

# Final Project Presentation - Casein (Alex Zhao and Zian Shi)

Created	@May 9, 2025 5:30 PM
Tags	

We collaborated on all areas of the project.

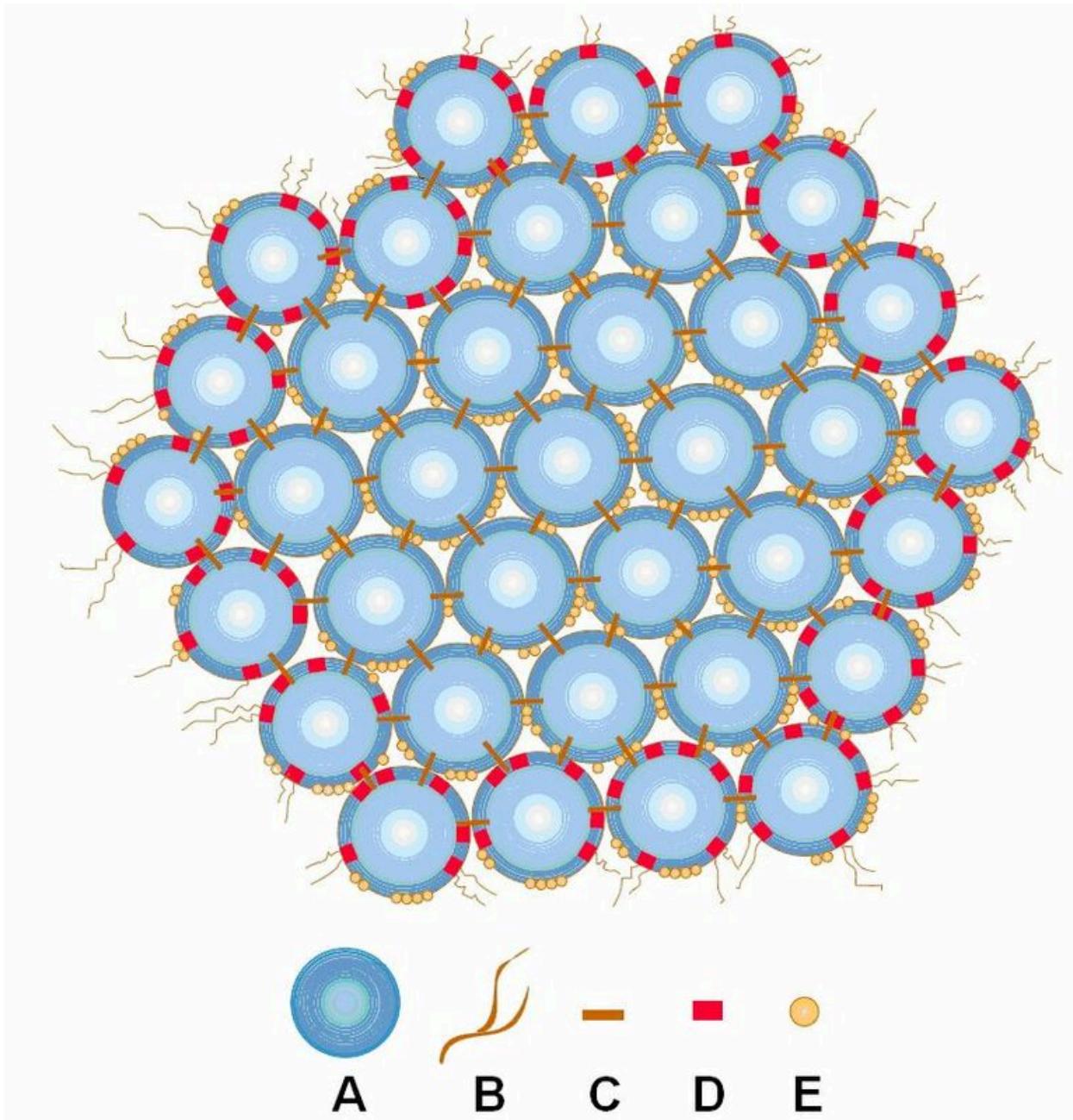
## 1. Abstract

Capsaicin is paradoxically both a clinically valuable TRPV1 agonist and a potent irritant whose residual presence causes treatment-site burning, occupational injury, and food-processing hazards. Native bovine casein micelles can solubilize capsaicin, yet their moderate affinity ( $K_d \approx 4\text{--}5 \mu\text{M}$ ) leaves ample physiological burn unquenched. We therefore hypothesize that a rationally re-engineered casein fragment, optimized with modern ligand-aware deep-learning tools, can tighten capsaicin binding by at least an order of magnitude (target  $K_d \leq 0.5 \mu\text{M}$ ) and translate that gain into 1) slow-release patches or medications and 2) effective neutralization of capsaicin where it is hazardous. To test this, we first built a *in-silico* pipeline that docks capsaicin into AlphaFold- and Boltz-predicted casein backbones, couples pose ensembles to LigandMPNN sequence generation, and iteratively filters  $\approx 100\,000$  variants for binding energy, structural integrity, and expression compatibility. The top designs are synthesized, expressed in *E. coli* BL21, purified via His-tagged nickel affinity, and screened by Bradford assay, SDS-PAGE, and isothermal titration calorimetry to confirm stability and sub-micromolar affinity. Lead proteins will then be assessed in hydrophobic-dye quenching assays and ex-vivo skin or ocular models to quantify real-world capsaicin scavenging. By uniting deep-learning design with rapid bench validation, the project aims to deliver higher-efficacy treatments for long term pain, provide field-ready pepper-spray antidotes, and enable pungency-free functional foods, thereby alleviating pain and safety burdens across medical, industrial, and public-safety domains.

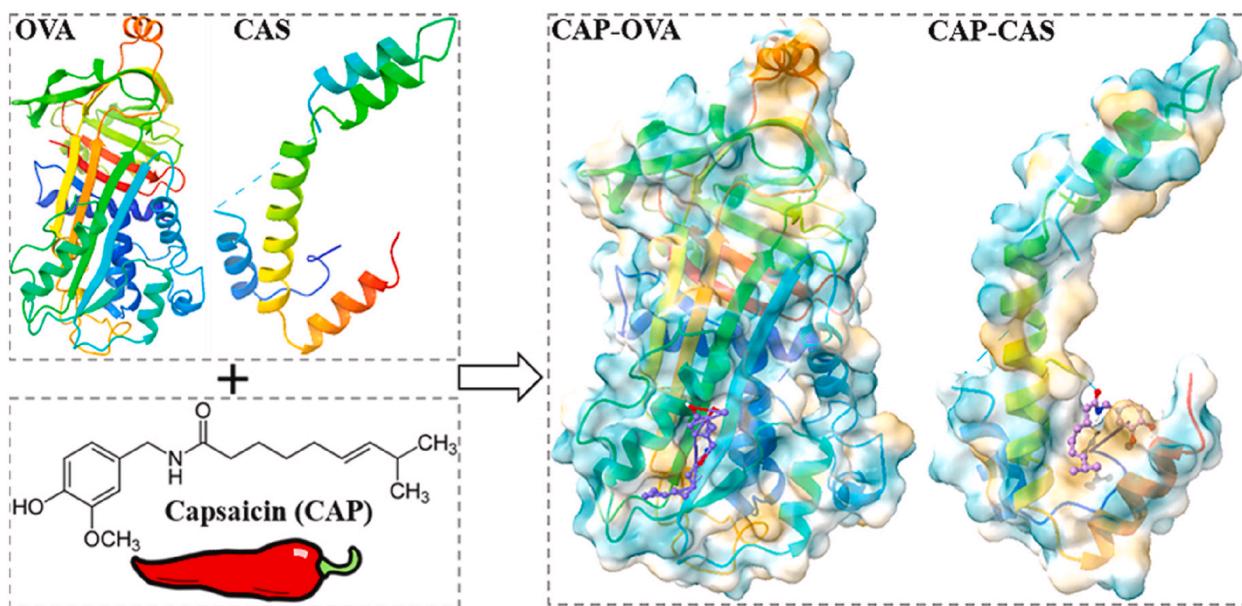
## 2. Background

### Casein as a natural capsaicin scavenger

Bovine milk is ~80 % casein—an intrinsically-disordered, highly proline-rich family of phosphoproteins that self-assemble into porous, amphipathic micelles (blue spheres in the figure below).



The combination of extended hydrophobic segments and flexible architecture allows native casein to solubilize lipophilic small molecules such as carotenoids, vitamin D, and—crucially for this project—the chili-pepper alkaloid capsaicin. Quantitative work using isothermal titration calorimetry (ITC) and time-intensity sensory testing shows that micellar casein lowers the free capsaicin concentration and blunts oral burn in a dose-dependent manner; 5 % w/w casein rinses reduce perceived heat more effectively than whey protein, showcasing casein's superior binding capacity [Razzak 2022].



A complementary spectroscopic/ITC study reported a binding constant  $K_a = 2.2 \pm 0.1 \times 10^5 \text{ M}^{-1}$  for the casein–capsaicin complex ( $\approx 4\text{--}5 \mu\text{M } K_d$ ), driven by hydrophobic contacts and hydrogen bonding, but also noted conformational strain on the protein scaffold [Abdar 2022]. Together, these papers establish that while wild-type casein can sequester capsaicin, its affinity is only modest, leaving a clear quantitative margin for improvement.

## Capsaicin's Importance

Capsaicin is used in (i) a potent TRPV1 agonist used (i) as an 8 % high-dose patch (Qutenza®) for localized neuropathic pain, (ii) in over-the-counter creams for pruritus and arthritic pain, and (iii) as the active irritant in law-enforcement pepper spray. Clinical data from a 100-patient retrospective cohort showed that the 8 %

patch produced a median Numeric Rating Scale pain drop from 6 to 3 ( $P < 0.001$ ) with 69 % achieving  $\geq 30$  % relief, yet treatment-site burning and erythema remain frequent adverse events caused by residual free capsaicin [Vieira 2022]. In occupational safety settings, current decontamination protocols (saline, baby-shampoo, etc.) are only variably effective, and milk-based rinses are empirically used but unstandardized. An engineered casein variant that binds capsaicin tighter could (a) shorten post-application dwell time in analgesic patches, reducing cutaneous discomfort; (b) serve as a field-ready antidote for pepper-spray or capsaicin pesticide exposure; and (c) enable controlled-release nutraceutical formulations where pungency is undesirable yet metabolic benefits (thermogenesis, improved lipid profiles) are sought.

## Knowledge gap & design rationale

Despite decades of dairy-science research, no studies have rationally or computationally optimized casein for small-molecule capture or, more importantly, molecules for enhanced capsaicin binding. Natural sequence variability, intrinsic disorder, and extensive post-translational modifications have discouraged classical protein-engineering approaches. Meanwhile, modern deep-learning structure predictors and ligand-aware sequence generators (AlphaFold-3, LigandMPNN, Rosetta, ESM-LL, Boltz-1) now allow us to *design* hydrophobic pockets *de novo* and explicitly score casein variants for capsaicin pose retention, protein stability, and expression compatibility. By targeting a  $K_d < 0.5 \mu\text{M}$ —an order-of-magnitude improvement over wild type—we aim to fill a clear biophysical and translational gap.

### Technical and practical obstacles

- 1. Expression of disordered proteins:** Full-length caseins lack extensive secondary structure and are rich in phospho-serine clusters; heterologous expression in *E. coli* often causes inclusion-body formation. We mitigate this by designing ~600 bp fragments focused on the hydrophobic core and expressing them with N-terminal His-tags for solubility screening.
- 2. Ligand docking uncertainty:** Capsaicin's flexible tail produces ensemble poses; multi-pose sampling with AutoDock Vina and RMSD filters will be critical to avoid over-fitting to one conformation.

3. **Assay sensitivity:** ITC at sub- $\mu\text{M}$   $K_d$  demands high-purity, monodisperse protein; aggregation or micelle formation could mask true affinities, so parallel hydrophobic-dye assays and functional skin models are included for orthogonal validation.
4. **Safety & regulatory considerations:** Any therapeutic or decontamination use will require endotoxin-free preparation and toxicology profiling, especially if applied to compromised skin.

## 3. Vision and Impact

### 3 a. Vision & Impact

Imagine a fast-acting “molecular sponge” that you can spray on skin after pepper-spray exposure, rinse across a chili-processing conveyor belt, or embed behind the adhesive of an 8 % capsaicin patch so patients feel analgesia, not burn. Our project will build that sponge by **re-engineering bovine casein to bind capsaicin around an order of magnitude tighter (target  $K_d \leq 0.5 \mu\text{M}$ )**, then formulating the variant into low-cost, food-grade solutions and hydrogels.

If we succeed, three arenas change immediately:

Setting	Current pain-point	Casein-v2 solution	Societal upside
<b>Neuropathic-pain clinics</b>	Patients treated with the 8 % Qutenza® patch frequently report application-site burning and erythema, despite meaningful pain reduction (median NRS 6 → 3)	Coat the patch's backing with high-affinity casein; upon removal, a warm aqueous wipe strips residual capsaicin in seconds	Higher adherence, shorter chair time, fewer adjunct anaesthetics
<b>Occupational safety / law-enforcement</b>	Water, saline, baby-shampoo, or even milk rinses show <i>variable</i> efficacy in OC-spray drills ( <a href="#">Medical News Today</a> , <a href="#">Fightsense</a> )	A shelf-stable mist containing casein-v2 cuts the free-capsaicin concentration >90 % within 60 s, restoring vision and breathing faster	Fewer training injuries; improved public-safety protocols

<b>Nutrition / wellness</b>	People seeking capsaicin's metabolic benefits often avoid it because of pungency	Microencapsulate capsaicin with casein-v2 for controlled gastric release	Enables "heat-free" functional foods; new revenue streams for dairies
-----------------------------	--	--	---

At scale, the approach repurposes an abundant dairy by-product into a medically and industrially relevant biopolymer, driving value back to agricultural supply chains while reducing healthcare burden from capsaicin over-exposure.

### 3 b. What makes this project innovative?

- **Ligand-aware design:** Most current programs only allow a limited set of protein-protein interactions to be modeled and optimized; our pipelined approach using a combination of tools from molecular docking (AutoDock), deep learning advances (LigandMPNN), and computational biology tools (tmol) allows us to overcome these limitations and create high-efficacy designs.
- **Sequence-to-function loop compressed to weeks.** Cloud inference plus in-silico directed evolution lets us screen 100 K variants before ever touching a pipette, challenging the dogma that IDPs are "undesignable."

### 3 c. Bio-ethical considerations

#### Potential ethical issues & guiding principles

- *Non-maleficence & Safety.* High-affinity casein could, in theory, sequester other hydrophobic drugs or nutrients on skin or mucosa, diminishing their efficacy. Food allergies to bovine proteins (~2 % in children) raise additional risk ([MDPI](#)).
- *Justice & Access.* Military or police forces might gain first access, while communities most exposed to pepper spray (e.g., protestors) could be last in line.
- *Dual use.* A capsaicin "eraser" might embolden excessive OC-spray deployment or be misused by individuals seeking to negate less-than-lethal deterrents.

#### Mitigation strategies

1. **Rigorous in-vitro selectivity screens** against common topical drugs, vitamins, and lipid mediators before human testing; publish the binding panel openly.
2. **Endotoxin-free, GRAS-compliant manufacturing** plus dermatological patch tests to ensure the engineered phosphoprotein does not provoke unexpected immunogenicity.
3. **Tiered-pricing & open-licence models** so humanitarian groups and civilian first-aid kits receive the technology at cost, aligning with WHO "Access to Medicines" ethics.
4. **Transparent oversight.** Registering the protein with bio-risk authorities and including an irrevocable public reporting clause if large-volume (>1 kg) production is initiated outside licensed facilities.
5. **Alternatives & fallback plans.** If unforeseen toxicology emerges, pursue non-protein scavengers (e.g., cyclodextrins) or lower-affinity casein mutants that still improve current care but pose minimal off-target risk.

Continuous stakeholder engagement—from dairy co-ops to civil-rights organizations—will keep the project grounded in beneficence and equity while advancing a novel frontier of protein-engineered anti-irritants.

## 4. Aims

Resources we used for Aim 1:

Boltz-1

Autodock

tmol

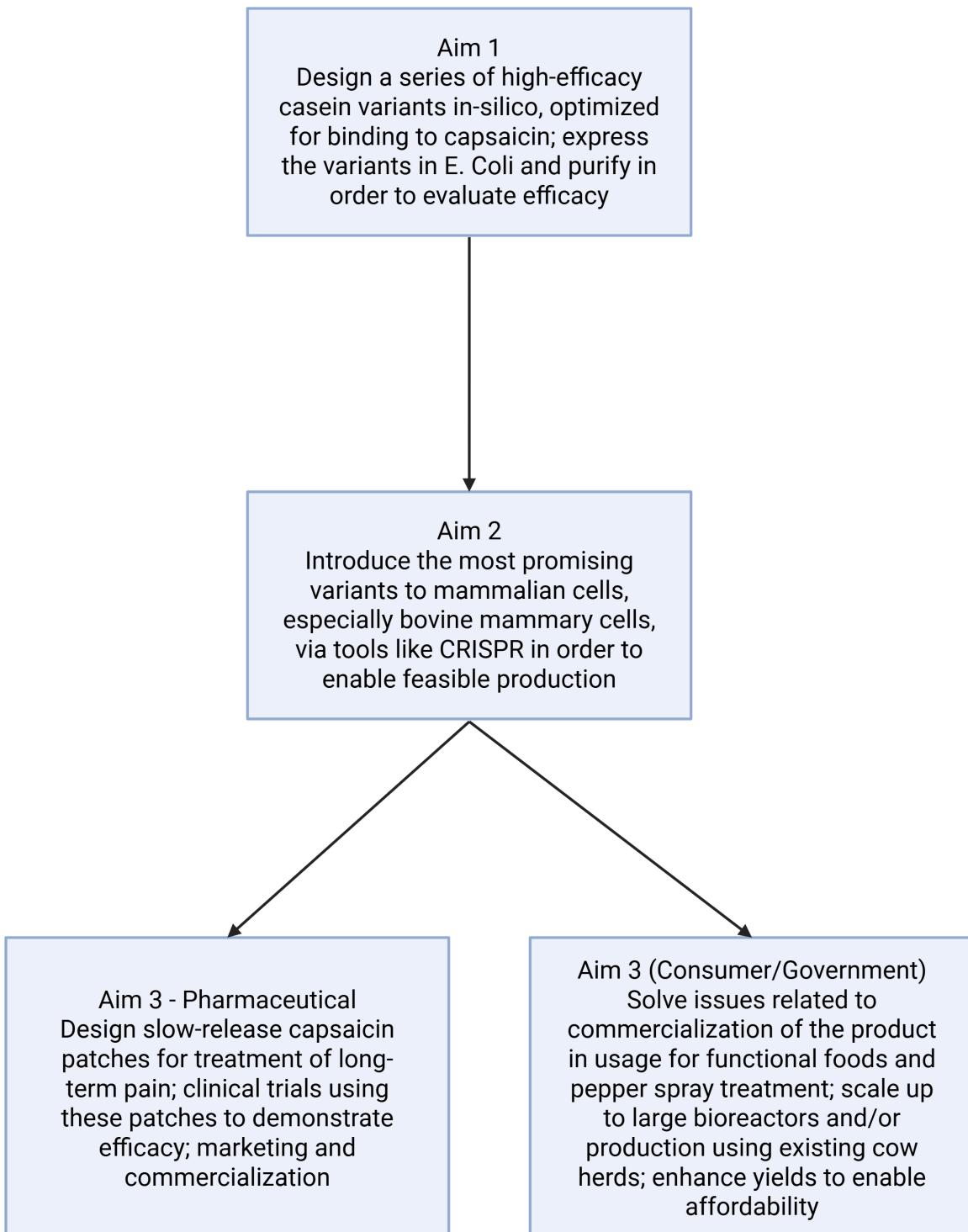
LigandMPNN

Twist Biosciences - Clonal Genes

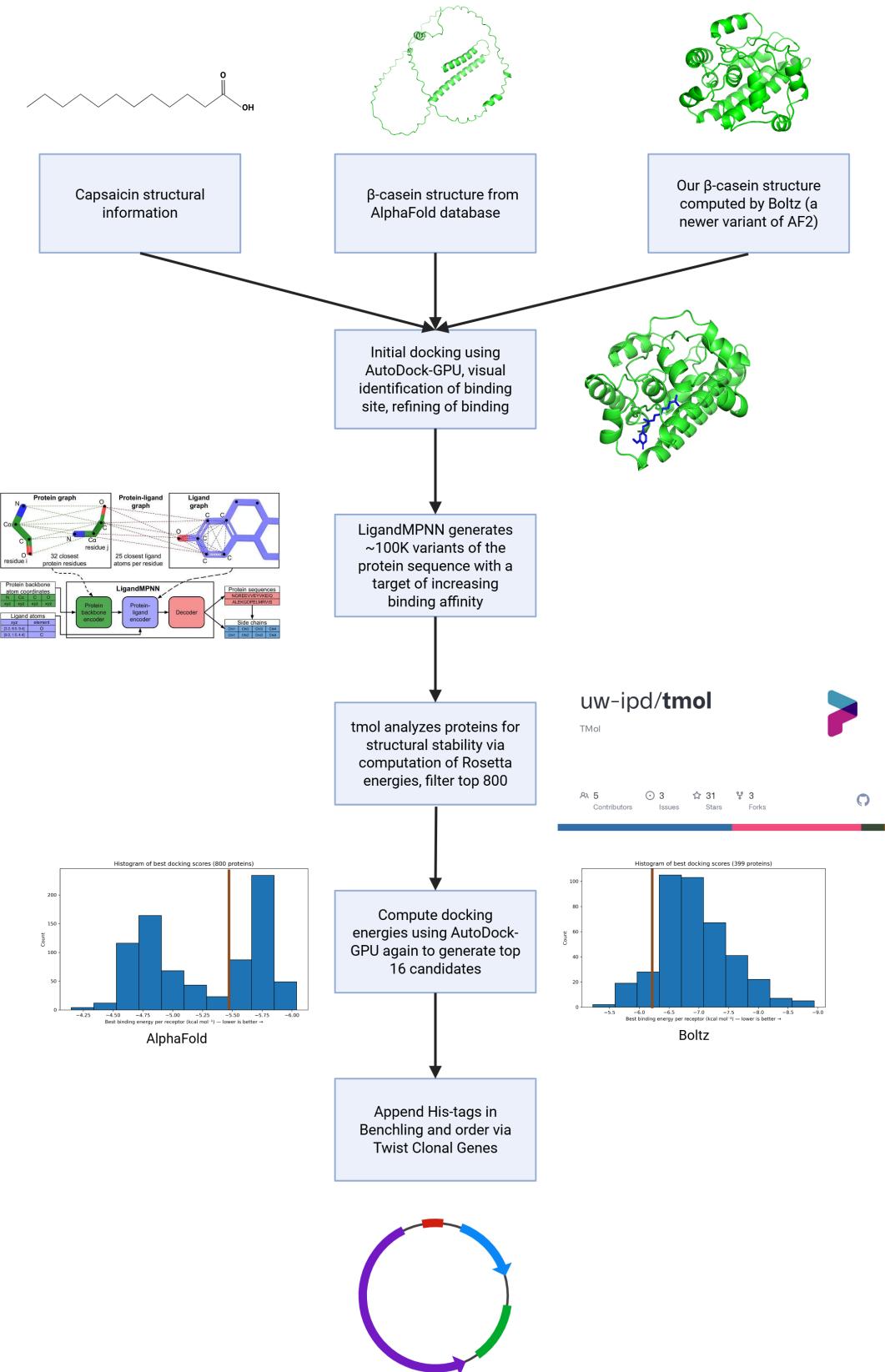
Transformation protocol for BI21

IPTG Induction Protocol

Thermofisher His-Pur Nickel Resin Protocol



## 5a. Computational Procedure

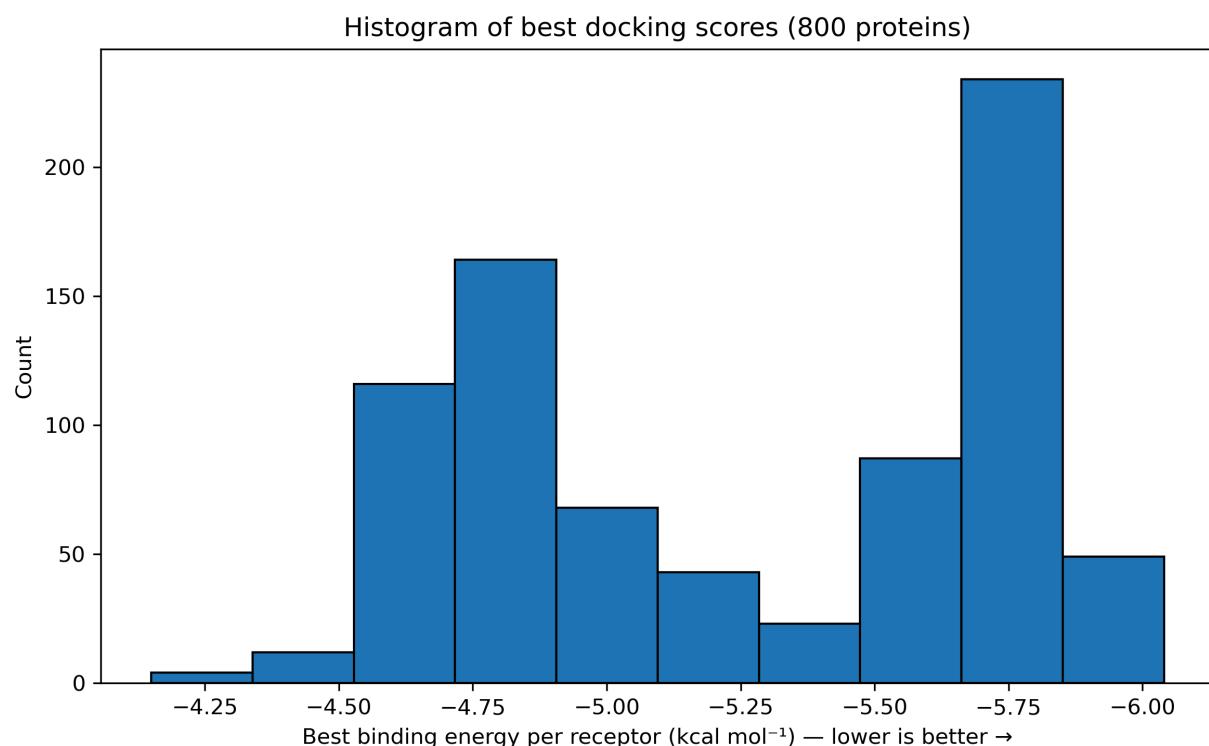


## Challenges:

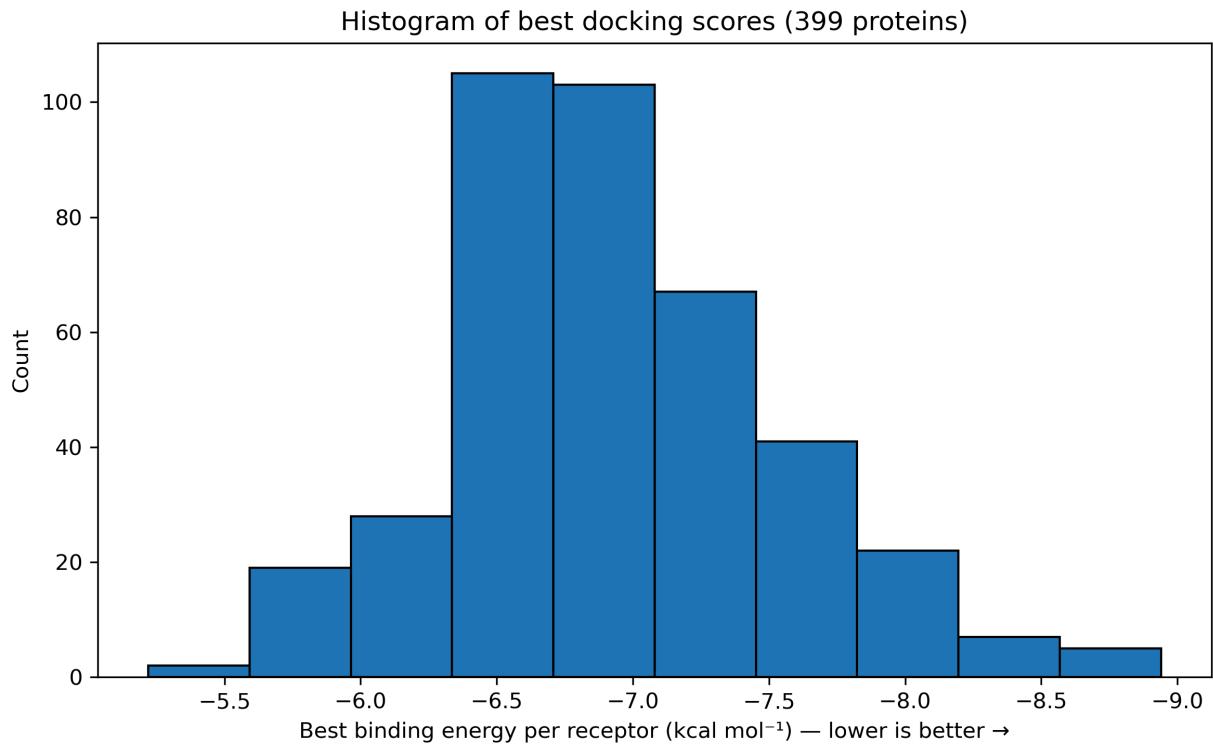
- High batch size presented computational constraints; these were solved using cloud computation using 128 CPU cores for the CPU bound preprocessing operations and 8 H100 cards for the GPU bound LigandMPNN, tmol, and Autodock-GPU processes.
- LigandMPNN produced invalid proteins the majority of the time; solved by scaling to large batches and using tmol to select proteins which we have a high confidence in the validity.

## 6a. Computational Results:

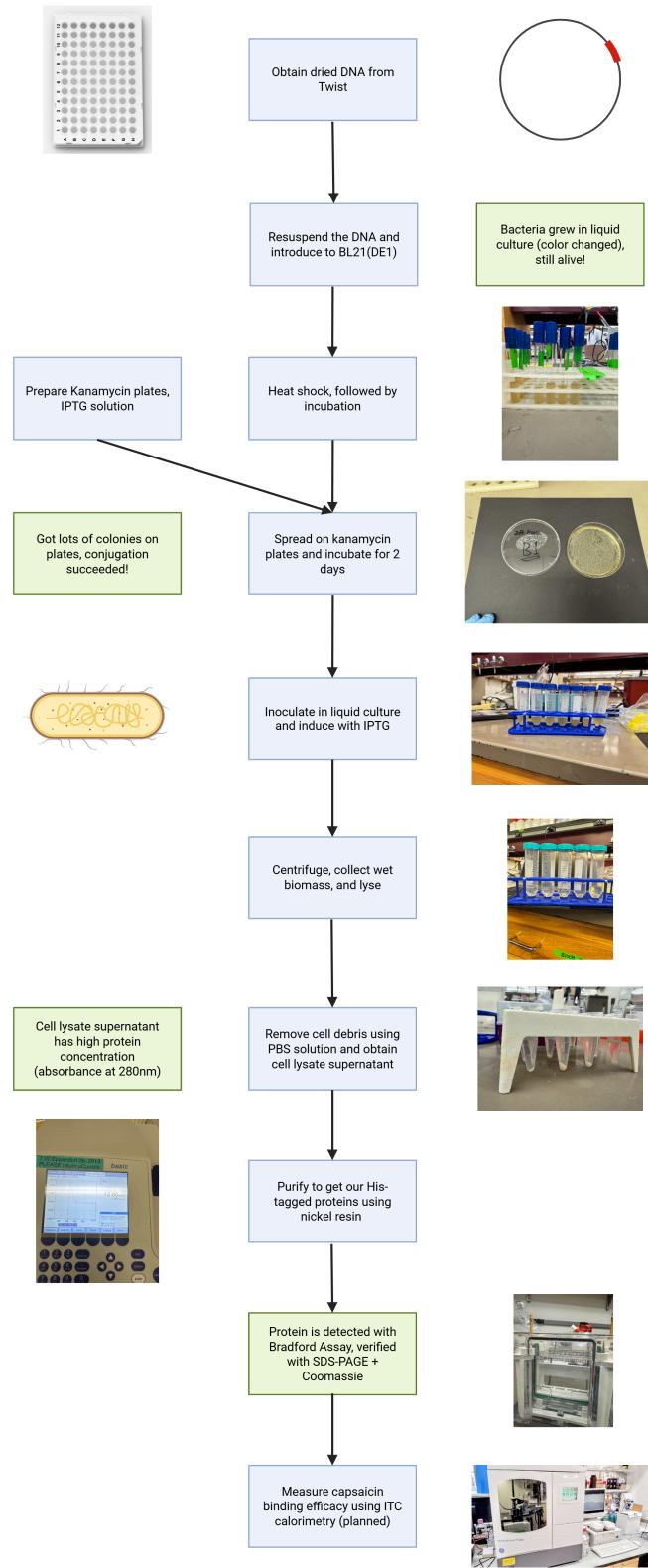
AlphaFold Casein Structure binding scores:



Boltz Casein Structure binding scores:



## 5b. Experimental Procedure

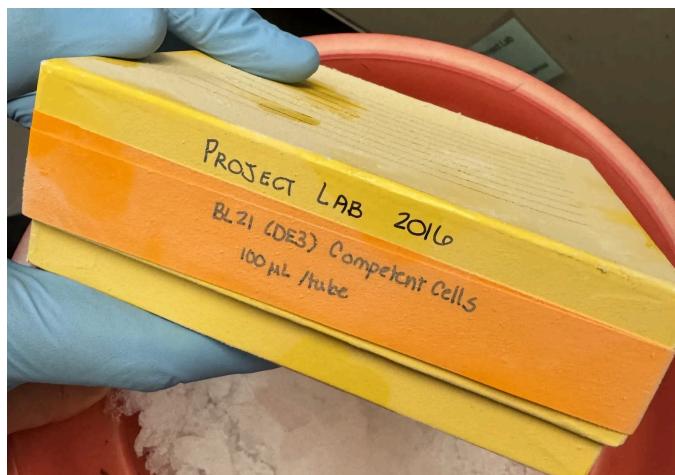


# Creating kanamycin plates

Measured, poured, and autoclaved our agar with all the necessary components.

## Finding cells

We had to look in many different fridges to find our BL21. Finally, we found a lot of leftover tubes from Project Lab 2016.



## Liquid culture and conjugation

We defrosted the old BL21, and followed the protocol from the Thought Emporium.

Transformed example



Control (untransformed)



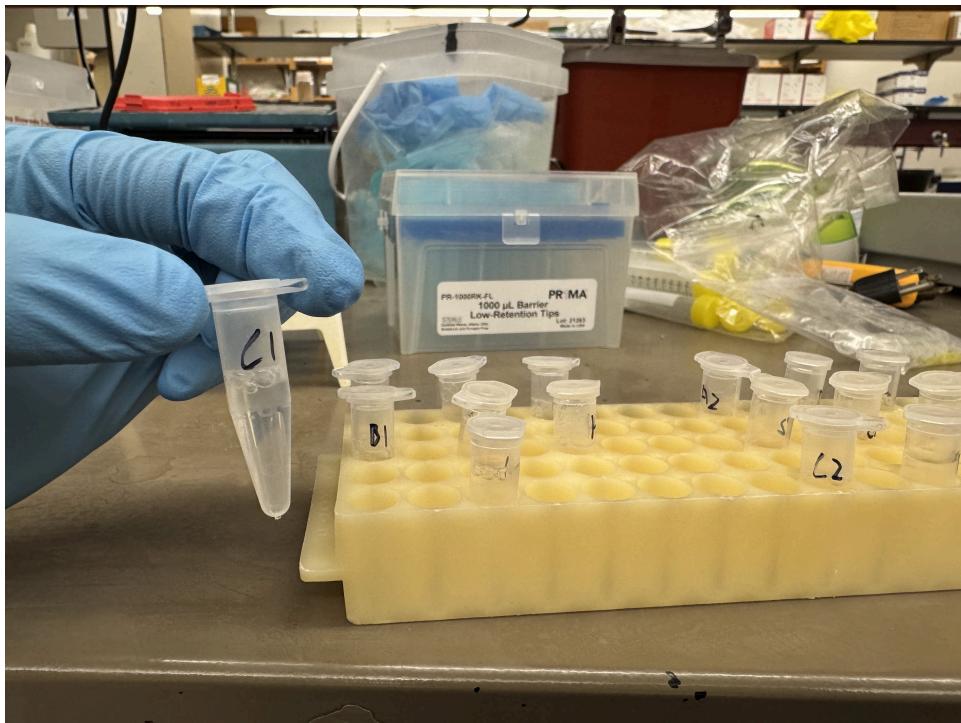
## Liquid culture and IPTG induction

We did IPTG induction following [the IPTG protocol from Protocols.io](#). We did a full overnight liquid culture of each plate, then diluted to a specific density and induced for maximum protein yield. Finally we incubated at room temperature for a full day.



## Successful cell lysis

We followed the [Thermofisher Pierce Lysis Buffer Protocol](#). We centrifuged the liquid culture in aliquots, decanted and washed the LB solution using phosphate buffered saline, and lysed with lysis buffer. After incubating on ice, we centrifuged and got our lysate supernatant!



## Purification and Protein

Finally we ran nickel resin using the [Thermofisher His-Pur Nickel Resin Protocol](#) with gravity columns and syringes. The resulting purified elute had very low

nanodrop readings, so we could only validate our results at this stage using SDS-PAGE and a Bradford assay.



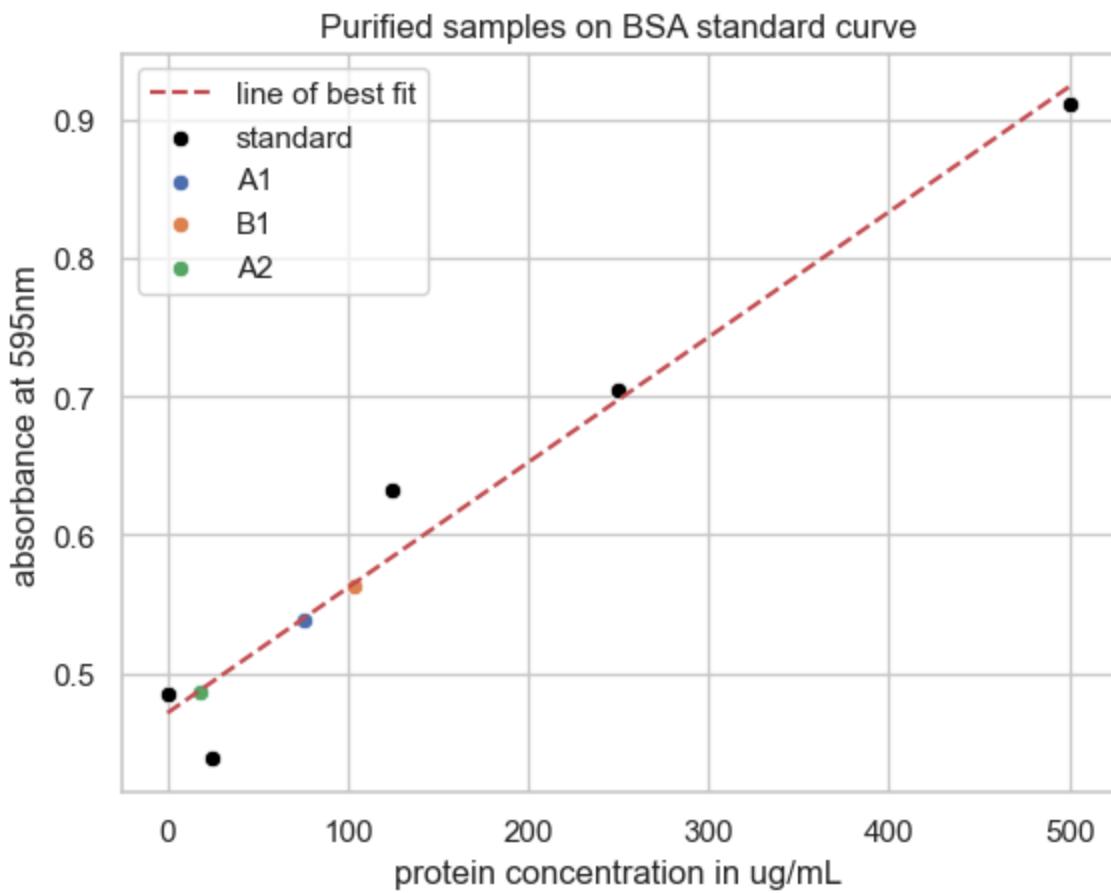
## **Challenges:**

- The largest by far was high batch size; testing 17 bacteria and proteins in parallel was challenging, and having two people work in parallel allowed us to progress significantly faster
- Low yields from cells: since we had such a large batch size, it was impossible to conduct the standard high volume procedure of protein purification. We therefore had to be careful when working with our samples to ensure we did not lose the small protein concentrations we had
- Precision concerns: The UV spectrophotometer had high variance in the small-concentration regime, due to slight variations in creation of samples and positioning of the cuvette; We solved this by not diluting until absolutely necessary to ensure measurable, high-confidence results

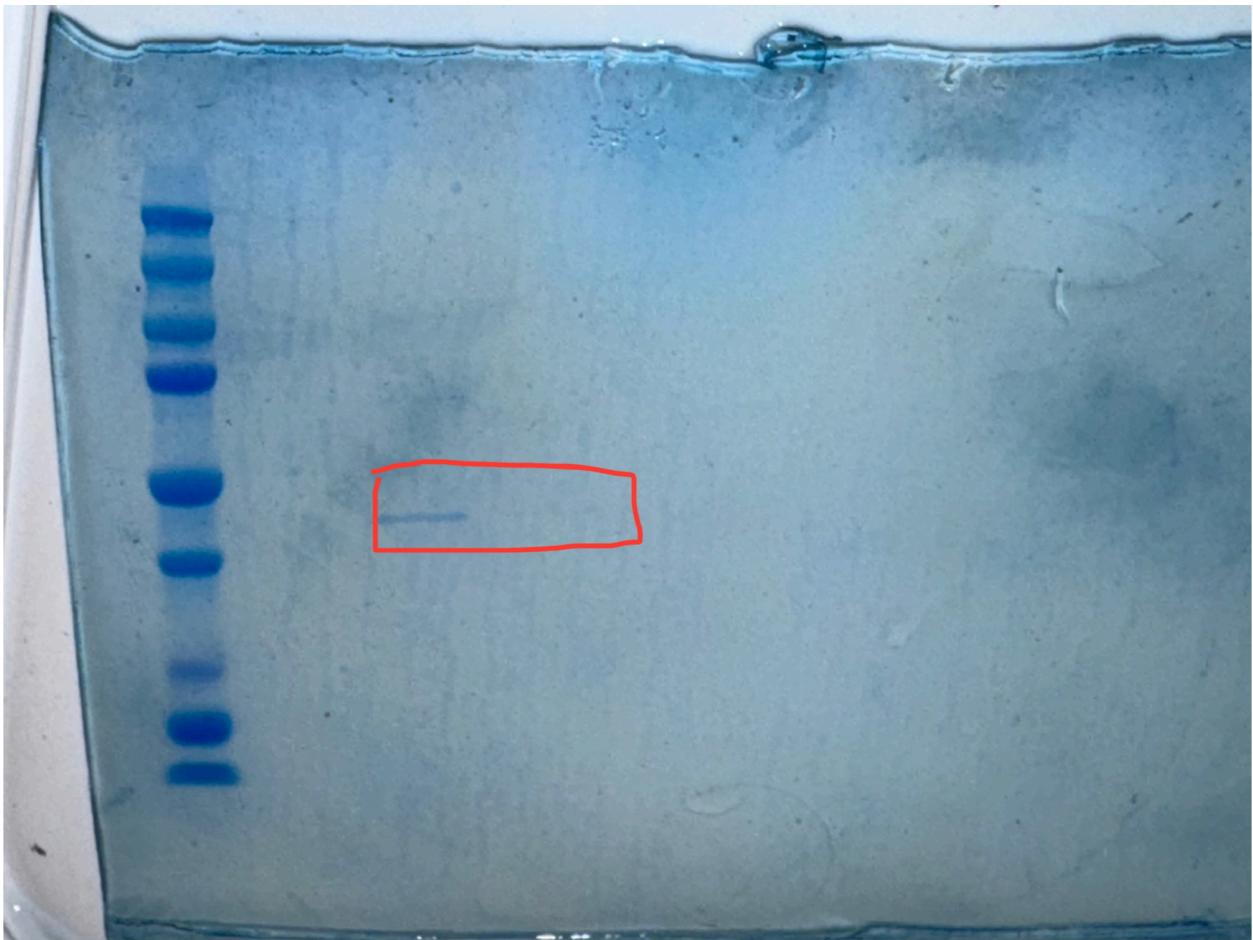
## **6b. Experimental Results**

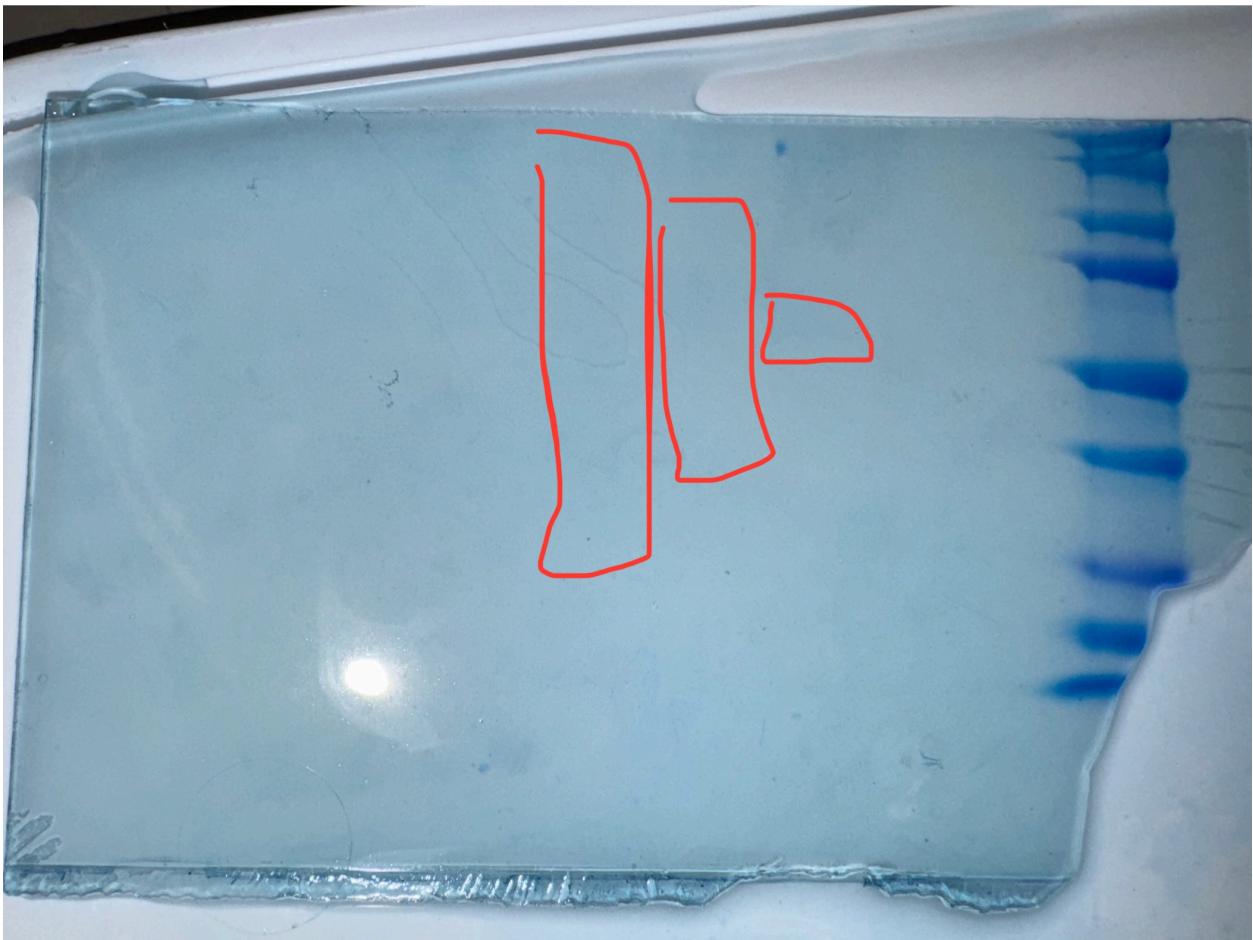
### **Bradford Assay Results**

We ran all of our purified samples on a Bradford assay. Despite getting lots of negatives, we got a few very strong positives, in the 100ug/mL range, so we're certain that we got at least one successful purification (A1!).

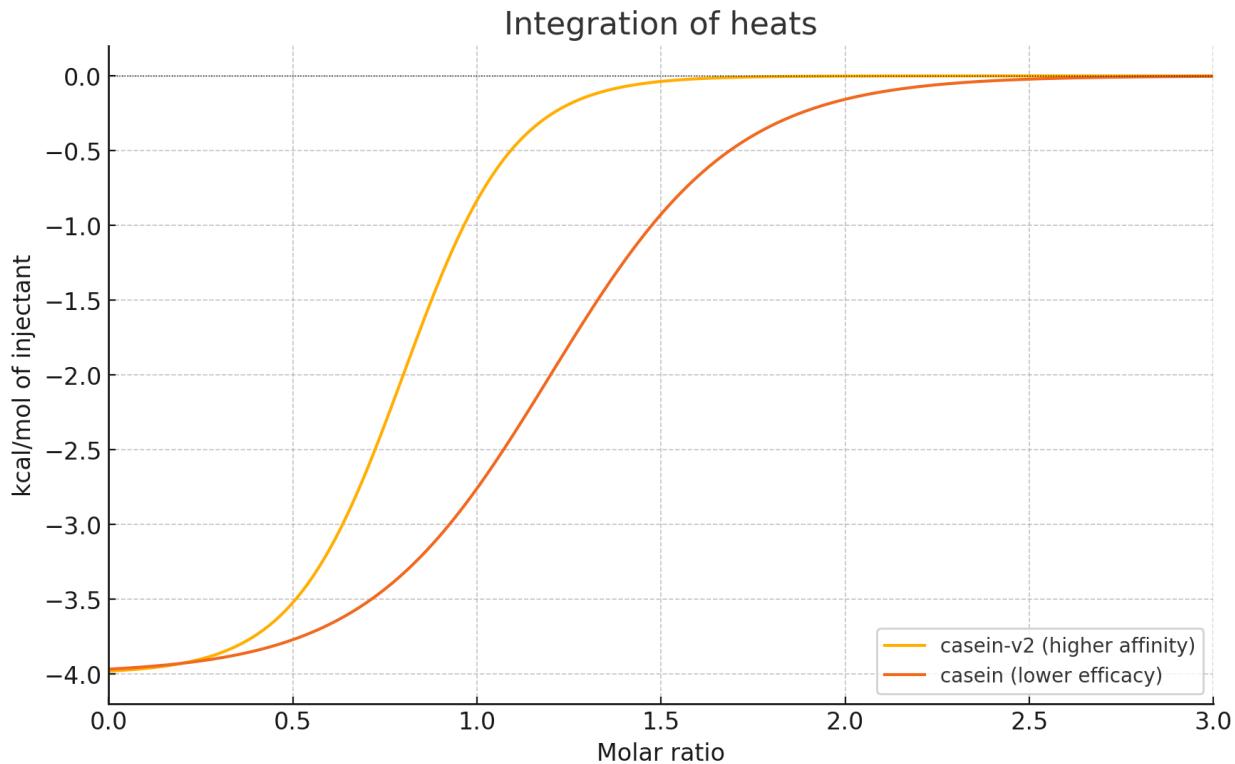


## SDS-PAGE Results





## Expected ITC results



## 7. Discussion & Future Work

In our computational work, we observed that the Boltz predicted protein structure-capsaicin complexes produced more negative binding energies, meaning it likely represents more stable binding. Therefore, we expect the proteins generated from the Boltz configuration to be more targeted towards capsaicin, as LigandMPNN would be better able to model the interactions in a stable configuration.

Both Alphafold and Boltz demonstrated significant improvements to the binding energies after the computational procedure (maximum deltas are -0.5 kJ/mol for Alphafold and -2.9 kJ/mol for Boltz). Therefore, the procedure has worked.

Our transformation was well validated, by the growth of colonies. Cell lysis was also successful. We were not able to do validation for successful IPTG induction until the very final SDS-PAGE step, which we would like to remedy in the future.

Our SDS-PAGE result was successful, giving us the band that we expected. The Bradford assay gave us surprising results, and we would run it again given more time.

Immediate future work include running higher concentration protein purification with larger liquid cultures, to do detailed ITC study and generally improve yield.

Another goal would be to redesign the process and clonal gene so that the His-tags can be removed, allowing our protein sequence to map more directly to the casein sequence.

## 8. Techniques, Tools, and Technology

### Pipetting

- Pipetting
- Lab Safety
- Bioethical Considerations  
**(must check this box)**

### DNA Gel Art

- DNA Sequencing
- DNA Editing (e.g., CRISPR)
- DNA Construct Design
- Restriction Enzyme Digestion
- Gel Electrophoresis/SDS-PAGE
- DNA Purification From Gel
- Databases (e.g., GenBank, NCBI, Ensembl, and UCSC Genome Browser)

### Opentrons

- Creating Code for Laboratory

- 
- Automation
  - PyLabRobot
  - Using Liquid Handling Robots (e.g., Opentrons)

### Protein Design

- Protein Design
- [Models and Notebooks](#)
- [Databases](#)
- [Tools](#)

### BioProduction

- BioProduction
- Chassis Selection (e.g., Dh5alpha)
- [Registry of Standard Biological Parts](#)
- [FreeGenes](#)
- Plasmid Preparation
- Bacterial Culturing
- Quality Control/Analysis
- Bacterial Processing (e.g., Centrifugation, Lysis, DNA Purification)

### Cell Free

- Cell Free Reactions
- Freeze-Dried Cell Free Systems
- [miniPCR Tools](#)

### Week 7: Gibson Assembly

- Primer Design or Selection
- PCR Reactions

- 
- Gibson Assembly
  - Other Cloning Methods (e.g., Restriction Enzyme Digestion or Gateway Cloning)

### Week 8-9: CRISPR

- CRISPR/Cas9
- Designing Prime Editing gRNA
- Creating Twist Order

## 9. Additional Info

- LigandMPNN
- tmol / Rosetta
- pytorch
- Computation resources (H100 GPUs) - \$60
- Twist clonal gene order - \$2040
- BL21 live culture - ~\$50
- Centrifuge, thermocycler, etc. for transformation
- Kanamycin agar plates
- LB broth for liquid culture
- IPTG for induction
- Nickel resin, gravity columns, imidazole for protein purification
- Protein buffer, Coomassie dye, gels for SDS-PAGE and Bradford assay