James Diggans, PhD
Twist Bioscience Corporation

Subcommittee on Oversight and Investigations
of the
Committee on Energy and Commerce
United States House of Representatives
Hearing on
Examining Biosecurity at the Intersection of AI and Biology

Response to Questions from Subcommittee Chair John Joyce

1. What is the current capability of AI systems to design, optimize, or stabilize viral or bacterial constructs with enhanced fitness?

Recent AI systems have demonstrated meaningful and concrete capabilities relevant to pathogen engineering, particularly in the areas of protein structure prediction[1] and generative sequence design[2]. Such tools have substantially reduced the expertise required to reason about how sequence changes in a pathogen's surface proteins might affect their function — including properties like receptor binding affinity and resistance to antibody neutralization. Separately, generative models trained on protein sequence data can now propose novel sequences likely to adopt a target structure or exhibit a specified function, and published work has demonstrated AI-assisted optimization of protein stability and binding properties[3]. These are real capabilities that have advanced considerably over the past several years.

The more important framing for risk assessment, however, may be less about what any single AI system can accomplish independently, and more about the cumulative effect of these tools on the overall technical burden of pathogen engineering. No current AI system can design an enhanced pathogen from end to end. The combination of structure prediction, sequence design, fitness modeling, and AI-assisted literature synthesis, however, means that each individual step in that process now requires much less specialized expertise than it would have five years ago. The barrier being lowered is not primarily the ability to *conceive* of a dangerous modification, but the practical difficulty of working through the many technical decision points required to design and execute one.

Here, significant limitations remain. The most important is the absence of large, well-labeled datasets linking sequence variation to in vivo fitness outcomes in humans including metrics like transmissibility and clinical severity. We have viral and bacterial sequence data at scale, and we increasingly have structural data, but the critical connection between sequence and real-world pathogen behavior in a human host is still poorly characterized, particularly for heavily modified

[1] https://www.nature.com/articles/s41586-024-07487-w
[2] https://www.biorxiv.org/content/10.1101/2025.09.18.676967v1
[3] https://www.science.org/doi/10.1126/science.abd7331

or novel constructs. This data gap meaningfully constrains what AI systems can reliably predict. Federal funding of national labs to generate such data to explore the relationship between known threat sequences and sequence modifications would provide very high long term value. The national labs are well-positioned, given their on-going national security role and compute resources, to use such data to produce defensive AIPD models that, given a novel sequence, could estimate the risk that sequence could carry out specific harmful functions of concern.

A further constraint, independent of AI capability, is the wet-lab execution requirement: AI tools can suggest sequences, but those sequences must still be synthesized, recovered as functional agents, and validated experimentally. These represent significant technical and resource barriers even as the computational predictions improve.

It is also worth noting that these considerations apply to bacterial constructs as well as viral ones. Bacterial genetics are in some respects better characterized than viral genetics, and there are published examples of AI-assisted optimization of properties relevant to pathogen fitness, including antimicrobial resistance profiles[4] and toxin expression. The bacterial case may in fact be more tractable for AI-assisted design today, given the depth and quality of the available training data.

2. How often do your customers request biological agents and toxins that are subject to Select Agent regulations?

As you know, the Federal Select Agent Program (FSAP) governs access to full, functional forms of controlled toxins and to viable stocks of controlled pathogens. We elect not to fulfill these orders and therefore do not require an FSAP license.

A subset of our customers order synthetic DNA encoding portions of these organisms or toxins. It is common for our customers studying such pathogens to order these individual components - we receive orders for such sequences multiple times a day, representing anywhere from 1-5% of our order stream. Synthetic DNA for these Select Agent components allows these customers to study the functions and capabilities of controlled pathogens and toxins safely and without risk to laboratory personnel. These efforts represent an incredibly valuable contribution to our growing understanding of pathogenicity and support the development of therapeutics and other methods for reducing the risk these agents pose. While these sequences are not controlled under FSAP, Twist Bioscience screens all double-stranded DNA sequences ordered to determine whether they are a portion of an organism deemed to be controlled for possession either domestically or internationally. If a controlled sequence or a portion of a controlled sequence is detected during our screening process, we stop the order process immediately, following up with the customer to verify their intended use and past publication record, as well as any licenses required to receive the sequence in compliance with the U.S. Department of Commerce Export Administration Regulations, the 2023 HHS Screening Framework Guidance, and the OSTP Framework for Nucleic Acid Synthesis Screening.

[4] https://www.sciencedirect.com/science/article/abs/pii/S0092867425008554

3. How can DNA synthesis providers balance customer privacy concerns with effective screening practices?

DNA synthesis providers have demonstrated that effective biosecurity screening and customer privacy protection are complementary and not competing objectives. Leading companies in the industry treat customer data with the same rigor applied to any sensitive proprietary information. Customer order data is encrypted both in transit and at rest, access is governed by strong authentication and audit controls, and cybersecurity programs are maintained to industry-recognized standards. Twist Bioscience, for example, holds ISO 27001 certification, the internationally recognized standard for information security management, providing independent validation that our data handling practices meet the rigorous requirements to maintain the trust our customers place in us.

3a. Can AI help reduce the number of false positives or negatives in sequencing screening programs?

Yes - one of AI's most significant advantages in biosecurity screening is its ability to detect patterns across large volumes of data that no human reviewer or static lookup table could assess in real time. Traditional sequence screening compares an ordered sequence against a database of known sequences of concern. This is effective but inherently backward-looking: it flags what has already been characterized as dangerous. AI can go further by analyzing contextual signals across thousands of orders over time, identifying anomalous ordering patterns, unusual combinations of sequences, or customer behavior that deviates from established norms. This closely parallels the maturation of the cybersecurity field, which now makes heavy use of AI to detect anomalous use patterns. This kind of data-scale pattern recognition can significantly reduce false positives by providing reviewers with richer context, allowing legitimate research orders from known institutions to be distinguished more reliably from orders that warrant closer scrutiny.

AI also offers a more fundamental improvement in how sequence similarity itself is assessed. Conventional alignment-based screening compares sequences letter by letter and can miss engineered variants that have been modified precisely to evade those comparisons while retaining dangerous biological function. Protein language models — AI systems trained on vast libraries of biological sequences — learn to represent sequences in terms of their predicted structural and functional properties, not just their literal composition. In this "latent space" representation, two protein sequences that look quite different at the sequence level but fold into similar three-dimensional shapes and bind the same biological targets will appear close to one another, making them detectable as related regardless of surface-level differences. This structural similarity approach is particularly valuable for screening because it is harder to engineer around: modifying a sequence to evade an alignment check is straightforward, but preserving dangerous function while simultaneously escaping a structure-aware AI model is substantially more difficult. Together, these capabilities represent the opportunity for a meaningful and practical advance in the accuracy of biosecurity screening programs.

4. Are current DNA printing capabilities a limiting factor to the threat risk?

We interpret this question as referring principally to benchtop DNA synthesis devices. These are tabletop instruments designed to produce synthetic DNA sequences within a laboratory setting. Current benchtop devices do impose meaningful technical constraints that limit their utility for the most serious misuse scenarios: they are generally restricted in the length and complexity of sequences they can produce, and synthesizing a dangerous pathogen genome can require tens of thousands of base pairs assembled with high fidelity. This remains well beyond what most commercially available benchtop instruments can accomplish today. In that narrow sense, current technological limitations do represent a near-term constraint on certain categories of risk. However, synthesis technology is advancing, and it would be a mistake to treat today's technical barriers as a durable policy safeguard.

The more reliable and sustainable hedge against misuse risk is the model that commercial synthesis providers operating biosecurity screening programs represent. Unlike benchtop devices, which currently operate largely outside any systematic screening framework, orders placed with commercial providers are subject to comprehensive review. This means that the centralized provider model provides oversight that is both universal and scalable: it does not depend on any particular technology remaining limited, and it does not create gaps for well-resourced actors who might work around individual technical constraints. As benchtop capabilities inevitably improve, policymakers will need to determine how to extend the screening norms that commercial providers like Twist already practice to the full landscape of synthesis access to ensure the US federal government is not relying on the temporary limitations of any particular instrument category to mitigate biosecurity threats.

4a. How likely is it that benchtop synthesis devices will soon be able to create longer DNA sequences?

Near-term advances in benchtop synthesis of long DNA sequences face significant and unresolved technical barriers. The most commercially advanced benchtop platform currently available is designed primarily for short oligonucleotide synthesis up to 120 base pairs and does not currently support the gene-length sequences that would be required for the most serious biosecurity threat scenarios. A small number of other companies are developing next-generation enzymatic synthesis platforms with longer-sequence ambitions but none has yet brought to market a device capable of synthesizing the multi-kilobase sequences that would represent a meaningful shift in the threat landscape. The core technical challenges are that errors accumulate with length in enzymatic synthesis and that assembling short fragments into longer constructs requires additional steps and expertise. These challenges are not close to being solved at the benchtop scale. For the near term, the more relevant policy concern is not that benchtop devices will soon match centralized synthesis capabilities, but that the benchtop devices already in distribution operate largely outside the biosecurity screening frameworks that commercial providers like Twist adhere to for every order.

5. Is the suitability screening of personnel to be in the Federal Select Agent Program sufficient to protect against an insider threat at a government laboratory?

Twist Bioscience is not licensed under the Federal Select Agent Program (FSAP) and so is not well-positioned to evaluate the specific adequacy of FSAP personnel suitability screening.

We can, however, speak to the broader challenge that this question reflects: namely, that verifying the legitimacy of individuals who wish to access dual-use biological materials or capabilities is genuinely difficult. U.S. government guidance specifically asks commercial DNA synthesis providers to screen not just sequences but the customers and institutional contexts behind each order: verifying that stated affiliations are plausible, that the combination of sequences ordered is consistent with a legitimate research purpose, and that orders are not inconsistent with the customer's known scientific profile. Like personnel vetting, it can be imperfect: a sufficiently determined and credentialed actor can misrepresent purpose in ways that are difficult to detect. Layered alongside sequence-level biosecurity screening and the pattern-detection capabilities discussed elsewhere in these responses, however, legitimacy review represents a meaningful contribution to the overall biosecurity architecture as it relates to nucleic acid synthesis. It is also imperative that governments and interested stakeholders, including but not limited to nucleic acid synthesis providers, contribute to additional layers of biosecurity protection.

6. What responsibilities do biotechnology and AI companies bear when developing tools that could be misused for biological harm?

Biotechnology and AI companies that develop tools with potential dual-use applications in the life sciences should bear a meaningful responsibility to invest in layered defenses against misuse, an approach borrowed from cybersecurity practice that recognizes no single control is sufficient. The foundational layer of any responsible development program is pre-release evaluation: before a tool (or technology) reaches users, developers should conduct rigorous assessments specifically designed to surface dual-use capabilities. This is not merely a reputational consideration; a company that releases a tool without understanding its biosecurity risk profile has made an architectural decision with potentially serious consequences. The results of those evaluations should directly shape what controls are applied and at what level of access the tool is made available, making pre-release evaluation the prerequisite on which all other layers depend.

With that foundation in place, companies should invest in a set of operational safeguards commensurate with the risk profile their evaluations reveal. For tools with significant dual-use potential, managed access (including requiring users to establish institutional affiliation, intended use, and accountability before gaining access to sensitive capabilities) provides a meaningful first filter. For AI models intended for open-weight release, responsible developers should give careful consideration to training data curation, recognizing that embedding detailed biosecurity-relevant technical knowledge in widely-distributed model weights creates a different and more persistent risk profile than providing equivalent capabilities through a managed API.

Finally, companies should implement alerting protocols capable of identifying patterns of use consistent with misuse attempts and routing those signals to both internal review and, where appropriate, relevant authorities. This ensures that the tool's operation generates actionable biosecurity intelligence rather than silent logs. Taken together, these layers represent a proportionate, technically grounded response to the dual-use challenge that does not require companies to choose between innovation and responsibility.

7. Should companies independently red team test high-risk AI biological models on a regular basis?

Yes, companies developing high-risk AI models with biological applications should conduct regular, independent red team evaluations, and the case for independence is as important as the case for regularity. In-house evaluation teams, however capable, often share the assumptions and blind spots of the developers who built the model. They may also face structural incentives that can subtly influence what risks they prioritize and how seriously they weigh what they find. Independent evaluators bring a genuinely adversarial perspective and approach the model as a potential bad actor would. Regular re-evaluation is equally important: models are frequently updated, fine-tuned, and deployed in contexts that differ from their original testing environment, and a biosecurity assessment that was accurate at launch may not reflect the model's capabilities six months later.

The design of these evaluations must also keep pace with the actual frontier of model capability and this is where the field currently faces a significant challenge. Biosecurity red-teaming frameworks are often anchored to known threats: they assess whether a model can reproduce documented pathogen characteristics, describe established synthesis routes, or explain recognized mechanisms of toxicity. Passing these benchmarks provides meaningful but incomplete assurance. A more consequential risk posed by increasingly capable AI systems may not be their ability to reproduce today's understanding of dangerous biology, but their potential to assist in designing novel threats. These may include engineered pathogens, novel delivery mechanisms, or attacks on agricultural systems, ecosystems, and other biological systems beyond human health for which no existing benchmark was designed and no current defense is calibrated. Red teaming programs that treat biosecurity evaluations as a compliance checkpoint rather than a living, capability-tracking practice will find themselves systematically behind the risk they are meant to assess. The standard should be not whether a model can articulate what is already known, but whether it provides meaningful assistance toward causing biological harms that have not yet been conceived.