Tanmoy Jana (Ph.D.)

1835 NORIEGA ST APT 312, 94122-4300 – San Francisco, CA, USA
+1 (858) 305 8033 • janatanmoy5@gmail.com
LinkedIn: tanmoy-jana-864b7a43 • GitHub: janatanmoy5
Google Scholar: 7i633QgAAAAJ • ResearchGate: Tanmoy-Jana
ORCID: 0000-0003-1321-2240

PROFESSIONAL EXPERIENCE

Dec 2024 – Present **Postdoctoral Researcher**
University of California San Francisco, San Francisco, USA
Advisor: **Rajalingam Raja, PhD**

Dec 2022 – Dec 2024 **Postdoctoral Researcher**
University of California San Francisco, San Francisco, USA
Advisors: **Prof. Laura Koth, Prof. Mark Ansel**

Sep 2022 – Dec 2022 **Visiting Mentored Researcher**
Salk Institute for Biological Studies, San Diego, USA
Advisor: **Prof. Uri Manor**

Sep 2021 – Sep 2022 **Postdoctoral Researcher**
University of California San Diego, San Diego, USA
Advisor: **Prof. Rodney Gabriel**

Jun 2020 – Aug 2021 **Postdoctoral Research Associate**
ICMR–National Institute for Research in Bacterial Infections (formerly National Institute of Cholera and Enteric Diseases), Kolkata, India
Advisor: **Dr. Santasabuj Das, Scientist G and Director**

Oct 2015 – Nov 2019 **Ph.D. Graduate Researcher**
Maulana Abul Kalam Azad University of Technology
Research conducted at Bose Institute, Kolkata, India
Advisors: **Prof. Sudipto Saha, Prof. Raja Banerjee**

Oct 2012 – Oct 2015 **Research Assistant**
Bose Institute, Kolkata, India
Advisor: **Prof. Sudipto Saha**

EDUCATION

Oct 2015 – Nov 2019 **Ph.D. (Technology) in Bioinformatics**
Maulana Abul Kalam Azad University of Technology, West Bengal, India
Thesis: *Computer-Aided Prediction of Potential Drug Candidates against Cancer using Protein–Protein Interaction Modulators*
Advisors: **Prof. Sudipto Saha, Prof. Raja Banerjee**

Aug 2010 – Jul 2012 **M.Sc. in Computer Science**
Vidyasagar University, West Bengal, India
Minor: Fuzzy Logic

May 2006 – Apr 2010 **B.Sc. in Computer Science**
Vidyasagar University, West Bengal, India
Minor: Physics and Mathematics

CERTIFICATIONS

Dec 2025 **Advanced Certifications in AI, Data Science & Research Ethics**
Massachusetts Institute of Technology (MIT), Cambridge, MA, USA
Includes: Data or Specimens Only Research; CITI Good Clinical Practice;
Essentials of Responsible AI; Big Data and Data Science Research Ethics

May 2023 **Single-Cell RNA-Seq Analysis**
Columbia University, New York, USA

Jun 2010 **J2EE (Java Enterprise Edition)**
NCVT, Kolkata, India

AWARDS

2015 – 2018 **ICMR Senior Research Scholarship Award**
2010 – 2012 **West Bengal Merit-Cum-Means Government Scholarship**

RESEARCH EXPERIENCE

Postdoctoral Researcher, University of California San Francisco, San Francisco, CA (Dec 2024 – Present)
Advisor: Rajalingam Raja, PhD
Developed a clinical platform for classification of cross-reactive epitope groups (CREGs) using HLA antibody binding patterns from patient sera. Evaluated immunogenicity of mismatched HLA epitopes to improve donor–recipient compatibility, graft survival, and reduce sensitization and rejection risk. Designed AI/ML-based tools for immunogenomics and transplantation analytics to optimize clinical decision-making and patient outcomes.

Postdoctoral Researcher, University of California San Francisco, San Francisco, CA (Dec 2022 – Dec 2024).
Advisors: Prof. Laura Koth and Prof. Mark Ansel.
Collaborators: UCSF CoLabs, UCSF–Genentech.
Conducted large-scale single-cell immunogenomic analysis of sarcoidosis using UCSF-derived clinical datasets, profiling more than 2.7 million peripheral blood mononuclear cells (PBMCs) from a cohort of 368 participants. Investigated cytokine- and chemokine-driven immune dysregulation, characterized immune cell heterogeneity ($CD4^+$/$CD8^+$ T cells, B cells, TH17, monocytes), and analyzed cell–cell interactions to identify disease-associated signatures and potential therapeutic targets.

Postdoctoral Researcher, University of California San Diego, San Diego, CA (Sep 2021 – Sep 2022).
Advisor: Prof. Rodney Gabriel.
Prediction of same-day hip replacement and laminectomy surgery using ACS National Surgical Quality Improvement Program data.

Postdoctoral Research Associate, ICMR–National Institute of Cholera and Enteric Diseases, Kolkata, India (Jun 2020 – Aug 2021).
Advisor: Dr. Santasabuj Das.
Developed predictive models to identify non-resistant antimicrobial drugs by integrating whole-genome sequencing data and comparative analysis of multi-drug resistant and sensitive strains, with a focus on virulence factor profiling.

Ph.D. Graduate Researcher, Maulana Abul Kalam Azad University of Technology, West Bengal, India (Oct 2015 – Nov 2019).
Research conducted at Bose Institute, Kolkata, India.
Advisors: Dr. Sudipto Saha and Prof. Raja Banerjee.
Discovery of oncogenic Protein–Protein Interaction (PPI) inhibitors through large-scale virtual screening and machine learning techniques. It was demonstrated that the pharmacophore and physicochemical features of small molecules being the inhibitors are critical determinants of their potential as oncogenic PPI inhibitors.

Research Assistant, Bose Institute, Kolkata, India (Oct 2012 – Oct 2015).
Advisor: Dr. Sudipto Saha.
Databases development: Ring-Hydroxylating Oxygenase compounds, allergy and asthma biomarkers, allergy and asthma biomarkers and functional sites and biochemical properties of c-Myc in normal and cancer cells.

RESEARCH INTERESTS

- **Immunogenomics and HLA Analytics:**
  Development of a clinical platform integrating AI and machine learning approaches to classify cross-reactive epitope groups (CREGs) through analysis of complex HLA antibody binding patterns from patient sera, while evaluating the immunogenic potential of mismatched HLA epitopes to improve donor–recipient compatibility, enhance graft survival, and reduce immune sensitization and rejection risk—thereby advancing transplantation outcomes and supporting high-quality, impactful patient care at UCSF.
- **AI & Machine Learning for Biomedical Data:**
  Development of predictive and prognostic models using electronic health records and multi-omics data to forecast disease outcomes; analysis of protein–protein interactions (PPIs) and signaling pathways; single-cell multi-omics approaches to characterize immune cell states and cell–cell communication in cancer and sarcoidosis.
- **Multimodal AI and Vision–Language Models:**
  Design and training of multimodal large language models integrating medical imaging (radiology) and clinical narratives (e.g., MIMIC datasets) to support clinical decision-making, risk stratification, and outcome prediction.
- **Gene Expression and Single-Cell Modeling:**
  Deep learning-based comparative analysis of disease versus healthy transcriptomes for immune and non-immune cell classification and functional state prediction.
- **AI-Driven Drug Discovery:**
  Application of deep learning for pharmacophore modeling, docking, and virtual screening of small-molecule libraries for therapeutic target identification and lead optimization.

VOLUNTEER JOURNAL REVIEWER

- **P**LOS One
- **J**ournal of Optimization Theory and Applications, Springer
- **J**ournal of the Indian Society of Remote Sensing, Springer
- **F**rontiers in Tropical Diseases
- **F**rontiers in Pharmacology

TECHNICAL SKILLS

- **Computational Biology & Bioinformatics:**
  Expertise in single-cell RNA-seq analysis and immune profiling using Seurat, Scanpy, Muon, and CellTypist; sequence and motif analysis using BLAST and Clustal Omega; experience with structural and assay datasets including PDB, LABScreen Single Antigen, and HistoTrac.
- **Machine Learning for Biomedical Data:**
  Development of predictive models for biological and clinical datasets using supervised and unsupervised learning, deep learning, and neural networks; experience with Vision–Language Models (VLMs), large language models (LLMs), and natural language processing (NLP) for biomedical data integration.
- **Genomics & Molecular Modeling:**
  Analysis of high-throughput genomics data including 10x Genomics and Cell Ranger; molecular docking and structural analysis using AutoDock Vina, Chimera, and PyMOL.

- **Data Science & Statistical Analysis:**
  Data preprocessing, feature engineering, and statistical modeling for large-scale biological datasets; experience with high-performance computing (HPC) for scalable analyses.
- **Programming & Workflow Development:**
  Development of reproducible bioinformatics pipelines using Python, R, C/C++, Java, and Unix shell scripting.
- **Web & Database Development:**
  Development of bioinformatics web tools and databases using PHP, JSP, JavaScript, HTML/XML, and MySQL.
- **Research & Reproducibility Tools:**
  Version control using Git; computational notebooks using Jupyter and RStudio; scientific reporting using LaTeX; figure preparation and data visualization.
- **Operating Systems:**
  Linux, macOS, and Windows.

CONFERENCES & PRESENTATIONS

2023 — **University of California San Francisco Pulmonary Retreat**
University of California, San Francisco, CA, USA
Presented a poster on **immune cell network inference in sarcoidosis** using single-cell analysis, emphasizing AI/ML-driven insights into disease mechanisms.

2018 — **International Conference on Bioinformatics (InCoB)**
Jawaharlal Nehru University, New Delhi, India
Presented a poster: **PPIMdb**, a curated database of small molecules modulating protein–protein interactions, highlighting data integration and computational screening approaches.

2013 — **Kolkata International Conference on Systems Biology (KOLSYSBIO)**
Saha Institute of Nuclear Physics, Bidhannagar, Kolkata, India
Presented a poster on **peptide-mediated protein–protein interactions**, focusing on computational strategies for systematic cataloguing and interaction prediction.

PUBLICATIONS

Peer-Reviewed Articles

- **Jana, T.**, Sarkar, D., Ganguli, D., Mukherjee, S. K., Mandal, R. S. and Das, S. ABDpred: Prediction of active antimicrobial compounds using supervised machine learning techniques. *Indian J. Med. Res.* **159**, 78–90 (2024).
- Majumdar, S., Bhattacharjee, S., **Jana, T.** and Saha, S. DAAB-V2: Updated database of allergy and asthma biomarkers. *Allergy* **45**, 1 (2021).
- Banerjee, K., **Jana, T.**, Ghosh, Z. and Saha, S. PSCRIdb: A database of regulatory interactions and networks of pluripotent stem cell lines. *J. Biosci.* **45**, 1 (2020).
- Sarkar, D., **Jana, T.** and Saha, S. LMDIPred: A web server for predicting linear peptide

sequences that bind to SH3, WW, and PDZ domains. *PLoS ONE* **13**, e0200430 (2018).

- Chakravorty, D., **Jana, T.**, Mandal, S. D., Seth, A., Bhattacharya, A. and Saha, S. MYCbase: A database of functional sites and biochemical properties of Myc in both normal and cancer cells. *BMC Bioinformatics* **18**, 1 (2017).
- **Jana, T.**, Ghosh, A., Mandal, S. D., Banerjee, R. and Saha, S. PPIMpred: A web server for high-throughput screening of small molecules targeting protein–protein interactions. *R. Soc. Open Sci.* **4**, 160501 (2017).
- Sircar, G., Saha, B., **Jana, T.**, Dasgupta, A., Bhattacharya, S. G. and Saha, S. DAAB: A manually curated database of allergy and asthma biomarkers. *Clin. Exp. Allergy* **45**, 1259–1261 (2015).
- Barman, R. K., **Jana, T.**, Das, S. and Saha, S. Prediction of intra-species protein–protein interactions in enteropathogens facilitating systems biology study. *PLoS ONE* **10**, e0143774 (2015).
- Sarkar, D., **Jana, T.** and Saha, S. LMPID: A manually curated database of linear motifs mediating protein–protein interactions. *Database* **2015**, bav101 (2015).
- Chakraborty, J., **Jana, T.**, Saha, S. and Dutta, T. K. Ring-hydroxylating oxygenase database: A database of bacterial aromatic ring-hydroxylating oxygenases in bioremediation and biocatalysis. *Environ. Microbiol. Rep.* **6**, 519–523 (2014).

Manuscripts Under Preparation

- **Jana, T.** and Gabriel, R. A. Machine learning approaches to predicting same-day discharge after total hip arthroplasty using ACS NSQIP data. *In preparation*.
- **Jana, T.**, Koth, L. and Ansel, M. Systematic computational mapping of immune cells from bronchoalveolar lavage and peripheral blood in sarcoidosis at single-cell resolution. *In preparation*.
- **Jana, T.**, Raghav, P. K., Gente, G. and Rajalingam, R. Predicting cross-reactive epitope groups in HLA class I and II using machine learning. *In preparation*.
- **Jana, T.**, Karmakar, J., Banerjee, R. and Saha, S. Supervised learning-driven prediction of small-molecule modulator activity against protein–protein interactions. *In preparation*.

LANGUAGES

English (Fluent); Bengali (Fluent); Hindi (Fluent)

REFERENCES

Rajalingam Raja, PhD
Professor
Department of Surgery
University of California, San Francisco
1100 Van Ness Ave, Suite 800
San Francisco, CA 94109
Email: rajalingam.raja@ucsf.edu

Prof. Mark Ansel
Professor
Department of Microbiology and Immunology
University of California, San Francisco
San Francisco, CA 94143
Email: mark.ansel@ucsf.edu

Prof. Laura Koth
Professor
School of Medicine
University of California, San Francisco
San Francisco, CA 94143
Email: laura.koth@ucsf.edu

Dr. Sudipto Saha
Associate Professor
Division of Bioinformatics
Bose Institute
Kolkata 700054, India
Email: ssaha4@gmail.com

Dr. Santasabuj Das
Scientist G
Clinical Medicine
ICMR–National Institute of Occupational Health
Ahmedabad, India
Email: santasabujdas@yahoo.com; santasabujdas@gmail.com