# Towards Hardware-Governed Benchtop DNA Synthesizers

Oraya Srimokla James Petrie

April 26, 2026

**Abstract**

Benchtop DNA synthesizers may soon enable bioweapon synthesis in individual labs without hardware-enforced controls. We propose a hardware design with three layers of defense: sequence screening, a regulator signature the device refuses to run without, and physical monitoring of the synthesis process. The first two reuse hardware primitives from AI chip governance. The third is novel, and addresses an attacker who submits a benign sequence and physically tampers with the device to produce a hazardous one instead.

## 1 Introduction

Benchtop DNA synthesizers are being developed to speed up biological R&D, by letting labs make DNA on demand instead of ordering it from centralized providers. In the next few years, these devices are expected to be capable of synthesizing significant portions of viral genomes (NTI 2023; Langenkamp 2024). The same technology is dual-use: it could also be used to make bioweapons. Today there is no mandatory sequence or customer screening for benchtop synthesizers beyond export controls (ACA 2025). Recent legislation in the US (U.S. Congress 2026) and the EU (European Commission 2025) is starting to change this, but will need hardware enforcement to be effective against motivated misuse.

In the best case, benchtop synthesizers could be designed so that they are useful for benign applications but unusable for dangerous ones. This report presents a preliminary design for how such a device could be built. The design uses three layers of defense: sequence screening, device authorization, and process integrity monitoring. Two different threats motivate this structure.

**The first threat is a malicious operator who submits a hazardous sequence directly.** This is a similar problem to the one faced by traditional mail-in DNA synthesis providers, who screen orders and customers before fulfilling them (Baum et al. 2024; IBBIS 2024). Benchtop synthesizers could use the same screening tools, but reliable enforcement on a device that sits in the user's lab requires dedicated hardware. Recent work on governing AI chips has developed the hardware primitives needed for this kind of enforcement in a related setting (Brass and Aarne 2024; Petrie 2024; Petrie et al. 2025; O'Gara et al. 2025). One contribution of this report is to show that these primitives transfer almost directly to benchtop synthesizers, which means that much of the hardware infrastructure for benchtop biosecurity is closer to being ready than it might appear.

**The second threat is process tampering: rather than requesting a hazardous sequence, the attacker modifies the device so that it produces one during synthesis.** This is the harder part of the problem, and it has no direct analogue in the AI chip setting. An operator with several hours of access to the inside of a synthesizer can interfere with reagent flows, mechanical components, or reaction conditions. For example: the operator submits a benign sequence $S$;

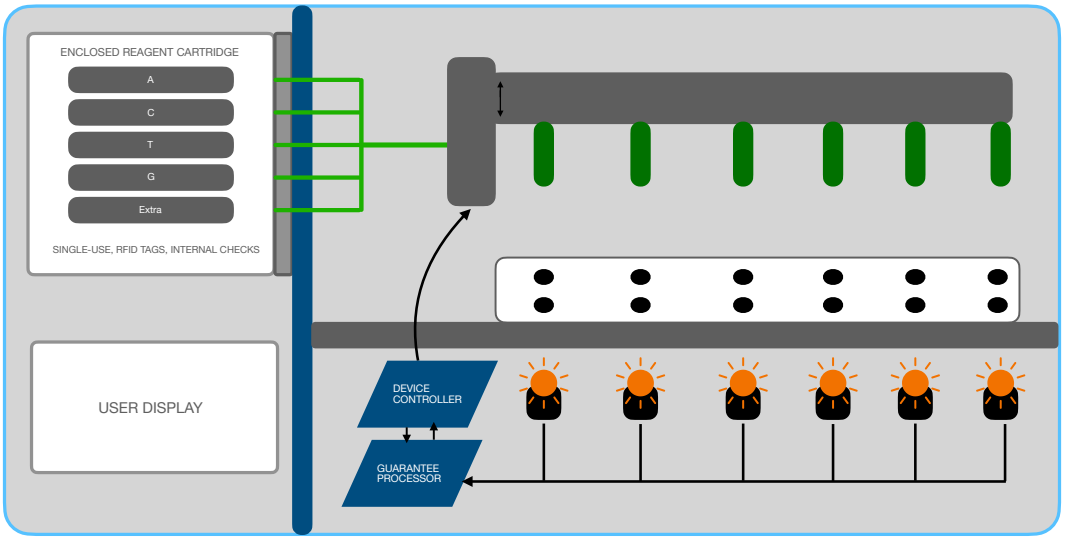


Figure 1: Mock layout of a hardware-governed benchtop synthesizer. The single-use, RFID-tagged reagent cartridge (left) feeds dispense lines into a microfluidic head that moves over an array of reaction wells, with in-line sensors (dot row) and heating/optical units (orange) reporting back to the device controller and guarantee processor. The blue outline indicates the tamper-evident enclosure.

the regulator signs a hash of $S$ authorizing synthesis; screening passes; synthesis begins. But the operator has pre-installed a small mechanical clamp on one reagent line, timed to block a few dispenses partway through the run. The device produces a sequence $S'$ that differs from $S$ by a few missing bases, chosen by the attacker so that $S'$ encodes a hazard while $S$ does not. The submitted sequence, the regulator's signature, and the screening database are all honest; the attack is on the physical execution of the synthesis. Neither cryptographic authorization nor algorithmic screening can prevent it.

Securing a complex physical process against actors with physical access is difficult in general, and the design presented here is not secure against highly sophisticated adversaries such as well-resourced state actors. The goal is to make circumvention infeasible for mid-resource attackers, such as skilled hobbyists or small teams with machine-shop access, who are the most likely misuse path. Most of the novel technical content in this report is in the process-integrity layer, which is the focus of Section 2.3.

## 2 Design

The design adds three things to a standard benchtop synthesizer: a *guarantee processor* that checks each run before it starts, *sensors* that measure the physical process as it happens, and *physical modifications* that make it harder to tamper with the device. The guarantee processor sits alongside the existing device controller. These additions support three mechanisms, applied in order for each

synthesis request: the sequence is screened against a hazard database; the device refuses to run unless it has a fresh regulator signature confirming that the screening passed; and the physical process is monitored during synthesis to detect attempts to produce a different sequence than the one approved.

### 2.1 Sequence screening

Every requested sequence is checked against a hazard database before synthesis begins. This is the same problem faced by mail-in DNA synthesis providers, and existing tools such as SecureDNA (Baum et al. 2024) and commec (IBBIS 2024) are designed for it. Current tools work well for homology-based screening of longer sequences, but have known weaknesses against AI-designed variants that are functionally hazardous but sequence-distant from known hazards (Wittmann et al. 2025) and against fragment-level requests that can be assembled into hazards downstream (Edison et al. 2026). The specific screening algorithm is out of scope for this report.

Screening can happen in two places. In the *remote* path, the device computes a privacy-preserving representation of the sequence, for example the oblivious hashes produced by SecureDNA's DOPRF (Baum et al. 2024), sends it to an external screening service, and receives back a signed attestation that those hashes passed screening. In the *local* path, screening is performed on the device itself by the *guarantee processor*, a general-purpose processor inside the synthesizer that carries an approved screening tool and a signed hazard database. The guarantee processor runs the screening tool and produces a signed record of the result, including the hash of the sequence screened and the version of the tool and database used. This is the same pattern used in flexHEG (Petrie et al. 2025) for policy checks on AI chips.

In both cases the output is an *evidence-of-safety bundle* that commits to a specific sequence (via hashes of some kind) and certifies that it passed screening under a specific policy. The bundle does not contain the plaintext sequence. Because the guarantee processor is general-purpose and updateable, newer screening algorithms can be deployed to the field as they become available, and experimental classifiers such as ML-based functional predictors can run in advisory mode alongside homology-based screening and be promoted to enforcement once their false-positive rates are acceptable.

### 2.2 The device only runs with recent approval

The device is designed to refuse to synthesize anything unless it has a recent authorization signature from the regulator. For each requested job, the guarantee processor first obtains an evidence-of-safety bundle as described in Section 2.1, then sends a request to the regulator containing the bundle, a freshly generated nonce, and the device ID. The regulator verifies the screening evidence and returns a signature over the whole request. The guarantee processor will only allow the device controller to proceed if a valid signature is received.

Before signing, the regulator can require additional checks:

1. **Operator identity.** The request can be accompanied by a signed exemption certificate from the operator's institutional biosafety authority, using SecureDNA's delegated PKI (SecureDNA 2024). This matches how approval already works for traditional mail-in synthesis and avoids building a new know-your-customer system.
2. **Hardware 2FA.** The operator presents a one-time code from a hardware key (e.g., YubiKey) tied to their exemption certificate.

3. **Device location.** An external service pings the device and uses round-trip times to estimate its location (Brass and Aarne 2024). The regulator can decline to sign if the device is in a restricted location.

Which of these are required is a policy decision.

The guarantee processor's job is to check that the authorization is valid *and* that the sequence the controller is about to synthesize is the one the authorization refers to. This consistency check is what makes the authorization meaningful: without it, an operator could obtain a signature for a benign sequence and then instruct the controller to run a different sequence. For remote screening, the guarantee processor recomputes the privacy-preserving hashes (e.g., DOPRF output) from the sequence about to be synthesized and checks that they match the ones in the authorization bundle. For local screening, it recomputes the sequence hash and checks that it matches the one in its own earlier signed record. In either case, a mismatch causes synthesis to halt.

Secure boot and firmware rollback protection on both the guarantee processor and the device controller prevent an attacker from loading modified firmware that would bypass the approval check. If these protections fail, the entire authorization mechanism fails with them.

A useful side effect of this process is that the regulator ends up holding a signed record of every approved sequence (as hashes), for every run. This is an external, authoritative record of what the device was authorized to make. Baker and Church (2024) and Sandbrink (2021) motivate strong record-keeping for synthesis attribution; the authorization provides this record as a byproduct of the approval flow, without requiring a separate on-device log.

### 2.3 Process integrity

The first two mechanisms ensure that the guarantee processor only lets the device controller synthesize approved sequences. They do not ensure that the controller and the wet components actually produce what was asked for. A malicious operator with access to the inside of the device can tamper with the physical process so that the controller runs the approved program but the chemistry produces something different. The clamp attack from Section 1 is one instance of a more general class; others include:

1. Swapping the order or identity of reagents mid-synthesis (e.g. exchanging cartridges).

2. Rotating or shifting the substrate so dispenses land in the wrong wells.

For any of these to produce a hazardous product rather than just a corrupted one, the operator picks an approved sequence that one of these edits can turn into a hazard (see Edison et al. 2026 for the cross-order version of this idea). None of these attacks are prevented by the authorization mechanism, because the submitted sequence is genuinely benign and the regulator's signature is honest. The approach here is defense in depth, with several independent mitigations.

**Make tampering physically difficult.** Fluid pathways inside the device should be as inaccessible as possible. Microfluidic cartridges with sealed internal channels remove most of the externally clampable tubing that a simple attack would target. For well-based enzymatic synthesis, single-use microfluidic plates (as used by Kilobaser (2024)) exploit the fact that the linker chemistry and cleavage molecules are consumed during a single run, which naturally prevents reuse. The enclosure should be tamper-evident, so that opening it leaves visible traces. An air pressure sensor inside the enclosure can detect whether it is opened during a run, and the guarantee processor can halt synthesis and optionally heat the sample to destroy it if the enclosure is breached.

**Make cartridges hard to replicate and tamper with.** Each cartridge carries a secure element that signs dispense operations, similar to the authentication chips used in inkjet printer cartridges. Asymmetric signatures are preferred over symmetric message authentication codes, so that a compromised synthesizer cannot forge cartridges that other devices will accept. A monotonic counter in the cartridge's secure element defeats attacks that reuse a genuine cartridge chip with refilled or substituted reagents. Cartridges can be given irregular physical shapes that make third-party replication harder, and individual reagent cartridges can be nested inside a sealed outer enclosure. The cartridge's authentication ID can be bound to the device's tamper-sensor state, so that opening the enclosure invalidates the ID even if the cartridge itself is not damaged.

**Verify reagent identity.** Cryptographic authentication of the cartridge proves that the cartridge is genuine, but not that the reagent inside matches the base it is labeled with. Spectroscopy of the reagent stream can close this gap. Near-infrared (NIR) spectroscopy measures molecular vibrations from absorbed light and is relatively cheap; Raman spectroscopy measures scattered laser light and is more accurate but more expensive. Either can be scheduled to verify the reagent in each channel against a reference spectrum at some fraction of dispense cycles (for example, every 10–20 bases) without substantially extending synthesis time. For column-based synthesizers, a UV-Vis diode array at the flow cell can monitor the spectrum of the DMT group released during each nucleotide addition, which confirms both that an addition occurred and which base was added. Existing benchtops generally do not record this spectrum at single-base resolution because they only need bulk yield; for process integrity it is worth doing so. For well-based enzymatic synthesizers, a scanning Raman or UV-Vis detector at the dispensing nozzle or under the well plate can confirm which base is being dispensed into which well.

**Verify reagent volume.** In-line flow meters at each reagent channel confirm that each commanded dispense delivered the expected volume of reagent, and differential pressure transducers upstream and downstream of each valve detect anomalous pressure signatures from a clamped or blocked line. Each sensor has its own secure element and signs its readings independently, so that a man-in-the-middle between sensor and controller is detectable. Before synthesis begins, the guarantee processor commits to the full reagent-channel-time schedule derived from the approved sequence. Sensor readings are checked against this commitment during execution. If an anomaly is detected, the guarantee processor halts the controller and optionally destroys the partial product.

**Protect the controller.** The device controller runs on an FPGA with an encrypted bitstream and secure boot, with protection against voltage glitching and firmware rollback. The FPGA logic is obfuscated so that an attacker who extracts the bitstream cannot easily understand or modify it. Communication between the controller, the guarantee processor, and the sensors is encrypted. The configurable parameters available to the operator should be minimized: the controller should have no operator-accessible settings beyond the sequence itself, so that there are no tuning knobs that could be used to divert synthesis. The controller logic should also be audited for flaws that could let the device produce DNA different from what was requested.

**Place the enforcement hardware where it is hard to remove.** The guarantee processor, the secure elements in the cartridges, and the FPGA controller need to be integrated into the device in a way that makes bypass by replacement difficult. Integration points that carry this property include the controller FPGA (bitstream will not run without a valid authorization), the reagent-dispensing actuation path (pumps and valves will not fire without a valid authorization), and the cartridge authentication interface (cartridges will not authenticate without a live guarantee processor). Requiring all of these to be satisfied means that a replacement-based bypass would have to rebuild a substantial fraction of the device.

None of these mitigations is individually sufficient. Each raises the cost of a successful attack, and each attack has to defeat several of them at once to succeed without detection.

## 3 Limitations and open problems

The design presented here has several limitations worth stating explicitly.

**Sophisticated physical attackers.** The attack surface of a benchtop synthesizer is large, and a well-resourced attacker with extended physical access can likely defeat any specific sensor or seal. The goal of the process-integrity layer is to raise the cost of a successful attack to a point where mid-resource attackers are excluded, not to provide an absolute guarantee. The device also cannot defend against an attacker who builds their own synthesizer from scratch; this design is only relevant if regulated devices become the easiest available option.

**Air-gapped operation.** Many laboratories operate benchtop devices without continuous network connectivity. In this case, the device can store a small number of pre-signed authorizations, each for a specific sequence hash, with strict expiry times. Synthesis fails closed if no valid authorization is available. The length of the expiry window is a policy question that trades off operational convenience against the time a stolen or diverted device remains usable.

**Compromised regulator keys.** If the regulator's signing key is leaked or misused, an attacker can authorize arbitrary syntheses. This risk can be mitigated by requiring signatures from multiple regulators on each authorization, for example a fraction of an approved set. The same approach was proposed for Offline Licensing of AI chips (Petrie 2024).

## 4 Conclusion

Benchtop DNA synthesizers will soon be capable of producing significant portions of viral genomes, and regulatory momentum in the US and EU is opening a window for mandatory screening. For this screening to be enforceable on devices that sit in a user's lab, the hardware has to do most of the work. The design sketched here uses three mechanisms: sequence screening, a regulator-signed approval that the device will not run without, and process monitoring during synthesis.

The first two mechanisms are close adaptations of recent work on governing AI chips (Brass and Aarne 2024; Petrie 2024; Petrie et al. 2025). The hardware infrastructure for benchtop biosecurity is closer to being ready than it might appear. The third mechanism, preventing an attacker with physical access from tampering with the synthesis process, is where the design has to do new work. The physical attack surface is large and no individual mitigation is sufficient, but defense in depth across cartridge authentication, in-line sensors, tamper-evident enclosures, and a protected controller can make successful attacks infeasible for mid-resource adversaries. Even a partial solution to the process-integrity problem would make misuse of benchtop synthesizers significantly harder than it is today.

## References

ACA (Nov. 2025). *Regulatory Gaps in Benchtop Nucleic Acid Synthesis Create Biosecurity Vulnerabilities*. Arms Control Association blog. URL: https://www.armscontrol.org/blog/2025-11-24/regulatory-gaps-benchtop-nucleic-acid-synthesis-create-biosecurity-vulnerabilities.

Baker, D. and G. Church (2024). "Protein design meets biosecurity". In: *Science* 383.6681, p. 349. DOI: 10.1126/science.ado1671.

Baum, C. et al. (2024). *A system capable of verifiably and privately screening global DNA synthesis*. arXiv: 2403.14023 [cs.CR]. URL: https://arxiv.org/abs/2403.14023.

Brass, A. and O. Aarne (2024). *Location Verification for AI Chips*. Tech. rep. Institute for AI Policy and Strategy (IAPS). URL: https://www.iaps.ai/research/location-verification-for-ai-chips.

Edison, R. G., S. Toner, and K. M. Esvelt (2026). "Assembling unregulated DNA segments bypasses synthesis screening: regulate fragments as select agents". In: *Nature Communications*. DOI: 10.1038/s41467-025-67955-3. URL: https://www.nature.com/articles/s41467-025-67955-3.

European Commission (Dec. 2025). *Proposal for a Regulation Establishing a Framework of Measures for Strengthening the Union's Biotechnology and Biomanufacturing Sectors (Biotech Act)*. COM(2025) 1022. URL: https://health.ec.europa.eu/document/download/ec1475b7-e3f9-409e-b927-fc7e69306a8c_en?filename=biotech_reg-com2025-1022_act_en.pdf.

IBBIS (2024). *The Common Mechanism for DNA Synthesis Screening*. International Biosecurity and Biosafety Initiative for Science. URL: https://ibbis.bio/our-work/common-mechanism/.

Kilobaser (2024). *Our Technology: Single-Use Microfluidic Chip*. Kilobaser GmbH, kilobaser.com. URL: https://kilobaser.com/technology.

Langenkamp, M. (2024). *Securing Benchtop DNA Synthesizers*. Institute for Progress. URL: https://ifp.org/securing-benchtop-dna-synthesizers/.

NTI (May 2023). *Benchtop DNA Synthesis Devices: Capabilities, Biosecurity Implications, and Governance*. Nuclear Threat Initiative, Bio Report. URL: https://www.nti.org/wp-content/uploads/2023/05/NTIBIO_Benchtop-DNA-Report_FINAL.pdf.

O'Gara, A. et al. (2025). *Hardware-Enabled Mechanisms for Verifying Responsible AI Development*. arXiv: 2505.03742 [cs.CR]. URL: https://arxiv.org/abs/2505.03742.

Petrie, J. (2024). *Near-Term Enforcement of AI Chip Export Controls Using a Firmware-Based Design for Offline Licensing*. arXiv: 2404.18308 [cs.CY]. URL: https://arxiv.org/abs/2404.18308.

Petrie, J. et al. (2025). *Flexible Hardware-Enabled Guarantees for AI Compute*. arXiv: 2506.15093 [cs.CY]. URL: https://arxiv.org/abs/2506.15093.

Sandbrink, J. B. (2021). "Increased cyber-biosecurity for DNA synthesis". In: *Nature Biotechnology* 39. Published online December 2020, pp. 22–24. DOI: 10.1038/s41587-020-00761-y. URL: https://www.nature.com/articles/s41587-020-00761-y.

SecureDNA (2024). *Exemption Certification System: Safe Access to Sequences of Concern*. SecureDNA documentation. URL: https://securedna.org/exemption_certification_system/.

U.S. Congress (Jan. 2026). *S.3741 – Biosecurity Modernization and Innovation Act of 2026*. 119th Congress, 2nd Session. URL: https://www.congress.gov/bill/119th-congress/senate-bill/3741/text.

Wittmann, B. J. et al. (2025). "Strengthening nucleic acid biosecurity screening against generative protein design tools". In: *Science* 390, pp. 82–87. DOI: 10.1126/science.adu8578.