonstrating its potential for direct, high-throughput readout of functional cargo delivery. One limitation is that QuART does not quantify untranslated or degraded mRNA, as non-internalized molecules are digested during processing. Future developments incorporating methods to concurrently measure nonfunctional mRNA could further refine LNP performance evaluation.

### 3.2. RNA barcoding

DNA barcoding requires co-loading of nanoparticles with the DNA barcode in addition to their cargo, which could potentially affect its structure, physical characteristics, therefore, *in vivo* performance. To address this problem, one approach is to directly use the cargo mRNA as a barcode molecule. Embedding the unique molecular identifier and a barcode region to the untranslated or to the non-coding region within the mRNA enables its quantification *via* deep sequencing in a pooled library.

Guimaraes et al. demonstrated the utility of RNA barcoding by analyzing a 16 member LNP library and quantified the delivery of different formulations within the same organ using deep sequencing [90]. Notably, they observed that LNPs with identical formulation can exhibit different *in vivo* delivery profiles depending on the barcode type used. Specifically, LNPs carrying barcoded mRNA and those co-loaded with DNA barcodes distributed differently. This underscores the importance of carefully selecting the barcoding strategy, as it may not only impact tracking fidelity but also affect the biological behavior of the LNPs.

Yet, RNA barcoding also has certain limitations when compared to DNA barcoding. RNA is inherently less stable than DNA and more susceptible to enzymatic degradation by nucleases in biological fluids and cellular environments, which can introduce variability in barcode recovery and reduce the tracking accuracy over time. Furthermore, as RNA barcodes must remain intact and detectable until tissue processing, differences in RNA stability, intracellular degradation rates, or clearance kinetics across tissues can complicate data interpretation. In addition, RNA barcoding may be influenced by sequence-dependent effects on the mRNA payload. Incorporating additional barcode sequences can unintentionally affect mRNA folding, translation efficiency, immune recognition, or overall stability. These sequence-driven effects may introduce biases in experimental outcomes, particularly in therapeutic applications where even minor modifications to the mRNA sequence can alter expression levels or immunogenicity.

### 3.3. Peptide barcoding

Building on the limitations of DNA barcoding, *i.e.*, reliance on PCR-based quantification and challenges in detecting functional translation, Rhym et al. introduced peptide barcoding approach for assessing LNP mediated mRNA delivery *in vivo* [94]. In this method, each LNP in a library is formulated with a unique mRNA that encodes a short, quantifiable peptide barcode fused to a carrier protein. Upon successful cytosolic delivery and translation, the barcode peptides are expressed and can be quantified using liquid chromatography coupled with tandem mass spectrometry (LC-MS/MS). Expression levels can also be assessed independently through epitope tags detected *via* ELISA.

By pooling all LNPs and administering them as a single dose in the same animal, functional delivery of mRNA results in peptide barcode expression, which is then quantified using LC-MS/MS. To minimize background signals from non-barcoded proteins, they used monomeric streptavidin (mSA) as a carrier protein, chosen due to its enrichment in tissues. The carrier protein was purified from cell lysates using biotinylated beads, and the immobilized peptide barcodes were subsequently enzymatically cleaved, filtered, and analyzed to determine their relative abundances. They applied this method to a library of 384 biodegradable ionizable lipids, designed with variations in biodegradable ester linkages and tail branching, to evaluate hepatic protein production. Using only eight mice (with 48 unique LNPs per mouse), they found that 43 out of 384 LNPs resulted in detectable peptide products. The top four hits were further validated individually using firefly luciferase (Fluc). After identifying a lead compound, they optimized its formulation by conducting a full factorial design experiment varying lipid molar ratios and ionizable lipid-to-mRNA weight ratios in a 24-member library, ultimately identifying three lead formulations.

A key advantage of this method is that it does not depend on reporter cell lines, making it potentially applicable to disease models. However, a limitation is that the system is not suitable for knockdown assays. Additionally, for this system to work, the cell types of interest must be able to translate functional protein from exogenously delivered mRNA. While the current study was conducted in hepatic cells, *i.e.*, the natural target of systemically administered LNPs, its applicability to other organs with lower LNP delivery efficiency remains to be tested. An indication that there is still room for improvement is their attempt to apply peptide-barcoded LNP delivery to the spleen, where quantification proved more challenging, highlighting the need for further refinement of the approach.

The major advantage of peptide barcoding is its direct measurement of functional protein production without requiring engineered reporter cell lines, making it adaptable to disease models. However, it is limited to applications where the cell type of interest is competent for translation of exogenous mRNA. For example, while successful in hepatic cells, the method yielded lower signals in spleen, suggesting that delivery or translation efficiency may vary by organ and must be optimized accordingly. Additional efforts have explored alternative carrier proteins, such as human erythropoietin (hEPO), to improve sensitivity and clinical relevance [95].

There are other emerging barcoding strategies such as chemical barcoding by using unique chemical tags (*e.g.*, fluorescent dyes, mass tags, or isotope-labeled compounds) are attached to nanoparticles, allowing them to be distinguished through spectroscopy, mass spectrometry, or chromatography [103]. To date, these approaches have been primarily demonstrated in inorganic or polymeric nanoparticle systems, and their applicability to LNPs has not yet been systematically explored. Unlike DNA or RNA barcodes, small-molecule barcodes do not rely on genetic amplification techniques but instead enable detection *via* spectroscopy, mass spectrometry (MS), or chromatography-based methods. Up to now, small molecule barcoding based screening is applied to polymeric nanoparticle libraries [104]. Recently, however, Merivaari et al. [105] introduced a form of lipid barcoding in which liposomes were encoded by incorporating non-endogenous lipids at low molar percentages across six formulations, enabling their multiplexed detection in ocular tissues using LC–MS. While this demonstrates that barcoding through lipid composition is feasible, its broader utility for LNPs remains to be systematically explored, as altering lipid composition may influence nanoparticle biodistribution and overall performance.

Co-delivery may alter colloidal stability, promote aggregation, or influence biodistribution in ways that differ from single-formulation administration. This quality check is typically performed by measuring the hydrodynamic size of the pooled LNPs; however, this parameter alone may not be sufficient to fully assess whether individual formulations retain their stability when combined. For this reason, we recommend performing follow-up validation experiments on hit candidates without pooling. Additional validation may also be necessary when barcodes are removed and only the therapeutic cargo is encapsulated, as the particle characteristics may differ from the barcoded constructs.

Another factor to consider is the potential impact of the protein corona. In multiplexed settings, the corona formed around each formulation may differ from that observed when particles are administered alone, potentially affecting cellular uptake, immune recognition, and clearance kinetics. Therefore, while barcoding offers clear benefits, including a substantial reduction in the number of animals required for screening, its limitations should be carefully evaluated in the context of downstream validation and translational relevance.

## 4. The nano-bio-interface: understanding and exploiting the protein corona

In this chapter, we shift focus toward a long-recognized yet underutilized concept in the design of LNPs: the protein corona. Although the interaction of nanocarriers with proteins has been investigated since mid 70's, and the concept of protein corona has attracted attention for over a decade, it is only recently that this concept began to be considered a critical design parameter. We highlight this highly dynamic interface as both a fundamental and complex layer of NP identity, with inaugural impact on improving targeting specificity and delivery efficacy.

### 4.1. The visible face of lipid nanoparticles

Depending on the route of administration, the first contact of LNPs with the body occurs in a complex biological environment. For instance, upon intravenous injection, LNPs are immediately exposed to blood, a highly intricate biofluid containing more than 3700 proteins in plasma [87], where individual protein concentrations span several orders of magnitude. These proteins rapidly adsorb onto the surface of LNPs and form a dynamic layer of biomolecules known as the protein corona, which plays a crucial role in defining the biological identity of LNPs (Fig. 3A). The protein corona structure is typically identified as a two-compartment system, the hard and the soft corona, which together referred to as the biomolecular corona.

The hard corona is composed of proteins that have large net binding energy of adsorption that causes relatively stable attachment to the LNP surface. These proteins interact directly with the NP surface through a combination of electrostatic and hydrophobic interactions, as well as van der Waals forces and hydrogen bonding. In contrast, the soft corona consists of a dynamic and loosely bound layer of proteins that exchange continuously with the surrounding medium. These proteins often do not contact the LNP surface directly but instead interact with the proteins in the hard corona. However, there is no clear physical boundary between hard and soft corona layers, and the two are best viewed as a continuum of binding affinities and residence times [106].

Given the diversity of LNP surface chemistries, and protein-protein interactions on the LNP surface, the composition of the corona can vary considerably between different formulations and even among individual nanoparticles within a single formulation (Fig. 3B).

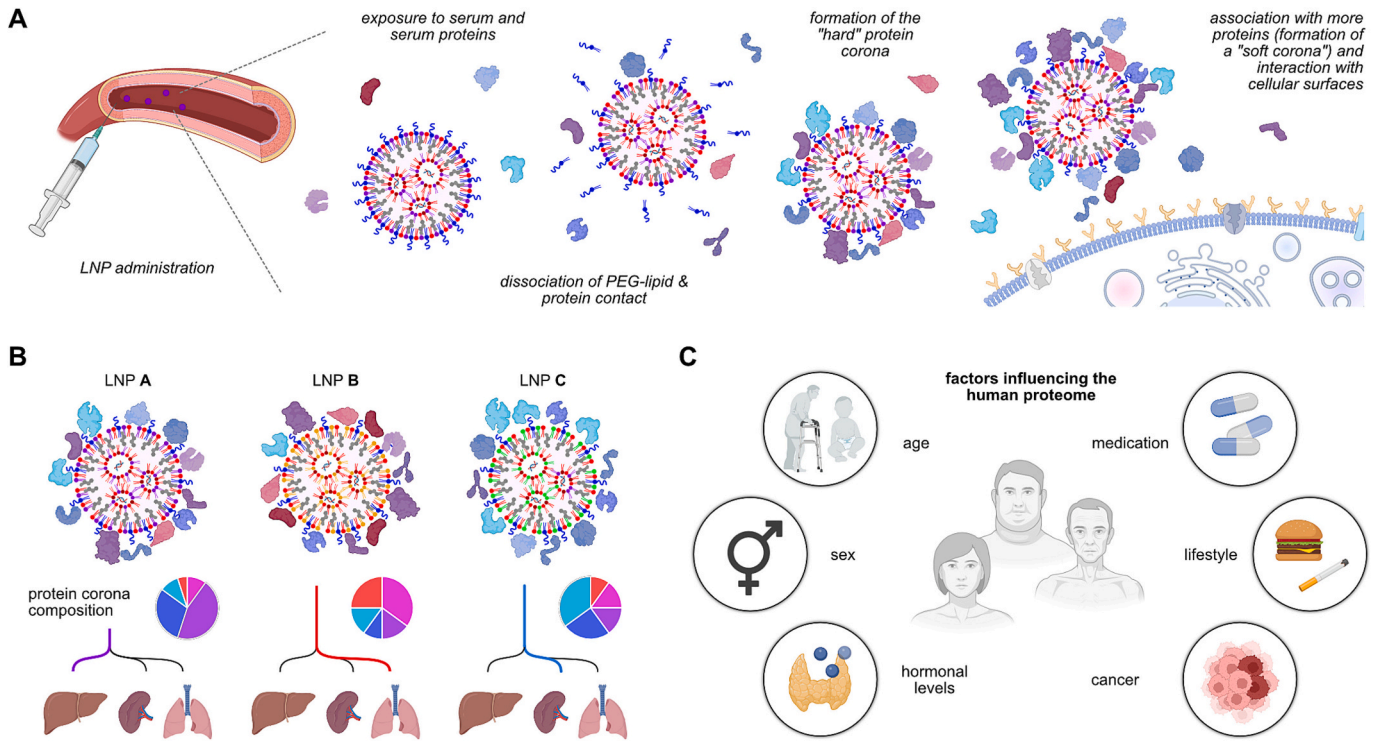


**Fig. 3.** Role and influencing factors of protein corona on LNPs. A) Typical journey of i.v. administered LNPs, from injection, PEG shedding, and the sequential formation of soft and hard protein coronas, to subsequent interactions with target cells. B) Influence of LNP formulation parameters (*e.g.*, lipid components and their ratios) on protein corona composition, and downstream effects of the corona on the organ/tissue tropism of LNPs. C) Host-intrinsic factors (age, sex, hormonal status, disease state), external exposures (such as medications), and lifestyle-related behaviors that shape the human proteome and thereby influence the protein composition of body fluids relevant to corona formation.

Importantly, proteins may undergo conformational changes upon adsorption, which may be further influenced by interactions with neighboring proteins within the corona. Consequently, the effective charge, hydrophobicity, and biological recognition elements displayed on the nanoparticle surface can differ substantially from the original formulation, making the corona highly challenging to predict or characterize in a reproducible way.

Thus, the biomolecular corona serves as the "visible face" of the LNP, the interface that cells, immune receptors, and tissues interact with. Certain plasma proteins are repeatedly found enriched in LNP coronas, among which apolipoprotein E (ApoE) has received particular attention due to its role in mediating hepatocyte uptake [107]. ApoE can facilitate interactions with hepatocyte receptors, *e.g.*, low-density lipoprotein receptor (LDL-R) and heparan sulfate proteoglycans (HSPGs), and is implicated in the rapid liver accumulation of LNPs following systemic administration. These interactions are closely tied to PEG-lipid desorption, a key process that exposes the NPs surface for protein binding. The shedding of PEG-lipids is influenced by the relative hydrophobicity of the anchor domain and PEG unit [108–110] which thereby modulates the timing and extent of ApoE adsorption. Once PEG is shed, ApoE binds to the exposed surface, enabling receptor-mediated uptake in the liver. Multiple studies have shown that liver tropism of systemically delivered LNPs is closely linked to the presence of ApoE-enriched coronas [109,111].

LNP structural features, notably surface-enriched phospholipids and cholesterol, have been implicated in promoting apolipoprotein adsorption, consistent with their physiological roles in lipid transport [112]. Interestingly, protein binding can induce rearrangements in LNP internal architecture, which further complicates the correlation between synthetic formulation and biological identity [113]. The protein corona rapidly forms upon exposure of the LNP to biological media and stabilizes over time, with the specific fingerprint being driven by affinity-based on/off-kinetics of competing proteins [106]. While the specific composition may evolve over time, particularly within the soft corona, many adsorbed proteins can remain associated with the particles until degradation and thereby significantly affect the critical window of biodistribution, clearance, and biological activity.

While much of the current literature has focused on the hard corona and its effect on LNP tropism, due to its relative stability and the availability of established isolation techniques, the role of the soft corona remains less understood. For example, a study using polystyrene nanoparticles in human plasma demonstrated that, while variations in the soft corona had minimal effect on cellular uptake, changes in the hard corona strongly influenced internalization [98]. Although this specific behavior might not directly apply to other kinds of nanostructures, such as LNPs, the analysis of both soft and hard protein corona should be considered in further studies to progress on the understanding of their role in organ-level biodistribution.

### 4.2. *Changes in protein corona in health and disease*

It is well established that the composition and properties of human plasma proteins vary significantly across different physiological and pathological conditions. Disease states such as cancer, diabetes, autoimmune disorders, and infections can alter not only the concentrations of individual plasma proteins but also their conformations and post-translational modifications, including glycosylation patterns (Fig. 3C). For example, during inflammation or infection, levels of acute-phase proteins such as C-reactive protein (CRP) increase markedly. Certain infections may also induce specific changes in glycosylation, potentially altering protein function and NP interactions. These alterations have direct implications for protein corona formation and, consequently, LNP performance.

In a study using hypercholesterolemic mice, inorganic model nanoparticles exhibited reduced interaction with complement proteins but increased apolipoprotein coverage, resulting in enhanced hepatic

accumulation [114]. Similarly, MC3-based LNPs formed distinct corona profiles when incubated in plasma derived from lean or obese rats, thereby confirming that physiological state significantly influences corona composition and, consequently, delivery efficiency [115]. Long-term medical interventions can also reshape plasma protein profiles. For example, prolonged hemodialysis has been associated with altered levels of several proteins, including apolipoproteins [116]. Although such shifts have not been systematically explored in the context of LNP protein coronas, they are expected to influence the biodistribution and efficacy of the therapeutic cargo. Hormonal fluctuations represent another underexplored factor with the potential to modulate the plasma proteome. Changes in hormone levels can alter the abundance of specific plasma proteins, thus indirectly shaping the composition of the protein corona. For instance, physiological changes during the menstrual cycle have been associated with increased NP accumulation in the ovaries, although a direct correlation with corona composition remains to be established [117].

Moreover, age, sex, and lifestyle factors contribute to inter-individual variability [118]. One study demonstrated that model liposomes were cleared more slowly in elderly patients compared to younger individuals, highlighting the influence of aging on NP pharmacokinetics [119]. Variations in plasma composition and hemodynamics associated with lifestyle factors can also affect protein corona formation. Additionally, sex-specific differences in NP biodistribution have been linked to distinct protein corona profiles, as summarized in a review by Schroeder and colleagues [120]. In another study, male plasma was found to contain higher levels of innate immune cytokines during infection, whereas female-derived cells exhibited stronger T cell activation responses [121].

As LNPs are increasingly applied in therapeutic contexts such as *in vivo* CAR-T cell generation and CRISPR-mediated gene editing, understanding the impact of disease-associated plasma states on protein corona formation becomes critical. Rational LNP design must account for this complexity, potentially through inclusion of metabolome- or condition-specific plasma in early-stage screening. Such strategies will be essential for developing next generation nanocarriers with enhanced reproducibility, safety, and translational relevance.

### 4.3. Techniques used to determine the protein corona

Currently, due to analytical limitations, protein corona characterization is predominantly performed under *ex vivo* conditions. Typically, NPs are mixed with the biological fluid in a test tube under controlled temperature and agitation for a defined duration, isolated from unbound proteins and excess media, and characterized. This approach does not fully recapitulate the *in vivo* conditions, where factors such as hemorheological properties and immune responses influence corona formation. In a model microfluidic setup, it was shown that the biomolecular corona formed in a dynamic environment differs from that formed under static conditions [122]. However, a key limitation remains the lack of effective methods for capturing NPs directly from blood. Robust high-precision assays and improved NP isolation techniques are urgently needed to enable more accurate characterization of the *in vivo* protein corona.

A landmark study by Hadjidemetriou et al. [123] investigated *in vivo* protein corona formation on liposomes in ovarian carcinoma patients. PEGylated, doxorubicin-loaded liposomes were intravenously administered, and the hard corona was isolated using two-step size exclusion chromatography followed by membrane ultrafiltration. The resulting protein corona was then analyzed. Their results revealed that the *in vivo* corona was molecularly more complex and compositionally distinct from coronas formed during *ex vivo* incubation in plasma, underscoring the limited predictive value of *ex vivo* data. Similar discrepancies between *in vivo* and *ex vivo* corona profiles have been confirmed for magnetic NPs by Landfester and Mailänder [124].

Most studies exploring the role of the protein corona in NP targeting focus primarily on the hard corona. This emphasis is largely due to the availability of established, relatively straightforward isolation methods. For example, a study using model polystyrene NP incubated in human plasma found that variations in the soft corona had minimal impact on cellular internalization, whereas changes in the hard corona significantly influenced uptake. Although the soft corona has received more attention, partly due to the emergence of gentler isolation techniques that preserve weakly bound proteins [125], its role in targeting specificity and *in vivo* performance remains underexplored.

Among isolation methods used in *ex vivo* corona characterization, centrifugation remains the most commonly applied technique. Other approaches include magnetic separation—limited to magnetic NPs—and affinity ligand-based capture, though ligand attachment can alter the native corona composition. Salvati and Sabirsh demonstrated that LNPs can be captured from plasma using anti-PEG antibody–conjugated magnetic beads [115], a technique that may have broader applicability in LNP corona studies.

Asymmetric flow field-flow fractionation (AF4) has also been employed to separate NP-protein complexes. AF4 is a chromatography-like separation technique that fractionates particles and macromolecules in a thin, ribbon shaped channel by applying a crossflow field perpendicular to the laminar carrier flow. Analytes are differentially retained according to their diffusion coefficients (and thus hydrodynamic sizes), leading to size-based separation without a stationary phase and with minimal shear force [126]. Each method presents specific advantages and limitations, which have been comprehensively reviewed by Francia et al. [12] Following isolation, liquid chromatography coupled with mass spectrometry (LC-MS) remains the current gold standard for qualitative and quantitative corona proteomics. This technique enables detailed profiling of protein classes, identification of the most abundant species, and inter-sample comparisons. However, LC-MS results are highly sensitive to experimental protocols, making cross-study comparisons challenging. Differences in sample preparation, instrument settings, and data analysis pipelines introduce significant variability, complicating efforts to correlate corona composition with NP formulation and biological fate.

Notably, different corona retrieval methods can yield divergent protein profiles [127]. Ashkarran et al. [128] highlighted substantial heterogeneity in protein corona datasets generated across various LC-MS workflows at different facilities. By standardizing sample preparation, instrumentation, and data processing methods, they increased the overlap in protein identifications from 11% to 40%, demonstrating the critical need for harmonized analytical protocols [129].

Despite increasing interest in the nano-bio-interface, reproducibility in protein corona characterization remains a major bottleneck for the field. Without standardized and reliable methodologies, it is difficult to identify reproducible patterns or establish design principles necessary for clinical translation of NP-based therapies. While variations between animal models and humans are inevitable, data from preclinical studies can still provide valuable guidance for formulation development. Addressing methodological reproducibility at this stage is therefore essential to advance our understanding of the nano-bio-interfaces and their role in linking formulation properties to biological performance.

### 4.4. Impact of administration route on nano-bio-interfaces

NPs administered *via* different routes encounter distinct biological environments, which in turn lead to the formation of divergent protein coronas and biological fates. For example, pulmonary delivery exposes particles to mucus, lung lining fluid, or other compartment-specific biofluids, each characterized by unique pH levels, ionic strengths, and protein compositions [130]. As with plasma, it is important to characterize NP–biofluid interactions that are specific to the intended administration route, such as interstitial fluid, cerebrospinal fluid (CSF), or respiratory tract lining fluid, as these shape the resulting corona composition. In the context of pulmonary delivery, coronas formed in

lung lining fluid have been studied; however, clear correlations between protein corona composition and *in vivo* outcomes remain to be established [131]. Schiffelers et al. compared protein corona profiles formed in plasma and cerebrospinal fluid (CSF), attributing differences in *in vitro* delivery efficiency to the distinct protein composition and abundance characteristic of each biological fluid [132]. Very recently, coronas formed in amniotic fluid were described for a panel of polymeric NPs, highlighting their potential implications in fetal therapeutics [133].

The dynamic nature of protein corona formation is further influenced by physiological variables such as pH, ionic composition, and fluid viscosity along the NP's transport pathway. These factors govern adsorption kinetics, protein exchange, and conformational rearrangements. Importantly, studies suggest that corona composition may shift qualitatively and quantitatively as particles traverse biological barriers, such as the blood–brain barrier [134]. These complexities underscore the importance of designing *in vitro* corona studies that more closely reflect the *in vivo* biofluid environments relevant to the intended application. Despite the widespread use of human serum for protein corona formation (primarily due to cost and accessibility) this practice may not reliably predict *in vivo* outcomes, particularly when tropism is assessed in murine models. Emerging evidence suggests that correlations between human serum–derived coronas and biodistribution in mice may be weak or inconsistent. This hypothesis is supported by two determining factors; first, the discrepancy between *in vitro versus in vivo* corona formation [124], and, second, the species dependent corona composition [135]. This challenge is further compounded by the dynamic and route-dependent evolution of the corona, as NPs encounter sequential, compositionally distinct biological milieus from the injection site to the target tissue. Recent investigations into corona formation across various biofluids reflect a growing recognition of the need for context-specific, administration route–matched experimental models, marking an important step toward improving the translational relevance of protein corona research.

### 4.5. Revealing protein corona-mediated tropism

In addition to the previously noted ApoE-mediated liver tropism—recently refined to highlight that a corona enriched in ApoE-containing high-density lipoproteins is a more accurate predictor of *in vivo* hepatocyte delivery than ApoE alone [115]—the protein corona has been identified as a key factor enabling extrahepatic delivery.

Recent mechanistic studies have validated these findings through different approaches. These include NP precoating with selected proteins, followed by mechanistic uptake studies in cells expressing the cognate receptors of the respective proteins [136,137]. Pre-coating of clusterin (ApoJ) has been shown to decrease the clearance of nanoparticles with low PEG density [138] while albumin precoating was associated with reduced liver tropism [139]. Biodistribution studies in knockout animal models aim to shed explanatory light on the role of specific plasma proteins on nanoparticle distribution. Here, either specific proteins or their cognate receptors are being knocked out, as for the ApoE / LDL receptor axis. Not only the role of ApoE in liver delivery, but also the role of ApoE in BBB crossing was established following this approach [107,138,140].

Notably, formulation chemistry can induce changes in the apparent pKa of LNP components, which in turn alters the protein corona profile and consequently shifts organ targeting (Fig. 3B). However, the underlying mechanisms are more complex, with other factors such as particle size, shape, or curvature also influencing specific protein adsorption [141,142].

The role of lipid chemistry in corona-mediated tropism was described by Mitchell and colleagues using a DNA barcode–based high-throughput *in vivo* screening approach, which revealed that changing the helper lipid identity from DOPE to DSPC lowered the affinity of LNPs for ApoE, leading to enhanced spleen delivery [143]. In another example, in the spleen-SORT formulation, Siegwart and colleagues introduced an anionic phospholipid as a fifth lipid component [137]. This modification enriched the corona with β2-glycoprotein I (β2-GPI), resulting in enhanced spleen tropism. Conversely, adding quaternary ammonium–containing lipids, such as TAP or EPC lipid variants, yielded a vitronectin (Vtn)-rich corona, promoting lung tropism. Crucially, the incorporation of such permanently cationic lipids increased the apparent pKa of the LNPs. Thereby electrostatic interactions with proteins in their proximity are altered, consequently increasing the abundance of proteins with an isoelectric point below the physiological pH, such as Vtn. These findings were further validated by precoating particles with selected proteins and demonstrating increased tropism *in vivo*. Confirmatory studies in cell culture and wild-type *versus* ApoE-knockout mice showed that lung and spleen targeting was independent of ApoE. However, the authors emphasize that it is unlikely a single protein alone determines biodistribution, as they found positive correlations with lung tropism for several proteins [144]. Beyond that, Vtn has been associated with other cationic materials, such as DC-cholesterol by Amiji's group [145], confirming targeting properties of Vtn-enriched coronae due to the Arg-Gly-Asp (RGD) sequence, recognizing the $\alpha_{v\beta3}$ integrin, overexpressed in pulmonary endothelium and tumor tissue [146,147].

Similarly, a recent formulation developed by the Mitchell lab, which successfully targeted the placenta, also exhibited β2-GPI enrichment in its protein corona and showed spleen tropism [86]. This formulation displayed slight surface electronegativity, consistent with the spleen-SORT particles where an anionic lipid component was incorporated.

Tuning LNP surface properties such as charge, PEG-lipid chain length, and PEG density also influences protein corona composition. For instance, LNPs with coronas enriched *in vitro*nectin demonstrated reduced hepatocyte transfection compared to those with ApoE-rich coronas [145]. Typically, these formulations are initially identified by differences in biodistribution or function. Subsequently, researchers return to examining the nano-bio-interface to understand the protein corona's role in driving tropism. An alternative and efficient approach involves first characterizing protein corona patterns *in vitro*, then validating these signatures *in vivo*, potentially reducing reliance on animal experiments and accelerating formulation screening. Incorporating such physicochemical and corona-derived parameters into personalized NP library screening may facilitate more predictable and efficient clinical translation of LNPs.

### 4.6. Protein corona as a design feature

The protein corona is not only a challenge to overcome, it can also be harnessed as a design variable to guide the biological fate of LNPs. Engineering the LNP surface to selectively attract beneficial plasma proteins has emerged as a promising strategy to enhance delivery performance. One such approach involves precoating NPs with specific proteins prior to administration, which has been shown to improve target cell interaction and uptake *in vitro* [148–150]. Although precoating strategies hold significant potential, the design of NPs must account for the specific binding affinities of target proteins, particularly in the context of complex biological fluids, to minimize competitive displacement and ensure stability of the intended corona upon administration.

Another strategy is to functionalize the LNP surface to actively “fish” for selected plasma proteins post-injection. This has been shown, for example, through albumin fishing, where albumin-binding peptides or albumin-binding domains are incorporated onto the NP surface [151]. Covalent conjugation methods, such as reacting maleimide-functionalized LNPs with the free sulfhydryl group on albumin, have also been employed. Similar strategies include the use of transferrin-binding peptides to enrich LNP coronas with transferrin. The role of cholesterol in shaping the protein corona has been highlighted through corona seeding strategies, where cholesterol-functionalized DNA nanostructures promoted apolipoprotein enrichment in the corona, thereby improving liver-targeted delivery [152]. In a study, Zhan et al. [153]

modified liposomes with short peptides derived from amyloid-β that selectively recruit apolipoproteins A1, E, and J from plasma. This corona engineering significantly enhanced blood–brain barrier penetration and improved targeting of brain tumors. In a related direction, the group of Roy Van Der Meel designed LNPs that promote ApoE adsorption to facilitate delivery across physiological barriers. They explored the design of apolipoprotein-based NPs, incorporating ApoA1 into LNP-like structures to exploit its natural role and target the bone marrow, specifically myeloid cells and hematopoietic stem and progenitor cells [154].

An alternative design strategy is to minimize protein adsorption altogether in order to prolong circulation time, as recently reviewed [136]. This is typically achieved by imparting "stealth" characteristics to LNPs through surface engineering. While polyethylene glycol (PEG) is the most widely used stealth material, alternatives such as zwitterionic polymers are being explored [155]. These surface modifications aim to reduce opsonization by plasma proteins such as complement factors, immunoglobulins, and vitronectin, and thereby limit recognition and clearance by the mononuclear phagocyte system.

As our understanding of how specific adsorbed proteins influence NP fate deepens, next-generation formulations may be designed to selectively recruit specific protein classes. The possibility of enriching corona components that engage distinct receptor-mediated uptake pathways opens new directions for activating passive targeting through biologically informed NP engineering. Importantly, increased cellular uptake does not always correlate with functional delivery. For instance, a corona enriched in vitronectin may enhance uptake, yet simultaneously impair endosomal escape, leading to lysosomal degradation and reduced mRNA translation [136]. Thus, protein corona–targeting relationships must be assessed not only in terms of biodistribution but also through their effects on intracellular trafficking and ultimate transgene expression.

### 4.7. *Toward predictive models for protein corona–guided targeting*

A central objective in NP research is to establish predictive relationships between physicochemical properties, corona composition, and *in vivo* behavior. Several groups have studied how tunable NP parameters, including size, shape, charge, and surface roughness, affect corona formation using highly controlled systems such as liposomes and gold NP [156]. It must, however, be noted that establishing such predictive relationships for more complex organic NPs, especially *in vivo*, remains challenging due to their intrinsic structural variability and dynamic surface interactions.

To move from empirical observations to predictive frameworks, Lazarovits et al. [157] presented a data-driven framework that links the dynamic, *in vivo*-evolved protein corona to NP biodistribution using supervised ML. Rather than attempting to control or model the corona composition *a priori*, they extract model gold NPs from rat blood at multiple time points post-injection and profiled their protein corona *via* quantitative MS. These time-resolved protein corona signatures serve as high-dimensional input features to train a neural network capable of accurately predicting key biological outcomes, namely, blood half-life and liver and spleen accumulation. The model captured the combinatorial nature of corona-mediated biological interactions and revealed that no single protein drives clearance; instead, complex and evolving protein patterns determine NP fate. By validating their model in blinded experiments and reducing input dimensionality through sparse feature selection, the authors substantially advanced the field toward predictive, corona-informed nanocarrier design for targeted delivery.

The integration of ML into corona analysis represents a promising next step. Predictive models trained on large proteomic datasets could eventually enable forecasting of circulation time, organ tropism, toxicity, and transfection efficiency based solely on corona composition [158]. However, such models require access to well-annotated NP libraries with systematically varied surface features and corresponding high-quality corona proteomics data. Recent efforts using DNA nanostructures with precisely defined shapes have demonstrated the feasibility of training ML models to predict corona composition [159]. While LNPs pose additional challenges due to their dynamic self-assembly and less controllable surface architecture, similar modeling approaches could, in principle, be developed, provided that the appropriate training data can be generated.

At present, we are beginning to understand the interaction of NPs with physiological microenvironments, but we are still far from being able to rationally and reproducibly manipulate these interactions to achieve tissue- and cell-specific delivery. Achieving this level of control will require a combination of mechanistic insight, system-level data integration, and predictive modeling, all grounded in a better understanding of the protein corona.

## 5. Perspective

The current landscape of LNP research appears divided between fast, cost-efficient HTP screening and the slower, mechanistically grounded pursuit of understanding. While bottom-up development strategies have enabled the rapid generation of large lipid libraries, traditional library screening is typically optimized for short-term identification of a lead formulation within a specific design space. These screens can yield empirical correlations between structure and function and inform subsequent iterations of design. However, obtaining SFRs does not necessarily provide mechanistic insight into the physicochemical or biological principles governing NP behavior. In many cases, these findings remain descriptive rather than explanatory.

This limitation underscores the need for more rational and interpretable approaches to library design, where systematic variation of structural features, potentially guided by ML, can help map functional outcomes across broader application domains. ML-based models offer the promise of generalization across datasets and the identification of transferable trends that may extend beyond human intuition. Most current datasets, however, are narrow, inconsistently annotated, or tailored to specific contexts, which limits model interpretability and generalizability. A key opportunity lies in expanding and standardizing experimental libraries so that they can be integrated across platforms, administration routes, or therapeutic indications. Cross-library analysis could enable the discovery of higher-order SFRs and support more robust prediction of biodistribution, transfection efficiency, and safety. However, this trend also introduces new risks. Shared design biases may skew ML outputs toward dominant but potentially suboptimal chemistries, while inconsistencies in experimental protocols may hinder data harmonization and introduce false correlations. These limitations call for critical reflection on how datasets are generated, curated, and shared. Yet, the fundamental discrepancy between descriptive SFRs and mechanistic understanding remains unresolved, even, or perhaps especially, in the context of ML-supported workflows, which may yield correlations that are statistically robust but biologically opaque.

Most critically, the incremental nature of knowledge accumulation in the LNP field continues to limit our ability to learn from failure. When non-performing formulations are discarded without systematic documentation, valuable negative data are lost. In contrast, well-structured libraries, regardless of immediate success, can serve as long-term resources for identifying mechanistic principles, optimizing delivery for novel payloads, or training more predictive models. As the field advances toward application-specific and personalized therapies, it will be essential to move beyond empirical discovery and establish a more mechanistically informed, data-integrated framework for NP design. In the era of personalized medicine, smart drug delivery, and nucleic acid–based therapeutics, precise understanding of nanoparticle structure, behavior, and fate is no longer a scientific luxury, it is a translational necessity.

## Statement for the use of LLMs

During the preparation of this work, the authors used LLMs to improve readability and language. The text was reviewed afterwards, and the authors take full responsibility for the content of the publication.

## CRediT authorship contribution statement

**Benjamin Winkeljann:** Writing – review & editing, Writing – original draft, Conceptualization. **Philipp Lapuhs:** Writing – review & editing, Writing – original draft. **María J. Alonso:** Writing – review & editing, Writing – original draft, Supervision. **Ceren Kimna:** Writing – review & editing, Writing – original draft, Conceptualization.

## Declaration of competing interest

BW has equity interests in RNhale GmbH. MJA has equity interests in Libera Bio. The authors declare no other competing interests.

## Acknowledgments

Figures include graphical elements created using BioRender (license held by BW). The authors thank Nora Martini (LMU München) and Matthias Schumacher (leon-nanodrugs GmbH) for providing snapshots from MD and CFD simulations, respectively, used in the illustrations presented herein.

## Data availability

No data was used for the research described in the article.

## References

[1] B.B. Mendes, et al., Nanodelivery of nucleic acids, Nat. Rev. Methods Primers 2 (2022) 24.
[2] A. Gupta, J.L. Andresen, R.S. Manan, R. Langer, Nucleic acid delivery for therapeutic applications, Adv. Drug Deliv. Rev. 178 (2021) 113834.
[3] Y. Eygeris, M. Gupta, J. Kim, G. Sahay, Chemistry of lipid nanoparticles for RNA delivery, Acc. Chem. Res. 55 (2022) 2–12.
[4] K. Paunovska, D. Loughrey, J.E. Dahlman, Drug delivery systems for RNA therapeutics, Nat. Rev. Genet. 23 (2022) 265–280.
[5] P.R. Cullis, P.L. Felgner, The 60-year evolution of lipid nanoparticles for nucleic acid delivery, Nat. Rev. Drug Discov. 23 (2024) 709–722.
[6] D. Adams, et al., Patisiran, an RNAi therapeutic, for hereditary transthyretin amyloidosis, N. Engl. J. Med. 379 (2018) 11–21.
[7] F.P. Polack, et al., Safety and efficacy of the BNT162b2 mRNA covid-19 vaccine, N. Engl. J. Med. 383 (2020) 2603–2615.
[8] L.R. Baden, et al., Efficacy and safety of the mRNA-1273 SARS-CoV-2 vaccine, N. Engl. J. Med. 384 (2021) 403–416.
[9] Gene, Cell, & RNA Therapy Landscape Report - Q1 *2025 Quarterly Data Report*. https://www.asgct.org/global/documents/asgct-citeline-q1-2025-report.aspx (2025).
[10] S. Park, M. Kim, J.W. Lee, Optimizing nucleic acid delivery systems through barcode technology, ACS Synth. Biol. 13 (2024) 1006–1018.
[11] J. Luo, et al., Nanocarrier imaging at single-cell resolution across entire mouse bodies with deep learning, Nat. Biotechnol. (2025) 1–14, https://doi.org/10.1038/s41587-024-02528-1.
[12] V. Francia, R.M. Schiffelers, P.R. Cullis, D. Witzigmann, The biomolecular corona of lipid nanoparticles for gene therapy, Bioconjug. Chem. 31 (2020) 2046–2059.
[13] P.K. Inguva, et al., Mechanistic modeling of lipid nanoparticle formation for the delivery of nucleic acid therapeutics, Biotechnol. Adv. 84 (2025) 108643.
[14] P. Joyce, et al., A translational framework to DELIVER nanomedicines to the clinic, Nat. Nanotechnol. 19 (2024) 1597–1611.
[15] P.K. Inguva, et al., Mechanistic Modeling of Lipid Nanoparticle Formation for the Delivery of Nucleic Acid Therapeutics, 2024, https://doi.org/10.48550/arXiv.2408.08577.
[16] R. van der Meel, F. Grisoni, W.J.M. Mulder, Lipid discovery for mRNA delivery guided by machine learning, Nat. Mater. 23 (2024) 880–881.
[17] Prediction of the Apparent pKa Value of Lipid Nanoparticles by Density Functional Theory, ACS Materials Au, 2026. https://pubs.acs.org/doi/10.1021/acsmaterialsau.4c00158.
[18] Grzetic, D. J., Hamilton, N. B. & Shelley, J. C. Coarse-grained simulation of mRNA-loaded lipid nanoparticle self-assembly. *Mol. Pharm.* doi:https://doi.org/10.1021/acs.molpharmaceut.4c00216 (2024) doi:https://doi.org/10.1021/acs.molpharmaceut.4c00216.
[19] M. Cornebise, et al., Discovery of a novel amino lipid that improves lipid nanoparticle performance through specific interactions with mRNA, Adv. Funct. Mater. 32 (2022) 2106727.
[20] M. Paloncýová, P. Čechová, M. Šrejber, P. Kührová, M. Otyepka, Role of ionizable lipids in SARS-CoV-2 vaccines as revealed by molecular dynamics simulations: from membrane structure to interaction with mRNA fragments, J. Phys. Chem. Lett. 12 (2021) 11199–11205.
[21] M.F.W. Trollmann, R.A. Böckmann, Decoding pH-Driven Phase Transition of Lipid Nanoparticles, 2025, https://doi.org/10.1101/2024.11.27.625717.
[22] M.F.W. Trollmann, R.A. Böckmann, MRNA lipid nanoparticle phase transition, Biophys. J. 121 (2022) 3927–3939.
[23] N.B. Hamilton, S. Arns, M. Shelley, I. Bechis, J.C. Shelley, Calculating apparent pKa values of ionizable lipids in lipid nanoparticles, Mol. Pharm. 22 (2025) 588–593.
[24] B.M. Bruininks, P.C. Souza, H. Ingolfsson, S.J. Marrink, A molecular view on the escape of lipoplexed DNA from the endosome, eLife 9 (2020) e52012.
[25] Siewert J. Marrink, H. Jelger Risselada, Serge Yefimov, D. Peter Tieleman, A. H. de Vries, The MARTINI Force Field: Coarse Grained Model for Biomolecular Simulations, ACS Publications, 2007, https://doi.org/10.1021/jp071097f.
[26] P.C.T. Souza, et al., Martini 3: A general purpose force field for coarse-grained molecular dynamics, Nat. Methods 18 (2021) 382–388.
[27] K.B. Pedersen, et al., The martini 3 lipidome: expanded and refined parameters improve lipid phase behavior, ACS Cent. Sci. 11 (2025) 1598–1610.
[28] L.R. Kjølbye, et al., Martini 3 Building Blocks for Lipid Nanoparticle Design, 2025, https://doi.org/10.26434/chemrxiv-2024-bf4n8-v2.
[29] K.M. Steinegger, et al., Molecular dynamics simulations elucidate the molecular organization of poly(beta-amino ester) based polyplexes for siRNA delivery, Nano Lett. 24 (2024) 15683–15692.
[30] J. Binder, et al., Closing the gap between experiment and simulation—A holistic study on the complexation of small interfering RNAs with polyethylenimine, Mol. Pharm. 21 (2024) 2163–2175.
[31] J. Binder, et al., Enhancing martini3 for protein self-interaction simulations, Eur. J. Pharm. Sci. (2025) 107068, https://doi.org/10.1016/j.ejps.2025.107068.
[32] J. Soni, S. Gupta, T. Mandal, Recalibration of MARTINI-3 parameters for improved interactions between peripheral proteins and lipid bilayers, J. Chem. Theory Comput. 20 (2024) 9673–9686.
[33] J.B. Simonsen, A perspective on bleb and empty LNP structures, J. Control. Release 373 (2024) 952–961.
[34] M.H.Y. Cheng, et al., Induction of bleb structures in lipid nanoparticle formulations of mRNA leads to improved transfection potency, Adv. Mater. 35 (2023) 2303370.
[35] M. Paloncýová, et al., Atomistic insights into organization of RNA-loaded lipid nanoparticles, J. Phys. Chem. B 127 (2023) 1158–1166.
[36] F. Sieber-Schäfer, et al., From bits to bonds: high-throughput virtual screening of ribonucleic acid nanocarriers using a combinatorial approach of machine learning and molecular dynamics, J. Am. Chem. Soc. 147 (2025) 44903–44915.
[37] Pure Substance and Mixture Viscosities Based on Entropy Scaling and an Analytic Equation of State, Industrial & Engineering Chemistry Research, 2026. https://pubs.acs.org/doi/full/10.1021/acs.iecr.7b04871.
[38] A. Jäger, L. Steinberg, E. Mickoleit, M. Thol, Residual entropy scaling for long-chain linear alkanes and isomers of alkanes, Ind. Eng. Chem. Res. 62 (2023) 3767–3791.
[39] C.M. Miles, et al., Unravelling the interactions between small molecules and liposomal bilayers via molecular dynamics and thermodynamic modelling, Int. J. Pharm. 660 (2024) 124367.
[40] Y. Qi, et al., Loading drugs into liposomes by temperature up-down cycle procedure with controllable results fitting prediction by mathematical and thermodynamic process, Mater. Sci. Eng. C 129 (2021) 112379.
[41] W.G. Chapman, K.E. Gubbins, G. Jackson, M. Radosz, SAFT: equation-of-state solution model for associating fluids, Fluid Phase Equilib. 52 (1989) 31–38.
[42] V. Papaioannou, et al., Group contribution methodology based on the statistical associating fluid theory for heteronuclear molecules formed from Mie segments, J. Chem. Phys. 140 (2014) 054107.
[43] J. Gross, G. Sadowski, Perturbed-chain SAFT: an equation of state based on a perturbation theory for chain molecules, Ind. Eng. Chem. Res. 40 (2001) 1244–1260.
[44] M.J.W. Evers, et al., State-of-the-art design and rapid-mixing production techniques of lipid nanoparticles for nucleic acid delivery, Small Methods 2 (2018) 1700375.
[45] D.F. Stieneker, Jet Impingement Reactor, 2023.
[46] M. Ishii, T. Hibiki, Thermo-Fluid Dynamics of Two-Phase Flow, Springer Science & Business Media, 2010.
[47] E.E. Michaelides, M. Sommerfeld, B. van Wachem, Multiphase Flows With Droplets and Particles, Third Edition, CRC Press, Boca Raton, 2022, https://doi.org/10.1201/9781003089278.
[48] S. Subramaniam, Lagrangian–Eulerian methods for multiphase flows, Prog. Energy Combust. Sci. 39 (2013) 215–245.
[49] D. Ramkrishna, A.W. Mahoney, Population balance modeling. Promise for the future, Chem. Eng. Sci. 57 (2002) 595–606.
[50] D.L. Marchisio, R.D. Vigil, R.O. Fox, Quadrature method of moments for aggregation–breakage processes, J. Colloid Interface Sci. 258 (2003) 322–334.
[51] A.R. Thiam, L. Forêt, The physics of lipid droplet nucleation, growth and budding, Biochim. Biophys. Acta Mol. Cell Biol. Lipids 1861 (2016) 715–722.
[52] M.R. Bhole, J.B. Joshi, D. Ramkrishna, CFD simulation of bubble columns incorporating population balance modeling, Chem. Eng. Sci. 63 (2008) 2267–2282.

[53] P. Arosio, S. Rima, M. Lattuada, M. Morbidelli, Population balance modeling of antibodies aggregation kinetics, J. Phys. Chem. B 116 (2012) 7066–7075.
[54] H. Shi, C.J. Mundy, G.K. Schenter, J. Chun, Incorporating the molecular-scale into a hydrodynamic description of confined aqueous systems, J. Chem. Phys. 163 (2025) 134708.
[55] K. Paunovska, 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 (2018) 2148–2157.
[56] A. Akinc, et al., A combinatorial library of lipid-like materials for delivery of RNAi therapeutics, Nat. Biotechnol. 26 (2008) 561–569.
[57] K.T. Love, et al., Lipid-like materials for low-dose, in vivo gene silencing, Proc. Natl. Acad. Sci. 107 (2010) 1864–1869.
[58] D. Chen, et al., Rapid discovery of potent siRNA-containing lipid nanoparticles enabled by controlled microfluidic formulation, J. Am. Chem. Soc. 134 (2012) 6948–6951.
[59] M. Jayaraman, et al., Maximizing the potency of siRNA lipid nanoparticles for hepatic gene silencing in vivo, Angew. Chem. Int. Ed. 51 (2012) 8529–8533.
[60] C.A. Alabi, et al., Multiparametric approach for the evaluation of lipid nanoparticles for siRNA delivery, Proc. Natl. Acad. Sci. 110 (2013) 12881–12886.
[61] K.A. Whitehead, et al., Degradable lipid nanoparticles with predictable in vivo siRNA delivery activity, Nat. Commun. 5 (2014) 4277.
[62] K.J. Kauffman, et al., Optimization of lipid nanoparticle formulations for mRNA delivery in vivo with fractional factorial and definitive screening designs, Nano Lett. 15 (2015) 7300–7306.
[63] B. Li, et al., An orthogonal array optimization of lipid-like nanoparticles for mRNA delivery in vivo, Nano Lett. 15 (2015) 8099–8107.
[64] L. Miao, et al., Delivery of mRNA vaccines with heterocyclic lipids increases anti-tumor efficacy by STING-mediated immune cell activation, Nat. Biotechnol. 37 (2019) 1174–1185.
[65] M.M. Billingsley, et al., Ionizable lipid nanoparticle-mediated mRNA delivery for human CAR t cell engineering, Nano Lett. 20 (2020) 1578–1589.
[66] K. Rajappan, et al., Property-driven design and development of lipids for efficient delivery of siRNA, J. Med. Chem. 63 (2020) 12992–13012.
[67] M.M. Billingsley, et al., Orthogonal design of experiments for optimization of lipid nanoparticles for mRNA engineering of CAR t cells, Nano Lett. 22 (2022) 533–542.
[68] Z. Chen, et al., Modular design of biodegradable ionizable lipids for improved mRNA delivery and precise cancer metastasis delineation in vivo, J. Am. Chem. Soc. 145 (2023) 24302–24314.
[69] Z. He, et al., A multidimensional approach to modulating ionizable lipids for high-performing and organ-selective mRNA delivery, Angew. Chem. Int. Ed. 62 (2023) e202310401.
[70] K. Lam, et al., Unsaturated, trialkyl ionizable lipids are versatile lipid-nanoparticle components for therapeutic and vaccine applications, Adv. Mater. 35 (2023) 2209624.
[71] B. Li, et al., Combinatorial design of nanoparticles for pulmonary mRNA delivery and genome editing, Nat. Biotechnol. 41 (2023) 1410–1415.
[72] R. Palanki, et al., Ionizable lipid nanoparticles for therapeutic base editing of congenital brain disease, ACS Nano 17 (2023) 13594–13610.
[73] W. Dong, et al., Multicomponent synthesis of imidazole-based ionizable lipids for highly efficient and spleen-selective messenger RNA delivery, J. Am. Chem. Soc. 146 (2024) 15085–15095.
[74] B. Li, et al., Accelerating ionizable lipid discovery for mRNA delivery using machine learning and combinatorial chemistry, Nat. Mater. 23 (2024) 1002–1008.
[75] Y. Xu, et al., AGILE platform: A deep learning powered approach to accelerate LNP development for mRNA delivery, Nat. Commun. 15 (2024) 6305.
[76] Y. Zhu, et al., Screening for lipid nanoparticles that modulate the immune activity of helper t cells towards enhanced antitumour activity, Nat. Biomed. Eng. 8 (2024) 544–560.
[77] H. Cui, et al., LUMI-lab: A Foundation Model-Driven Autonomous Platform Enabling Discovery of New Ionizable Lipid Designs for mRNA Delivery, 2025, https://doi.org/10.1101/2025.02.14.638383.
[78] X. Han, et al., Plug-and-Play Assembly of Biodegradable Ionizable Lipids for Potent mRNA Delivery and Gene Editing In Vivo, 2025, https://doi.org/10.1101/2025.02.25.640222.
[79] T. Chen, C. Guestrin, XGBoost: A scalable tree boosting system, in: Proceedings of the 22nd ACM SIGKDD International Conference on Knowledge Discovery and Data Mining, 2016, pp. 785–794, https://doi.org/10.1145/2939672.2939785.
[80] Q. Cheng, et al., Selective organ targeting (SORT) nanoparticles for tissue-specific mRNA delivery and CRISPR–Cas gene editing, Nat. Nanotechnol. 15 (2020) 313–320.
[81] W. Wang, et al., Artificial intelligence-driven rational design of ionizable lipids for mRNA delivery, Nat. Commun. 15 (2024) 10804.
[82] Witten, J. *et al.* Artificial intelligence-guided design of lipid nanoparticles for pulmonary gene therapy. *Nat. Biotechnol.* doi:https://doi.org/10.1038/s41587-024-02490-y (2024) doi:https://doi.org/10.1038/s41587-024-02490-y.
[83] Mehradfar, A. *et al.* LANTERN: A Machine Learning Framework for Lipid Nanoparticle Transfection Efficiency Prediction. Preprint at doi:10.48550/arXiv.2507.03209 (2025).
[84] A.R. Hanna, D.A. Issadore, M.J. Mitchell, High-throughput platforms for machine learning-guided lipid nanoparticle design, Nat. Rev. Mater. (2025) 1–15, https://doi.org/10.1038/s41578-025-00831-0.
[85] R. Zenhausern, et al., Lipid nanoparticle screening in nonhuman primates with minimal loss of life, Nat. Biotechnol. (2025) 1–9, https://doi.org/10.1038/s41587-025-02711-y.
[86] K.L. Swingle, et al., Placenta-tropic VEGF mRNA lipid nanoparticles ameliorate murine pre-eclampsia, Nature 637 (2025) 412–421.
[87] L. Xue, et al., High-throughput barcoding of nanoparticles identifies cationic, degradable lipid-like materials for mRNA delivery to the lungs in female preclinical models, Nat. Commun. 15 (2024) 1884.
[88] R. El-Mayta, et al., A nanoparticle platform for accelerated in vivo oral delivery screening of nucleic acids, Adv. Ther. 4 (2021) 2000111.
[89] C.D. Sago, et al., High-throughput in vivo screen of functional mRNA delivery identifies nanoparticles for endothelial cell gene editing, Proc. Natl. Acad. Sci. 115 (2018) E9944–E9952.
[90] P.P.G. Guimaraes, et al., Ionizable lipid nanoparticles encapsulating barcoded mRNA for accelerated in vivo delivery screening, J. Control. Release 316 (2019) 404–417.
[91] H. Ni, et al., Piperazine-derived lipid nanoparticles deliver mRNA to immune cells in vivo, Nat. Commun. 13 (2022) 4766.
[92] Sanchez, A. J. D. S. *et al.* Universal barcoding predicts in vivo ApoE-independent lipid nanoparticle delivery. *Nano Lett.* doi:https://doi.org/10.1021/acs.nanolett.2c01133 (2022) doi:https://doi.org/10.1021/acs.nanolett.2c01133.
[93] S.G. Huayamares, et al., High-throughput screens identify a lipid nanoparticle that preferentially delivers mRNA to human tumors *in vivo*, J. Control. Release 357 (2023) 394–403.
[94] L.H. Rhym, R.S. Manan, A. Koller, G. Stephanie, D.G. Anderson, Peptide-encoding mRNA barcodes for the high-throughput in vivo screening of libraries of lipid nanoparticles for mRNA delivery, Nat. Biomed. Eng. 7 (2023) 901–910.
[95] U. Odunze, et al., RNA encoded peptide barcodes enable efficient in vivo screening of RNA delivery systems, Nucleic Acids Res. 52 (2024) 9384–9396.
[96] Haley, R. M. *et al.* Lipid nanoparticles for in vivo lung delivery of CRISPR-Cas9 ribonucleoproteins allow gene editing of clinical targets. *ACS Nano* doi:https://doi.org/10.1021/acsnano.4c16617 (2025) doi:https://doi.org/10.1021/acsnano.4c16617.
[97] J.E. Dahlman, et al., Barcoded nanoparticles for high throughput in vivo discovery of targeted therapeutics, Proc. Natl. Acad. Sci. 114 (2017) 2060–2065.
[98] C.D. Sago, et al., Modifying a commonly expressed endocytic receptor retargets nanoparticles in vivo, Nano Lett. 18 (2018) 7590–7600.
[99] C. Dobrowolski, et al., Nanoparticle single-cell multiomic readouts reveal that cell heterogeneity influences lipid nanoparticle-mediated messenger RNA delivery, Nat. Nanotechnol. 17 (2022) 871–879.
[100] H. Kim, et al., Lipid nanoparticle-mediated mRNA delivery to CD34+ cells in rhesus monkeys, Nat. Biotechnol. (2024) 1–8, https://doi.org/10.1038/s41587-024-02470-2.
[101] M.Z.C. Hatit, et al., Species-dependent in vivo mRNA delivery and cellular responses to nanoparticles, Nat. Nanotechnol. 17 (2022) 310–318.
[102] K.C. Wang, et al., A reverse transcription nucleic-acid-based barcoding system for in vivo measurement of lipid nanoparticle mRNA delivery, ACS Bio Med. Chem. Au 5 (2025) 35–41.
[103] Z. Yang, et al., Digital barcodes for high-throughput screening, Chem. Bio. Eng. 1 (2024) 2–12.
[104] K. Vaidya, et al., Pooled nanoparticle screening using a chemical barcoding approach, Angew. Chem. Int. Ed. 64 (2025) e202420052.
[105] A. Merivaara, et al., Barcode lipids for absolute quantitation of liposomes in ocular tissues, J. Control. Release 370 (2024) 1–13.
[106] K.A. Dawson, Y. Yan, Current understanding of biological identity at the nanoscale and future prospects, Nat. Nanotechnol. 16 (2021) 229–242.
[107] A. Akinc, et al., Targeted delivery of RNAi therapeutics with endogenous and exogenous ligand-based mechanisms, Mol. Ther. 18 (2010) 1357–1364.
[108] T. Suzuki, et al., PEG shedding-rate-dependent blood clearance of PEGylated lipid nanoparticles in mice: faster PEG shedding attenuates anti-PEG IgM production, Int. J. Pharm. 588 (2020) 119792.
[109] B.L. Mui, et al., Influence of polyethylene glycol lipid desorption rates on pharmacokinetics and pharmacodynamics of siRNA lipid nanoparticles, Mol. Ther. Nucleic Acids 2 (2013).
[110] S.C. Wilson, et al., Real time measurement of PEG shedding from lipid nanoparticles in serum via NMR spectroscopy, Mol. Pharm. 12 (2015) 386–392.
[111] A. Akinc, et al., The onpattro story and the clinical translation of nanomedicines containing nucleic acid-based drugs, Nat. Nanotechnol. 14 (2019) 1084–1087.
[112] J.A. Kulkarni, et al., On the formation and morphology of lipid nanoparticles containing ionizable cationic lipids and siRNA, ACS Nano 12 (2018) 4787–4795.
[113] F. Sebastiani, et al., Apolipoprotein e binding drives structural and compositional rearrangement of mRNA-containing lipid nanoparticles, ACS Nano 15 (2021) 6709–6722.
[114] H. Tang, et al., Cholesterol modulates the physiological response to nanoparticles by changing the composition of protein corona, Nat. Nanotechnol. 18 (2023) 1067–1077.
[115] K. Liu, et al., Multiomics analysis of naturally efficacious lipid nanoparticle coronas reveals high-density lipoprotein is necessary for their function, Nat. Commun. 14 (2023) 4007.
[116] S.S. Acharya, D.M. Dimichele, Rare inherited disorders of fibrinogen, Haemophilia 14 (2008) 1151–1158.
[117] M. Poley, et al., Nanoparticles accumulate in the female reproductive system during ovulation affecting cancer treatment and fertility, ACS Nano 16 (2022) 5246–5257.
[118] M. Mahmoudi, N. Bertrand, H. Zope, O.C. Farokhzad, Emerging understanding of the protein corona at the nano-bio interfaces, Nano Today 11 (2016) 817–832.
[119] D. Crivellari, et al., Adjuvant pegylated liposomal doxorubicin for older women with endocrine nonresponsive breast cancer who are NOT suitable for a “standard

chemotherapy regimen": the CASA randomized trial, Breast Edinb. Scotl. 22 (2013) 130–137.
[120] M. Poley, et al., Sex-based differences in the biodistribution of nanoparticles and their effect on hormonal, immune, and metabolic function, Adv. NanoBiomed Res. 2 (2022) 2200089.
[121] S. Vadakedath, et al., Immunological aspects and gender bias during respiratory viral infections including novel coronavirus disease-19 (COVID-19): A scoping review, J. Med. Virol. 93 (2021) 5295–5309.
[122] L. Digiacomo, et al., The biomolecular corona of gold nanoparticles in a controlled microfluidic environment, Lab Chip 19 (2019) 2557–2567.
[123] M. Hadjidemetriou, et al., The human in vivo biomolecule corona onto PEGylated liposomes: a proof-of-concept clinical study, Adv. Mater. Deerfield Beach Fla 31 (2019) e1803335.
[124] J. Simon, et al., Unraveling the in vivo protein corona, Cells 10 (2021) 132.
[125] C. Weber, J. Simon, V. Mailänder, S. Morsbach, K. Landfester, Preservation of the soft protein corona in distinct flow allows identification of weakly bound proteins, Acta Biomater. 76 (2018) 217–224.
[126] F. Quattrini, G. Berrecoso, J. Crecente-Campo, M.J. Alonso, Asymmetric flow field-flow fractionation as a multifunctional technique for the characterization of polymeric nanocarriers, Drug Deliv. Transl. Res. 11 (2021) 373–395.
[127] C. Pisani, et al., Experimental separation steps influence the protein content of corona around mesoporous silica nanoparticles, Nanoscale 9 (2017) 5769–5772.
[128] A.A. Ashkarran, et al., Measurements of heterogeneity in proteomics analysis of the nanoparticle protein corona across core facilities, Nat. Commun. 13 (2022) 6610.
[129] A.A. Ashkarran, H. Gharibi, S.M. Modaresi, A.A. Saei, M. Mahmoudi, Standardizing protein corona characterization in nanomedicine: a multicenter study to enhance reproducibility and data homogeneity, Nano Lett. 24 (2024) 9874–9881.
[130] J. Sund, H. Alenius, M. Vippola, K. Savolainen, A. Puustinen, Proteomic characterization of engineered nanomaterial-protein interactions in relation to surface reactivity, ACS Nano 5 (2011) 4300–4309.
[131] N.V. Konduru, et al., Protein corona: implications for nanoparticle interactions with pulmonary cells, Part. Fibre Toxicol. 14 (2017) 42.
[132] D. van Straten, et al., Biofluid specific protein coronas affect lipid nanoparticle behavior in vitro, J. Control. Release 373 (2024) 481–492.
[133] A.Y. Lynn, et al., Investigation of the protein corona and biodistribution profile of polymeric nanoparticles for intra-amniotic delivery, Biomaterials 320 (2025) 123238.
[134] A. Cox, et al., Evolution of nanoparticle protein corona across the blood–brain barrier, ACS Nano 12 (2018) 7292–7300.
[135] K. Yang, C. Reker-Smit, M.C.A. Stuart, A. Salvati, Effects of protein source on liposome uptake by cells: Corona composition and impact of the excess free proteins, Adv. Healthc. Mater. 10 (2021) 2100370.
[136] E. Voke, et al., Protein corona formed on lipid nanoparticles compromises delivery efficiency of mRNA cargo, Nat. Commun. 16 (2025) 8699.
[137] S.A. Dilliard, Q. Cheng, D.J. Siegwart, On the mechanism of tissue-specific mRNA delivery by selective organ targeting nanoparticles, Proc. Natl. Acad. Sci. 118 (2021) e2109256118.
[138] N. Bertrand, et al., Mechanistic understanding of in vivo protein corona formation on polymeric nanoparticles and impact on pharmacokinetics, Nat. Commun. 8 (2017) 777.
[139] K. Ogawara, et al., Pre-coating with serum albumin reduces receptor-mediated hepatic disposition of polystyrene nanosphere: implications for rational design of nanoparticles, J. Control. Release 100 (2004) 451–455.
[140] J. Kreuter, et al., Apolipoprotein-mediated transport of nanoparticle-bound drugs across the blood-brain barrier, J. Drug Target. 10 (2002) 317–325.
[141] M. Lundqvist, et al., Nanoparticle size and surface properties determine the protein corona with possible implications for biological impacts, Proc. Natl. Acad. Sci. 105 (2008) 14265–14270.
[142] Y.-W. Yin, Y.-Q. Ma, H.-M. Ding, Effect of nanoparticle curvature on its interaction with serum proteins, Langmuir 40 (2024) 15205–15213.
[143] R. Zhang, et al., Helper lipid structure influences protein adsorption and delivery of lipid nanoparticles to spleen and liver, Biomater. Sci. 9 (2021) 1449–1463.
[144] S.A. Dilliard, et al., The interplay of quaternary ammonium lipid structure and protein corona on lung-specific mRNA delivery by selective organ targeting (SORT) nanoparticles, J. Control. Release 361 (2023) 361–372.
[145] D. Chen, N. Parayath, S. Ganesh, W. Wang, M. Amiji, The role of apolipoprotein-and vitronectin-enriched protein corona on lipid nanoparticles for in vivo targeted delivery and transfection of oligonucleotides in murine tumor models, Nanoscale 11 (2019) 18806–18824.
[146] G. Caracciolo, et al., Selective targeting capability acquired with a protein corona adsorbed on the surface of 1,2-Dioleoyl-3-trimethylammonium propane/DNA nanoparticles, ACS Appl. Mater. Interfaces 5 (2013) 13171–13179.
[147] B. Singh, C. Fu, J. Bhattacharya, Vascular expression of the αvβ3-integrin in lung and other organs, Am. J. Physiol.-Lung Cell. Mol. Physiol. 278 (2000) L217–L226.
[148] J. Simon, et al., Exploiting the biomolecular corona: pre-coating of nanoparticles enables controlled cellular interactions, Nanoscale 10 (2018) 10731–10739.
[149] Z. Zhong, et al., Quantitative analysis of protein corona on precoated protein nanoparticles and determined nanoparticles with ultralow protein corona and efficient targeting in vivo, ACS Appl. Mater. Interfaces 13 (2021) 56812–56824.
[150] Q. Peng, et al., Preformed albumin corona, a protective coating for nanoparticles based drug delivery system, Biomaterials 34 (2013) 8521–8530.
[151] H. Li, et al., Regulation of protein corona on liposomes using albumin-binding peptide for targeted tumor therapy, J. Control. Release 355 (2023) 593–603.
[152] K.-R. Kim, J. Kim, J.H. Back, J.E. Lee, D.-R. Ahn, Cholesterol-mediated seeding of protein corona on DNA nanostructures for targeted delivery of oligonucleotide therapeutics to treat liver fibrosis, ACS Nano 16 (2022) 7331–7343.
[153] Z. Zhang, et al., Brain-targeted drug delivery by manipulating protein corona functions, Nat. Commun. 10 (2019) 3561.
[154] S.R.J. Hofstraat, et al., Nature-inspired platform nanotechnology for RNA delivery to myeloid cells and their bone marrow progenitors, Nat. Nanotechnol. 20 (2025) 532–542.
[155] S. Schöttler, K. Landfester, V. Mailänder, Controlling the stealth effect of nanocarriers through understanding the protein corona, Angew. Chem. Int. Ed. Eng. 55 (2016) 8806–8815.
[156] S. Palchetti, et al., Nanoparticles-cell association predicted by protein corona fingerprints, Nanoscale 8 (2016) 12755–12763.
[157] J. Lazarovits, et al., Supervised learning and mass spectrometry predicts the in vivo fate of nanomaterials, ACS Nano 13 (2019) 8023–8034.
[158] Y. Duan, et al., Prediction of protein corona on nanomaterials by machine learning using novel descriptors, NanoImpact 17 (2020).
[159] J. Huzar, R. Coreas, M.P. Landry, G. Tikhomirov, AI-based prediction of protein corona composition on DNA nanostructures, ACS Nano 19 (2025) 4333–4345.