Biosensing and Biomedicine
https://www.sciltp.com/journals/bab

Perspective

From *Ex Vivo* to *In Vivo*: The Next Generation of CAR-T Therapy

Yuhan Jiang and Ke-Fei Xu *

State Key Laboratory of Digital Medical Engineering, School of Biological Science and Medical Engineering, Southeast University, No. 2 Southeast University Road, Nanjing 211189, China
* Correspondence: kf.xu@seu.edu.cn

How To Cite: Jiang, Y.; Xu, K-F. From *Ex Vivo* to *In Vivo*: The Next Generation of CAR-T Therapy. *Biosensing and Biomedicine* **2026**, *1*(1), 6. https://doi.org/10.53941/bab.2026.100006

Received: 14 May 2026
Revised: 31 May 2026
Accepted: 13 October 2026
Published: 27 October 2026

Abstract: Chimeric antigen receptor T cell (CAR-T) therapy has transformed hematological cancer treatment, but its broader clinical application remains constrained by the exorbitant costs and complex procedures of *ex vivo* manufacturing. To overcome these limitations, *in vivo* CAR-T engineering strategies are emerging as compelling alternatives, enabling direct reprogramming of T cells within the patients. Recent studies indicate that *in vivo* CAR-T therapy is evolving from a delivery-centered concept into an integrated engineering framework that should coordinate cell-selective targeting, controllable CAR expression, T-cell fitness, and long-term safety. Looking forward, this field is likely to move beyond vector innovation alone toward multidimensional platform design, which is comprehensively discussed in this perspective.

Keywords: CAR-T therapy; targeted delivery; immune therapy; multidimensional platform design

1. Introduction

Over the past decade, chimeric antigen receptor T cell (CAR-T) therapy has revolutionized the landscape of immuno-oncology, particularly in the treatment of hematological malignancies. This breakthrough has fundamentally altered the therapeutic paradigm for relapsed or refractory B-cell malignancies, achieving remarkable complete remission rates. However, as CAR-T therapies transition from specialized academic centers into broader clinical and commercial application, the inherent limitations (e.g., exorbitant costs and complex procedures) of traditional *ex vivo* manufacturing have emerged as a major translational bottleneck. Currently, all approved CAR-T products rely on a highly individualized and labor-intensive supply chain [1]. This process entails leukapheresis to isolate patient-derived T cells, followed by *ex vivo* activation and viral transduction, expansion in bioreactors, and reinfusion under complex cold chain logistics [2]. Therefore, conventional CAR-T therapy imposes an extreme financial burden, which ranges from $300,000 to $500,000 per dose, while necessitating vein-to-vein intervals of several weeks that may permit rapid disease progression during the manufacturing window [3,4]. Furthermore, patients must undergo highly toxic lymphodepleting chemotherapy that can prevent host immune rejection of the modified CAR-T cells and create a supportive ecological niche for expansion of CAR-T cells. These limitations have facilitated a shift towards *in vivo* CAR-T cell engineering. While the complex *ex vivo* CAR-T therapy process requires approximately 7–22 days, *in vivo* CAR-T therapy aims to reprogram endogenous T cells directly within the patients by simply injecting the targeted delivery vehicles delivering CAR-encoding mRNA, DNA templates, or genome-editing components, potentially generating functional CAR-T cells within 3 days (Figure 1).

By reducing the requirement for intensive chemotherapeutic conditioning and extensive *ex vivo* manipulation, *in vivo* CAR-T engineering strategies extend the clinical application of CAR-T therapy beyond blood cancers to autoimmune diseases such as systemic lupus erythematosus. Current development of *in vivo* CAR-T engineering

Copyright: © 2026 by the authors. This is an open access article under the terms and conditions of the Creative Commons Attribution (CC BY) license (https://creativecommons.org/licenses/by/4.0/).
Publisher's Note: Scilight stays neutral with regard to jurisdictional claims in published maps and institutional affiliations.

is no longer simply divided between nanoparticles delivering transient messenger RNA (mRNA) payloads and viral vectors (such as adeno-associated virus (AAV) and lentivirus) delivering integrating DNA payloads. In addition, implant platforms offer an alternative strategy for localized and sustained delivery of CAR-encoding viral vectors or nanoparticles (Table 1). Instead, recent studies have proposed a broader set of design questions: how CAR constructs can be precisely delivered to the target cells, how long CAR expression should persist, and how engineered T cells behave after they are generated directly within the host [5,6].

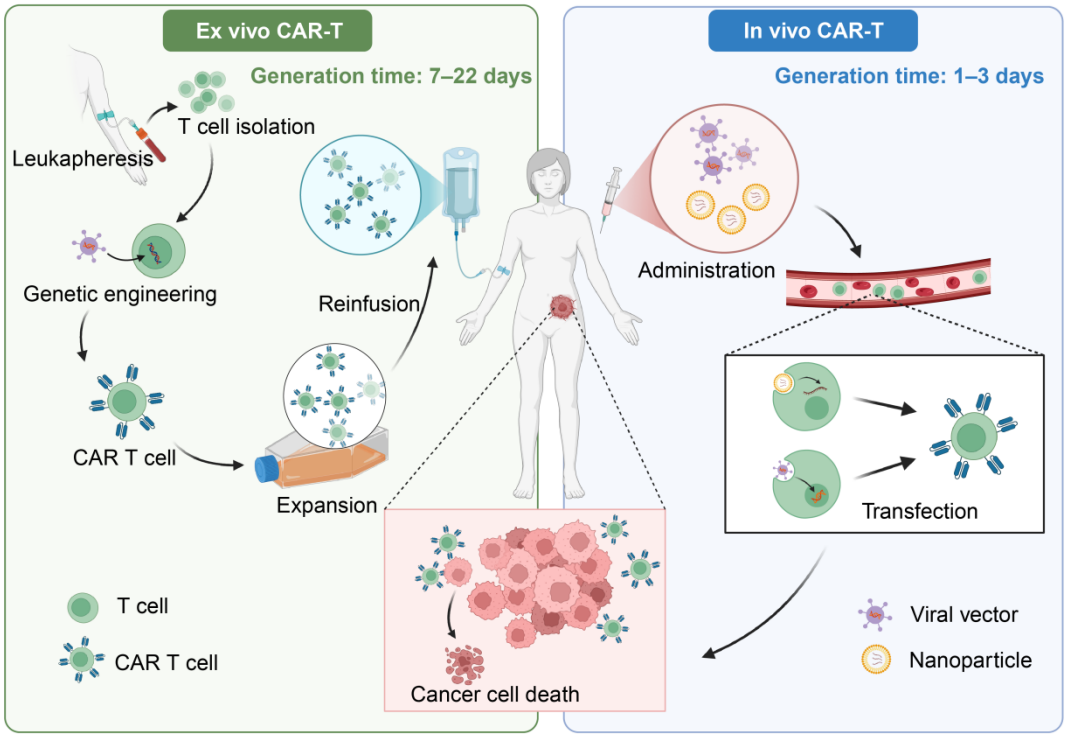


Figure 1. Schematic illustrating the processes of *ex vivo* and *in vivo* CAR-T therapies in cancer treatment.

Table 1. Comparison of three main platforms of *in vivo* CAR-T therapy.

Dimension	Virus-Based Vector	Nanoparticle Carrier	Implant Platforms
Platform carrier type	Lentivirus; Adeno-associated virus; Retrovirus	Lipid nanoparticles; Polymeric nanoparticles	Implants loaded with CAR virus vectors or nanoparticles
Advantages	Durable CAR expression; Suitable for settings requiring long-term immune surveillance	Scalable and chemically tunable; Low immunogenicity and genotoxicity; Suitable for transient CAR expression	Localized and sustained release; Reduced systemic exposure
Limitations	Random or semi-random integration for integrating vectors; Off-target transduction; High immunogenicity and redosing barriers	Short persistence for mRNA payloads; Hepatic exposure and off-target uptake	More invasive; Spatially restricted; Limited applicability to disseminated disease

Abbreviations: CAR: chimeric antigen receptor.

2. Recent Advances

Recently, a series of studies have expanded the conceptual framework of *in vivo* CAR-T therapy, transforming it from a proof-of-concept strategy into a programmable platform for immune engineering. For example, Hunter et al. developed a CD8-targeting lipid nanoparticle (LNP) capable of delivering CAR mRNA to specific T cell subsets [7]. Specifically, this LNP platform combined optimized CAR mRNA payload that could enhance CAR expression and functionality with a new ionizable lipid (termed L829) that was able to reduce liver delivery, enhance mRNA delivery to immune cells, and reduce reactogenicity. Moreover, the LNP platform was further functionalized with anti-CD8 antibodies to preferentially engineer $CD8^+$ T cells, thereby limiting unwanted $CD4^+$ T cell activation that may contribute to cytokine-release toxicity. This LNP platform enabled transient and non-integrating CAR expression in $CD8^+$ T cells, demonstrating that endogenous T cells can be rapidly

reprogrammed without leukapheresis, viral integration, or *ex vivo* expansion. This approach produced functional CAR-T cells capable of tumor control in humanized mice and deep B cell depletion in cynomolgus monkeys. Importantly, the reconstituted B-cell compartment after depletion was predominantly naïve, suggesting an "immune reset" that may be especially relevant for autoimmune diseases. The conceptual advance of this work is not durable integration, but controllable immune redirection: CAR expression is initiated by LNP administration, declines as the mRNA payload is cleared, and potentially be adjusted through repeat dosing.

Notably, Gattinoni et al. shifted the discussion from how CAR-T cells are delivered to what kind of CAR-T cells can sustain long-term functionality after *in vivo* establishment [8]. In patients with relapsed B cell malignancies after allogeneic hematopoietic stem cell transplantation, researchers compared donor-derived stem cell memory CAR-T cells with standard CAR-T cells. In details, stem cell memory CAR-T cells supported stronger expansion, longer persistence, and complete responses at lower doses without lymphodepletion. Mechanistically, CAR-modified stem-cell memory T (T_{SCM}) cells did not simply expand transiently and then disappear. Instead, they repopulated the stem-like compartment over time and maintained long-term engraftment through waves of clonal succession. These findings suggest that the efficacy of CAR-T therapy is not merely a delivery problem, but also a cell-fate problem. Consequently, future *in vivo* engineering strategies must generate not only CAR-expressing cells, but CAR-expressing cells with the appropriate memory phenotype, differentiation trajectory, persistence, and resistance to exhaustion.

In 2026, Nyberg et al. engineered a two-vector system that shifted *in vivo* CAR-T generation from transient expression or random DNA integration towards site-specific T cell engineering [9]. The platform combined enveloped delivery vehicles (EDVs) that delivered Clustered Regularly Interspaced Short Palindromic Repeats–CRISPR-associated protein 9 (CRISPR–Cas9) ribonucleoproteins (RNPs) with AAVs that delivered a DNA donor template, aiming to achieve stable and cell-specific transgene expression through *in vivo* integration of large DNA payloads. Notably, this strategy was designed to integrate a promoterless CAR transgene into the endogenous T cell receptor alpha constant (TRAC) locus, allowing CAR expression to be driven by the native T cell-specific TRAC promoter. To improve T cell-specific delivery, the researchers evolved an AAV variant (i.e., AAV-hT7) for human T cell tropism and identified CD7 as an important host factor for AAV-hT7 uptake. They also engineered anti-CD3 EDVs by incorporating the mutated glycoprotein G from vesicular stomatitis virus (VSVG) and an anti-CD3 single-chain variable fragment (scFv) that conferred both targeting and T cell activation. After CRISPR–Cas9 and AAV-mediated homology-directed repair (HDR), CAR expression became more uniform and physiologically regulated, potentially reducing ectopic expression, tumor cell antigen masking, and heterogeneous CAR expression associated with random integration.

Collectively, these studies highlight the major advantages of *in vivo* CAR-T strategies over conventional *ex vivo* manufacturing for shifting T cell engineering from *ex vivo* cell-manufacturing workflows into the patient's body. More importantly, *in vivo* CAR-T is not a single technological route, but a flexible engineering framework which can match the engineering strategy to the therapeutic need: transient immune reset for autoimmune diseases, durable immune surveillance for cancer, and precise control of cell state or genome editing when requiring long-term efficacy and safety.

3. Challenges

Despite rapid progress, emerging evidence suggests that several unresolved barriers may constitute fundamental and enduring challenges for the broader clinical translation of *in vivo* CAR-T therapy. The first challenge is selective, efficient, and repeatable delivery to endogenous T cells. Notably, Hunter et al. have demonstrated transient and non-integrating CAR expression along with deep B cell depletion in non-human primates. However, this short-course strategy may be insufficient for immunosuppressive tumor microenvironment that requires sustained immune surveillance and prolonged CAR-T activity [10]. Although compact repeat dosing has been demonstrated, the long-term feasibility of iterative administration requires systematic evaluation of immunogenicity, antidrug antibodies, cytokine activation, hepatic exposure, biodistribution, and potential cumulative toxicity. Therefore, future solutions require programmable dosing schedules, transient immune premedication, and tissue- or subset-selective targeting ligands that reduce off-target delivery while preserving T cell engineering efficiency.

Another major challenge is that the efficacy of *in vivo* CAR-T therapy depends not only on delivery efficiency, but also on the functional fitness and cell state of the engineered T cells. Current *in vivo* delivery systems do not yet fully control which state of T cells is generated after engineering. Future platforms may need to combine CAR delivery with strategies that preserve or bias engineered cells toward memory-like states, including calibrated CAR signalling, cytokine support, modulation of exhaustion, and differentiation pathways.

Additionally, the foundational reliance on CRISPR–Cas9 to induce DNA double-strand breaks (DSBs) presents a clinical transition risk. In the context of multiplex editing–such as knocking out inhibitory receptors (e.g., programmed cell death protein 1 (PD-1)) alongside integrating the CAR gene to overcome solid tumor microenvironments–simultaneous DSBs frequently lead to genomic instability, including large chromosomal deletions and translocations, emphasizing the urgent need for DSB-free tools. Therefore, the future of *in vivo* CAR-T engineering may depend on next-generation genomic processors, specifically base editors (BE) and prime editors (PE). BE enable DSB-free single-base conversion and may support multiplex knockout of endogenous receptors or immune checkpoints by installing stop codons. PE and emerging gene-writing systems may further enable template-guided sequence rewriting without DSBs [11]. However, full-length CAR transgene insertion remains technically challenging because of payload size, editing efficiency, and delivery constraints.

4. Outlook

The transition from *ex vivo* to *in vivo* CAR-T therapy has rapidly moved from academic theory to early industrial translation [12]. A prime example of this industrial validation occurred when Kite Pharma, a subsidiary of Gilead Sciences and a titan in conventional cell therapy, executed a massive acquisition of the *in vivo* CAR-T startup Interius BioTherapeutics. This strategic buyout implies that *in vivo* engineering may broaden access to current technologies. As Interius executives stated, "This marks a pivotal step for Interius and the future of *in vivo* therapy, which has the potential to reduce treatment timelines, broaden access to care and improve outcomes for patients with aggressive or advanced disease".

Given the rapid progress of *in vivo* CAR-T therapy, future development should move beyond feasibility demonstration toward more precise and context-specific design principles. In particular, platform design should integrate T-cell subset selection with disease-related therapeutic requirements. For example, $CD8^+$ T cells may be preferred for rapid cytotoxicity with reduced inflammatory risk, whereas $CD4^+$ T cells may provide helper functions. Likewise, naïve, stem-like, or central memory T-cell subsets may support persistence and exhaustion resistance. Furthermore, disease type should also guide platform selection. Specifically, autoimmune diseases may favor transient mRNA-LNP platforms that induce controllable immune reset while potentially reducing the risk of prolonged B-cell aplasia. By contrast, hematological malignancies may require more durable but safety-controlled systems, such as viral vectors. Solid tumors may require genome-editing or combinatorial delivery platforms that engineer memory-like T-cell subsets for tumor infiltration, persistence, exhaustion resistance, and tumor microenvironment remodeling.

Ultimately, the next generation of *in vivo* CAR-T therapy will not be defined by a single vector, but by the integration of selective delivery, controllable CAR expression, durable T cell fitness, and safe genome engineering. By eliminating the *ex vivo* manufacturing bottleneck from the patient and transitioning the engineering process into the human body's natural bioreactor, *in vivo* CAR-T therapy promises to broaden access to cellular medicines and reshape the future of immune therapy.

Author Contributions

Y.J.: data curation, writing—original draft preparation; visualization, investigation, software, validation; K.-F.X.: conceptualization, methodology, software, supervision, writing—reviewing and editing. Both authors have read and agreed to the published version of the manuscript.

Funding

This work was supported by China Postdoctoral Science Foundation (2024M760443), Natural Science Foundation of Jiangsu Province (BK20251337), Postdoctoral Fellowship Program of China Postdoctoral Science Foundation (CPSF) (GZB20250672), and Jiangsu Funding Program for Excellent Postdoctoral Talent (2025ZB293).

Institutional Review Board Statement

Not applicable.

Informed Consent Statement

Not applicable.

Data Availability Statement

Not applicable.

Conflicts of Interest

The authors declare no conflict of interest.

Use of AI and AI-Assisted Technologies

During the preparation of this work, the authors used Chat GPT to polish language. After using this tool, the authors reviewed and edited the content as needed and took full responsibility for the content of the published article.

References

1. Singh, N.; Orlando, E.; Xu, J.; et al. Mechanisms of resistance to CAR T cell therapies. *Semin. Cancer Biol.* **2020**, *65*, 91–98. https://doi.org/10.1016/j.semcancer.2019.12.002.
2. Parayath, N.N.; Stephan, M.T. *In Situ* Programming of CAR T Cells. *Annu. Rev. Biomed. Eng.* **2021**, *23*, 385–405. https://doi.org/10.1146/annurev-bioeng-070620-033348.
3. Alhallak, K.; Sun, J.; Muz, B.; et al. Nanoparticle T cell engagers for the treatment of acute myeloid leukemia. *Oncotarget* **2021**, *12*, 1878–1885. https://doi.org/10.18632/oncotarget.28054.
4. Ellis, G.I.; Sheppard, N.C.; Riley, J.L. Genetic engineering of T cells for immunotherapy. *Nat. Rev. Genet.* **2021**, *22*, 427–447. https://doi.org/10.1038/s41576-021-00329-9.
5. Li, Y.R.; Zhu, Y.; Yang, L. Clinical technology advances driving *in vivo* CAR therapy. *Trends Biotechnol.* **2026**, *44*, 1821–1824. https://doi.org/10.1016/j.tibtech.2025.11.006.
6. Aalipour, A.; Barreiro, A.; Garmilla, A.; et al. *In vivo* CAR T cell engineering: Design principles and open questions. *Trends Cancer* **2026**, *12*, 434–446. https://doi.org/10.1016/j.trecan.2026.01.007.
7. 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. https://doi.org/10.1126/science.ads8473.
8. Gattinoni, L.; Inchingolo, G.; Harrer, D.C.; et al. Distinct *in vivo* dynamics of donor-derived stem cell memory CAR T cells post-allogeneic HSCT relapse. *Cell* **2026**, *189*, 3686–3700.e10. https://doi.org/10.1016/j.cell.2026.03.047.
9. Nyberg, W.A.; Bernard, P.L.; Ngo, W.; et al. *In vivo* site-specific engineering to reprogram T cells. *Nature* **2026**, *652*, 712–721. https://doi.org/10.1038/s41586-026-10235-x.
10. Zugasti, I.; Espinosa-Aroca, L.; Fidyt, K.; et al. CAR-T cell therapy for cancer: Current challenges and future directions. *Signal Transduct. Target. Ther.* **2025**, *10*, 210. https://doi.org/10.1038/s41392-025-02269-w.
11. Chauhan, V.P.; Sharp, P.A.; Langer, R. Engineered prime editors with minimal genomic errors. *Nature* **2025**, *646*, 1254–1260. https://doi.org/10.1038/s41586-025-09537-3.
12. An, N.; Wang, D.; Zhang, P.; et al. *In vivo* generation of anti-BCMA CAR-T cells in relapsed or refractory multiple myeloma: A phase 1 study. *Nat. Med.* **2026**, *32*, 1257–1266. https://doi.org/10.1038/s41591-026-04244-6.