



Accelerate AI for Drug Discovery

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

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 \$2 billion 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.

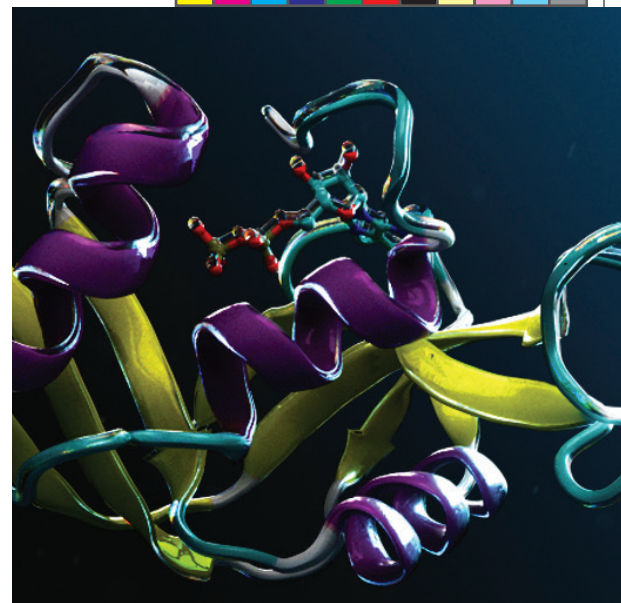
To unlock the full potential of AI for drug discovery, the industry needs:

- > 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.

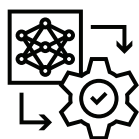
Benefits for AI Drug Discovery Researchers and Developers

- > 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.

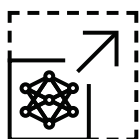
Benefits



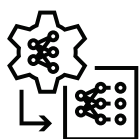
Instantly run and deploy AI: Experience and use biomolecular AI models with microservices optimized and accelerated by NVIDIA, then deploy anywhere with NVIDIA NIM.



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.

“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.”

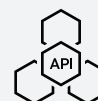
Chris James Langmead, Director of Digital Biologics Discovery, **Amgen**

Getting started with NVIDIA BioNeMo NIMs for drug discovery



Step 01

Discover models on ai.nvidia.com.



Step 02

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



Step 03

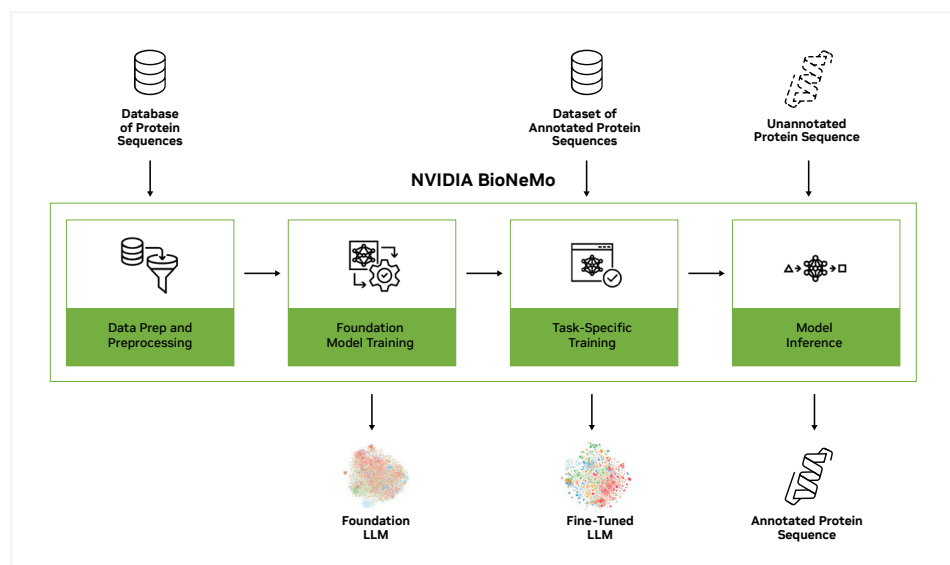
Apply to self-host NIM on the model page at build.nvidia.com/explore/healthcare.

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.

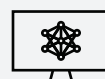
Train with NVIDIA BioNeMo's Framework

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.

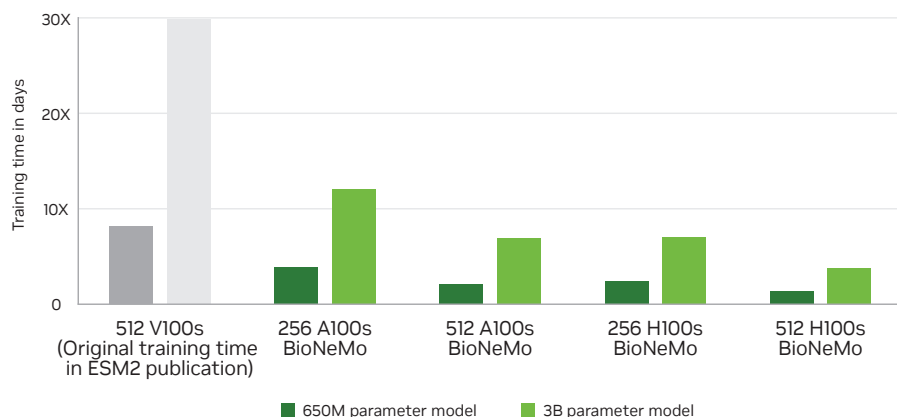


Step 03

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

Train Larger Models in Days Instead of Weeks and Scale Near Linearly

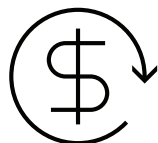
ESM2: time to train 1T tokens (1.19B sequences)



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.



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.

Ready to Get Started?

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/contact-healthcare/

© 2024 NVIDIA Corporation and affiliates. All rights reserved. NVIDIA, the NVIDIA logo, BioNeMo, DGX, and NVIDIA Triton are trademarks and/or registered trademarks of NVIDIA Corporation and affiliates in the U.S. and other countries. Other company and product names may be trademarks of the respective owners with which they are associated. 3352261. SEP24





Accelerate AI for Drug Discovery

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



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 \$2 billion 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.

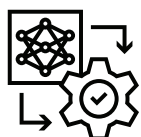
Benefits for AI Drug Discovery Researchers and Developers

- > 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.

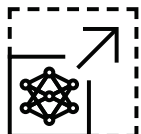
Benefits



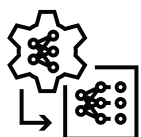
Instantly run and deploy AI: Experience and use biomolecular AI models with microservices optimized and accelerated by NVIDIA, then deploy anywhere with NVIDIA NIM.



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.

“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**

NVIDIA NIM to Deploy AI for Drug Discovery Applications

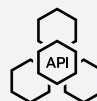
NVIDIA NIM is a set of optimized cloud-native microservices designed to shorten time-to-market and accelerate the deployment of AI models anywhere—across cloud, data center, and GPU-accelerated workstations. The NVIDIA API catalog includes easy access to NIM microservices for drug discovery, including 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.

Getting started with NVIDIA BioNeMo NIMs for drug discovery



Step 01

Discover models on ai.nvidia.com.



Step 02

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



Step 03

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

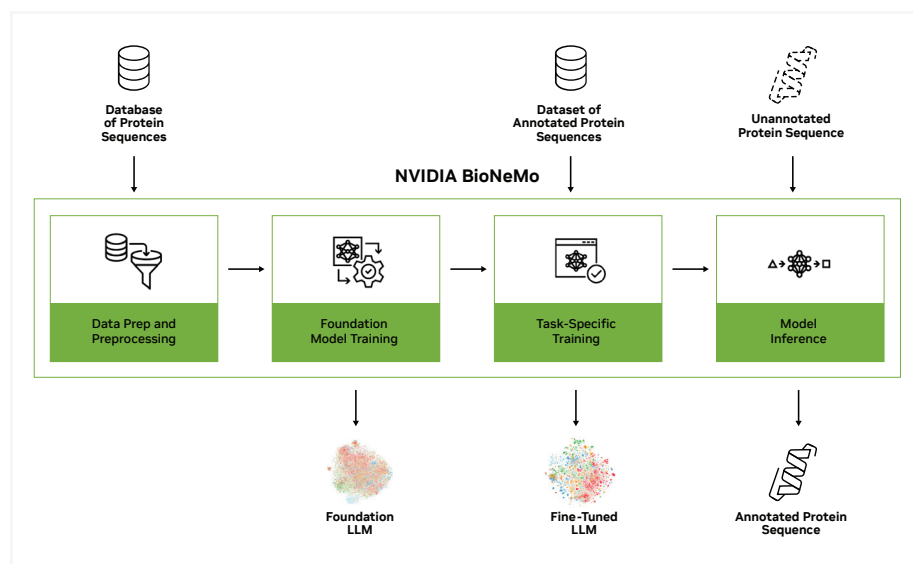


Building, Customizing, and Fine-Tuning Biomolecular AI Models With BioNeMo

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.

Train with NVIDIA BioNeMo's 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

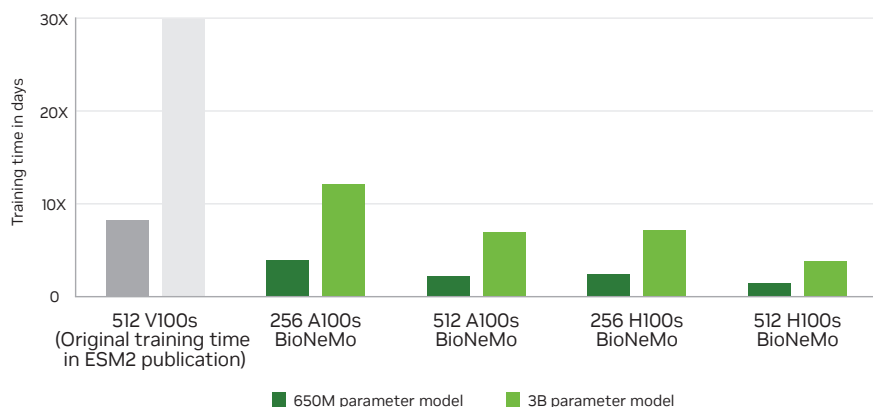
Begin building or training models with your proprietary or public data of choice.





Train Larger Models in Days Instead of Weeks and Scale Near Linearly

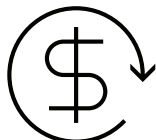
ESM2: time to train 1T tokens (1.19B sequences)



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.



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.

Ready to Get Started?

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/contact-healthcare/

© 2024 NVIDIA Corporation and affiliates. All rights reserved. NVIDIA, the NVIDIA logo, BioNeMo, DGX, and NVIDIA Triton are trademarks and/or registered trademarks of NVIDIA Corporation and affiliates in the U.S. and other countries. Other company and product names may be trademarks of the respective owners with which they are associated. 3352261. SEP24

