Rafael Enrique Velásquez Rojas

Bogotá, Colombia
r.velasquez@uniandes.edu.co
Tel: +57 311 327 0462

Website · LinkedIn · GitHub
Google Scholar · ORCID
Nationality: Colombian

EDUCATION

Jan 2025 – Expected Apr. 2027 — **M.Sc. in Biomedical Engineering**
Universidad de los Andes, Bogotá, Colombia
M.Sc. thesis in progress – advisor Prof. Pablo Arbeláez.

Jan 2020 – Apr. 2024 — **B.Sc. in Physics**
Universidad de los Andes, Bogotá, Colombia
Thesis on *information transmission in genetic networks* – advisor Prof. Juan Manuel Pedraza.

Aug 2020 – Oct. 2026 — **B.Sc. in Biomedical Engineering**
Universidad de los Andes, Bogotá, Colombia
Design project: *BiliScan* – non-invasive, skin-tone-independent qualitative detection of neonatal jaundice using digital image processing and deep learning.

RESEARCH EXPERIENCE

2024 – Present — **Drug Discovery as Next Geometric Token Prediction**
Prof. Pablo Arbeláez, Universidad de los Andes
- Designed a geometric tokenization of 3D protein–ligand structures and trained an autoregressive model to predict binding affinity by next-token prediction, transferring sequence-model machinery to 3D molecular data and improving MSE by about 20% on BindingDB.

2025 – Present — **Topology-Aware Segmentation of Tubular Structures**
Prof. Pablo Arbeláez, Universidad de los Andes; E. Puyol-Antón, King's College London
- Developed topology-aware objectives for 3D segmentation of thin tubular structures, a bipartite surface-matching metric and a surface-distance loss on centerline and signed-distance representations that penalize broken or spurious branches invisible to volumetric overlap.
- Benchmarked nnU-Net, SwinUNETR and NexToU on ImageCAS and ASOCA, and am extending the method to a multi-dataset benchmark (Medical Tubathlon) toward a generalist model for tubular structures.

2023 – 2026 — **Information in Gene-Regulatory Networks**
Prof. Juan Manuel Pedraza, Universidad de los Andes
- Built stochastic models (Gillespie and Langevin) of transcriptional regulatory motifs as noisy information channels and carried out the full theoretical and computational analysis, deriving an information-versus-cost trade-off that helps account for the observed prevalence of feed-forward loops. Submitted to Physical Review E.

2025 – 2026 — **Simulation-Based Inference of Bursty Gene Expression**
Prof. Pablo Arbeláez, Universidad de los Andes
- Showed that neural simulation-based inference of transcriptional bursting fails when trained on cheap Gaussian simulators, because the chemical Langevin approximation cannot represent discrete, zero-inflated counts, while a moment-matched negative-binomial surrogate restores calibrated posteriors.
- Introduced a Fano-factor diagnostic that predicts, before inference, where a cheap simulator can be trusted and where exact (Gillespie) simulation is needed. Accepted at the LatinX in AI Workshop, NeurIPS 2026.

2024 – 2026 — **Flow Matching for Crystal Structure Prediction**
Prof. Pablo Arbeláez, Universidad de los Andes
- Developed OPAL-CSP, a flow-matching generative model that samples a crystal's lattice and fractional atomic coordinates from its chemical composition, built on an SE(3)-equivariant graph transformer with a wrapped coordinate path that respects periodic boundary conditions.

– Co-designed OPAL-Bench, a polymorph-aware evaluation framework whose physical-plausibility metrics (formation energy, SOAP, XRD) reveal that geometric accuracy does not imply physical validity. Accepted at NeurIPS 2026 (Evaluations and Datasets Track).

2025 **Antimicrobial Peptide Classification (ESCAPE)**
Prof. Pablo Arbeláez, Universidad de los Andes
– Compiled and harmonized public peptide databases into a standardized multilabel benchmark for antimicrobial activity, and built a dual-encoder model fusing a sequence transformer and a vision transformer over 3D-derived distance maps by cross-attention. NeurIPS 2025 (Datasets and Benchmarks Track).

2023 – 2025 **Stochastic Modelling of Ageing**
Prof. A. Rutenberg, Dalhousie University (Mitacs Globalink)
– Modelled ageing as stochastic damage and repair dynamics on a network of health attributes, during a Mitacs Globalink internship.

PUBLICATIONS & PREPRINTS

[1] **Rafael Velásquez***, Juan Manuel Pedraza. *Cost-information trade-offs partially rationalize the feed-forward loop census*. Submitted to **Physical Review E.** [paper] (*corresponding author)

[2] **Rafael Velásquez**, Pablo Arbeláez. *Predicting When Cheap Simulators Break Neural Inference of Bursty Gene Expression*. **LatinX in AI (LXAI) Workshop, NeurIPS 2026.**

[3] Sergio Rincón*, Gabriel González*, Nicolás Andrade, **Rafael Velásquez**, Sebastián Ojeda, Juanita Puentes, Nicolás Aparicio, Paula Cárdenas, Pablo Arbeláez. *Quantitative Assessment of Crystal Structure Prediction*. **NeurIPS 2026, Evaluations & Datasets Track.** [paper] (*equal contribution)

[4] **Rafael Velásquez**, Esther Puyol-Antón, Pablo Arbeláez. *Beyond Volume Overlap: Surface Matching for Topology-Aware Coronary Artery Segmentation*. **STACOM Workshop, MICCAI 2026.** [paper]

[5] Sebastián Ojeda*, **Rafael Velásquez**, Nicolás Aparicio, Juanita Puentes, Paula Cárdenas, Nicolás Andrade, Gabriel González, Sergio Rincón, Carolina Muñoz-Camargo, Pablo Arbeláez. *A Standardized Benchmark for Multilabel Antimicrobial Peptide Classification*. **NeurIPS 2025, Datasets & Benchmarks Track.** [paper] (*corresponding author)

TALKS & POSTERS

Posters presented at **NeurIPS 2025** (San Diego, peptide benchmark [5]), **MICCAI 2026 / STACOM** (Strasbourg, coronary segmentation [4]), and **NeurIPS 2026** (Sydney, crystal-structure prediction [3] and bursty gene-expression inference [2]).

TECHNICAL SKILLS

Machine learning	flow matching, diffusion, transformers, equivariant and graph neural networks, multimodal and autoregressive models
Statistical modelling	statistical physics, stochastic simulation (Gillespie, Langevin), information theory, simulation-based inference
Biomedical data	3D medical image segmentation, gene-regulatory network modelling, molecular and structural data
Tools	Python, PyTorch, Linux, HPC, Git

SCHOLARSHIPS & AWARDS

2026	**RISE-MICCAI 2026 Travel Grant.** One of only 5–10 awarded; supported MICCAI 2026, Strasbourg.
2026	**LatinX in AI (LXAI) Travel Award.** To attend and present at NeurIPS 2026, Sydney.
2026	**High-Impact Mobility Fund (FMAI).** Uniandes co-financing to present high-impact papers at international conferences.
2025	**1st Place – LISA 2025 Challenge (Tasks 2a & 2b).** 1st of 94 teams, Low-field Pediatric Brain MRI Segmentation Challenge (RISE-MICCAI, MICCAI 2025); led to the RISE-MICCAI Paper-Lead Mentorship Program.
2023	**Mitacs Globalink Research Internship Award.** Highly competitive (under 5% acceptance); research internship at Dalhousie University, Canada.
	Silver Medal – iGEM Design League. Interdisciplinary systems & synthetic biology team.