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 205Tl 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, BenevolentAI   2018–2019

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 Biochemistry and Molecular Biophysics   2008–2014

Demonstrated conformational selection in DNA methyltransferase AlkB

Developed multiple quantum NMR spin relaxation experiments

Advisor: Professor Arthur G. Palmer, III

Publications

Dicks, S., Heumos, L., Jimenez, S., Angerer, P., Gold, I., Virshup, I., Fischer, F., May, L., Nolet, C. J., Gill, M., Boerries, M., Theis, F.. "Accelerating single-cell analysis with GPU-enabled RAPIDS-singlecell." **In preparation** (2025).

Sevgen, E., Moller, J., Lange, A., Parker, J., Quigley, S., Mayer, J., Srivastava, P., Gayatri, S., Hosfield, D., Korshunova, M., Livne, M., Gill, M., Ranganathan, R., Costa, A.B., Ferguson, A.L.. "ProT-VAE: Protein transformer variational autoencoder for functional protein design." **Proceedings of the National Academy of Sciences** 122 (2025), doi: 10.1073/pnas.2408737122.

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

Haukioja, R., Haemel, N., Tretina, K., Tadimeti, N., Costa, A.B.. "BioNeMo Framework: a modular, high-performance library for AI model development in drug discovery." **arXiv** (2024), doi: 10.48550/arXiv.2411.10548.

Strauss, M.T., Bludau, I., Zeng, W.F., Voytik, E., Ammar, C., Schessner, J., Illango, R., Gill, M.L., Meier, F., Willems, S., Mann, M.. "AlphaPept, a modern and open framework for MS-based proteomics." **Nature Communications** 15 (2024), doi: 10.1038/s41467-024-46485-4.

Reidenbach, D., Livne, M., Illango, R.K., Gill, M.L., Israeli, J.I.. "Improving small molecule generation using mutual information machine." **arXiv** (2022), doi: 10.48550/arXiv.2208.09016.

Gill, M.L.. "The rise of the machines in chemistry." **Magnetic Resonance in Chemistry** 60, 1044–1051 (2022), doi: 10.1002/mrc.5304.

Gill, M.L., Hsu, A., Palmer, A.G.. "Detection of chemical exchange in methyl groups of macromolecules." **Journal of Biomolecular NMR** 73, 443–450 (2019), doi: 10.1007/s10858-019-00240-w.

Tong, M., Pelton, J., Gill, M.L., Zhang, W., Picart, F., Seeliger, M.. "Survey of solution dynamics in Src kinase reveals cross talk between the ligand binding and regulatory sites." **Nature Communications** 8, 2160 (2017), doi: 10.1038/s41467-017-02240-6.

Gill, M.L., Byrd, R.A., Palmer, A.G.. "Dynamics of GCN4 facilitate DNA interaction: a model-free analysis of an intrinsically disordered region." **Physical Chemistry and Chemical Physics** 18, 5839–5849 (2016), doi: 10.1039/c5cp06197k.

Sun, S., Gill, M.L., Li, Y., Huang, M., Byrd, R.A.. "Efficient and generalized processing of multidimensional NUS NMR data: the NESTA algorithm and comparison of regularization terms." **Journal of Biomolecular NMR** 62, 105–117 (2015), doi: 10.1007/s10858-015-9923-x.

Gill, M.L., Byrd, R.A.. "Dynamic activation of apoptosis: conformational ensembles of cIAP1 are linked to a spring-loaded mechanism." **Nature Structural Molecular Biology** 21, 1022–1023 (2014), doi: 10.1038/nsmb.2925.

Ergel, B., Gill, M.L., Brown, L., Yu, B., Palmer, A.G., Hunt, J.F.. "Protein dynamics control the progression and efficiency of the catalytic reaction cycle of AlkB." **Journal of Biological Chemistry** 289, 29584–29601 (2014), doi: 10.1074/jbc.M114.575647.

Gill, M.L., Palmer, A.G.. "Local isotropic diffusion approximation for coupled internal and overall molecular motions in NMR spin relaxation." **Journal of Physical Chemistry B** 118, 11120–11128 (2014), doi: 10.1021/jp506580c.

Gill, M.L., Palmer, A.G.. "Multiplet-filtered and gradient-selected zero-quantum TROSY experiments for $^{13}C^{1}H_3$ methyl groups in proteins." **Journal of Biomolecular NMR** 51, 245–251 (2011), doi: 10.1007/s10858-011-9533-1.

Ramsey, J.D., Gill, M.L., Kamerzell, T.J., Price, E.S., Joshi, S.B., Bishop, S.M., Oliver, C.N., Middaugh, C.R.. "Using empirical phase diagrams to understand the role of intramolecular dynamics in immunoglobulin G stability." **Journal of Pharmaceutical Sciences** 98, 2432–2447 (2009), doi: 10.1002/jps.21619.

Gill, M.L., Strobel, S.A., Loria, J.P.. "Crystallization and characterization of the thallium form of the Oxytricha nova G-quadruplex." **Nucleic Acids Research** 34, 4506–4514 (2006), doi: 10.1093/nar/gkl616.

Gill, M.L., Strobel, S.A., Loria, J.P.. "^{205}Tl NMR methods for the study of monovalent metal binding sites in nucleic acids." **Journal of the American Chemical Society** 127, 16723–16732 (2005), doi: 10.1021/ja055358f.

Beach, H., Cole, R., Gill, M.L., Loria, J.P.. "Conservation of μs-ms enzyme motions in the apo- and substrate-mimicked state." **Journal of the American Chemical Society** 127, 9167–9176 (2005), doi: 10.1021/ja0514949.

Adams, P.L., Stahley, M.R., Gill, M.L., Kosek, A.B., Wang, J., Strobel, S.A.. "Crystal structure of a group I intron splicing intermediate." **RNA** 12, 1867–1887 (2004), doi: 10.1038/nature02642.

Wiethoff, C.M., Gill, M.L., Koe, G.S., Koe, J.G., Middaugh, C.R.. "A fluorescence study of the structure and accessibility of plasmid DNA condensed with cationic gene delivery vehicles." **Journal of Pharmaceutical Sciences** 92, 1272–1285 (2003), doi: 10.1002/jps.10391.

Wiethoff, C.M., Gill, M.L., Koe, G.S., Koe, J.G., Middaugh, C.R.. "The structural organization of cationic lipid-DNA complexes." **Journal of Biological Chemistry** 277, 44980–44987 (2002), doi: 10.1074/jbc.M207758200.

Silchenko, S., Sippel, M.L., Kuchment, O., Benson, D.R., Mauk, A.G., Altuve, A., Rivera, M.. "Hemin is kinetically trapped in cytochrome b5 from rat outer mitochondrial membrane." **Biochemical and Biophysical Research Communications** 273, 467–472 (2000), doi: 10.1006/bbrc.2000.2968.

Patents	Automated FEP Path Generation and Optimization by Molecule Foundation Models and Generative AI	2025

Peng, Y., Kucukbenli, E., Zhou, G., Gill, M.L., Livne, M., Korshunova, M., Rvachov, T., Israeli, Y.
US Patent Application No. 19/208,119; Filing Date: 2025/05/14

	Extraction of Informative Embeddings from Encoder-Decoder Models	2024

Livne, M., Gill, M.L.
US Patent Application No. 18/957,301; Filing Date: 2024/11/22

	Contrastive Framework for Unified Generative and Discriminative Representation Learning	2024

Livne, M., <u>Gill, M.L.</u>
US Patent Application No. 18/957,294; Filing Date: 2024/11/22

Efficient Data Loading for Deep Learning Workloads   2024
Darabi, S., Korshunova, M., Grewal, J., <u>Gill, M.L.</u>, 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., <u>Gill, M.L.</u>, Kucukbenli, E., Israeli, Y.
US Patent Application No. 18/807,805; Filing Date: 2024/08/16

Guardrails for Conditional Molecular Generation   2024
Korshunova, M., Peng, Y., Zhou, G., Rvachov, T., <u>Gill, M.L.</u>, Kucukbenli, E., Israeli, Y.
US Patent Application No. 18/807,808; Filing Date: 2024/08/16

Training-Time Guardrails for Molecular Generation   2024
Korshunova, M., Peng, Y., Zhou, G., Rvachov, T., <u>Gill, M.L.</u>, Kucukbenli, E., Israeli, Y.
US Patent Application No. 18/807,811; Filing Date: 2024/08/16

Guardrails for Instruction-Tuned Molecular Generation   2024
Korshunova, M., Peng, Y., Zhou, G., Rvachov, T., <u>Gill, M.L.</u>, Kucukbenli, E., Israeli, Y.
US Patent Application No. 18/807,814; Filing Date: 2024/08/16

End to End Deep Learning Workflow for In Silico Molecule Design   2024
Stern, A., <u>Gill, M.L.</u>, Stepniewska-Dziubinska, M., Grzegorzek, 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., <u>Gill, M.L.</u>, 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   2023
use of generative AI in drug discovery**, 6th RSC-BMCS / RSC-
CICAG Artificial Intelligence in Chemistry Conference
Invited Keynote, September 5, 2023, Cambridge, UK
Slides · Program

**Exploring Molecular Space and Accelerating Drug Discovery on   2021
the GPU with Clara Discovery**, Gates Foundation Grand

Challenges: Applications of Artificial Intelligence in Machine Learning
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

Artificial intelligence driven drug discovery, NYC R Conference   2019
Invited Presentation, May 10, 2019, New York, NY
Slides

Panel: Careers in data science, Tri-Institutional Career Symposium   2019
Panel, April 9, 2019, Memorial Sloan Kettering Cancer Center, The Rockefeller University and Weill Cornell Medicine, New York, NY
Program

Machine learning for target identification and lead optimization in drug discovery, New York Area Group for Informatics and Modeling   2019
Alix Lacoste and Michelle Gill
Invited Presentation, February 26, 2019, New York, NY
Abstract

Accelerating the journey from data to medicine, NeurIPS   2018
Amir Saffari, Dan Neil, Alix Lacoste, and Michelle Gill
Expo Talk, 2018, Montreal, Canada
Abstract

Artificial intelligence as a catalyst for scientific discovery, JupyterCon   2018
Invited Keynote, 2018, New York, NY
Slides · Video · Abstract

From structural biology to AI: a holistic approach to studying molecular machines, Brookhaven National Laboratory   2018
Invited Presentation, 2018, Upton, NY
Slides

Efficient image search and identification: the making of Wine-O.AI, SciPy Conference   2017
Selected Presentation, 2017, Austin, TX
Slides · Video · Code

Development of 205Tl NMR methods for the direct study of monovalent metal ions and ligands in nucleic acids, Ph.D. Thesis Defense, Yale University

2006

Thesis Defense, 2006, New Haven, CT

Slides · Thesis

Awards

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

Judge, Preliminary and Final Rounds, NYC STEM Fair, New York, NY

2022

Evaluated submissions in the biochemistry track

Program Chair, PyData NYC, New York, NY

2018

Responsible for conference content, proposal review, speaker selection, and scheduling

Machine Learning Symposium Co-Chair, SciPy Conference, Austin, TX

2018

Co-chair of machine learning / deep learning symposium

Proposal Reviewer, JupyterCon

2018

Reviewer, Journal of Open Source Software (JOSS)

2018–2019