# Tanmoy Sanyal

*Scientist, Protein Design & Structural Biology*

San Francisco Bay Area, CA

tanmoy.7989@gmail.com • (805) 637-0375 • tanmoy7989.github.io • LinkedIn • GitHub

Computational scientist with over a decade of experience developing physics-based and machine learning methods for protein design and structural biology, with a track record spanning academic research, biotech, and large pharma. Specialized in MLOps, multispecific molecule design, and integrative structural modeling. Demonstrated ability to translate rigorous scientific methods into scalable, cloud-native production systems.

## Education

**University of California, Santa Barbara**
*Ph.D., Chemical Engineering* — Graduate Emphasis in Computational Science & Engineering (HPC) *2013–2018*

**Indian Institute of Technology, Kharagpur**
*Integrated M.Tech & B.Tech (Hons.), Chemical Engineering* *2008–2013*

## Experience

**Amgen Inc.** — South San Francisco, CA

*Principal Protein Design Scientist* *Apr 2024–Present*
*Senior Protein Design Scientist* *Oct 2023–Mar 2024*
*Protein Design Scientist* *Jan 2022–Oct 2023*

- **Multispecifics design platform:** Developed cloud-native, highly parallelized ODE solvers for pharmacokinetic modeling of bispecific molecules; deployed via ML regression models achieving a **50× speedup**.
- **Statistical mechanics of multispecifics:** Translated statistical mechanical models of superselectivity (originally developed for multivalent ligands) to multispecific molecules; currently undergoing experimental validation and benchmarking on bispecific candidates.
- **MLOps architecture:** Built Amgen's first end-to-end, cloud-native MLOps pipeline for supervised biophysical property prediction from protein sequences, now used across **8+ antibody design projects**. Mentored junior data scientists in thermostability prediction workflows.
- **In-silico display library design:** Developed linear-programming algorithms to optimize degenerate codon libraries using protein language models. Set an internal record for fastest request-to-delivery (∼3 days) for a 1M+ diverse antibody library; applied to engineer cross-reactivity across 2 therapeutic areas.
- **Biologics design:** Designed mini-binders for a rare disease target using Rosetta-ddG and the AlphaFold, RFDiffusion, and ProteinMPNN workflow. Developed an active-learning pipeline (Gaussian process regression) yielding a **20**× affinity improvement over 3 design-and-screen iterations.
- **AlphaFold cloud deployment:** Collaborated with AWS to re-tool open-source AlphaFold-2 into a scalable cloud-native version supporting ∼100k sequences; now in regular use for variant design.

**Novo Nordisk Research Center Seattle** — Seattle, WA

*Protein Design Scientist* *2022–2023*

- **Peptide modifications for half-life extension:** Developed a computational method for systematically scanning PEGylation sites on peptides for half-life protraction. Published and open-sourced the method with a benchmark on GLP-1 analogs relevant to anti-obesity therapeutics.

**University of California, San Francisco** — Sali Lab

*Postdoctoral Scholar* *Jan 2019–Jan 2022*

- **Nanobody biophysics & vaccine design:** Developed a protein–protein docking score for nanobody interactions with the SARS-CoV-2 Spike protein using chemical crosslinking and escape mutation data; identified vulnerable and resilient epitopes across variants of concern. Also performed MD simulation studies elucidating the effects of framework mutations on distal CDR3 residues.

- **Protein structure modeling from crosslinks:** Developed graph-sampling algorithms to extract domain boundaries from chemical crosslinking data; published the first coarse-grained structure of the full SMC5/6 protein complex.
- **Whole-cell digital twin:** Developed Bayesian networks combining pharmacokinetic, Brownian dynamics, and enzyme-pathway models of the pancreatic beta-cell into a proof-of-concept multimodal digital twin. Served as software lead in the Pancreatic Beta-Cell Consortium (UCSF/USC).

**University of California, Santa Barbara** — Shell Lab

*Ph.D. Researcher* *2013–2018*

- **Coarse-graining for phase separation:** Introduced the local density potential for structurally and thermodynamically accurate coarse-grained MD simulations of liquid–liquid phase separation.
- **Protein folding via variational inference:** Developed backbone models for template-free folding of 200+ residue protein domains using MD simulations, and for studying self-assembly in amyloidogenic peptides.
- **HPC cluster administration:** Set up and administered a 144-core Linux (Rocks 6.2) cluster; wrote utilities for automated resource sharing among lab members.

## Computational Skills

| | |
|---|---|
| **Protein Design** | Rosetta, RFDiffusion, ProteinMPNN, AlphaFold, ESM2, Amplify (protein language models) |
| **Molecular Modeling** | OpenMM, LAMMPS, GROMACS, IMP, Modeller, UCSF ChimeraX, PyMOL |
| **ML / MLOps** | PyTorch, PyTorch-Lightning, TensorFlow-Probability, PyMC3, MLFlow, AWS Batch |
| **Languages** | Python (+Cython, 10k+ lines), C++ (∼4k lines), Fortran-90 (∼1k lines), Bash |

## Open-Source Software

- **nnprotscan** — Computational scanning of chemical modifications in peptides.
- **nbspike** — Integrative epitope prediction for nanobody–antigen interactions.
- **IMP-raindrops** — Crosslink-guided domain discovery within the IMP framework.
- **PDB-Dev deposition tools** — Structural deposition utilities for the SMC5/6 complex.
- **LAMMPS local density potential** — Plugin implementing the local density coarse-graining potential.

## Selected Publications

- N. Thomas, **T. Sanyal**, P. Greisen Jr. & K. Diebler. Structure-Based Computational Scanning of Chemical Modification Sites in Biologics. *ACS Omega*, 9(34), 36787–36794, 2024.
- F.D. Mast, P.C. Fridy, N.E. Karen, J. Wang, E.Y. Jacobs, J.P. Olivier, **T. Sanyal**, et al. Highly synergistic combinations of nanobodies that target SARS-CoV-2 and are resistant to escape. *eLife*, 10: e73027, 2021.
- B. Raveh, L. Sun, K.L. White, **T. Sanyal**, et al. Bayesian metamodeling of complex biological systems across varying representations. *PNAS*, 118(35) e2104559118, 2021.
- Y. Yu, S. Li, Z. Ser, **T. Sanyal**, et al. Integrative analysis reveals unique structural and functional features of the Smc5/6 complex. *PNAS*, 118(19) e2026844118, 2021.
- **T. Sanyal**, J. Mittal and M. Scott Shell. A hybrid, bottom-up, structurally-accurate, Go-like coarse-grained protein model. *J. Chem. Phys.*, 151(4): 044111, 2019.
- **T. Sanyal** and M. Scott Shell. Coarse-grained models using local-density potentials optimized with the relative entropy: Application to implicit solvation. *J. Chem. Phys.*, 145, 034109, 2016.

## Service & Peer Review

*Reviewer* (2019–Present): *Structure*; *Journal of Physical Chemistry*; *Proteins: Structure, Function and Bioinformatics*; *Review of Scientific Instruments*; *Rapid Reviews COVID-19*; *Life*.