Accelerate AI for drug discovery

Build, customize, and deploy biomolecular models at scale with NVIDIA BioNeMo

Lenovo | NVIDIA

Challenges within the current drug discovery landscape

Improving the speed and quality of early preclinical drug discovery pipelines is directly related to unlocking new therapies that can improve patient outcomes and save lives. Traditional drug discovery is time-consuming and expensive, taking 10 years and billions of invested dollars to bring a new drug to market on average, with only about 10 percent of drug development programs successfully making it to market. As the biopharmaceutical industry adopts computer-aided drug discovery methods to address molecular modeling, high-throughput screening, molecular docking, and virtual screening, AI models have been a key tool to accelerate these processes.

To unlock the full potential of AI for drug discovery, the industry needs:

- A domain-specific platform for developing, customizing, and deploying biomolecular AI models at scale
- Scalable, distributed GPU computing for training generative AI without the overhead of setting up and maintaining the necessary infrastructure
- Deployment of generative AI into production, where optimized runtimes and autoscaling can generate, predict, and screen the near-infinite number of potential drug-target combinations

NVIDIA BioNeMo™

NVIDIA BioNeMo — included with **NVIDIA AI Enterprise** — is a platform that accelerates and simplifies the building, customization, and deployment of AI models for drug discovery. This ensures scalable, easy model development and deployment for drug discovery applications. BioNeMo provides researchers and developers a fast and easy way to build and integrate state-of-the-art AI applications across the entire drug discovery pipeline — from target identification to lead optimization — with AI workflows for 3D protein structure prediction, small molecule de novo design, virtual screening, docking, and property prediction. It also equips researchers and developers with scaling recipes, an optimized framework, and pretrained models that accelerate large-scale model training.

Whether running on **NVIDIA DGX™ Cloud, NVIDIA DGX** systems, or any cloud service provider, customers get an end-to-end platform for drug discovery that can scale effortlessly. **NVIDIA NIM™**, which provides optimized inference microservices for deploying AI **models** at scale, helps application developers accelerate their journey to AI-powered drug discovery with quick and easy access to models for protein structure prediction, generative chemistry, and molecular docking prediction.

Key challenges for AI drug discovery

- Scaling: Advancing from proof of concept to enterprise deployment requires effective scaling through efficient use of resources to ensure manageability, availability, and infrastructure cost.
- Performance: High performance with multi-node, multi-GPU support is critical for AI.
- Application Development and Deployment: Pulling together disparate products into an end-to-end AI solution for drug discovery is difficult.

Benefits for AI drug discovery researchers and developers

- Accelerated AI training, customization, and deployment.
- Streamlined path for developing and deploying AI-powered drug discovery applications.
- Pretrained state-of-the-art biomolecular models.
- Access to NVIDIA AI domain experts for hands-on support.

Lenovo

NVIDIA

Smarter technology for all

Lenovo

Benefits

Instantly run and deploy AI:
Experience and use biomolecular AI

Innovate faster with customizable, state-of-the-art AI:
Use your proprietary datasets to train and fine-tune AI models for drug discovery.

Build with scalable and seamless AI microservices:
Achieve the ultimate flexibility in experimenting with and building enterprise-grade AI workflows on scalable, managed infrastructure.

Streamline model training with validation-in-the-loop:
Evaluate the quality of a model's embeddings and stay in control of what your models learn.

NVIDIA NIM to deploy AI for drug discovery applications

NVIDIA NIM is a set of optimized cloud-native microservices designed to shorten time-to-market and accelerate the deployment of AI models anywhere — across cloud, data center, and GPU-accelerated workstations. The NVIDIA API catalog includes easy access to NIM microservices for drug discovery, including MolMIM for generative chemistry, ESMFold for protein structure prediction, and DiffDock to help researchers understand how drug molecules interact with targets. These models can be tested using NVIDIA-managed APIs in the API catalog or downloaded as self-hosted models that can run on premises or any major cloud provider — reducing development time, complexity, and the total cost of the workflow. With NVIDIA NIM, application developers can easily use composable microservices to build flexible drug discovery workflows for their own applications.

Workstations are an excellent resource for increasing local computing capacity and for training and validating AI models for industry-specific workflows. Lenovo offers a wide range of workstations designed to include a user-defined number of advanced NVIDIA RTX GPUs so performance can be customized and scaled for each use case.

Getting started with NVIDIA BioNeMo NIM microservices for drug discovery

01 Discover models on **ai.nvidia.com**.

02 Try and prototype with APIs on **build.nvidia.com**.

03 Apply to self-host NIM on the model page at **build.nvidia.com/explore/healthcare**.

❝ The ease of multi-node training with BioNeMo and the ability to use larger batch sizes within DGX Cloud enabled us to achieve our three-month objectives in just four weeks."

Chris James Langmead, Director of Digital Biologics Discovery, Amgen
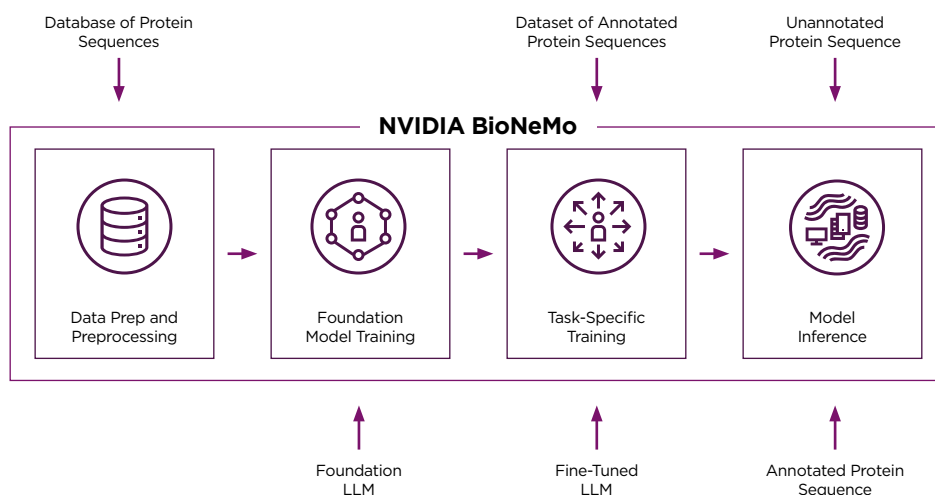
Smarter technology for all

Building, customizing, and fine-tuning biomolecular AI models with BioNeMo

BioNeMo Framework can be downloaded as a container and run on products from any cloud provider to seamlessly scale with dataset sizes, model sizes, and training tasks on multi-GPU and multi-node infrastructure. This significantly reduces time and cost for model development. Optimized and accelerated throughout the entire stack, it offers the quickest path to scale, speed, and experimentation in drug discovery, letting researchers train and customize models using their proprietary data without the need to write elaborate code.

AI biomolecular model training

Train a foundation model and fine-tune it to perform task-specific predictions.

Effortlessly train large models and scale operations, whether using NVIDIA DGX Cloud, an AI cloud platform for enterprise developers, or on-premise DGX systems. The portability of the NVIDIA BioNeMo Framework container caters to diverse drug discovery and research needs, so users can deploy it anywhere. The end-to-end nature of the framework also enables the creation of custom inference pipelines using included scripts for on-premises deployment.

Database of Protein Sequences → NVIDIA BioNeMo

Dataset of Annotated Protein Sequences

Unannotated Protein Sequence

NVIDIA BioNeMo

Data Prep and Preprocessing → Foundation Model Training → Task-Specific Training → Model Inference

Foundation LLM

Fine-Tuned LLM

Annotated Protein Sequence

Train with NVIDIA BioNeMo framework
Get going in three easy steps

Step 01
Go to NGC and download the BioNeMo Framework container.

Step 02
Deploy the container on your cloud service provider of choice.

Step 03
Begin building or training models with your proprietary or public data of choice.

Lenovo | NVIDIA

Lenovo

Smarter technology for all

Train larger models in days instead of weeks and scale near linearly

ESM2: time to train 1T tokens (1.19B sequences)
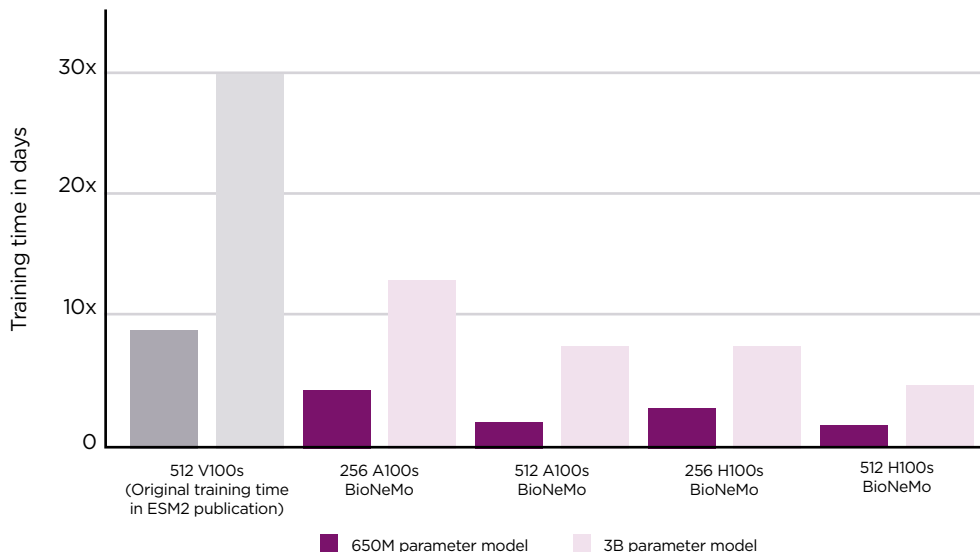
Example training times (in days) of ESM2 for different GPUs with BioNeMo Framework. Original published model training time shown for reference in grey. Models trained with BioNeMo were trained on 1 trillion tokens, equivalent to 1.19 billion protein sequences.

Getting started with NVIDIA BioNeMo NIM microservices for drug discovery

Best-of-breed platform for AI development:

The BioNeMo platform integrates AI software with cloud infrastructure, providing a comprehensive solution tailored for the accelerating drug discovery processes.

Scalable computing at predictable cost:

Regardless of your deployment choice, the BioNeMo platform offers a straightforward, predictable cost model for AI development infrastructure.

Your own AI expert:

Get access to experts who know AI development and have worked on your problems before.

Smarter technology for all

Building an AI architecture begins with a technology partner like Lenovo, experienced in bridging best practices for cross-industry compliance, performance, security, and data management. This approach is especially important to meet the unique needs for professionals working in healthcare and life sciences.

In addition to powerful workstations, Lenovo offers flexible AI infrastructure, including servers, high-performance computing (HPC), edge devices, and storage — each tailored to your needs today and scalable as your AI implementation evolved.

AI technology requires more system power than ever, putting an increased strain on energy and cooling demands. Lenovo solutions are engineered to manage these demands. For instance, Lenovo's Neptune® Liquid Cooling enables performance and reduces energy consumption by up to 40 percent while providing maximum performance and reliability for critical AI systems. Together, these features ensure Lenovo AI systems can operate optimally without overheating or unanticipated downtime.

The Lenovo ThinkStation P series portfolio offers a range of workstations.
Choose the performance level best suited for your organization's needs.

P3: Intel CPU, 1 NVIDIA RTX GPU*

P5: Intel CPU, 2 NVIDIA RTX GPUs*

P7: Intel CPU, 3 NVIDIA RTX GPUs*

P8: AMD CPU, 3 NVIDIA RTX GPUs*

PX: Intel CPU, 4 NVIDIA RTX GPUs*

Performance →

*NVIDIA RTX GPU models can vary depending on processor and/or customer choice.

To learn more about Lenovo workstations and AI infrastructure, visit:
www.lenovo.com/Pharma

To learn more about NVIDIA BioNeMo, visit:
www.nvidia.com/bionemo

To learn more about NVIDIA DGX platform, visit:
www.nvidia.com/dgx

Connect with an expert at:
www.nvidia.com/contacthealthcare/

Smarter technology for all

Lenovo | NVIDIA

Lenovo