

# Michelle Lynn Gill, PhD

Applied Research Manager, NVIDIA Virtual Cell Team

GitHub | LinkedIn | Personal Website

## Overview

I am an Applied Research Manager on NVIDIA's Virtual Cell team. Our team uses data, deep learning, and HPC to develop models that simulate cellular states. My focus is on the benchmarks and development of supporting tools that help evaluate and improve these models.

Previously I was a Senior Machine Learning Engineer and Data Scientist at BenevolentAI and a member of Arthur Palmer's research group where I developed and applied nuclear magnetic resonance (NMR) spin relaxation experiments to understand how enzyme dynamics are critical to biological function.

## Education

**Ph.D., Molecular Biophysics & Biochemistry** 2003–2006  
Yale University, New Haven, CT

Thesis: Development of 2D NMR methods for the direct study of monovalent metal ions and ligands in nucleic acids

Advisors: J. Patrick Loria and Scott Strobel

**M.Phil., Molecular Biophysics & Biochemistry** 2001–2003  
Yale University, New Haven, CT

**B.S., Biochemistry** 1997–2001  
University of Kansas, Lawrence, KS  
Highest Distinction and Honors (Summa Cum Laude)

## Experience

**Applied Research Manager, NVIDIA, Virtual Cell Team** 2022–Present  
Lead benchmarking efforts for virtual cell models, including internal and external collaborations

Developed benchmarking framework and tooling for model assessment and reproducibility

Collaboration with Chan Zuckerberg Initiative on cz-benchmarks for evaluating virtual cell models

**Scientific Lead, BioNeMo, NVIDIA** 2019–2023  
R&D Manager and Scientific Lead for Clara Discovery, NVIDIA's platform for accelerating drug discovery

Developed BioNeMo for pre-training and fine tuning of large language models for cheminformatics and proteomics

**Senior AI and Deep Learning Scientist, NVIDIA** 2019–2023

Focused on proteomics including deep learning models to predict peptide spectral matches with >95% F1

Led team using GCNNs to predict molecular properties

**Senior Data Scientist and Machine Learning Engineer,**

2018-2019

BenevolentAI

Matrix factorization and GCNNs for drug mechanism importance in knowledge graphs

3D CNNs for ligand pose and affinity prediction

**Senior Deep Learning Consultant, NVIDIA**

2017-2018

Assisted clients in pharmaceutical and materials science in utilizing deep learning

**Scientist, National Cancer Institute, NIH**

2014-2016

Developed parallelized, compressed sensing methods for NMR data reconstruction

**Postdoctoral Research Fellow, Columbia University, Department of**

2008-2014

Biochemistry and Molecular Biophysics

Demonstrated conformational selection in DNA methyltransferase AlkB

Developed multiple quantum NMR spin relaxation experiments

Advisor: Professor Arthur G. Palmer, III

**Publications**

S. Dicks, L. Heumos, S. Jimenez, P. Angerer, I. Gold, I. Virshup, F.

Fischer, L. May, C. J. Nolet, M. Gill, M. Boerries, and F. Theis

**"Accelerating single-cell analysis with GPU-enabled RAPIDS-singlecell"** In preparation (2025).

E. Sevgen, J. Moller, A. Lange, J. Parker, S. Quigley, J. Mayer, P. Srivastava, S. Gayatri, D. Hosfield, M. Korshunova, M. Livne, M. Gill, R. Ranganathan, A.B. Costa, and A.L. Ferguson **"ProT-VAE: Protein transformer variational autoencoder for functional protein design"** *Proceedings of the National Academy of Sciences* (2025).

P. St. John, D. Lin, P. Binder, M. Greaves, V. Shah, J. St. John, A. Lange, P. Hsu, R. Illango, A. Ramanathan, A. Anandkumar, D.H. Brookes, A. Busia, A. Mahajan, S. Malina, N. Prasad, S. Sinai, L. Edwards, T. Gaudet, C. Regep, M. Steinegger, B. Rost, A. Brace, K. Hippe, L. Naef, K. Kamata, G. Armstrong, K. Boyd, Z. Cao, H.Y. Chou, S. Chu, A. Santos Costa, S. Darabi, E. Dawson, K. Didi, C. Fu, M. Geiger, M. Gill, D.J. Hsu, G. Kaushik, M. Korshunova, S. Kothén-Hill, Y. Lee, M. Liu, M. Livne, Z. McClure, J. Mitchell, A. Moradzadeh, O. Mosafi, Y. Nashed, S. Paliwal, Y. Peng, S. Rabhi, F. Ramezanghorbani, D. Reidenbach, C. Ricketts, B.C. Roland, K. Shah, T. Shimko, H. Sirelkhatim, S. Srinivasan, A.C. Stern, D. Toczydlowska, S.P. Veccham, N.A.E. Venanzi, A. Vorontsov, J. Wilber, I. Wilkinson, W.J. Wong, E. Xue, C. Ye, X. Yu, Y. Zhang, G. Zhou, B. Zandstein, A. Chacon, P. Sohani, M. Stadler, C. Hundt, F. Zhu, C. Dallago, B. Trentini, E. Kucukbenli, T. Rvachov, E. Calleja, J. Israeli, H. Clifford, R. Haukioja, N. Haemel, K. Tretina, N. Tadimetri, and A.B. Costa

**"BioNeMo Framework: a modular, high-performance library for AI model development in drug discovery"** [arXiv](#) (2024).

M.T. Strauss, I. Bludau, W.F. Zeng, E. Voytik, C. Ammar, J. Schessner, R. Illango, M.L. Gill, F. Meier, S. Willems, and M. Mann **"AlphaPept, a modern and open framework for MS-based proteomics"** [Nature Communications](#) (2024).

D. Reidenbach, M. Livne, R.K. Illango, M.L. Gill, and J.I. Israeli **"Improving small molecule generation using mutual information machine"** [arXiv](#) (2022).

M.L. Gill, A. Hsu, and A.G. Palmer **"Detection of chemical exchange in methyl groups of macromolecules"** [Journal of Biomolecular NMR](#) (2019).

M.L. Gill, R.A. Byrd, and A.G. Palmer **"Dynamics of GCN4 facilitate DNA interaction: a model-free analysis of an intrinsically disordered region"** [Physical Chemistry and Chemical Physics](#) (2016).

S. Sun, M.L. Gill, Y. Li, M. Huang, and R.A. Byrd **"Efficient and generalized processing of multidimensional NUS NMR data: the NESTA algorithm and comparison of regularization terms"** [Journal of Biomolecular NMR](#) (2015).

B. Ergel, M.L. Gill, L. Brown, B. Yu, A.G. Palmer, and J.F. Hunt **"Protein dynamics control the progression and efficiency of the catalytic reaction cycle of AlkB"** [Journal of Biological Chemistry](#) (2014).

M.L. Gill, and A.G. Palmer **"Local isotropic diffusion approximation for coupled internal and overall molecular motions in NMR spin relaxation"** [Journal of Physical Chemistry B](#) (2014).

M.L. Gill, and A.G. Palmer **"Multiplet-filtered and gradient-selected zero-quantum TROSY experiments for  $^{13}\text{C}^{1}\text{H}_3$  methyl groups in proteins"** [Journal of Biomolecular NMR](#) (2011).

|         |   |      |
|---------|---|------|
| Patents | <b>Automated FEP Path Generation and Optimization by Molecule Foundation Models and Generative AI</b>   | 2025 |
|         | Peng, Y., Kucukbenli, E., Zhou, G., <a href="#">Gill, M.L.</a> , Livne, M., Korshunova, M., Rvachov, T., Israeli, Y.<br>US Patent Application No. 19/208,119; Filing Date: 2025/05/14 |      |
|         | <b>Extraction of Informative Embeddings from Encoder-Decoder Models</b>   | 2024 |
|         | Livne, M., <a href="#">Gill, M.L.</a><br>US Patent Application No. 18/957,301; Filing Date: 2024/11/22  |      |
|         | <b>Contrastive Framework for Unified Generative and Discriminative Representation Learning</b>  | 2024 |
|         | Livne, M., <a href="#">Gill, M.L.</a><br>US Patent Application No. 18/957,294; Filing Date: 2024/11/22  |      |
|         | <b>Efficient Data Loading for Deep Learning Workloads</b>   | 2024 |

Darabi, S., Korshunova, M., Grewal, J., [Gill, M.L.](#), Morkisz, P.  
US Patent Application No. 18/885,428; Filing Date: 2024/09/13

**Guardrails for Molecular Generation** 2024

Korshunova, M., Peng, Y., Zhou, G., Rvachov, T., [Gill, M.L.](#), Kucukbenli, E., Israeli, Y.  
US Patent Application No. 18/807,805; Filing Date: 2024/08/16

**End to End Deep Learning Workflow for In Silico Molecule Design** 2024

Stern, A., [Gill, M.L.](#), Stepniewska-Dziubinska, M., Grzegorzec, T., Nowaczynski, A.,  
Toczydlowska, D., Israeli, Y., Ribalta Lorenzo, P.  
US Patent Application No. 18/412,168; Filing Date: 2024/01/12

**Small Molecule Generation Using Machine Learning Models** 2023

Livne, M., Reidenbach, D., [Gill, M.L.](#), Ilango, R., Israeli, Y.  
US Patent Application No. 18/450,745; Filing Date: 2023/08/16

## Presentations

**Scientific Discovery: From the Lab Bench to the GPU**, Andy Byrd 2024  
Retirement Symposium

Invited Keynote, April 19, 2024, Institute for Bioscience & Biotechnology Research,  
University of Maryland, NIST, Bethesda, MD  
[Slides](#) · [Program](#)

**Scientific Discovery: From the Lab Bench to the GPU**, PyData NYC 2023  
Invited Keynote, November 2, 2023, New York, NY  
[Slides](#) · [Video](#) · [Abstract](#)

**NVIDIA BioNeMo: A framework and service for development and use of generative AI in drug discovery**, 6th RSC-BMCS / RSC-CICAG Artificial Intelligence in Chemistry Conference 2023  
Invited Keynote, September 5, 2023, Cambridge, UK  
[Slides](#) · [Program](#)

**Exploring Molecular Space and Accelerating Drug Discovery on the GPU with Clara Discovery**, Gates Foundation Grand Challenges: Applications of Artificial Intelligence in Machine Learning 2021  
Invited Talk, November 10, 2021, Virtual

**Accelerating Drug Discovery with Clara Discovery and MegaMolBART**, 4th RSC-BMCS / RSC-CICAG Artificial Intelligence in Chemistry 2021  
Invited Talk, September 27, 2021, Virtual  
[Slides](#)

**Real Time, GPU-Accelerated Analysis and Visualization in the Life Sciences**, Ken Kennedy Institute Data Science Conference 2020  
[Michelle Gill](#) and Avantika Lal  
Invited Keynote, October 26-27, 2020, Virtual  
[Slides](#) · [Abstract](#) · [Program](#)

|  |      |
|--|------|
| <b>Artificial intelligence driven drug discovery</b> , NYC R Conference  | 2019 |
| Invited Presentation, May 10, 2019, New York, NY   |      |
| <a href="#">Slides</a>   |      |
| <b>Machine learning for target identification and lead optimization in drug discovery</b> , New York Area Group for Informatics and Modeling               | 2019 |
| Alix Lacoste and <a href="#">Michelle Gill</a>   |      |
| Invited Presentation, February 26, 2019, New York, NY  |      |
| <a href="#">Abstract</a>   |      |
| <b>Accelerating the journey from data to medicine</b> , NeurIPS  | 2018 |
| Amir Saffari, Dan Neil, Alix Lacoste, and <a href="#">Michelle Gill</a>  |      |
| Expo Talk, 2018, Montreal, Canada  |      |
| <a href="#">Abstract</a>   |      |
| <b>Artificial intelligence as a catalyst for scientific discovery</b> , JupyterCon   | 2018 |
| Invited Keynote, 2018, New York, NY  |      |
| <a href="#">Slides</a> · <a href="#">Video</a> · <a href="#">Abstract</a>  |      |
| <b>From structural biology to AI: a holistic approach to studying molecular machines</b> , Brookhaven National Laboratory                                  | 2018 |
| Invited Presentation, 2018, Upton, NY  |      |
| <a href="#">Slides</a>   |      |
| <b>Efficient image search and identification: the making of Wine-O.AI</b> , SciPy Conference   | 2017 |
| Selected Presentation, 2017, Austin, TX  |      |
| <a href="#">Slides</a> · <a href="#">Video</a> · <a href="#">Code</a>  |      |
| <b>Development of 2D/3D NMR methods for the direct study of monovalent metal ions and ligands in nucleic acids</b> , Ph.D. Thesis Defense, Yale University | 2006 |
| Thesis Defense, 2006, New Haven, CT  |      |
| <a href="#">Slides</a> · <a href="#">Thesis</a>  |      |

|               |  |           |
|---------------|--|-----------|
| <b>Awards</b> | NIH Postdoctoral Research Fellowship (F32 GM089047)    | 2009-2012 |
|               | Global BCG Strategy Olympics, Winning Team             | 2007      |
|               | NSF Graduate Research Fellowship                       | 2002-2006 |
|               | Barry M. Goldwater Scholar                             | 2000-2001 |
|               | Outstanding Undergraduate Honors Research Thesis       | 2001      |
|               | Kansas Board of Regents Full Tuition Merit Scholarship | 1997-2001 |

## Service

|   |      |
|---|------|
| <b>Judge, Preliminary and Final Rounds</b> , NYC STEM Fair, New York, NY<br>Evaluated submissions in the biochemistry track               | 2022 |
| <b>Program Chair</b> , PyData NYC, New York, NY<br>Responsible for conference content, proposal review, speaker selection, and scheduling | 2018 |
| <b>Machine Learning Symposium Co-Chair</b> , SciPy Conference, Austin, TX<br>Co-chair of machine learning / deep learning symposium       | 2018 |