

In Vivo CAR-T Cell Engineering: Translating Vector Pharmacology into Therapeutic Performance

Jiaqiang Ren *, David Stroncek and Anh Dinh

Department of Transfusion Medicine and Center for Cellular Engineering, Clinical Center, National Institutes of Health, Bethesda, MD 20892, USA

* Correspondence: jiaqiang.ren@nih.gov

How To Cite: Ren, J.; Stroncek, D.; Dinh, A. *In Vivo* CAR-T Cell Engineering: Translating Vector Pharmacology into Therapeutic Performance. *Translational Insights* 2026, 1(1), 19. <https://doi.org/10.53941/ti.2026.100019>

Received: 11 June 2026

Revised: 27 August 2026

Accepted: 2 September 2026

Published: 10 September 2026

Abstract: *In Vivo* CAR-T cell engineering delivers chimeric antigen receptor (CAR) molecules transgenes directly to endogenous T cells using lentiviral vectors or targeted lipid nanoparticles, bypassing the manufacturing processes of *ex vivo* cellular therapies. Lentiviral vectors provide integration-based expression for sustained tumor control, whereas lipid nanoparticles enable transient, repeat-dosable therapies for autoimmune disease. First-in-human clinical trials have confirmed proof-of-concept: CAR-T cells expand, deplete target cells, and generate clinical responses, including negative minimal residual disease in multiple myeloma and induce B-cell depletion in systemic lupus erythematosus. However, persistence remains substantially shorter than *ex vivo* CAR T cell products. This review presents platform architecture, preclinical data, and emerging clinical evidence. We propose a Quality-by-Design (QbD) framework—a systematic, science- and risk-based development approach that begins with predefined product objectives and links critical quality attributes to material attributes and process controls—integrating vector pharmacology with immune parameters and discuss development pathways that emphasize platform–indication matching, innate immune evasion, and adjunctive T-cell conditioning to address the persistence gap and safety control.

Keywords: chimeric antigen receptor; CAR; CAR-T; *in vivo* CAR-T; quality by design

1. Introduction

Chimeric antigen receptor (CAR) T-cell therapies have achieved durable remissions in relapsed/refractory B-cell malignancies and multiple myeloma [1]. Conventional autologous CAR-T treatment is time-consuming and costly. *In vivo* engineering eliminates patient specific *ex vivo* cell manufacture, and lower total treatment cost remains a projected potential. *In vivo* engineering addresses these barriers by delivering CAR transgenes to endogenous T cells via systemically administered vectors, thereby eliminating the need for laboratory-based cell manufacturing.

In vivo CAR-T engineering integrates principles from *in situ* cell reprogramming in gene therapy and regenerative medicine with clinical experiences from *ex vivo* CAR-T therapies. In gene therapy and regenerative medicine, *in situ* (*in vivo*) reprogramming aims to directly convert resident cells within their native tissue to restore function, avoiding the need for cell harvest, *ex vivo* expansion, and reinfusion or transplantation [2–4]. *In vivo* CAR-T extends this approach to the immune system by converting circulating or tissue-resident T cells into CAR-T cells within the patients.

This strategy combines: (1) construct design, manufacturing control, and safety principles from *ex vivo* CAR-T development, where vector selection, CAR architecture, and critical quality attributes (CQA) have been extensively characterized; and (2) delivery, targeting, and transient reprogramming strategies from *in situ* gene



and cell engineering in regenerative medicine. The vector pharmacology along with the host immune parameters, and tissue microenvironments collectively control the performance of CAR-T cells, such as generation, expansion, and persistence.

Two *in vivo* T-cell targeted platforms have entered clinical trial: lentiviral vectors (LVV) that provide integration-based persistent expression, and lipid nanoparticles (LNP) that deliver messenger RNA (tLNP-mRNA) based transient expression [1,5–8]. Both platform have provided proof-of-concept (POC) data in first-in-human (FIH) studies, including ESO-T01 (LVV for multiple myeloma) [7], HN2301 (tLNP for systemic lupus erythematosus (SLE)) [8], and recent KLN-1010 (LVV for multiple myeloma) [9].

Unlike the *ex vivo* manufacturing processes that are controlled by unit operations, *in vivo* CAR-T cells are generated in the patients' body. Besides vector pharmacology, the patient's immune system, e.g., patient-specific T-cell fitness (exhaustion state, differentiation status), controls the T-cell activation (if any), transduction, and expansion. Vector pharmacokinetics (clearance by mononuclear phagocytes, enzymatic degradation, etc.), immune activation (innate sensing, antibody neutralization) are also determinants of therapeutic performance.

The central translational challenge is the reduced persistence relative to *ex vivo* products—weeks versus months. While transient CAR-T persistence may be advantageous for autoimmune indications, long-term persistence is associated with durable responses for oncologic indications. This persistence gap necessitates platform–indication matching and engineering strategies that account for immune heterogeneity as determinants of CAR expression, expansion kinetics, and therapeutic potential.

Acute toxicity caused by *in vivo* vectors represents a significant constraint in dose optimization. In the absence of clearly defined efficacy thresholds, toxicity-mediated dose limitations result in narrow therapeutic windows. For example, the ESO-T01 trial reported acute toxicity requiring tocilizumab intervention [7], although KLN-1010 demonstrated an improved safety profile [9]. In an early-phase clinical trial, the tLNP platform also produced acute toxicity in HN2301 trial participants [8]. The L829 formulation of tLNP-CD20 CAR mRNA caused acute toxicity, with one non-human primate (NHP) developing severe hemophagocytic lymphohistiocytosis [10].

2. Platform Architectures and Mechanistic Principles

Platform selection for *in vivo* T-cell engineering is guided by the core product requirements. Key considerations include which T-cell subset to target (CD4-positive or CD8-positive cells, or defined populations such as naïve and central-memory cells). Additional factors include the desired expression profile (onset kinetics and durability), tissue tropism (emphasizing T cell tropism while limiting hepatic uptake and tropism with other tissues), feasibility of repeat dosing, and safety requirements (including acute toxicity, on-target but off-tumor side effects, cytokine-release events, and insertional mutagenesis risk for integrating vectors).

In vivo-generated CAR-T cells engineered with lentiviral vectors exhibit delayed expression kinetics, with CARs expressed about 4–8 days after administration, peaking after antigen-driven expansion (around 10–20 days), and persisting for weeks to months. In high antigen burden settings, the expansion of *in vivo* CAR-T cells functionally mirrors that of *ex vivo* manufactured CAR-T cells post-infusion (Figure 1a). In contrast, targeted lipid nanoparticle (tLNP) platforms deliver *in vitro*-transcribed CAR-encoding mRNA, resulting in rapid but transient expression. CAR protein can be detected within roughly 4–6 h, reaches maximal levels at around 12–24 h, and declines toward baseline by about 48–72 h as the mRNA payload is degraded and cleared (Figure 1a). These different temporal profiles are related to the difference of the biological activities of the platforms modalities: lentiviral vectors mediate stable genomic integration and durable expression of the CAR transgene, whereas tLNP-delivered mRNA remains cytoplasmic and results in temporary translation without genome modification (Figure 1b). Furthermore, tLNP-driven CAR expression is limited by endogenous ribonuclease activity and dilution with successive cell division.

2.1. Lentiviral Vectors: Integration-Based Durability

LVV enables stable genomic integration, with T-cell selectivity achieved through envelope pseudotyping. By utilizing the modular architecture of paramyxovirus glycoproteins (such as measles or Nipah virus), researchers can decouple the receptor-binding head domain from the fusion-triggering stalk domain. Replacing the head domain with targeting moieties such as single-chain variable fragments (scFvs) or designed ankyrin repeat proteins (DARPs) allow for precise cell entry without losing fusion competence. Unlike pH-dependent systems, these vectors enter via direct plasma membrane fusion, which minimizes off-target transduction, avoids endosomal degradation or macrophage sequestration, and enhances the transduction of quiescent cells [7,11–15].

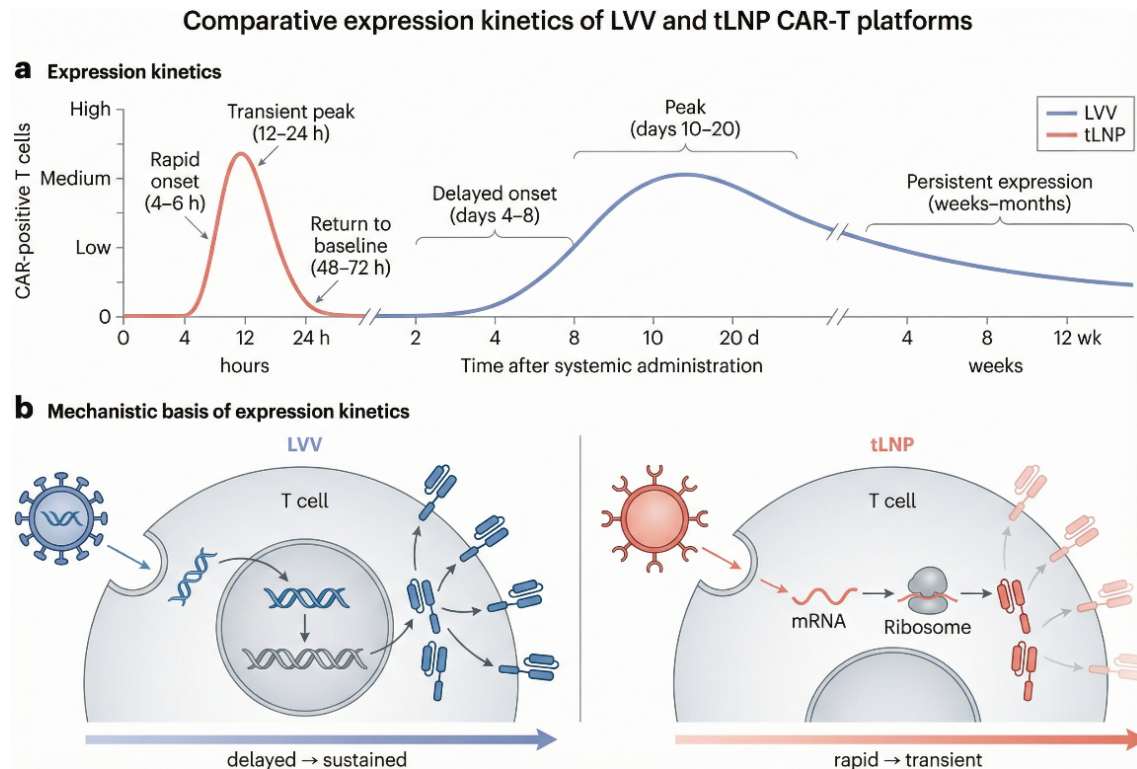


Figure 1. Comparative expression kinetics of lentiviral vector and tLNP-mRNA CAR-T cell platforms. Time course of CAR expression following systemic administration of lentiviral vectors (blue) versus targeted lipid nanoparticles (red). Panel (a) shows schematic representation of CAR-positive T-cell frequency over time. Lentiviral vectors exhibit delayed onset (days 4–8), peak expression at days 10–20, and persistence extending weeks to months. Targeted lipid nanoparticles show rapid onset (4–6 h), peak at 12–24 h and return to baseline by 48–72 h. Panel (b) illustrates the mechanistic basis for the distinct kinetics. LVV integrates the CAR molecules into the T cell genome and results in delayed and sustained CAR expression. tLNP-mRNA delivers mRNA into T cell cytoplasm and result in transient translation, yielding rapid and self-limiting CAR expression.

The selection of targets is based on the applications. Specifically, CD3-targeted vectors, such as TR66-LV, provide uniform transduction for both CD4⁺ and CD8⁺ T-cells [12]; CD8-specific vectors have the potential to treating special T-cell lymphomas. By delivering anti-CD4 CARs exclusively to CD8⁺ cells, researchers can eliminate malignant CD4⁺ cells while preventing fratricide among the engineered effector cells. This strategy was validated in a preclinical mouse model of angioimmunoblastic T-cell lymphoma (AITL), where it achieved 80% survival and a 70% depletion of neoplastic CD4⁺/PD-1 cells [13]. Lentiviral vectors targeting L-selectin (CD62L) to preferentially transduce naïve and early memory T cells have also been developed [14].

Advanced engineering platforms enhance functionality. The Delivery to Intended REcipient Cells Through Envelope Design (DIRECTED) system achieves programmable targeting of lentiviral vectors to specific cell types through antibody recruitment or covalent immobilization on vesicular stomatitis virus glycoprotein G (VSV-G) variants with ablated tropism and retained fusion capability [16]. The VivoVec platform incorporates T-cell activation (anti-CD3) and costimulatory molecules (CD80, CD58) directly onto viral particles, which are pseudotyped with Cocal fusion glycoproteins that resist serum inactivation in humans and improve *in vivo* persistence [17,18]. The vector also includes a rapamycin-activated cytokine receptor system, which promotes selective expansion of CAR T cells while suppressing non-transduced immune cells in the presence of exogenous rapamycin, thereby eliminating the need for lymphodepletion. In non-human primates (NHP), VivoVec resulted in the production of 65% anti-CD20 CAR-positive T cells 10 days following intranodal administration and sustained B-cell depletion more than 70 days post-administration [17]. Integration site analysis confirmed transgenes remained sequestered in injected and downstream lymph nodes for months, with minimal off-target tissue distribution.

The KLN-1010 platform utilizes a third-generation lentiviral vector modified with VSV-G fusogen and an anti-CD3 antibody on the envelope. This architecture enables viral entry specifically through CD3 on T cells [9]. KLN-1010 was used to generate *in vivo* anti-B-cell maturation antigen (BCMA) CAR T cells to treat relapsed and refractory multiple myeloma without lymphodepletion in a Phase I clinical trial (NCT07075185) [9]. Preliminary

findings indicate early MRD-negative responses, persistent memory CAR T cells, and toxicities comparable to those of *ex vivo* anti-CAR T cells.

Multiple designs incorporate immune evasion features, such as CD47 overexpression that suppresses phagocytosis and MHC-I knockout in producer cells to reduce immunogenicity [7,19]. In non-human primates, these modifications resulted in increased serum persistence, high on-target transduction, and minimal off-target biodistribution [11,17,20].

2.2. Lipid Nanoparticles: Transient Programmable Expression

Current lipid nanoparticle formulations incorporate ionizable lipids that facilitate endosomal escape [10], polyethylene glycol (PEG) lipids for stability, and surface-conjugated antibodies (anti-CD3, anti-CD8, anti-CD4, or anti-CD5) that direct biodistribution toward T-cells. Antibody functionalization may be achieved using N-succinimidyl-S-acetylthioacetate (SATA)-maleimide chemistry to enhance T cell binding and endocytosis [5]. Optimal ionizable cationic lipids exhibit acid dissociation constant (pKa) values of 6.2–6.5, thus maintaining neutral surface charge at physiological pH (avoiding non-specific immune recognition) while becoming positively charged in acidic endosomal environments (pH < 6.0) to facilitate membrane fusion and cargo release [21].

Strategic design of ionizable cationic lipids

Ionizable cationic lipid chemistry is a primary determinant of lipid nanoparticle (LNP) tropism, potency, and safety. For instance, the L829 ionizable lipid was designed to minimize hepatic sequestration and accelerate clearance from non-target tissues when formulated as an antibody-decorated particle (e.g., anti-CD8 moiety). Steric shielding can further reduce engagement with hepatocyte uptake pathways, including LDL receptor-mediated internalization. Collectively, these features were reported to reduce off-target liver delivery and improve relative T-cell targeting compared with benchmark ionizable lipids [10]. Shenzhen MagicRNA has reported an LNP formulation incorporating the ionizable lipid ILB-3132 to systemically deliver CD19 CAR-encoding cargo for the treatment of SLE [8]. Although transient expression has a safety advantage through reversibility, a key benchmark for lipid nanoparticle platforms will be potentially enabling repeat dosing on clinically practical schedules (e.g., every couple of weeks) rather than relying on frequent re-administration at approximately 72-h intervals [22].

Payload architecture

Five-component formulations (ionizable lipid, helper lipids, PEG-lipid, cholesterol, and targeting moiety) are defined by particle size and polydispersity index, encapsulation efficiency, nitrogen-to-phosphate ratio, and surface antibody density. Systematic modification of lipid structure—including hydrophobic chain length, branching, structural isomerism, and degree of saturation—led to increased T-cell transfection efficiency compared with unoptimized formulations [23]. Optimizing lipid composition and formulation parameters can improve circulation lifetime, gene expression, mRNA integrity, and transfection efficiency [5,24].

mRNA optimization

To reduce innate RNA sensing and improve translation, many therapeutic mRNA formulations replace uridine with N1-methylpseudouridine [25]. This approach mirrors structural optimization in SARS-CoV-2 vaccines that overcome inherent mRNA instability and immunogenicity. In addition, rigorous purification removes double-stranded RNA (dsRNA) contaminants to less than 0.1%—a critical process parameter that prevents cytokine storms and ensures translational efficiency [26]. Furthermore, mRNA sequence optimization, including diverse untranslated region sequences and codon usage combinations, has been shown to augment CAR activity [10,27]. Manufacturing efforts involve microfluidic mixing to leverage a high-speed combination of organic lipid and aqueous mRNA phases to achieve uniform size distribution and reproducibility [28].

Targeting ligands

CD8-targeted lipid nanoparticle systems can achieve high specificity. In humanized NSG-PBMC mouse model, CD8-targeted tLNP encoding anti-CD19 CAR mRNA generated specific and efficient CD8+ CAR T cells, leading to rapid and profound B cell depletion in blood and tissues [29]. In cynomolgus monkeys, CD8-targeted tLNP encoding surrogate anti-CD20 CAR mRNA produced robust and transient CAR-engineering of up to 80% of CD8+ T cells resulting in peripheral B cell depletion and profound B cell depletion in tissues [30].

CD3-targeted lipid nanoparticles were most effective for *in vivo* T-cell targeting compared with CD5-targeted or CD7-targeted formulations [5]. CAR T cells produced *in vivo* using lipid nanoparticles targeting CD8 demonstrated polyfunctional cytokine production and antigen-specific killing exceeding 90% *in vitro* [23].

Unlike integrating vectors, mRNA systems avoid insertional mutagenesis, enable temporal control through dosing schedules, and permit iterative CAR clinical trial design. CAR expression drives rapid antigen-induced

activation and cytotoxicity, but because the RNA payload is non-integrating, CAR expression decays with rapid T-cell division [1,8]. Repeated dosing maintains the desired cytotoxicity without inducing T-cell dysfunction, validating the feasibility of this dosing strategy [23]. However, it has been suggested that CAR constructs should be optimized specifically for transient expression to achieve rapid therapeutic effects, which differs from designs that have been optimized for long-term persistence in *ex vivo* applications [31].

2.3. Innate and Adaptive Immune Responses to Systemic Vector Administration

Both lentiviral and lipid nanoparticle platforms elicit innate and adaptive immune responses that affect safety, limit systemic exposure, and constrain effective dosing. Innate immune responses, in particular, remain a significant barrier to *in vivo* gene therapy development. Clinical trials using systemic administration of lentivirus remain limited, but data from adeno-associated virus (AAV) studies provide insights into immune responses.

Mechanisms of early immune recognition and vector clearance

Upon systemic administration, LVV and LNPs may trigger rapid innate immune responses that significantly limit transduction efficiency and generate side effects.

Within minutes to hours: Complement activation drives opsonization and uptake by the mononuclear phagocyte system—hepatic Kupffer cells, splenic macrophages, and circulating monocytes—via Fc, complement, and scavenger receptors [32]. In studies of CD3-targeted nanocarriers, approximately 20% of both macrophages and monocytes were positive for the carriers, indicating that phagocytic activity remains a major contributor to particle clearance regardless of targeting [32]. Simultaneously, membrane-associated TLR2/TLR4 on monocytes and macrophages detect viral capsid proteins and envelope glycoproteins, triggering MyD88-dependent NF- κ B activation and secretion of IL-6, IL-1 β , and TNF- α [33,34] (Figure 2a).

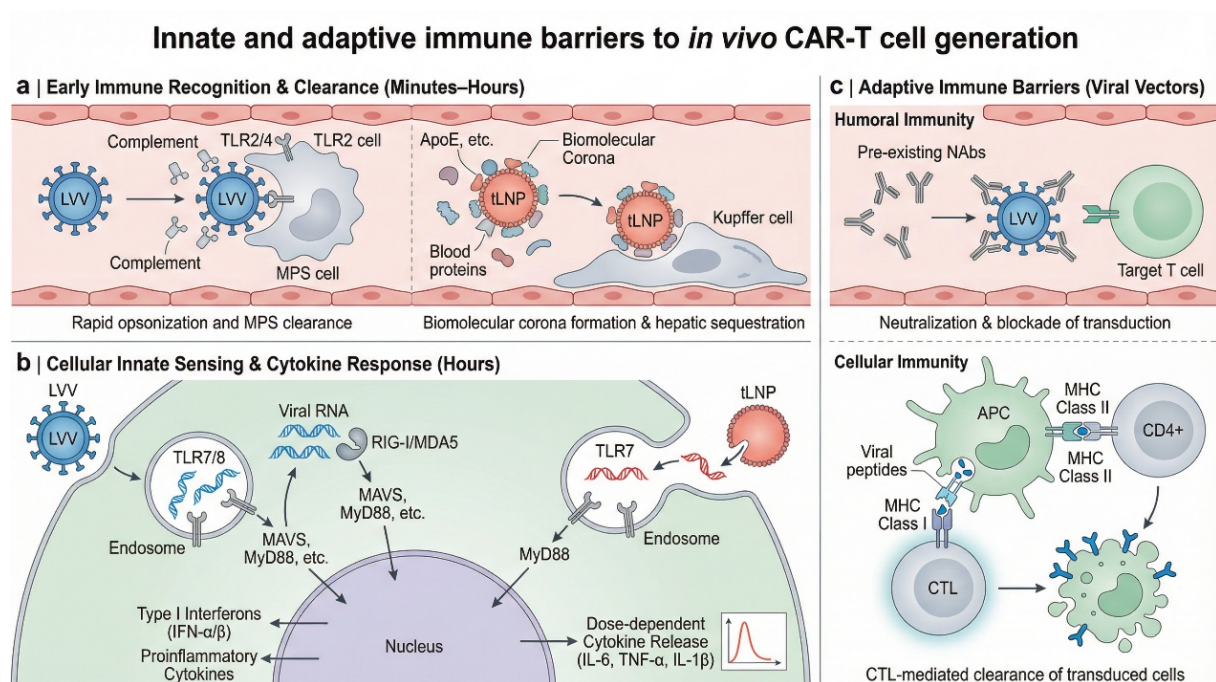


Figure 2. Innate and adaptive immune barriers to *in vivo* CAR-T cell generation. Panel (a) Early immune recognition and vector clearance. LVVs and tLNPs cause rapid innate immune responses following administration. Within minutes, complements opsonize LVVs, priming them for receptor-mediated uptake and clearance by the mononuclear phagocyte system (MPS). tLNPs adsorb plasma proteins to establish a biomolecular corona enriched with apolipoproteins (such as ApoE), which facilitates their sequestration and clearance. Panel (b) Cellular innate sensing and cytokine response. Following endocytosis, internalized vectors are sensed by intracellular pattern recognition receptors (PRRs). The RNAs of LVV are detected by endosomal Toll-like receptors (TLR7/8) and cytosolic sensors (RIG-I/MDA5), which activate signaling through MyD88 and MAVS adaptors, respectively, culminating in robust production of type I interferons (IFN- α/β) and proinflammatory cytokines. For tLNP, the ionizable lipids and mRNA trigger endosomal TLR7, leading to release of proinflammatory cytokines such as IL-6, TNF- α , and IL-1 β . Panel (c) Adaptive immune barriers to viral vectors. Pre-existing neutralizing antibodies can bind to the viral capsid, blocking cellular entry and neutralizing the vector. Professional antigen-presenting cells

can process viral capsid proteins and present derived peptides on MHC class I and II molecules, leading to activation of CD8⁺ cytotoxic T lymphocytes, which can identify and eliminate successfully transduced cells.

Upon systemic administration, LNPs interact with blood components to form a biomolecular corona comprised of apolipoproteins, complement elements, and immunoglobulins. Corona composition—specifically Apolipoprotein E (ApoE) adsorption—facilitates natural liver targeting via LDL receptors, while hepatic Kupffer cells further sequester nanoparticles, reducing therapeutic efficiency [35,36]. This is modulated by PEG shedding, defined as the dissociation of PEG-lipids during circulation [21]. Designing T-cell-targeted LNPs for *in vivo* CAR-T therefore requires strategic lipid engineering to overcome inherent hepatotropism. Optimized ionizable lipid chemistries (such as L829) reduce ApoE binding and hepatocyte uptake [10], while LNP surface antibody conjugation redirects biodistribution from liver toward T cells (Figure 2a).

Within hours: Endosomal Toll-like receptors (TLR) become dominant sensors. For RNA-based vectors (LVV, mRNA-lipid nanoparticle), plasmacytoid dendritic and myeloid cell TLR7/8 recognize single-stranded RNA and GU/U-rich motifs, resulting in the activation of MyD88-dependent pathways that drive IFN- α/β and proinflammatory cytokine production [33,34]. Viral components and impurities derived from preparations containing residual producer-cell components can trigger these TLR and cytosolic sensors, leading to the release of IL-6, TNF- α , and type I interferons [33,34]. Cytosolic RNA sensors (RIG-I and MDA5) recognize short dsRNA/5'-triphosphate RNA and long dsRNA, respectively, interacting with the mitochondrial adaptor MAVS to activate TBK1 and drive type I interferon responses [37,38] (Figure 2b). When LVVs are administered intravenously, they induce IFN- α/β production in the liver and spleen within hours. This production of cytokines is correlated with sharp declines in proviral DNA and reduced transduction [39].

Whole transcriptome analysis has revealed that CAR-negative cells exhibit prominent antiviral defense signatures with upregulation of interferon-induced transmembrane protein 1, 2, and 3 (IFITM1/2/3) restriction factors [40]. Rapamycin treatment downmodulates these factors, thereby enhancing CD8- and CD4-LVV transduction [41,42]. In humanized mice with pronounced myeloid engraftment, CAR-T cell generation efficiency decreased significantly compared to standard models, directly attributable to macrophage-mediated vector consumption. CD47-high and MHC-I-low envelope engineering can overcome this barrier, improving transduction 2- to 3-fold [19].

Targeted lipid nanoparticles activate ionizable lipid-triggered complement, endosomal TLR7/8 (particularly from unmodified uridine or dsRNA contaminants), and hepatic Kupffer cells, which release IL-6, TNF- α , and IL-1 β [43,44]. In non-human primate studies, cytokine elevations followed dose-dependent kinetics. Transient increases peaked within 6–24 h at lower LNP doses but produced sustained elevations and immune effector cell-associated hemophagocytic lymphohistiocytosis-like syndrome at doses of 1.5–2.0 mg/kg. One animal receiving three 1.5 mg/kg doses developed a moribund state with hepatitis, cytopenias, coagulopathy, hyperferritinemia (exceeding 10,000 ng/mL), and cytokine storm [10].

Adaptive immune barriers

Experience from AAV preclinical and clinical studies also provides insight into immune responses to LVV. Pre-existing humoral immunity is a critical barrier, particularly for viral vectors (Figure 2c). A substantial proportion of the human population harbors neutralizing antibodies against common viral serotypes from prior natural infections [45]. Even at low titers, pre-existing neutralizing antibodies can significantly reduce transduction efficiency by binding to the vector capsid and blocking cellular entry or promoting rapid clearance via Fc receptor-mediated phagocytosis [46,47].

Beyond humoral immunity, adaptive T-cell responses play a central role in clearance. Professional antigen-presenting cells (APC) that take up vector particles process and present capsid-derived peptides on MHC class I and class II, activating capsid-specific CD8⁺ cytotoxic T lymphocytes and CD4⁺ helper T cells (Figure 2c) [48–50]. Capsid-specific cytotoxic T lymphocytes can directly lyse transduced cells, as observed in clinical trials where transient elevation of liver enzymes has been linked to cytotoxic T lymphocyte-mediated destruction of transduced hepatocytes, thereby eliminating the source of available transgene and reducing the effective vector dose over time [51].

2.4. Patient T-Cell Fitness: The Variable Substrate

There is a limited understanding of the quality of *in vivo*-engineered CAR-T cells, but data from *ex vivo*-manufactured CAR-T cells regarding *in vivo* expansion and cytotoxicity can be informative. Pre-infusion CTL019 CAR-T cell products enriched for inhibitory/exhaustion phenotypes (PD-1 with LAG-3 or TIM-3 on CD8⁺ CAR-T cells) are associated with poor clinical responses. CAR-T cell expression of these receptors is widely interpreted as indicating reduced cell fitness, which contributes to inferior expansion and persistence [52]. A large multicenter

clinical dataset also links robust *in vivo* expansion with a CAR T-cell phenotype enriched for central memory and depleted for exhausted PD-1⁺LAG-3⁺TIM-3⁺ CAR⁺CD8⁺ populations at peak expansion [53]. A study that showed patient-derived CAR-T cells are highly sensitive to stimulation conditions, with increased PD-1/TIM-3 co-expression and substantial phenotype shifts and CAR expression changes, demonstrated a clear relationship between stimulation dose, the phenotype of the T cells prior to activation, and the functionality of the resulting CAR-T cell products [54]. Heavily pretreated patients have high percentages of exhausted T cells, and the impact of these cells on the quality attributes of *in vivo*-engineered T cells and associated pharmacological effects warrants further study. Product development strategies may benefit from baseline immune profiling (exhaustion markers, differentiation status, inflammatory signatures), refined patient selection criteria, and potentially preparative conditioning to optimize the T-cells before vector administration.

3. Preclinical Validation: Cross-Species Data

Preclinical studies involving rodents, humanized mice, and NHP have revealed consistent pharmacological patterns, enabling the translation of *in vivo* gene therapy to first-in-human trials.

3.1. Lentiviral Vector Platforms

Early CD8-targeted LVV studies in humanized mice established proof-of-concept [55,56]. Single-dose administration (2×10^6 transduction units, intraperitoneal) yielded CD19 CAR⁺CD8⁺ T cells in 60% of animals (10–35% of peripheral CD8⁺ cells), achieved B-cell depletion (70%), and completely eliminated systemic Nalm-6 tumors (88%). Plasma cytokine elevations (IL-6, IFN- γ , GM-CSF) mirrored *ex vivo* CAR-T cell cytokine release syndrome (CRS) signatures, with two high-responder animals developing CRS-like symptoms [55].

The LVV SG299 platform demonstrated selective biodistribution in NHP: quantifiable vector copy number (VCN) persisted in CD8-enriched peripheral blood mononuclear cells through day 34 and in spleen and lymph nodes at day 90, with undetectable VCN in liver, testis, kidneys, and heart. *In vitro* analysis confirmed high VCN in isolated CD8⁺ T cells with no detection in CD4⁺ T cells or B cells [11,20].

The delivery of a CD20 CAR transgene by the INT2104 LVV platform achieved B-cell depletion by day 7 and complete tumor ablation across a 15-fold dose range in humanized mice [57]. In cynomolgus macaques, single doses reduced peripheral B cells by day 4 (84% and 53%) with CAR expression restricted to CD3⁺ T cells and NK cells in spleen and liver—no detection in B cells or hepatocytes.

KLN-1010, an anti-BCMA iGPS vector, produced anti-tumour activity in preclinical multiple myeloma models. For translational NHP studies, the platform used a surrogate anti-CD20 CAR and a single dose produced complete peripheral B-cell depletion lasting more than two months without lymphodepletion [58,59].

3.2. Targeted Lipid Nanoparticle Platforms

In murine models, tLNP-delivered CAR mRNA produces rapid, transient CAR expression over hours to days [60–62]. In cynomolgus monkeys, repeat dosing with CD8-targeted tLNP generated transient CAR-positive lymphocytes and rapid B-cell depletion [10].

In comprehensive NHP studies, the administration of 1.5–2.0 mg/kg across three doses of CD8-targeted L829-LNPs encoding anti-CD20 CAR mRNA resulted in up to 85% CAR expression in circulating CD8⁺ T cells and 95% in CD8⁺ NK cells [10,29]. Peripheral B cells declined sharply within 6 h, approaching undetectable levels by 24 h, with >95% depletion in the blood, spleen, and lymph nodes. B-cell recovery began around day 21 and returned to near-baseline levels by day 35. Reconstituted populations exhibited predominantly naïve phenotype (IgD⁺CD27⁺CD10⁺ transitional) with marked memory depletion, consistent with the immune reset observed with *ex vivo* CAR-T [10,63]. A two-dose (every 72 h) regimen achieved a pharmacological profile comparable to three doses with lower inflammatory biomarkers, supporting scheduling flexibility [10].

Optimized CD8-targeting systems demonstrated that T cells maintained cytotoxicity under repeated transduction without dysfunction. CAR expression remained limited to CD8-positive cells (T and NK) with no detection in CD4-positive T cells, B cells, or granulocytes [23,64]. LNP-transduced T cells from peripheral blood mononuclear cells of a patient with lymphoma showed generally increased transduction levels, with CAR expression greatest in central memory and effector memory phenotypes and decreased in naïve T cells [64].

Anti-CD4-targeted tLNP studies provide complementary biodistribution data. In C57BL/6 mice, anti-CD4-conjugated LNPs were concentrated in the spleen, with blood activity falling over 24 h. Spleen luciferase signal was approximately 7-fold higher than the non-targeted LNPs, and CD3-positive T-cell luminescence was approximately 33-fold higher, with reduced liver signal compared to controls, demonstrating lymphoid-favored distribution [65].

3.3. Mechanistic Convergence and Dose Rationale

Despite fundamental mechanistic differences, LVV and tLNP platforms share important features. First, optimal therapeutic performance depends on the availability of sufficient high-quality T cells. T-cell exhaustion negatively impacts therapeutic outcomes by reducing CAR expression, proliferative capacity, and long-term persistence. Early innate immune responses that can trigger acute, dose-limiting toxicities are a primary safety concern. Despite these challenges, the available pharmacokinetic and pharmacodynamic (PK/PD) data in pre-clinical studies facilitated the establishment of human dosing.

The minimal anticipated biological effect level (MABEL) framework was used to establish the FIH dosing for the L829-tLNP platform. In cynomolgus monkeys, MABEL was defined as the threshold for near-complete B-cell depletion. Pharmacokinetic analysis confirmed dose-proportional exposure and rapid lymphoid distribution. While body-surface-area (BSA) scaling suggested a human equivalent dose (HED) of 0.032 mg/kg, a 10-fold safety factor was applied to mitigate safety uncertainties. This resulted in a starting dose of approximately 0.003 mg/kg administered on days 0 and 3 [10].

For ESO-T01, dose selection was based on the safe and effective dose in humanized mice, as well as doses from similar products in NHP studies, and the starting dose was set at one-tenth of the human-equivalent dose calculated from the effective mouse dose by body weight. Four patients received a single intravenous dose of 2.0×10^8 transduction units [7].

4. Clinical Translation: Early-Stage Validation and Persistence as A Translational Challenge

4.1. tLNP-mRNA in Autoimmune Disease: Immune Reset

Autologous *ex vivo* CD19-directed CAR-T cell therapy demonstrates robust efficacy in severe or refractory autoimmune disorders, including SLE, systemic sclerosis, and myasthenia gravis. Clinical data indicate that a single infusion can induce high remission rates and the clinical responses can be sustained for 2.5 to 4 years, and persisted well beyond the typical 3- to 6-month window of B-cell reconstitution [63,66]. Mechanistic analyses revealed that the eradication of autoreactive memory B-cells and the subsequent clearance of pathogenic autoantibodies were followed with immune system reconstitution with naive, non-autoreactive B-cell populations. These findings confirmed that CAR-mediated deep B-cell depletion establishes a critical antigen-free window, which facilitates the regeneration of a new immune repertoire.

The HN2301 study evaluated CD8-targeted tLNP with anti-CD19 CAR in 5 patients with refractory SLE [8]. Transient CAR expression peaked within 24 h and returned to baseline by 72 h, successfully driving rapid (>90%) peripheral B-cell depletion. This treatment induced early clinical improvements, including reduced proteinuria, declining anti-dsDNA titers, and normalized complement. However, only 40% of patients maintained prolonged responses. These results suggest that durable disease control likely requires deeper tissue penetration to eradicate pathogenic B-cells sequestered within tertiary lymphoid structures [22].

Quantitative comparisons reveal discrepancies in the kinetics of *in vivo* tLNP-mRNA and *ex vivo* autologous CAR-T therapies. Rather than exhibiting sustained cellular expansion, *in vivo* mRNA delivery drives rapid, dose-dependent bursts of CAR expression. Consequently, circulating CAR-T cell numbers peak transiently and return to baseline levels within 3 to 5 days. By contrast, *ex vivo* manufactured products generally show a single, broader expansion wave, often peaking around days 7–14 followed by a more gradual contraction over 4–12 weeks. Despite these divergent persistence profiles, both approaches can mediate rapid, near-complete B-cell depletion (>95% within about 1–7 days). However, the duration of pharmacodynamic effects differs in that *in vivo* tLNP-mRNA strategies are commonly associated with earlier B-cell recovery, with partial reconstitution emerging within 1–2 months, whereas *ex vivo* CAR-T therapies more often sustain depletion for 2–4 months or longer.

4.2. Lentiviral Vectors in Multiple Myeloma: Persistent Presence

First-in-human data from a Phase 1 study (NCT06691685) evaluated ESO-T01, a nanobody-targeted, immune-shielded LVV delivering anti-BCMA CAR, in four adults with heavily pretreated relapsed/refractory multiple myeloma [7]. Each patient received a single intravenous infusion at 2.0×10^8 transduction units without lymphodepletion—approximately one-tenth of the human equivalent dose derived from effective mouse models. All patients developed acute inflammatory reactions within day 1 (chills and fever lasting 6–18 h). Grade 3 CRS occurred in 75% (patients 1, 2, 4), often requiring vasopressor support and management with glucocorticoids. A milder grade 1 CRS peak occurred days 8–12, coinciding with maximal CAR-T expansion. Patient 4 developed grade 1 ICANS on day 8, which was controlled by glucocorticoids.

Viral titers in the peripheral blood peaked within 12 h and fell below detection by 48 h. CAR-T cells were first detected days 4–8, expanding to peak between days 10–17. $C_{\max}$ and AUC_{0-28d} were comparable to commercial BCMA CAR-T products (e.g., Idecabtagene Vicleucel, Ciltacabtagene Autoleucel). CAR-T cells demonstrated broad biodistribution. They were detected in bone marrow, tumor tissues, pleural effusions, and cerebrospinal fluid. Clinical efficacy data showed that patient 1 achieved a stringent complete response by month 2. Patient 2 achieved a stringent complete response with complete lesion resolution by day 28. Patients 3 and 4 achieved partial response and MRD negativity in bone marrow by day 28.

The clinical landscape for patients with multiple myeloma was enhanced by the recent inMMycAR Phase 1 study of KLN-1010, which targets BCMA [9]. Unlike ESO-T01, where grade 3 CRS occurred in the majority of patients, KLN-1010 recipients experienced only grade 2 CRS and no ICANS, despite high peak CAR+ frequencies (up to 72% of CD3⁺ lymphocytes on day 15). All KLN-1010 patients (n = 4) achieved MRD negativity at 10⁻⁵ or 10⁻⁶ sensitivity within one month, with responses deepening to very good partial response by month 3. In contrast to rapid ESO-T01 decline, KLN-1010 CAR-T cells remained detectable through month 3, predominantly memory-phenotype T cells associated with durable remissions [9].

4.3. Persistence as a Translational Challenge

While initial CAR-T cell expansion parameters ($C_{\max}$ and AUC_{0-28d}) aligned with *ex vivo* therapies, ESO-T01 CAR transgene levels declined substantially faster. CAR copies fell below the quantitative PCR (qPCR) limit of detection (LOD) by approximately day 60 in most patients, whereas benchmark *ex vivo* BCMA CAR T cell products (e.g., Ciltacabtagene Autoleucel) often remain detectable beyond 150 days (Figure 3). The comparison of *in vivo* and *ex vivo* CAR-T cells demonstrated fundamental distinctions. Ciltacabtagene Autoleucel showed CAR transgene peaks at 10,000–100,000 copies/ μ g DNA within 2–3 weeks, followed by a biphasic decline with detectable levels commonly persisting beyond 150 days and often >1000 copies/ μ g DNA at ≥ 12 months [67,68]. ESO-T01 demonstrated substantially shorter persistence (peak 1000–10,000 copies/ μ g DNA, undetectable by day 60) [7], reflecting their biological distinctions despite assay sensitivity differences.

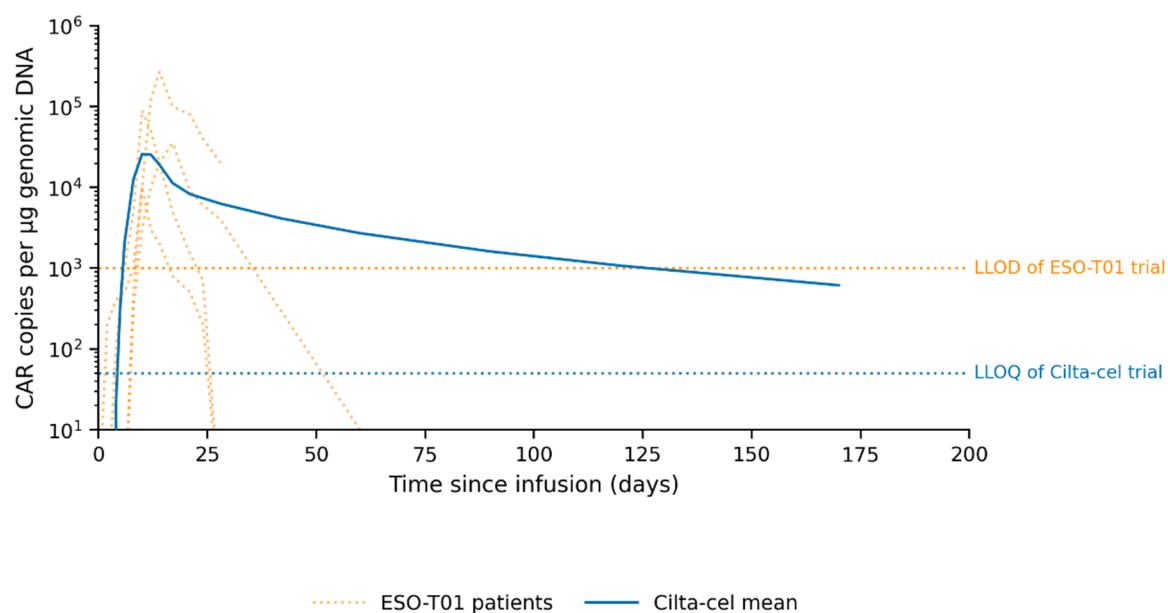


Figure 3. Comparative persistence of *ex vivo* CAR-T cells (Cilta-cel) and *in vivo* CAR-T (ESO-T01) cells. Comparison of CAR levels in peripheral blood following *ex vivo* Cilta-cel (blue curve, shown as population mean) versus *in vivo* ESO-T01 (orange curves, individual patients). Cilta-cel exhibits peak expansion between weeks 2 and 3 post-infusion, with cellular persistence extending beyond 150 days (lower limit of quantification [LLOQ]: 50 copies/ μ g DNA). ESO-T01 demonstrates a more transient persistence, with transgene levels falling below the lower limit of detection (LLOD: 1000 copies/ μ g DNA; represented by the dashed line) by day 60 in the majority of patients. Notably, different assay sensitivities preclude a direct quantitative comparison of absolute CAR concentrations between these platforms. The x-axis spans 200 days (approximately 28 weeks) to illustrate long-term longitudinal profiles. Data adapted from Berdeja et al. for cilta-cel [67] and Xu et al. for ESO-T01 [7].

In vivo tLNP-generated anti-CD19 CAR-T cells exhibit short but discrete expansion bursts, with a rapid rise within 6–12 h after dose and return to baseline within 3–5 days [8]. However, benchmark *ex vivo* autologous CD19-targeted CAR-T cells demonstrate broader and more sustained expansion peaks occurring days 7–14, followed by gradual contraction over 4–12 weeks [64,67]. Both platforms achieve rapid and near-complete B-cell depletion (>90–95% reduction) within the first week. However, the durations of depletion are quite different: *in vivo* tLNP-mRNA CAR-T cell therapy showed earlier B-cell reconstitution beginning at 1–2 months, whereas *ex vivo* autologous CAR-T cell cohorts often exhibit 2–4 months or more of sustained B-cell depletion before reconstitution (Figure 4). The relatively limited persistence with *in vivo* platforms reflects fundamental biological differences between the transient, mRNA-driven CAR expression that rises and decays after delivery and the robust antigen-driven expansion and longer-term maintenance typically achieved by *ex vivo*-manufactured CAR-T cell products after infusion.

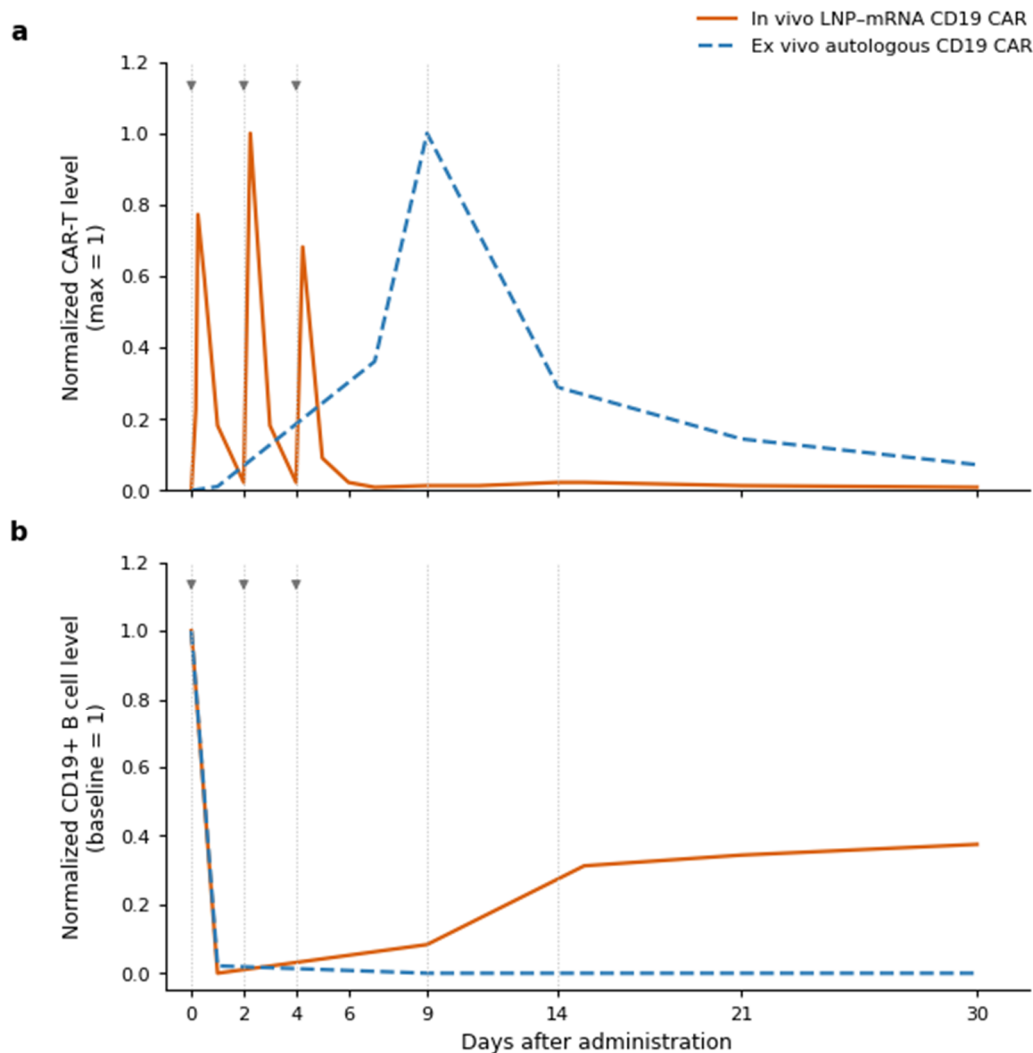


Figure 4. CAR-T cell expansion and peripheral B-cell dynamics following *in vivo* tLNP-CAR (HN2301) and *ex vivo* autologous CD19 CAR-T therapy. Panel (a) Normalized CAR-T cell expansion kinetics Comparison of the HN2301 (orange line) and *ex vivo* autologous CAR T-cell therapy (blue line) in patients with SLE. The HN2301, administered on days 0, 2, and 4 (indicated by arrows), elicits rapid and transient peaks of CAR-T cells that resolve within several days. By contrast, *ex vivo* CAR-T cells show a broad expansion peak at 7–14 days with a subsequent slow contraction phase. All CAR-T levels are normalized to the maximum value for each platform. Data adapted from Wang et al. for HN2301 [8] and Mackensen et al. for autologous CD19 CAR-T [63]. Panel (b) Corresponding dynamics of peripheral CD19⁺ B cells. Comparing *in vivo* lipid nanoparticle-mRNA HN2301 (orange line) with *ex vivo* autologous (blue line) CAR-T cell platforms in SLE patients. Both platforms achieve rapid and near-complete B-cell depletion (>90–95%). However, B-cell reconstitution begins earlier (at 1–2 months) following treatment with the HN2301 products than with the *ex vivo* CD19 CAR-T cells, which mediates sustained depletion for 2–4

months or longer. B-cell counts are normalized relative to baseline values, with curves depicting median values. Data adapted from Wang et al. for HN2301 [8] and Mackensen et al. for autologous CD19 CAR-T cells [63].

Taken together, these differences in clinical performance likely arise from a combination of intrinsic biology and how each modality is produced and delivered. *Ex vivo* manufacturing has several advantages: T cells are selected, activated, and expanded under tightly controlled culture conditions, enabling the delivery of a large, functional cell dose at infusion and allowing phenotype shifting, including enrichment for memory-like subsets and other attributes linked to persistence. Moreover, most *ex vivo* CAR-T cell clinical protocols use lymphodepletion regimen that is able to eliminate endogenous lymphocytes and trigger upregulation of key cytokines, such as IL-7 and IL-15. Therefore, the lymphodepletion regimen provides a permissive microenvironment that drives the engraftment and expansion, even the durable persistence of infused CAR-T cells.

By contrast, *in vivo* CAR-T cell therapies take place within patients' immune system. Gene transduction efficiency is highly variable and influenced by baseline T-cell fitness, innate immune sensing, and rapid vector clearance. Most of the *vivo* CAR-T therapies skipped lymphodepletion regimen to preserve the T cells, thus bypassing the advantages that lymphodepletion introduce. Consequently, *in vivo* CAR-T therapies typically yield smaller initial CAR⁺ cells that must expand substantially *in vivo*, which frequently delays peak therapeutic activity and imposes practical ceilings on both the magnitude and durability of the clinical response [6,32]. Therefore, while current *in vivo* CAR-T therapies can elicit rapid and potent cytotoxicity, they have yet to consistently match the deep immune reset and durable persistence achieved by *ex vivo* CAR-T cell products.

Ex vivo and *in vivo* CAR-T therapies are distinct modalities rather than manufacturing platforms. *Ex vivo* products deliver characterized cellular products after cell collection, activation, genetic modification, expansion, release testing, and usually lymphodepleting conditioning, whereas *in vivo* platforms deliver vectors whose pharmacokinetics, cell transduction, target cell recognition, and host dependent expansion jointly determine the resulting cellular exposure. These differences affect dose definition, product testing especially potency testing, redosing, logistics, etc. A comprehensive head-to-head analysis of each pharmacological and manufacturing dimension would therefore substantially broaden the scope of this review. This manuscript focuses on the vector pharmacology, therapeutic performance, and the proposed Quality-by-Design framework.

These biological differences are likely to translate into distinct clinical positioning (Table 1). Indications in which durable CAR-T cell persistence is central to long-term disease control, such as relapsed/refractory multiple myeloma or heavily pretreated lymphoma where prolonged tumor surveillance is desirable, may continue to favor *ex vivo*–manufactured products. By contrast, clinical contexts in which benefit is driven largely by early, intensive cytoreductive activity, such as eradication of minimal residual disease, frontline debulking strategies, or transient immune reset in autoimmune disease, may be particularly well matched to emerging *in vivo* platforms, for which potent yet time-limited effector activity could be sufficient.

Table 1. High-level comparison of *ex vivo* and *in vivo* CAR engineering modalities.

Modality	Durability	Efficacy	Manufacturing and Cost	Safety and Key Considerations
<i>Ex vivo</i> LVV engineered CAR-T cells	- Stable genomic integration; - CAR-T cell persistence may extend from months to years.	- Clinically established in B-cell malignancies and myeloma; - Activity in solid tumors remains limited.	- Patient-specific apheresis collection, activation, transduction, expansion, and release testing; - Long lead time, complex logistics, and high cost.	- Usually requires lymphodepletion; - CRS, ICANS, cytopenias and infection risk; - Integration and replication-competent lentivirus (RCL) require long-term follow-up.
<i>Ex vivo</i> non-viral vector-engineered CAR-T cells, Including LNP (RNA)	- Non-integrating and transient; - Typically, lasting days after infusion.	- Mainly preclinical or early-stage; - Suited to applications in which temporary CAR activity is desirable.	- Still requires individualized cell collection and processing; - Viral vector manufacture is not required and may shorten production duration.	- Lower insertional risk and self-limiting expression; - Cell stress, transfection heterogeneity, and repeat dosing may constrain efficacy.
<i>In vivo</i> LVV-engineered CAR-T cells	- Stable integration in endogenous lymphocytes; - Potentially durable, although clinical persistence remains under further development.	- Preclinical and early clinical evidence supports B-cell or myeloma cell targeting; - Outcome depends on transduction, expansion, and host cell fitness.	- Off-the-shelf vector eliminates cell collection and laboratory manufacturing; - LVV production and control, potency testing, and cold chain logistics remain complex.	- Off-target transduction, insertional mutagenesis, vector immunity, and cytokine-mediated toxicities require control; - Redosing may be limited.
<i>In vivo</i> LNP (mRNA)-engineered CAR-T cells	- Transient CAR expression over days; - Repeated dosing is feasible; - Pharmacodynamic effects may outlast expression.	- Preclinical and early clinical proof-of-concept; including rapid B-cell depletion; - Durability is generally shorter than integrating approaches.	- Scalable off-the-shelf manufacture without leukapheresis; - RNA quality, targeting ligands, nanoparticle consistency and biodistribution need tight control.	- No genomic integration and time-limited exposure; - Innate sensing, complement system activation, or infusion reactions; - Off-target delivery and cumulative redosing risk require monitoring.

Abbreviations: LVV, lentiviral vector; LNP, lipid nanoparticle; CRS, cytokine release syndrome; ICANS, immune effector cell-associated neurotoxicity syndrome.

5. Quality-by-Design Framework for *In Vivo* CAR-T Cells

Building on these platform-dependent advances and constraints, the development pathway for *in vivo* CAR-T cell platforms will depend not only on biological performance, but also on the extent to which manufacturing and clinical implementation can be optimized to deliver predictable exposure–response relationships. In this context, Quality-by-Design (QbD) provides a useful framework that can be adapted from similar systemic gene-delivery modalities.

At the therapeutic implementation level, QbD for *in vivo* CAR-T cell therapy could directly benefit from regulatory and clinical precedents established for intravenously-administered AAV gene therapies (e.g., HEMGENIX, ROCTAVIAN, ELEVIDYS, ZOLGENSMA, and BEQVEZ). In particular, inclusion of corticosteroid-based immunomodulation in the treatment protocol, baseline screening for anti-vector immunity, predefined assessments of vector shedding, and structured long-term follow-up to evaluate durability and late toxicities constitute a translational blueprint for managing systemically delivered genetic payloads that must function within the immunologically complex hosts.

At the product level, however, *in vivo* CAR-T cell technologies will require an expanded QbD framework that goes beyond conventional release and characterization attributes such as potency, purity, and stability. Because immune sensing, innate activation, transduction efficiency, and *in vivo* exposure kinetics can directly constrain cellular yield, expansion, and persistence, these parameters should be treated as desired critical quality drivers rather than downstream clinical variables. Embedding such biological factors into early-stage design and vector engineering, alongside pharmacologic exposure targets, may improve reproducibility across patients, strengthen risk mitigation strategies, and better align development packages with regulatory expectations for systemically *in vivo* CAR-T cells.

5.1. Quality Target Product Profile (QTPP)

The Quality Target Product Profile (QTPP) for *in vivo* CAR-T cell platforms should integrate disease- and platform-specific parameters. At the clinical level, the QTPP should define the intended patient population and treatment intent (e.g., intent-for-cure, consolidation therapy, or as a bridge to standard-of-care or subsequent therapies), together with the expected temporal profile and durability of CAR expression and specific tolerability targets, including acceptable thresholds for cytokine release and organ-specific toxicities.

At the platform level, QTPPs should specify mechanistic and pharmacologic design targets:

Target-cell specificity: define the T-cell subset(s) to be engineered, the intended transduction efficiency, and acceptable boundaries for off-target lineage modification.

Expression kinetics: characterize the temporal profile dictated by platform biology (e.g., LVV platform with delayed onset and multi-week persistence versus tLNP-mRNA for rapid onset and transient expression).

Exposure ceiling and therapeutic window: pre-specify quantitative exposure targets, integrating CAR-expression area under the curve, peak CAR⁺ T-cell fraction, and cytokine profiles to balance efficacy against risks such as acute toxicity, CRS, and neurotoxicity (Section 5.3).

Vector-specific safety profile: For LVVs, safety assessments are prioritized on integration-site mapping, vector copy number (VCN) limitations, and comprehensive genotoxicity surveillance. For LNP therapies, safety assessments are prioritized on the monitoring of acute responses, complement activation, and hepatotoxicity.

Redosing feasibility: tLNP therapies are amenable to redosing, but LVV therapies redosing is constrained by significant barriers, including neutralizing anti-vector immunogenicity and cumulative insertional mutagenesis.

5.2. Critical Quality Attributes (CQA)

The CQA of *in vivo* CAR-T cell products link their quality attributes to biological activities, immune responses, and clinical risks.

CQA for tLNP-mRNA platforms

For tLNP mRNA therapies, CQAs integrate platform specific factors and critical good manufacturing practice (GMP) requirements.

Identity tLNP-mRNA products require rigorous physicochemical identification on the payload and the delivery vehicle. For mRNA cargo, the assessments include sequence fidelity, transcript capping, poly(A) tail length, and the precise incorporation of modified nucleotides (e.g., N1-methylpseudouridine). The delivery vehicle nanoparticle formulation needs strict control over lipid composition, including the ionizable lipid pKa, helper/PEG-lipid ratios, as well as the surface density and orientation of targeting antibodies.

Strength mRNA content per dose in the tLNP-mRNA products should be measured within the context of preserved transcript integrity and capping efficiency.

Purity the purity of tLNP-mRNA products requires strict control over the biophysical properties of the nanoparticle and the clearance of impurities. The critical biophysical properties include tightly controlled particle size, low polydispersity, high encapsulation efficiency, and a near-neutral surface charge. Process-related impurities such as residual double-stranded RNA (dsRNA), truncated transcripts, and organic solvents must be minimized to avoid triggering unwanted innate immune responses.

Potency assay development is encouraged at early phase of clinical development. It will be helpful to increase our understanding of the tLNP-mRNA products using well-developed assays, such as *in vitro* cell-based assays that quantitatively correlate the mRNA dose with CAR expression and downstream T-cell effector functions, e.g., targeted cytotoxicity and cytokine release. These assays are expected to reflect the expression magnitude and temporal profiles that are consistent with the intended exposure window.

Safety includes thresholds for complement activation and innate immune stimulation (e.g., TLR7/8 activation) and maintenance of residual dsRNA below immunogenic thresholds. Safety also includes standard controls for bioburden, endotoxin, and residual DNA template that are appropriate for parenteral gene-modifying products [69]. Collectively, these attributes anchor immune activation ceilings and hepatotoxicity limits defined in the QTPP.

CQA for LVV platforms

For LVV-based *in vivo* CAR-T cell platforms, CQAs are similarly derived from the QTPP but must additionally address genomic integration and vector–host interface risks.

Identity is confirmed through vector genome sequence integrity, transgene cassette configuration, and envelope composition, including the intended pseudotype and any engineered surface modifications that influence tropism or immune evasion.

Strength is quantified by infectious titer (e.g., transduction units/mL or transduction units/kg) and particle-to-infectious unit ratio.

Purity requires tight control of process-related impurities, vector heterogeneity and envelope integrity. The process-related impurities include host cell DNA and proteins, residual plasmids, and manufacturing reagents. Vector heterogeneity can be monitored by the ratio of full-to-empty particles. Moreover, envelope integrity, specifically the capability of incorporation of fusogen with the viral membrane, is critical for efficient target-cell fusion and reducing unintended off-target interactions.

Potency development is encouraged starting from early phase clinical development. The development can be established through robust *in vitro* cell-based assays that can reflect the correlation between vector dose and CAR expression and subsequent effector functions, e.g., target-specific cytotoxicity and cytokine release. Different from LVV for *ex vivo* CAR-T cells, the potency of *in vivo* LVV must consider the confirm the ligand specificity.

Safety requires rigorous documentation of the absence of replication-competent lentivirus (RCL), vector copy number (VCN) and integration site distribution within target cells. Furthermore, vectors must be systematically evaluated for complement sensitivity and off-target transduction. These parameters are monitored alongside standard controls for sterility, endotoxin, and adventitious agents as is done with gene therapy products.

5.3. Exposure-Based Dosing Strategy

To address the challenges with dose optimization for *in vivo* CAR-T cell platforms that are frequently constrained by an early, innate immune-driven toxicity window, we propose a conceptual exposure–response framework. Following systemic administration, a rapid increase in vector exposure temporally coincides with acute inflammation triggered by innate sensing of the delivery vehicle and the associated impurities. This initial inflammatory phase generates an acute toxicity ceiling, the maximum tolerated threshold that limits dose escalation before therapeutic CAR-T cell expansion is fully achieved.

For LVV *in vivo* CAR-T product, systemic viral titers peak shortly after infusion and are predominantly cleared within 48 h. CAR expression and subsequent T-cell activity typically emerge between days 4 and 8, then undergo a gradual expansion that peaks over 10 to 20 days. This kinetics of initial vector exposure and later CAR-T cell expansion generate two temporally risk windows. The first represents an acute toxicity window, mostly occurred in the initial 1 to 2 days, during which the LVV can provoke acute inflammation through innate immune system. The acute toxicity ceiling can be a dose-limiting factor, precluding the LVV doses that are necessary to achieve the desired therapeutic CAR-T cell expansion. A second, delayed risk window coincides with maximal antigen-driven CAR-T cell expansion, mostly reflected by effector cell-mediated complications, such as CRS and ICANS (Figure 5).

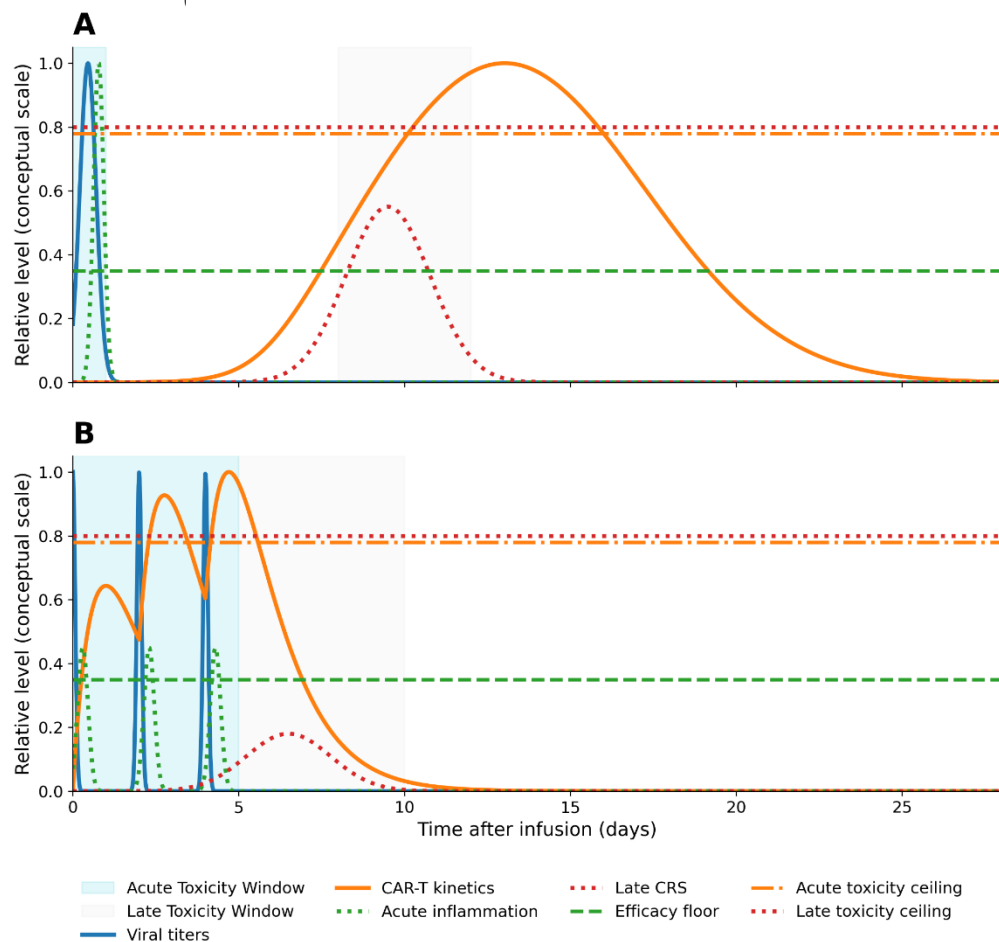


Figure 5. Conceptual exposure–response and safety windows for *in vivo* CAR-T cells. Panel (a) Lentiviral vector. Vector titers (solid blue line) peak immediately post-infusion and are mostly cleared within ~48 h. CAR+ T cells (solid orange line) emerged on days ~4–8, then gradually expanded to peak levels (days ~10–20). An initial acute toxicity window (light blue box) reflects acute inflammation (dotted green line), which establishes an acute toxicity ceiling (dashed orange line) that limits dose escalation. A late toxicity window (grey box) aligns with maximal CAR-T cell expansion and the corresponding emergence of CRS and ICANS (red dotted line), indicating a late toxicity ceiling (dashed red line). Panel (b) tLNP–mRNA. Each administration yields a brief vector peak (solid blue line) and a rapid and transient CAR-T cell pulse (solid orange line; peak ~12–24 h after dosing). The acute inflammatory responses (green dotted lines) constitute the acute toxicity window (light blue box). The dose escalation is limited by the acute toxicity ceiling (dashed orange line). Normalized, schematic representations illustrate the temporal dynamics of vectors, CAR-T cell kinetics, and associated toxicities. Across both platforms, horizontal dashed lines denote the therapeutic efficacy floor (green), the acute toxicity ceiling (orange), and the late toxicity ceiling (red).

tLNP–mRNA products exhibit a distinct kinetics yet similar constraint by early innate immune response. Each administration yields a pulse accompanied by a rapid, transient surge in CAR expression, which typically peaks between 12 and 24 h before declining over subsequent days. Each dose has the potential to induce its own acute inflammatory response within this early toxicity window. Thus, optimization of dose escalation relies primarily on maintaining innate immune response below the acute toxicity ceiling across repeated administrations. In summary, the early innate immune response frequently defines the practical upper limits of dosing, thereby establishing the acute toxicity window as a central limiting factor for rational dose selection, step-up regimens, and redosing strategies (Figure 5).

Traditional mg/kg or transduction units/kg dosing inadequately capture *in vivo* CAR-T cell pharmacology because the administered vector triggers a dynamic, patient-specific “manufacturing process” rather than delivering a fixed dose. Pharmacokinetic and pharmacodynamic data—CAR expression kinetics, cytokine release profile, and target-cell depletion—are only observable after the dose has been given, which means that dose individualization cannot be simply calculated in advance but must be determined by on-treatment readouts.

Early *in vivo* CAR-T cell clinical trials, such as ESO-T01, KLN-1010, HN2301, and clinical trials of other related platforms, have largely relied on conventional stepwise dose escalation with dose adjustments driven post hoc by toxicity signals in initial cohorts. This predominantly toxicity-driven strategy can expose early patients to risk and underutilize pharmacokinetic and pharmacodynamic data. Looking ahead, dosing strategies may benefit from frameworks informed by *in vivo* exposure and pharmacodynamic response. In this view, conventional dose metrics could be complemented by quantitative measures such as peak CAR⁺ T-cell frequency, CAR-expression area under the curve, innate cytokine kinetics, antigen-driven expansion magnitude for integrating lentiviral platforms, cumulative CAR exposure across repeated tLNP administrations, and the kinetics of target-cell depletion. Using these exposure and response measurements to guide dosing could make results easier to compare across patients, help plan repeat dosing more sensibly, and better balance effectiveness with side effects.

Because key PK and PD information only becomes available after patients have received some treatment, split-dose or step-up dosing may be a good fit for *in vivo* CAR-T approaches. Giving a small “starter” portion of the total planned dose can provide early readouts—such as initial CAR expression, early CAR⁺ T-cell expansion, cytokine levels, and the first signs of target-cell depletion. Subsequent portions (or step-up doses) can then be adjusted using prespecified rules based on these early signals, rather than waiting for toxicity or lack of response to drive changes. This exposure-guided approach is somewhat similar to therapeutic drug monitoring, but it relies on cellular and immune readouts instead of drug concentrations. Incorporating split or step-up dosing with clear escalation and de-escalation criteria into model-informed development could offer a practical route to more individualized dosing, while reducing dependence on toxicity-driven dose modifications.

6. Future Directions

6.1. Safety Considerations

For LVV-based CAR-T platforms, permanent proviral integration raises concerns about insertional mutagenesis and delayed secondary malignancies [70,71]. Rare cases of secondary T-cell neoplasms expressing the CAR transgene have been described after *ex vivo* CAR-T cell administration, although a direct causal link to the CAR construct or vector remains unresolved. FDA acknowledged the risk of T-cell malignancies after CAR-T therapy, and they also emphasized that benefits of CAR-T therapy continue to outweigh the risks in approved indications. A meta-analysis of 5,517 lymphoma and myeloma patients estimated an overall second primary malignancy frequency of approximately 6% and found no increase relative to standard-of-care cohorts; T-cell malignancies represented a small minority of events (1.5%) [72]. In the ZUMA-7 study, 8/170 patients in the CAR-T arm developed a second primary cancer, compared to 3/161 patients in the standard of care arm. In TRANSFORM study, 3 CAR-T patients versus 7 standards of care patients developed second malignancies [73]. Nevertheless, controlling expansion dynamics remains a central challenge for LVV platforms, as illustrated by early *in vivo* programs that reduced vector doses after observing unexpectedly vigorous and prolonged CAR-T cell expansion [22].

For tLNP-mRNA platforms, safety considerations are driven by the combined effects of formulation, RNA payload, and CAR immunopharmacology [1]. Lipid nanoparticles can trigger acute infusion reactions, complement activation, and liver-targeted uptake, while the mRNA payload and any residual dsRNA can strongly activate innate immune sensors, and the encoded CAR can produce CAR-T cells that induce deep and rapid depletion of target cells with accompanying cytokine-mediated toxicities. Key development challenges therefore include minimizing off-target hepatic exposure, attenuating innate immune activation from both the lipid and RNA components, and ensuring that the intensity and kinetics of CAR activity remain within clinically manageable limits.

Both LVV and tLNP-mRNA *in vivo* platforms have common safety issues. Off-target delivery can induce CAR expression in non-target cells, which may mask the target epitopes [74]. Achieving sufficient levels of CAR-T cells usually needs high vector dosages that increase the risk of systemic inflammation, including complement-mediated hypersensitivity or infusion-related toxicities [6]. Furthermore, CAR-T cell-associated adverse events, e.g., CRS, ICANS, and delayed cytopenias, often appear several days post-administration, complicating timely risk mitigation. CAR-T cells generated from *in vivo* platforms cannot be fully characterized in the way of *ex vivo* CAR-T cells, increasing the difficulty in risk assessment. All these issues emphasize the need for novel cell tracking and monitoring strategies for early detection of safety signals.

6.2. Platform Selection Based on Indications

Although *in vivo* CAR-T cells may not recapitulate the long-term persistence of *ex vivo* CAR-T cells in the near future, they have their own advantages in addressing clinical needs by enhancing patient accessibility: *in vivo* platforms have the potential to reduce production costs and expand supply as off-shelf product and bypassing

resource-intensive GMP manufacturing of cellular product. Moreover, the capacity for programmable dosing and iterative construct optimization facilitates a transition toward more modular, pharmacologically adaptive treatments.

The current persistence gap reinforces a central principle: available *in vivo* platforms are optimized for potent early cytotoxicity rather than long-term persistence. As a result, persistence-driven indications such as late-stage multiple myeloma or heavily pretreated lymphoma—where durable CAR-T cell survival is critical for long-term disease control—are likely to remain the domain of *ex vivo* products in the near future, whereas exposure-driven settings such as measurable residual disease (MRD) eradication, frontline cytoreduction, or immune reset in autoimmunity are more aligned with transient but intense CAR activity induced by *in vivo* CAR-T cell platforms. Therefore, LVV-based *in vivo* CAR-T cells may be preferentially suited to single-dose applications in earlier-stage malignancies where T-cell fitness is preserved and weeks to months of CAR expression can be leveraged without prohibitively high integration-related risk. In contrast, tLNP-based approaches are well-positioned for programmable, multidose immune modulation, such as repeated short pulses of CAR-T cell expression for autoimmune disease, MRD control via brief but deep target depletion, frontline cytoreduction, and potentially non-oncologic indications such as tissue fibrosis where transient cellular reprogramming is sufficient and long-term engraftment may be undesirable.

6.3. Innate Immune Evasion as a Design Objective

For *in vivo* CAR-T cell therapies, the innate immune system is a major determinant of effective dose, because rapid complement activation, phagocytic clearance, and interferon responses can both trigger acute toxicity and reduce vector exposure before CAR-T cell generation occurs. Design strategies that reduce innate immune recognition may be as important as improving transduction efficiency, because attenuating early clearance and inflammatory cascades could allow a given administered dose to produce higher and potentially longer-lasting *in vivo* CAR expression.

LVV engineering priorities

For LVV-based *in vivo* CAR-T cell therapy, key priorities include envelopes with reduced complement activation (e.g., lower C1q binding and functional incorporation of complement regulators such as CD55/CD59), CD47-based signals that prolong circulation [19], and modulation of MHC-I to reduce innate and adaptive recognition while preserving safety. In addition, pH-independent paramyxovirus-derived entry mechanisms (e.g., measles or related F/H glycoprotein systems), which mediate direct plasma-membrane fusion and have been used to retarget vectors to lymphoid cells, providing an entry pathway that can reduce reliance on endocytic/phagocytic uptake routes that favor hepatic/splenic sequestration [75].

tLNP-mRNA engineering priorities

For tLNP-mRNA platforms, innate immune system evasion is driven largely by nucleic acid and lipids. Priorities include TLR-silent mRNA (high levels of N1-methylpseudouridine incorporation, codon optimization, Cap-1 structures, and rigorous dsRNA removal), together with lipid nanoparticle compositions that minimize complement activation and Kupffer cell uptake through ionizable lipids with tuned pKa [23], complement-inhibitory or self-mimicking surfaces, and antibody-mediated targeting to T cells. Additional measures include anti-inflammatory or biodegradable lipid components and strategies to mitigate accelerated blood clearance—such as careful regulation of PEG density or use of alternative hydrophilic coatings, thereby supporting repeat dosing without steep loss of exposure.

In this context, innate immune evasion may be a decisive factor in narrowing the persistence gap between *in vivo* and *ex vivo* CAR-T cell therapies. Vectors that avoid early complement deposition, phagocytic clearance, and excessive interferon signaling are more likely to achieve higher peak CAR-T cell expression, extended duration of functional CAR-T cell activity, and improved feasibility of split-dose or redosing regimens, enabling exposure-guided, model-informed dosing strategies without prohibitive toxicity.

6.4. Adjunctive T-Cell Conditioning Strategies

Narrowing the persistence gap between *in vivo* and *ex vivo* CAR-T cell therapy will require not only vector optimization but also deliberate conditioning of the host T cells. In practice, this means treating T cell fitness as a manipulable co-factor, using adjunctive regimens to create a permissive niche for CAR-T cell generation, expansion, and long-term function.

In *ex vivo* CAR-T cell therapy, lymphodepletion conditioning is routinely used to deplete regulatory T cells, eliminate cytokine sinks, induce homeostatic cytokines (IL-7, IL-15), and create expansion space [76]. In contrast, many *in vivo* CAR-T cell strategies are designed to omit intensive lymphodepletion in order to preserve

endogenous T cells and reduce conditioning-related toxicities. In the *in vivo* CAR-T cell setting, the risk–benefit profile of introducing low-intensity lymphodepletion (e.g., cyclophosphamide–fludarabine regimens) or alternative conditioning regimens is not yet well established and will benefit from systematic assessment in carefully designed clinical studies.

Adjunct cytokine support—using IL-7, IL-15 or IL-15 superagonists (e.g., N-803)—could enhance the survival and expansion of naïve and memory T-cell subsets, thereby increasing the magnitude and potentially prolonging the durability of CAR-T cell responses generated *in vivo* [12,13].

Metabolic interventions provide an additional approach for modulating T-cell fitness. Pharmacological modulation of mTOR signaling via agents such as rapamycin, or the utilization of metabolic regulators like metformin, can polarize T cells toward favorable memory-like phenotypes. In preclinical models, such interventions were able to attenuate host antiviral restriction pathways, thereby enhancing lentiviral transduction efficiency without impairing effector cytotoxic function [42].

Immune checkpoint inhibition, particularly anti-PD-1 monoclonal antibodies, can attenuate inhibitory signaling cascades within exhausted or chronically stimulated T-cells, thereby preserving CAR-T cell effector potency over extended durations [77].

6.5. Redosing as a Strategic Differentiator

For LVV-based *in vivo* CAR-T cell therapies, redosing is constrained by anti-vector immunity. This host immune response, characterized by neutralizing antibodies and T-cell-mediated reactivities against viral components, significantly compromises subsequent transduction efficiency while increasing the risk of systemic immunotoxicity. Experiments such as envelope switching, peri-dosing immunosuppression, and further envelope engineering are being explored to enable repeat administration. However, these strategies remain largely unvalidated in the clinical setting, so LVV-based platforms are likely to function as single-dose therapies in the near future.

tLNP-mRNA systems are intrinsically more amenable to redosing because they lack viral proteins and typically induce transient expression. Emerging clinical experience (e.g., multi-dose regimens in the HN2301 study), together with NHP studies showing maintenance of CAR expression across multiple administrations, and *in vitro* data demonstrating preserved cytotoxicity with repeated transduction, support the feasibility of repeated tLNP dosing without major loss of efficacy.

However, repeated tLNP dosing may be limited by accelerated blood clearance (ABC), whereby an initial nanoparticle exposure primes immune mechanisms that hasten the removal of subsequent doses. First described for PEGylated liposomes, ABC is often linked to anti-PEG IgM-mediated complement activation, opsonization, and rapid uptake by the mononuclear phagocyte system (including hepatic Kupffer cells), which reduces systemic exposure and target delivery [78]. Similar ABC-like kinetics with redosing have been reported for mRNA LNPs, particularly for PEG-lipid-containing formulations, motivating PEG-alternative surface chemistries to preserve the effectiveness of repeat-dose exposure [79]. For tLNP-mRNA CAR-T cell development that may rely on clinically practical redosing, ABC could therefore reduce delivered mRNA and CAR expression, increase inter-patient variability, and, in some settings, contribute to increased hypersensitivity or inflammatory reactions [80]. Accordingly, development of tLNP-mRNA-based CAR-T cell therapy may benefit from prospectively tracking ABC biomarkers (e.g., anti-PEG titers, complement activation markers and biodistribution shifts) and incorporating mitigation strategies, such as alternative hydrophilic coronas, optimized PEG chemistry/density, interval selection and peri-dosing immunomodulation, into redosing strategies and safety monitoring.

6.6. Host Immune Parameters

In vivo CAR-T cell therapy requires an additional layer of host-bioreactor control because the CAR-T cells are ‘manufactured’ within the patient’s immune system rather than by controlled *ex vivo* processes. As a result, the immune system effectively becomes part of the production environment, and patient-specific immunologic parameters can act as critical control factors that shape vector pharmacology, CAR-T cell yield and the risk of inflammatory toxicities. The immunological parameters comprise both innate and adaptive host responses. Innate responses (e.g., complement activation, Toll-like receptors) establish a tolerability threshold for vector exposure. Pre-existing adaptive responses comprising neutralizing antibodies and memory T-cell responses limit redosing feasibility and undermine the durability of CAR expression. Notably, pre-existing host antiviral states may also reduce transduction efficiency as evidenced that elevated baseline type I interferon levels correlate with impaired LVV delivery [19]. Furthermore, baseline T-cell fitness, such as the expression of exhaustion markers (e.g., PD-

1, LAG-3, TIM-3), the proportion of T-cell memory subsets, and the patient's prevailing lymphopenia status, plays pivotal roles in controlling the generation of sufficiently functional CAR-T cells.

The host immune parameters mirror the critical process parameters (CPP) on *ex vivo* CAR-T cells; the prospective profiling of these parameters and their systematic integration into the QTPP and CQA represent a reasonable strategy to control the quality of the *in vivo* CAR-T cells. In practice, some factors can be considered: eligibility criteria linked to T-cell fitness; acceptable ranges of pre-existing anti-vector immunity; thresholds for innate activation signatures.

6.7. Solid Tumors, Tumor Microenvironment, and Alternative Immune-Cell Targets

Both *ex vivo* and *in vivo* CAR-T therapies remain at an early stage for solid tumors. Major barriers include heterogeneous and spatially variable antigen expression, antigen loss and immune escape, inefficient trafficking and tissue penetration, stromal exclusion, and suppressive signals within the tumor microenvironment (TME) [81]. These constraints may be particularly important for current *in vivo* platforms because the number, phenotype, and duration of CAR-expressing cells are determined within the same hostile tissue and immune environment that the therapy must overcome.

One potential strategy for addressing antigen heterogeneity is epitope spreading. CAR-mediated tumor cell killing can release additional tumor antigens and inflammatory signals, enabling dendritic cell cross-presentation and recruitment of endogenous polyclonal T-cell responses. Preclinical studies using cytokine- or chemokine-armed CAR-T cells have demonstrated bystander responses and epitope spreading in antigen heterogeneous solid tumor model [82,83]. However, whether *in vivo* CAR-T cells can reliably induce new tumor antigen or neoantigen responses remains unknown. This possibility should therefore be considered a testable development objective rather than an established attribute of *in vivo* CAR-T platforms.

The interaction between CAR-T cells and the TME is bidirectional. Hypoxia, nutrient competition, suppressive metabolites, checkpoint ligands, tumor-associated macrophages, regulatory T cells, and cancer-associated fibroblasts can impair CAR-T-cell trafficking, proliferation, cytotoxicity, and persistence. Conversely, CAR-mediated cytotoxicity and cytokine production can alter antigen availability, dendritic cell activation, myeloid-cell polarization, and recruitment of endogenous effector cells. A new strategy is to target FAP-positive fibroblasts or other mesenchymal stromal components rather than malignant cells directly. *In vivo* generation of transient FAP-directed CAR-T cells reduced cardiac fibrosis in a non-malignant mouse model [62], demonstrating the feasibility of stromal-cell targeting while not yet establishing efficiency against tumor stroma.

Other immune cell lineages are also attractive substrates for *in vivo* CAR engineering. NK cells combine CAR-directed recognition with intrinsic cytotoxicity, and targeted lentiviral delivery has generated both CAR-T and CAR-NK cells in preclinical models [57]. Macrophages offer tumor-tropic trafficking, phagocytosis, antigen presentation, and the capacity to remodel suppressive myeloid environments; myeloid tropic LNPs have delivered CAR mRNA to myeloid cells, for example, intravenous TROP2-CAR mRNA-LNP inhibited tumour growth in a mouse xenograft model and was expressed predominantly in myeloid cells in tumour-bearing mice and cynomolgus monkeys [84]. These approaches remain preclinical, and their translation will require lineage specific delivery, control of functional persistence and polarization, avoidance of off-target programming, and cell type appropriate safety and potency assays.

7. Conclusions

In vivo CAR-T cells have entered early clinical trials with FIH studies confirming mechanistic principles. Currently *in vivo* CAR-T cells achieve shorter persistence than *ex vivo* products. It is recommended that the selection of platforms be based on the PK and PD kinetics: LVVs for integration-based cytotoxicity in oncology; tLNP-mRNA for rapid and repeat-dosable modulation in autoimmune disease.

Optimizing *in vivo* CAR-T platforms requires integrating host immunological parameters into Quality-by-Design (QbD) frameworks to achieve CQAs. Transitioning from conventional mass-based dosing to exposure-driven strategies that are guided by CAR kinetics and pharmacodynamic biomarkers offers a viable pathway to mitigate clinical heterogeneity. Next-generation molecular engineering should benefit from prioritizing innate immune evasion, incorporating complement-resistance, TLR-silent architectures, and macrophage-evasion strategies to overcome the exposure limits.

As technology matures, *in vivo* CAR-T cell therapy will likely coexist with *ex vivo* CAR-T cell products in a complementary landscape: *ex vivo* for persistent disease control in late-line oncology; *in vivo* LVV therapies for intermediate-duration applications in earlier-line oncology; and *in vivo* tLNP-mRNA therapies for programmable

immune modulation in autoimmune disease and combination regimens. This integration will likely increase global cell therapy access.

Author Contributions

J.R. conceived and designed the work, analyzed and interpreted the data, and drafted or substantively revised the manuscript. D.S. contributed substantive revisions to the manuscript, and A.D. also provided substantive revisions. All authors have reviewed and approved the final submitted version.

Funding

This research was supported by the Intramural Research Program of Clinical Center, National Institutes of Health.

Institutional Review Board Statement

Not applicable.

Informed Consent Statement

Not applicable.

Data Availability Statement

This research was supported in part by the Intramural Research Program of the National Institutes of Health (NIH). The contributions of the NIH author(s) are considered Works of the United States Government. The findings and conclusions presented in this paper are those of the author(s) and do not necessarily reflect the views of the NIH or the U.S. Department of Health and Human Services.

Conflicts of Interest

The authors declare no conflict of interest.

Use of AI and AI-Assisted Technologies

HHS Enterprise ChatGPT Pro and HHS Enterprise Claude Sonnet 4.5 were used to assist with wording and formatting to enhance the clarity and readability of this manuscript. All AI generated text was reviewed, verified, and revised by the authors for accuracy and relevance. Nano Banana Pro was used to generate preliminary concept images for Figures 1 and 2, which were subsequently refined and finalized by the authors. HHS Enterprise ChatGPT Pro and Google Colab were used to create initial concept images for Figures 3–5. The authors further refined and finalized all figures.

References

1. Bot, A.; Scharenberg, A.; Friedman, K.; et al. *In vivo* chimeric antigen receptor (CAR)-T cell therapy. *Nat. Rev. Drug Discov.* **2025**, *25*, 116–137.
2. Asci, E.; Åkerström, F.; Sreekumar Nair, M.; et al. *In vivo* dendritic cell reprogramming for cancer immunotherapy. *Science* **2024**, *386*, eadn9083.
3. Smith, T.T.; Stephan, S.B.; Moffett, H.F.; et al. *In Situ* programming of leukaemia-specific T cells using synthetic DNA nanocarriers. *Nat. Nanotechnol.* **2017**, *12*, 813–820.
4. Heinrich, C.; Spagnoli, F.M.; Berninger, B. *In vivo* reprogramming for tissue repair. *Nat. Cell Biol.* **2015**, *17*, 204–211.
5. Short, L.; Holt, R.A.; Cullis, P.R.; et al. Direct *in vivo* CAR T cell engineering. *Trends Pharmacol. Sci.* **2024**, *45*, 406–418.
6. Li, Y.-R.; Zhu, Y.; Halladay, T.; et al. *In vivo* CAR engineering for immunotherapy. *Nat. Rev. Immunol.* **2025**, *25*, 725–744.
7. Xu, J.; Liu, L.; Parone, P.; et al. *In-vivo* B-cell maturation antigen CAR T-cell therapy for relapsed or refractory multiple myeloma. *Lancet* **2025**, *406*, 228–231.
8. Wang, Q.; Xiao, Z.X.; Zheng, X.; et al. *In vivo* CD19 CAR T-cell therapy for refractory systemic lupus erythematosus. *New Engl. J. Med.* **2025**, *393*, 1542–1544.
9. Harrison, S.; Ho, P.J.; Lim, S.-L.; et al. Minimal residual disease (MRD)-negative outcomes following a novel, *in vivo* gene therapy generating anti-B-cell maturation antigen (BCMA) chimeric antigen receptor (CAR)-T cells in patients with relapsed and refractory multiple myeloma (RRMM): Preliminary results from inMMycAR, the first-in-human phase 1 study of KLN-1010. *Blood* **2025**, *146*, LBA-1.

10. Hunter, T.L.; Bao, Y.; Zhang, Y.; et al. *In vivo* CAR T cell generation to treat cancer and autoimmune disease. *Science* **2025**, *388*, 1311–1317.
11. Amatya, S.; Green, J.; Chu, C.-S.; et al. Enabling Potent T Cell Delivery and Mitigating Key Off-Target Concerns of *in vivo* CAR T Lentiviral Vectors with the Targeted Paramyxovirus Fusogen System. *Blood* **2024**, *144*, 2047.
12. Frank, A.M.; Braun, A.H.; Scheib, L.; et al. Combining T-cell-specific activation and *in vivo* gene delivery through CD3-targeted lentiviral vectors. *Blood Adv.* **2020**, *4*, 5702–5715.
13. Krug, A.; Saidane, A.; Martinello, C.; et al. *In vivo* CAR T cell therapy against angioimmunoblastic T cell lymphoma. *J. Exp. Clin. Cancer Res.* **2024**, *43*, 262.
14. Kapitza, L.; Ho, N.; Kerzel, T.; et al. CD62L as target receptor for specific gene delivery into less differentiated human T lymphocytes. *Front. Immunol.* **2023**, *14*, 1183698.
15. Buchholz, C.J.; Mühlebach, M.D.; Cichutek, K. Lentiviral vectors with measles virus glycoproteins—dream team for gene transfer? *Trends Biotechnol.* **2009**, *27*, 259–265.
16. Streibinger, D.; Frangieh, C.J.; Friedrich, M.J.; et al. Cell type-specific delivery by modular envelope design. *Nat. Commun.* **2023**, *14*, 5141.
17. Nicolai, C.J.; Parker, M.H.; Qin, J.; et al. *In vivo* CAR T-cell generation in nonhuman primates using lentiviral vectors displaying a multidomain fusion ligand. *Blood* **2024**, *144*, 977–987.
18. Michels, K.R.; Sheih, A.; Hernandez, S.A.; et al. Preclinical proof of concept for VivoVec, a lentiviral-based platform for *in vivo* CAR T-cell engineering. *J. ImmunoTherapy Cancer* **2023**, *11*, e006292.
19. Ho, N.; Agarwal, S.; Milani, M.; et al. *In vivo* generation of CAR T cells in the presence of human myeloid cells. *Mol. Ther. Methods Clin. Dev.* **2022**, *26*, 144–156.
20. Chavan, H.; Amatya, S.; Berlfein, V.; et al. In Vitro and *in vivo* Specificity and Biodistribution of a Novel CD8-Targeted Fusosome. *Blood* **2023**, *142*, 3630.
21. Haghighi, E.; Abolmaali, S.S.; Dehshahri, A.; et al. Navigating the intricate in-vivo journey of lipid nanoparticles tailored for the targeted delivery of RNA therapeutics: A quality-by-design approach. *J. Nanobiotechnology* **2024**, *22*, 710.
22. Marshall, A. *In Vivo CAR-Ts Captivate Big Pharma*; Nature Publishing Group: New York, NY, USA, 2025.
23. Lemgart, V.; Richards, J.; Strauss, L.; et al. In Vitro and *In vivo* Activity of a Lipid Nanoparticle System for the *in vivo* Generation of CAR T Cells. *Blood* **2024**, *144*, 3421.
24. Cheng, M.H.Y.; Leung, J.; Zhang, Y.; et al. Induction of bleb structures in lipid nanoparticle formulations of mRNA leads to improved transfection potency. *Adv. Mater.* **2023**, *35*, 2303370.
25. Kariko, K.; Buckstein, M.; Ni, H.; et al. Suppression of RNA recognition by Toll-like receptors: the impact of nucleoside modification and the evolutionary origin of RNA. *Immunity* **2005**, *23*, 165–175.
26. Daniel, S.; Kis, Z.; Kontoravdi, C.; et al. Quality by Design for enabling RNA platform production processes. *Trends Biotechnol.* **2022**, *40*, 1213–1228.
27. Adams, G.; Soldevila, F.; Matsuda, D.; et al. *In vivo* engineering of CAR T cells using a novel targeted LNP-mRNA technology. *J. Immunother. Cancer* **2023**, *11* (Suppl. S1), A1–A731.
28. Ma, Y.; VanKeulen-Miller, R.; Fenton, O.S. mRNA lipid nanoparticle formulation, characterization and evaluation. *Nat. Protoc.* **2025**, *20*, 2618–2651.
29. Hunter, T.; Soldevila, F.; Zhang, Y.; et al. A novel product candidate (CPTX2309) for *in vivo* mRNA engineering of anti-CD19 CAR T cells utilizing novel CD8-targeted lipid nanoparticles. *Arthritis Rheumatol.* **2024**, *76*, 0019.
30. Aghajanian, H.; Frantz, A.M.; Riener, R.; et al. Design and Preclinical Development of a Novel In Vivo Chimeric Antigen Receptor (CAR) Product for B-Cell Involved. *Cytotherapy* **2025**, *27*, S161.
31. Mullard, A. *In vivo* CAR T cells move into clinical trials. *Nat. Rev. Drug Discov.* **2024**, *23*, 727–730.
32. Michels, A.; Ho, N.; Buchholz, C.J. Precision medicine: *In vivo* CAR therapy as a showcase for receptor-targeted vector platforms. *Mol. Ther.* **2022**, *30*, 2401–2415.
33. Annoni, A.; Gregori, S.; Naldini, L.; et al. Modulation of immune responses in lentiviral vector-mediated gene transfer. *Cell. Immunol.* **2019**, *342*, 103802.
34. Wang, Y.; Shao, W. Innate immune response to viral vectors in gene therapy. *Viruses* **2023**, *15*, 1801.
35. Hosseini-Kharat, M.; Bremmell, K.E.; Prestidge, C.A. Why do lipid nanoparticles target the liver? Understanding of biodistribution and liver-specific tropism. *Mol. Ther. Methods Clin. Dev.* **2025**, *33*, 101436.
36. Tsoi, K.M.; MacParland, S.A.; Ma, X.-Z.; et al. Mechanism of hard-nanomaterial clearance by the liver. *Nat. Mater.* **2016**, *15*, 1212–1221.
37. Chan, C.-P.; Jin, D.-Y. Cytoplasmic RNA sensors and their interplay with RNA-binding partners in innate antiviral response: Theme and variations. *Rna* **2022**, *28*, 449–477.
38. Kouwaki, T.; Nishimura, T.; Wang, G.; et al. RIG-I-like receptor-mediated recognition of viral genomic RNA of severe acute respiratory syndrome coronavirus-2 and viral escape from the host innate immune responses. *Front. Immunol.* **2021**, *12*, 700926.

39. Brown, B.D.; Sitia, G.; Annoni, A.; et al. *In vivo* administration of lentiviral vectors triggers a type I interferon response that restricts hepatocyte gene transfer and promotes vector clearance. *Blood* **2007**, *109*, 2797–2805.
40. Adabi, E.; Charitidis, F.T.; Thalheimer, F.B.; et al. Enhanced conversion of T cells into CAR T cells by modulation of the MAPK/ERK pathway. *Cell Rep. Med.* **2025**, *6*, 101970.
41. Chaturvedi, V.; Green, J.; Zipp, G.; et al. Preclinical Efficacy of a Novel CD8-Targeted Fusosome for *In vivo* CAR T Cell Therapy. *Blood* **2023**, *142*, 7121.
42. Charitidis, F.T.; Adabi, E.; Ho, N.; et al. CAR Gene Delivery by T-cell Targeted Lentiviral Vectors is Enhanced by Rapamycin Induced Reduction of Antiviral Mechanisms. *Adv. Sci.* **2023**, *10*, 2302992.
43. Nelson, J.; Sorensen, E.W.; Mintri, S.; et al. Impact of mRNA chemistry and manufacturing process on innate immune activation. *Sci. Adv.* **2020**, *6*, eaaz6893.
44. Moghimi, S.M.; Simberg, D. Pro-inflammatory concerns with lipid nanoparticles. *Mol. Ther.* **2022**, *30*, 2109–2110.
45. Iroanya, G.I.; Subramanyam, P.N.; Wells, K.D.; et al. Pre-Existing Anti-Adeno-Associated Virus Immunity in Gene Therapy: Mechanisms, Challenges, and Potential Solutions. *Hum. Gene Ther.* **2025**, *36*, 1463–1480.
46. Muhuri, M.; Maeda, Y.; Ma, H.; et al. Overcoming innate immune barriers that impede AAV gene therapy vectors. *J. Clin. Investig.* **2021**, *131*, e143780.
47. Gorovits, B.; Fiscella, M.; Havert, M.; et al. Recommendations for the development of cell-based anti-viral vector neutralizing antibody assays. *AAPS J.* **2020**, *22*, 24.
48. Rogers, G.L.; Shirley, J.L.; Zolotukhin, I.; et al. Plasmacytoid and conventional dendritic cells cooperate in crosspriming AAV capsid-specific CD8⁺ T cells. *Blood J. Am. Soc. Hematol.* **2017**, *129*, 3184–3195.
49. Mingozzi, F.; Maus, M.V.; Hui, D.J.; et al. CD8⁺ T-cell responses to adeno-associated virus capsid in humans. *Nat. Med.* **2007**, *13*, 419–422.
50. Chen, J.; Wu, Q.; Yang, P.; et al. Determination of specific CD4 and CD8 T cell epitopes after AAV2-and AAV8-hF. IX gene therapy. *Mol. Ther.* **2006**, *13*, 260–269.
51. Nathwani, A.C.; Tuddenham, E.G.; Rangarajan, S.; et al. Adenovirus-associated virus vector-mediated gene transfer in hemophilia B. *New Engl. J. Med.* **2011**, *365*, 2357–2365.
52. Fraietta, J.A.; Lacey, S.F.; Orlando, E.J.; et al. Determinants of response and resistance to CD19 chimeric antigen receptor (CAR) T cell therapy of chronic lymphocytic leukemia. *Nat. Med.* **2018**, *24*, 563–571.
53. Magni, M.; Jonnalagadda, S.; Bonifazi, F.; et al. CAR T expansion and systemic inflammation: Diverging impacts on large B-cell lymphoma therapy in the multicenter CART SIE study. *Haematologica* **2025**, *110*, 2661.
54. Zhang, D.K.; Adu-Berchie, K.; Iyer, S.; et al. Enhancing CAR-T cell functionality in a patient-specific manner. *Nat. Commun.* **2023**, *14*, 506.
55. Pfeiffer, A.; Thalheimer, F.B.; Hartmann, S.; et al. *In vivo* generation of human CD 19-CAR T cells results in B-cell depletion and signs of cytokine release syndrome. *EMBO Mol. Med.* **2018**, *10*, e9158.
56. Agarwal, S.; Weidner, T.; Thalheimer, F.B.; et al. *In vivo* generated human CAR T cells eradicate tumor cells. *Oncoimmunology* **2019**, *8*, e1671761.
57. Andorko, J.I.; Russell, R.M.; Schnepf, B.C.; et al. Targeted *in vivo* delivery of genetic medicines utilizing an engineered lentiviral vector platform results in CAR T and NK cell generation. *Mol. Ther.* **2025**, *33*, 4937–4952.
58. Beura, E.T.; Latimer, H.J.; Wong, D.; et al. T cell-specific *in vivo* transduction with preclinical candidate KLN-1010 generates BCMA directed CAR T cells with potent anti-multiple myeloma activity. *Cancer Res.* **2024**, *84* (Suppl. S6), 48.
59. Contrastano, S.G.; Beura, E.T.; Wong, D.; et al. Potent *in vivo* Transduction by iGPS Particles Generates CAR T Cells with Durable Anti-Tumor Activity. In *Molecular Therapy*; Cell Press 50 Hampshire St, Floor 5: Cambridge, MA, USA, 2023; Volume 31, pp. 48–49.
60. Billingsley, M.M.; Gong, N.; Mukalel, A.J.; et al. *In vivo* mRNA CAR T cell engineering via targeted ionizable lipid nanoparticles with extrahepatic tropism. *Small* **2024**, *20*, 2304378.
61. Zhang, N.; Seo, J.W.; Robinson, E.; et al. Development of an *in situ* CAR-T cell protocol through optical and PSMA-targeted PET imaging. *Proc. Natl. Acad. Sci. USA* **2025**, *122*, e2504950122.
62. Rurik, J.G.; Tombácz, I.; Yadegari, A.; et al. CAR T cells produced *in vivo* to treat cardiac injury. *Science* **2022**, *375*, 91–96.
63. Mackensen, A.; Müller, F.; Mougiakakos, D.; et al. Anti-CD19 CAR T cell therapy for refractory systemic lupus erythematosus. *Nat. Med.* **2022**, *28*, 2124–2132.
64. Sawyer, A.J.; Lemgart, V.; Kuhlman, W.; et al. Efficient generation of active CAR-T cells from healthy donors and lymphoma patients. *Cancer Res.* **2025**, *85* (Suppl. S8), 6119.
65. Tombácz, I.; Laczkó, D.; Shah Nawaz, H.; et al. Highly efficient CD4⁺ T cell targeting and genetic recombination using engineered CD4⁺ cell-homing mRNA-LNPs. *Mol. Ther.* **2021**, *29*, 3293–3304.
66. Mougiakakos, D.; Krönke, G.; Völkl, S.; et al. CD19-targeted CAR T cells in refractory systemic lupus erythematosus. *New Engl. J. Med.* **2021**, *385*, 567–569.

67. Berdeja, J.G.; Madduri, D.; Usmani, S.Z.; et al. Ciltacabtagene autoleucel, a B-cell maturation antigen-directed chimeric antigen receptor T-cell therapy in patients with relapsed or refractory multiple myeloma (CARTITUDE-1): A phase 1b/2 open-label study. *Lancet* **2021**, *398*, 314–324.
68. Wu, L.S.; Su, Y.; Li, C.; et al. Population-based cellular kinetic characterization of ciltacabtagene autoleucel in subjects with relapsed or refractory multiple myeloma. *Clin. Transl. Sci.* **2022**, *15*, 3000–3011.
69. Cameau, E.; Zhang, P.; Ip, S.; et al. Process & analytical insights for GMP manufacturing of mRNA lipid nanoparticles. *Cell Gene Ther. Insights* **2022**, *8*, 621–635.
70. Harrison, S.J.; Nguyen, T.; Rahman, M.; et al. CAR+ T-cell lymphoma post ciltacabtagene autoleucel therapy for relapsed refractory multiple myeloma. *Blood* **2023**, *142*, 6939.
71. Ghilardi, G.; Fraietta, J.A.; Gerson, J.N.; et al. T cell lymphoma and secondary primary malignancy risk after commercial CAR T cell therapy. *Nat. Med.* **2024**, *30*, 984–989.
72. Tix, T.; Alhomoud, M.; Shouval, R.; et al. Second primary malignancies after CAR T-cell therapy: A systematic review and meta-analysis of 5517 lymphoma and myeloma patients. *Clin. Cancer Res.* **2024**, *30*, 4690–4700.
73. Tang, L.; Cai, D.; Yan, X.; et al. Second-line chimeric antigen receptor T-cell therapy versus standard of care in relapsed or refractory large B-cell lymphoma: A systematic review and meta-analysis. *Cancer* **2026**, *132*, e70317.
74. Ruella, M.; Xu, J.; Barrett, D.M.; et al. Induction of resistance to chimeric antigen receptor T cell therapy by transduction of a single leukemic B cell. *Nat. Med.* **2018**, *24*, 1499–1503.
75. Smith, E.C.; Popa, A.; Chang, A.; et al. Viral entry mechanisms: The increasing diversity of paramyxovirus entry. *FEBS J.* **2009**, *276*, 7217–7227.
76. Canelo-Vilaseca, M.; Sabbah, M.; Di Blasi, R.; et al. Lymphodepletion chemotherapy in chimeric antigen receptor-engineered T (CAR-T) cell therapy in lymphoma. *Bone Marrow Transplant.* **2025**, *60*, 559–567.
77. Grosser, R.; Cherkassky, L.; Chintala, N.; et al. Combination immunotherapy with CAR T cells and checkpoint blockade for the treatment of solid tumors. *Cancer Cell* **2019**, *36*, 471–482.
78. Tenchov, R.; Sasso, J.M.; Zhou, Q.A. PEGylated lipid nanoparticle formulations: Immunological safety and efficiency perspective. *Bioconjugate Chem.* **2023**, *34*, 941–960.
79. Hassanel, D.N.B.; Ju, Y.; Takanashi, A.; et al. Influence of hydrophilic polymers on the accelerated blood clearance of mRNA lipid nanoparticles upon repeated administration. *Nanoscale Horiz.* **2025**, *10*, 3396–3404.
80. Stavnsbjerg, C.; Christensen, E.; Münter, R.; et al. Accelerated blood clearance and hypersensitivity by PEGylated liposomes containing TLR agonists. *J. Control. Release* **2022**, *342*, 337–344.
81. Sterner, R.C.; Sterner, R.M. CAR-T cell therapy: Current limitations and potential strategies. *Blood Cancer J.* **2021**, *11*, 69.
82. Jin, C.; Ma, J.; Ramachandran, M.; et al. CAR T cells expressing a bacterial virulence factor trigger potent bystander antitumor responses in solid cancers. *Nat. Biomed. Eng.* **2022**, *6*, 830–841.
83. Adachi, K.; Sakoda, Y.; Kaisho, T.; et al. CAR-T cells co-expressing IL-7 and CCL19 promote epitope spreading to enhance antitumor immunity. *Cancer Immunol. Immunother.* **2026**, *75*, 68.
84. Argueta, S.; Wang, Y.; Zhao, H.; et al. *In vivo* programmed myeloid cells expressing novel chimeric antigen receptors show potent anti-tumor activity in preclinical solid tumor models. *Front. Immunol.* **2024**, *15*, 1501365.