The feasibility of a whole-cell model of the human platelet

Steve Haigh

21 May 2026

Quick recap

What I evaluated

- Aiming to build a computer model of the platelet.
- Has been done before in prokaryotes and yeast but not in other eukaryotic cells.
- Lots of modelling done in platelets but not a whole-cell attempt.
- Lots of options e.g., COPASI, V-Cell and others.
- Landed on wcEcoli project from Covert lab — capable of scaling to whole cell, code quality excellent, something I could work with and expand.

What I did

- Forked wcEcoli (i.e., took my own working copy).
- Pruned all *E. coli* biology.
- Used AI to build tests to verify I didn't remove too much.
- Rebuilt cell contents and processes from the platelet literature.
- Kept the simulation engine, state partitioning, listener / analysis framework.
- ~5,000 lines of working code I didn't have to write.
- Emailed Prof. Covert — "sounds fun, good luck".

Method to build a calibrated model

For each mechanism, a five-step process:

1. **Anchor paper.** E.g., Dolan & Diamond 2014 supplied the validation experiment (Fig. 4 Ca^{2+} transients ± extracellular Ca^{2+}) and resting-state targets (100 nM cyt, 250 µM DTS).
2. **Literature review for kinetics.** Primary sources: deYoung–Keizer 1992 for IP_3R, Caride 2007 for PMCA, Hoover & Lewis 2011 for SOCE, Mazet 2020 for the PI cycle…
3. **Species enumeration.** Every Ca^{2+}-binding or -gating protein state, with a compartment tag (DTS, Cyt etc.). ~50 species for the calcium pathway.
4. **Copy numbers** from Burkhart 2012 platelet proteomics.
5. **Rate constants.** Primary-source values, units normalised, every value carrying a citation comment in code.

Then: does the resting state hold? Do the timescales line up? Do the unit tests still pass?

Build code from source data — start with the data

Start with the published data (usually PDF, sometimes CSV).

Biophysical Journal 106(9) 2049–2060

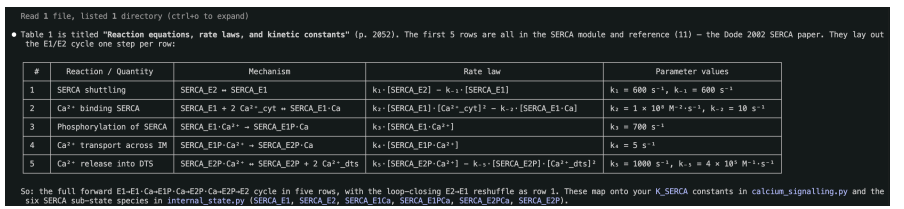
TABLE 1 Reaction equations, rate laws, and kinetic constants

	Reaction / Quantity	Mechanism	Rate Law / Rule	Parameter Values	Ref.
SERCA	**SERCA shuttling**	$SERCA_{E2} \leftrightarrow SERCA_{E1}$	$k_1 \cdot [SERCA_{E2}] - k_{-1} \cdot [SERCA_{E1}]$	$k_1 = 600\ s^{-1}$, $k_{-1} = 600\ s^{-1}$	(11)
	Ca^{2+}_{cyt} binding SERCA	$SERCA_{E1} + 2\ Ca^{2+}_{cyt} \leftrightarrow SERCA_{E1}{\cdot}Ca^{2+}$	$k_1 \cdot [SERCA_{E1}] \cdot [Ca^{2+}_{cyt}]^2 - k_{-1} \cdot [SERCA_{E1}{\cdot}Ca^{2+}]$	$k_1 = 1 \times 10^{15}\ M^{-2}{\cdot}s^{-1}$, $k_{-1} = 10\ s^{-1}$	(11)
	Phosphorylation of SERCA	$SERCA_{E1}{\cdot}Ca^{2+} \leftrightarrow SERCA_{E1}P{\cdot}Ca^{2+}$	$k_1 \cdot [SERCA_{E1}Ca^{2+}] - k_{-1} \cdot [SERCA_{E1}P{\cdot}Ca^{2+}]$	$k_1 = 700\ s^{-1}$, $k_{-1} = 5\ s^{-1}$	(11)
	Ca^{2+} transport across IM	$SERCA_{E1}P{\cdot}Ca^{2+} \leftrightarrow SERCA_{E2}P{\cdot}Ca^{2+}$	$k_1 \cdot [SERCA_{E1}P{\cdot}Ca^{2+}] - k_{-1} \cdot [SERCA_{E2}P{\cdot}Ca^{2+}]$	$k_1 = 600\ s^{-1}$, $k_{-1} = 50\ s^{-1}$	(11)
	Ca^{2+}_{dts} release into DTS	$SERCA_{E2}P{\cdot}Ca^{2+} \leftrightarrow SERCA_{E2}P + 2\ Ca^{2+}_{dts}$	$k_1 \cdot [SERCA_{E2}P{\cdot}Ca^{2+}] - k_{-1} \cdot [SERCA_{E2}P] \cdot [Ca^{2+}_{dts}]^2$	$k_1 = 1000\ s^{-1}$, $k_{-1} = 4 \times 10^9\ M^{-2}{\cdot}s^{-1}$	(11)
	SERCA dephosphorylation	$SERCA_{E2}P \leftrightarrow SERCA_{E2}$	$k_1 \cdot [SERCA_{E2}P] - k_{-1} \cdot [SERCA_{E2}]$	$k_1 = 500\ s^{-1}$, $k_{-1} = 1\ s^{-1}$	(11)
IP_3R	**IP_3R inhibition**	$IP_3R_n + Ca^{2+}_{cyt} \leftrightarrow IP_3R_{i1}$	$[IP_3R_n] \cdot ((k_1 \cdot L_1 + I_2) \cdot [Ca^{2+}_{cyt}] / (L_1 + [Ca^{2+}_{cyt}] \cdot (1 + L_1 / L_3)))$	$k_1 = 0.64\ s^{-1}{\cdot}\mu M^{-1}$, $L_1 = 0.12\ \mu M$, $I_2 = 1.7\ s^{-1}$	(12)
			$- [IP_3R_{i1}] \cdot (k_{-1} + I_{-2})$	$L_3 = 0.025\ \mu M$, $k_{-1} = 0.04\ s^{-1}$, $I_{-2} = 0.8\ s^{-1}$	
	IP_3R binding IP_3	$IP_3R_n + IP_3 \leftrightarrow IP_3R_o$	$[IP_3R_n] \cdot [IP_3] \cdot ((k_2 \cdot L_3 + I_4 \cdot [Ca^{2+}_{cyt}]) / (L_3 + [Ca^{2+}_{cyt}] \cdot (1 + L_3 / L_1)))$	$k_2 = 37.4\ s^{-1}{\cdot}\mu M^{-1}$, $I_4 = 1.7\ s^{-1}{\cdot}\mu M^{-1}$	(12)
			$- [IP_3R_o] \cdot ((k_{-2} + I_{-4} \cdot [Ca^{2+}_{cyt}]) / (1 + [Ca^{2+}_{cyt}] / L_5))$	$k_{-2} = 1.4\ s^{-1}$, $I_{-4} = 2.5\ \mu M^{-1}{\cdot}s^{-1}$, $L_5 = 54.7\ \mu M$	
	IP_3R activation	$IP_3R_o + Ca^{2+}_{cyt} \leftrightarrow IP_3R_a$	$[IP_3R_o] \cdot ((k_4 \cdot L_5 + I_6) \cdot [Ca^{2+}_{cyt}] / (L_5 + [Ca^{2+}_{cyt}]))$	$k_4 = 4\ s^{-1}{\cdot}\mu M^{-1}$, $I_6 = 4707\ s^{-1}$	(12)
			$- [IP_3R_a] \cdot (L_1 \cdot (k_{-4} + I_{-6}) / (L_1 + [Ca^{2+}_{cyt}]))$	$k_{-4} = 0.54\ s^{-1}{\cdot}\mu M^{-1}$, $I_{-6} = 11.4\ s^{-1}$	(12)
	IP_3R inhibition	$IP_3R_a + Ca^{2+}_{cyt} \leftrightarrow IP_3R_{i2}$	$[IP_3R_a] \cdot (k_1 \cdot L_1 + I_2) \cdot [Ca^{2+}_{cyt}] / (L_1 + [Ca^{2+}_{cyt}]) - [IP_3R_{i2}] \cdot (k_{-1} + I_{-2})$		(12)
	IP_3R closing	$IP_3R_o \leftrightarrow IP_3R_s$	$[IP_3R_o] \cdot (k_3 \cdot L_5 / (L_5 + [Ca^{2+}_{cyt}])) - [IP_3R_s] \cdot k_{-3}$	$k_3 = 11\ s^{-1}{\cdot}\mu M^{-1}$, $k_{-3} = 29.8\ s^{-1}$	(12)
	Channel open probability (P_{oIP3R})		$(0.9 \cdot IP_3R_a / IP_3R_{total} + 0.1 \cdot pIP_3R_o / IP_3R_{total})^4$		(12)
	IM potential (ψ_{IM})		$RT / zF \cdot \ln (Ca^{2+}_{cyt} / Ca^{2+}_{dts})$		
	Ca^{2+} release from DTS via IP_3R	$Ca^{2+}_{dts} \leftrightarrow Ca^{2+}_{cyt}$	$N_{IP3R} \cdot P_{oIP3R} \cdot \gamma_{IP3R} \cdot e \cdot (V_{IM} - \psi_{IM})$	$\gamma_{IP3R} = 10\ pS$	(30)
	Ca^{2+}_{dts} binding STIM1	$STIM1 + Ca^{2+}_{dts} \leftrightarrow STIM1{\cdot}Ca^{2+}_{dts}$	$k_1 \cdot [STIM1] \cdot [Ca^{2+}_{dts}] - k_{-1} \cdot [STIM1{\cdot}Ca^{2+}]$	$k_{-1}/k_1 = 200\ \mu M$	(1)
	STIM1 dimerization	$2\ STIM1 \leftrightarrow STIM1_2$	$k_1 \cdot [STIM1] \cdot [STIM1] - k_{-1} \cdot [STIM1_2]$	$k_1 = 9 \times 10^8\ s^{-1}{\cdot}M^{-1}$, $k_{-1} = 3.5\ s^{-1}$	This study
	Puncta formation factor (θ_p)		$\alpha \cdot [Ca^{2+}_{cyt}]^n / ([Ca^{2+}_{cyt}]^n + K_M^n) + 0.01$	$K_M = 100$ - $300\ nM$, $n = 0.5$ - 3, $\alpha = 0.2$	This study
	$STIM1_2$ puncta entry ($(STIM1_2)_p$)	$STIM1_2 \leftrightarrow (STIM1_2)_p$	$[STIM1_2] \cdot \theta_p$		
	STIM in puncta binding Orai1	$(STIM1_2)_p + Orai1 \leftrightarrow (STIM1_2)_p{\cdot}Orai1$	*see equations below*		

Just a small excerpt.

Build code from source data — AI extraction

Ask AI to extract — get it to tell you what it's doing to sanity check (but it is *very* good at this).



Read 1 file, listed 1 directory (ctrl+o to expand)

● Table 1 is titled **"Reaction equations, rate laws, and kinetic constants"** (p. 2052). The first 5 rows are all in the SERCA module and reference (11) — the Dode 2002 SERCA paper. They lay out the E1/E2 cycle one step per row:

#	Reaction / Quantity	Mechanism	Rate law	Parameter values
1	SERCA shuttling	SERCA_E2 ⇌ SERCA_E1	$k_1 \cdot [SERCA\_E2] - k_{-1} \cdot [SERCA\_E1]$	$k_1 = 600\ s^{-1}$, $k_{-1} = 600\ s^{-1}$
2	Ca^{2+} binding SERCA	SERCA_E1 + 2 Ca^{2+}_cyt ⇌ SERCA_E1·Ca	$k_2 \cdot [SERCA\_E1] \cdot [Ca^{2+}\_cyt]^2 - k_{-2} \cdot [SERCA\_E1{\cdot}Ca]$	$k_2 = 1 \times 10^8\ M^{-2} s^{-1}$, $k_{-2} = 10\ s^{-1}$
3	Phosphorylation of SERCA	SERCA_E1·Ca^{2+} → SERCA_E1P·Ca	$k_3 \cdot [SERCA\_E1{\cdot}Ca^{2+}]$	$k_3 = 700\ s^{-1}$
4	Ca^{2+} transport across IM	SERCA_E1P·Ca^{2+} → SERCA_E2P·Ca	$k_4 \cdot [SERCA\_E1P{\cdot}Ca^{2+}]$	$k_4 = 5\ s^{-1}$
5	Ca^{2+} release into DTS	SERCA_E2P·Ca^{2+} ⇌ SERCA_E2P + 2 Ca^{2+}_dts	$k_5 \cdot [SERCA\_E2P{\cdot}Ca^{2+}] - k_{-5} \cdot [SERCA\_E2P] \cdot [Ca^{2+}\_dts]^2$	$k_5 = 1000\ s^{-1}$, $k_{-5} = 4 \times 10^3\ M^{-1} s^{-1}$

So: the full forward E1→E1·Ca→E1P·Ca→E2P·Ca→E2P→E2 cycle in five rows, with the loop-closing E2→E1 reshuffle as row 1. These map onto your K_SERCA constants in calcium_signalling.py and the six SERCA sub-state species in internal_state.py (SERCA_E1, SERCA_E2, SERCA_E1Ca, SERCA_E1PCa, SERCA_E2PCa, SERCA_E2P).

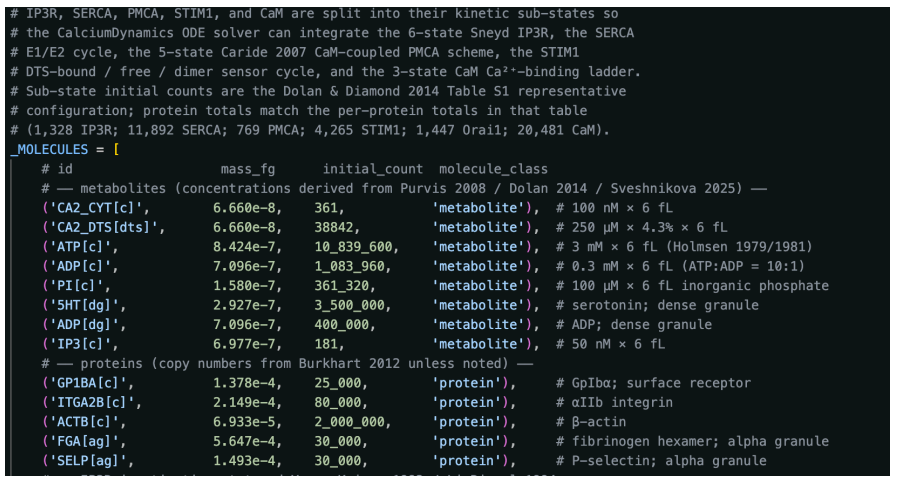
> E.g. I asked: "Please read Dolan and Diamond 2014 and tell me what the first 5 lines of table 1 show."

Simply not feasible to do this by hand at large scale; AI could process 100s of papers in a day. The only challenge is verification, but this is also amenable to AI assistance.

Build code from source data — AI writes the code

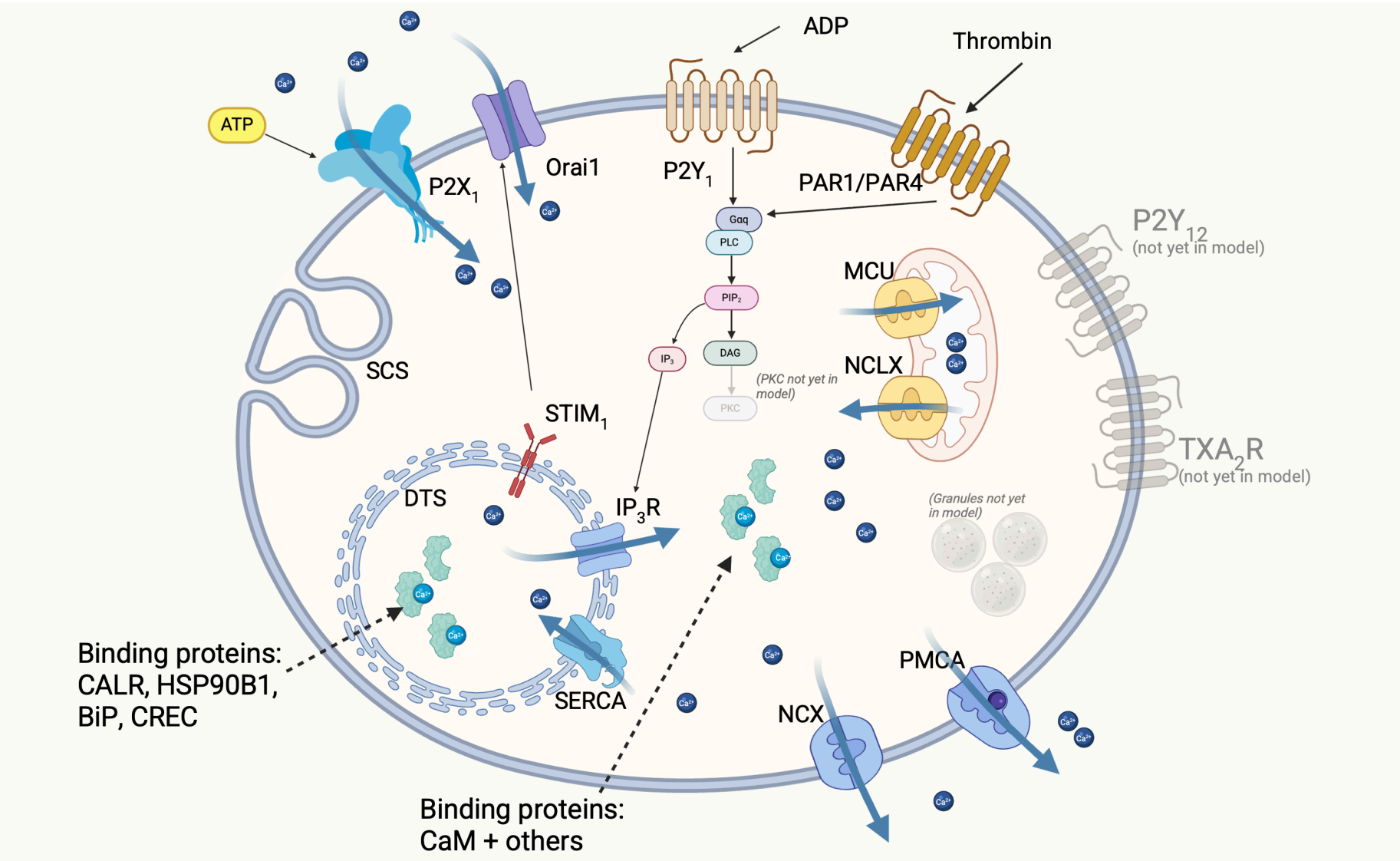
Then get AI to build the code, using the *wcEcoli* code as a template.

The code works well, the architecture is not perfect — needs some work to make it more usable for future expansion — but it's an excellent start.



```
# IP3R, SERCA, PMCA, STIM1, and CaM are split into their kinetic sub-states so
# the CalciumDynamics ODE solver can integrate the 6-state Sneyd IP3R, the SERCA
# E1/E2 cycle, the 5-state Caride 2007 CaM-coupled PMCA scheme, the STIM1
# DTS-bound / free / dimer sensor cycle, and the 3-state CaM Ca²⁺-binding ladder.
# Sub-state initial counts are the Dolan & Diamond 2014 Table S1 representative
# configuration; protein totals match the per-protein totals in that table
# (1,328 IP3R; 11,892 SERCA; 769 PMCA; 4,265 STIM1; 1,447 Orai1; 20,481 CaM).
_MOLECULES = [
    # id                  mass_fg      initial_count  molecule_class
    # ── metabolites (concentrations derived from Purvis 2008 / Dolan 2014 / Sveshnikova 2025) ──
    ('CA2_CYT[c]',        6.660e-8,    361,           'metabolite'),  # 100 nM × 6 fL
    ('CA2_DTS[dts]',      6.660e-8,    38842,         'metabolite'),  # 250 µM × 4.3% × 6 fL
    ('ATP[c]',            8.424e-7,    10_839_600,    'metabolite'),  # 3 mM × 6 fL (Holmsen 1979/1981)
    ('ADP[c]',            7.096e-7,    1_083_960,     'metabolite'),  # 0.3 mM × 6 fL (ATP:ADP = 10:1)
    ('PI[c]',             1.580e-7,    361_320,       'metabolite'),  # 100 µM × 6 fL inorganic phosphate
    ('5HT[dg]',           2.927e-7,    3_500_000,     'metabolite'),  # serotonin; dense granule
    ('ADP[dg]',           7.096e-7,    400_000,       'metabolite'),  # ADP; dense granule
    ('IP3[c]',            6.977e-7,    181,           'metabolite'),  # 50 nM × 6 fL
    # ── proteins (copy numbers from Burkhart 2012 unless noted) ──
    ('GP1BA[c]',          1.378e-4,    25_000,        'protein'),     # GpIbα; surface receptor
    ('ITGA2B[c]',         2.149e-4,    80_000,        'protein'),     # αIIb integrin
    ('ACTB[c]',           6.933e-5,    2_000_000,     'protein'),     # β-actin
    ('FGA[ag]',           5.647e-4,    30_000,        'protein'),     # fibrinogen hexamer; alpha granule
    ('SELP[ag]',          1.493e-4,    30_000,        'protein'),     # P-selectin; alpha granule
```

What did I build? (so far)



Spotting errors in the source

The Purvis 2008 k_3 story.

- Asked AI to check values it had extracted.
- AI spotted that a value quoted in Purvis & Diamond 2008 (a well-known platelet modelling paper) was wrong — units were wrong and value was 100× too large.
- AI found the original study via the references and I verified for myself; AI was correct, there was a typo.
- Very strong use-case for AI — tracking values and assertions across papers. Also very useful for helping to review data: I got the AI to print values it had found, with reference to the primary source and the code, so I could easily review.

Validation

Criteria

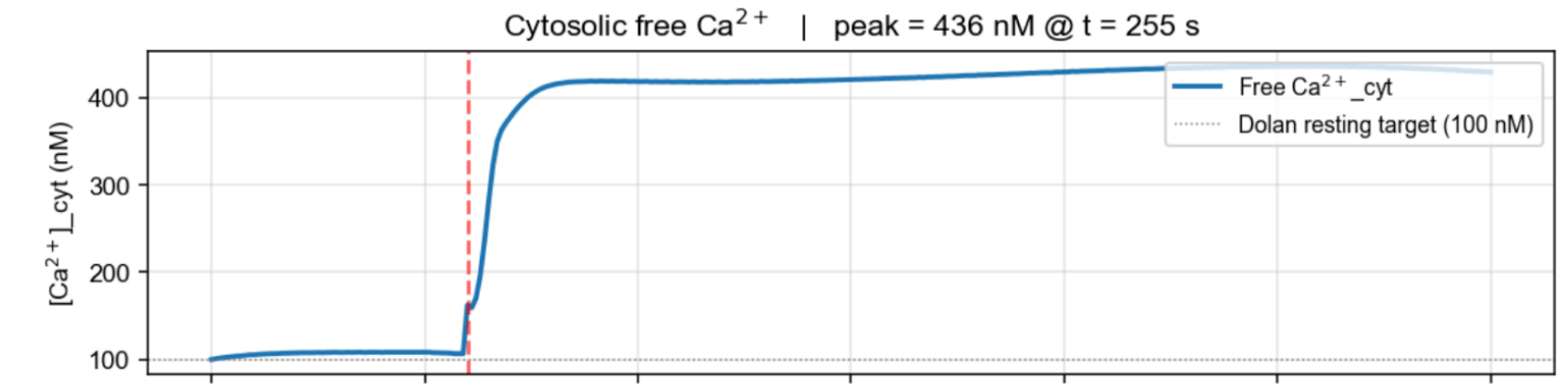
Compared output to existing models. Success criteria:

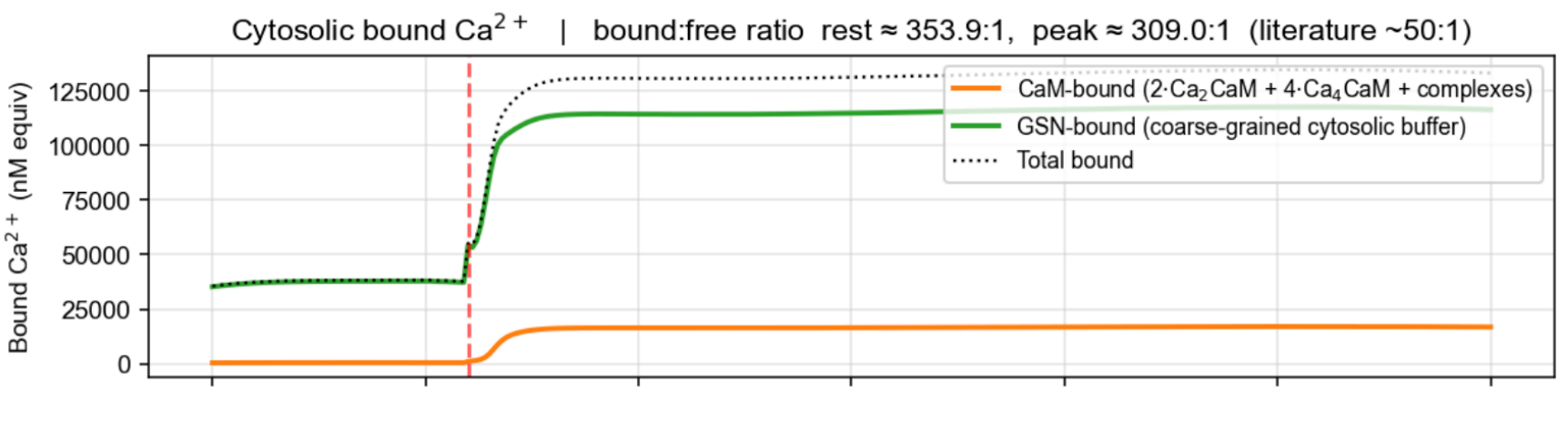
- Resting cytosolic Ca^{2+} within 100 ± 10 nM band
- Resting DTS Ca^{2+} in the literature range
- IP_3-stimulated transient peak height in band
- Transient duration matches shape of previous model(s)

Model state at 14 May

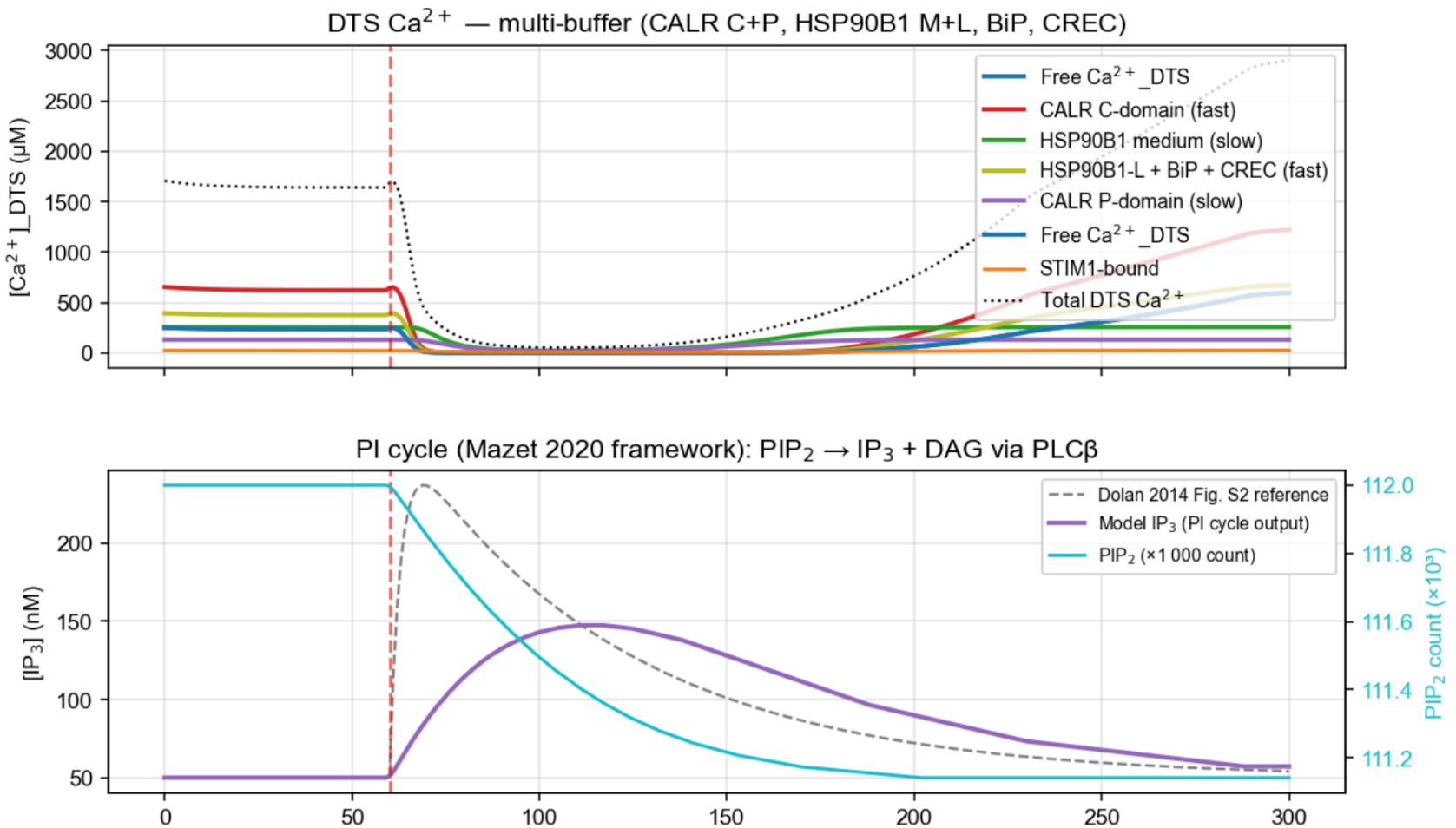
- Resting cyt Ca^{2+}: 104 nM ✓
- Resting DTS Ca^{2+}: 235 µM ✓
- Resting IP_3: 50 nM ✓
- Resting Gαq-active: 100 of 5000 ✓
- 21 / 21 unit tests pass
- Driven by physiological agonists — 1 nM thrombin, 10 µM ADP. No hand-fitted IP_3 forcing.

Model stimulated with ADP / thrombin — cytosolic Ca^{2+}





Model stimulated with ADP / thrombin — DTS Ca^{2+}



Where it goes from here

Near-term biology

- Granule release — the model currently stops at the Ca^{2+} peak.
- Implement $P2Y_{12}$, TXA2R receptors and more, depending on time.
- Run experiments e.g. when $P2Y_{12}$ / G_i are wired up, add clopidogrel as an *in silico* drug experiment.

Wider impact

- Platelet–CTC interactions. A calibrated platelet model could in principle be coupled to a tumour-cell model.
- Methodology generalises to other single-cell calibration problems — if pathways are well-quantified, no reason not to try modelling them in a cancer cell.
- Multiple cell–cell interactions — thrombus formation?

Blue sky goal

- I'd like to get to a point where a biologist with limited coding experience could build a model with AI help. Very long shot: a system where a user drops primary data and text instructions into an app and the model builds itself.