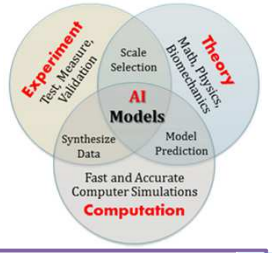
# Machine Learning in Multiscale Modeling and Validation of In Vitro Experiments of Blood Flow and Platelet Mediated Thrombosis Initiation

Peng Zhang[1], Jawaad Sheriff[1], Prachi Gupta[1], Changnian Han[1], Peineng Wang[1], Ziji Zhang[1], Marvin J. Slepian[1,2], Yuefan Deng[1], **Danny Bluestein[1]**

[1]Stony Brook University, Stony Brook, NY, [2]University of Arizona, Tucson, AZ, USA




## Project Summary

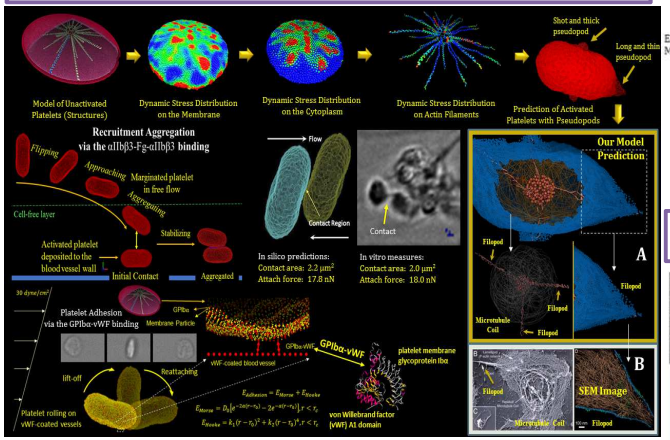
**MULTISCALE PROBLEM:** incorporated coarse-grained molecular dynamics (CGMD) and dissipative particle dynamics (DPD) to describe platelet mechanotransduction induced by blood flow in cardiovascular pathologies which may initiate thrombosis[1-6]. Because of the complex nature of this dynamic process that covers many spatio-temporal scales, such modeling may become computationally prohibitive even when employing the strongest HPC resources available. We have developed three Machine Learning (ML) approaches for optimizing the modeling efficiency and its experimental validation for significantly reducing the computational costs to achieve both: ***Reduced computational costs without loss of accuracy*.**

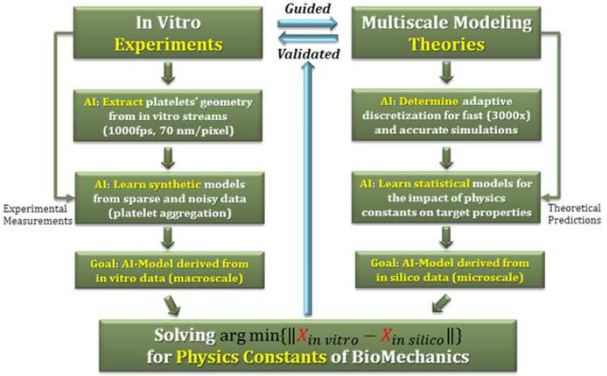
**ML-1: Machine Learning for Synthesizing Sparse In Vitro Data (V&V)**

*Multiscale Challenge: Sparse in-vitro data to validate in-silico predictions*
*Goal: To accurately predict the contact area for aggregated platelets under low shear stresses (0~10 dyne/cm$^2$)*
*Application: Used in our multiscale modeling to validate the in silico results*

**ML- 2: Machine Learning for Adaptive Discretization in Massive Multiscale**

*Multiscale Challenge: No effective algorithm for adapting temporal scales to the diverse spatial scales*
*Goal: Cut computing times for different simulation phases by 20~75%*
*Application: 'Coarse-Graining' in time scales to cut unnecessary computational steps without losing significant accuracy*

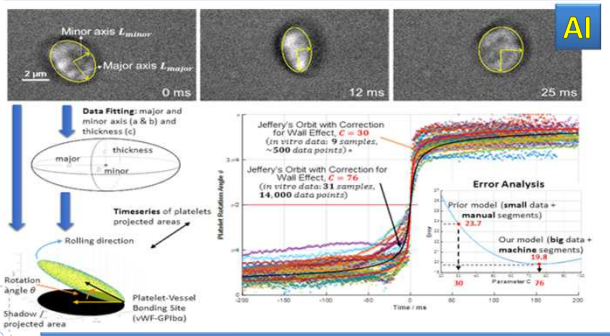


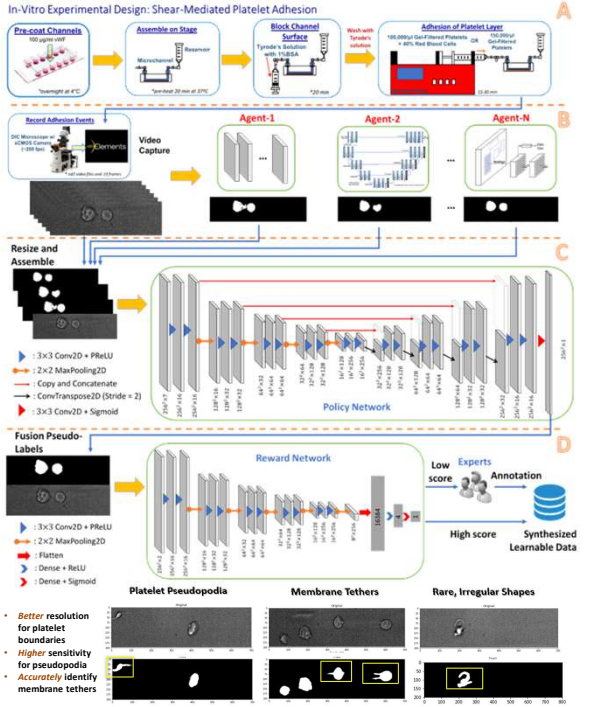

## Three Machine Learning AI-Models and Results

- **Multiscale Modeling:** A MD model describes intracellular components such as cytoskeleton, cytoplasm and protrusible actin filaments at molecular scales. A CGMD model describes (1) dynamic interaction between deformable platelets and blood flows; (2) the inter-platelet interaction: the recruitment aggregation through the αIIbβ3-Fibrinogen binding and the flipping of adhered platelets on vWF-coated blood vessels through the GPIbα-vWF binding. Last, in silico predictions are made through computer simulations and learned to produce to a prediction model by machine learning methods.
- **In Vitro Experiments:** Platelet flowing, flipping, recruitment aggregation and adhesion experiments are conducted under a variety of stresses. Platelet geometrical data are segmented with attention to boundaries: pseudopods and membrane tethers. Last, in vitro measures are learned to produce a synthetic model by machine learning methods.
- **Interfacing In Vitro and In Silico** (between macroscale and nanoscale): Formulate a multi-objective optimization problem by combining the in-vitro synthetic AI-model and the in-silico prediction AI-model. Solving the problem yields the physics constants for unveiling underlying mechanisms.
- **Machine Learning Models in this Framework:**
  - **Extract** platelets' discrete geometrical data from in vitro images.
  - **Learn** a synthetized model from sparse and noisy data.
  - **Adapt** temporal scales to diverse spatial scales under different dynamics.

- **Prediction Model for** Platelet Adhesion Under Flow Modeling In-Vitro Data Statistics of in-vitro results and Comparison (31 samples, 14,000 data points).




- **A Semi-Unsupervised Machine Learning System for Platelet Segmentation at Submicron Resolution**

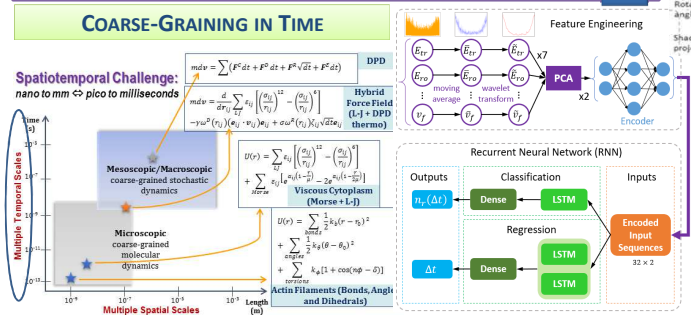



- **Addressing the CPMS Ten Simple Rules**

| | |
|---|---|
| **Rule 1. Define context** | Models are designed to reflect platelet properties and dynamics found in disease- and device-associated blood flow |
| **Rule 2. Use appropriate data** | Parameters and input variables are based on published and in-house in vitro observations. If any parameters cannot be validated , other model variables are monitored to ensure accurate reflection of platelet biology |
| **Rule 3. Evaluate within context** | Simulations are performed under physiological and pathological shear stresses relevant to blood vessels and blood-recirculating cardiovascular devices, with appropriate blood properties (i.e. viscosity, temperature). |
| **Rule 4. List limitations explicitly** | Numerical simulations are accurate in the context of published data and in-house in vitro observations. We do not make Further limitations are the capacity of the software to model biological observations and \HPC resources |
| **Rule 5. Version control** | All experimental data are traced by their creation date and generators. All DPD-CGMD files track the creation date. |
| **Rule 6. Document adequately** | Simulation codes/model markups are tracked and shared among the simulation group. All experimental data are stored in a video/spreadsheet database and shared among all team members via Stony Brook's Google Drive service |
| **Rule 7. Disseminate broadly** | We are exploring sharing simulation software and data/experimental data broadly via the Google Cloud Platform. These items are also presented during regular meetings and national/international conferences. |
| **Rule 8. Get independent reviews** | Our algorithms and experimental data will be shared with fellow IMAG researchers with similar work (i.e. Drs. Alber and Karniadakis) for independent evaluation. |
| **Rule 9. Test competing implementations** | We test the efficiency of various iterations of our DPD and CGMD codes to select the most appropriate model parameters. Due to the uniqueness of our approach, we do not have an external algorithm for direct comparison. |
| **Rule 10. Conform to standards** | While there are no set standards for our platelet-based experiments, we follow commonly followed practices for blood/platelet preparation, microscopy, and statistical analysis as published in relevant experimental journals. |

**MULTIPLE TIME SCALE ALGORITHM FOR MULTIPLE SPACE SCALE MODEL:**
Machine Learning-Guided Time Stepping Algorithms: MSM + RNN framework
- Multiple Time Stepping (**MTS**) : **Four-Level** Integrator Algorithm
- Recurrent Neural Network (**RNN**) : **Data-Driven** Online Learning Algorithm

### COARSE-GRAINING IN TIME







**Covered Spatial-Temporal Scales: Length: *nm – mm* Time: *ps – ms***

**NEURAL NETWORKS WE EMPLOYED:** Deep Neural Network (DNN); Convolutional Neural Network (CNN), Recurrent Neural Network (RNN), U-Net, Mark R-CNN.
**CONCLUSIONS:** Our computationally affordable, highly resolved, and validated multiscale modeling framework provides a potentially predictive platform to describe flow shear-induced activation, aggregation, and adhesion down to the nanoscales. **Our novel machine learning models can be used to validate simulation predictions and also can improve the modeling efficiency.** Ongoing simulations and experiments evaluate aggregation events with multiple platelets and incorporate GPIbα-vWF interactions for adhesion at moderate to high shear stresses. Our framework can be used to test new anti-platelet therapeutic approaches that increase platelet shear resistance. **We're building a AI-based approach for convergence of experiment, theory and computational sciences.**

**ACKNOWLEDGEMENTS:**
- **NIH** (NHLBI U01 **HL131052**, Bluestein, Danny)
- **NIH** (NHLBI R21 **HL096930-01**, Bluestein, Danny)
- **NIH** (NIBIB Quantum **U01EB012487**, Bluestein, Danny)
- **XSEDE** (**DMS140019** on TACC Stampede, Zhang, Peng)
- **XSEDE** (**DMS150011** on SDSC Comet, Zhang, Peng)

**PUBLICATIONS:**
[1] Zhang, P., et al, *Cellular and Molecular Bioengineering*, 7:552-574, 2014.
[2] Gao, C., et al, *Journal of Computational Physics*, 335:812-827, 2017.
[3] Zhang, P., et al, *Journal of Biomechanics*, 50:26-33, 2017.
[4] Zhang, N., et al, *Journal of Computational Physics*, 257:726-736, 2014
[5] Pothapragada, S., et al, *Int. J. Numer Biomed Engng*, 31:1-16, 2015
[6] Gupta, P., et al, *Cellular and Molecular Bioengineering*, 12:327-343, 2019