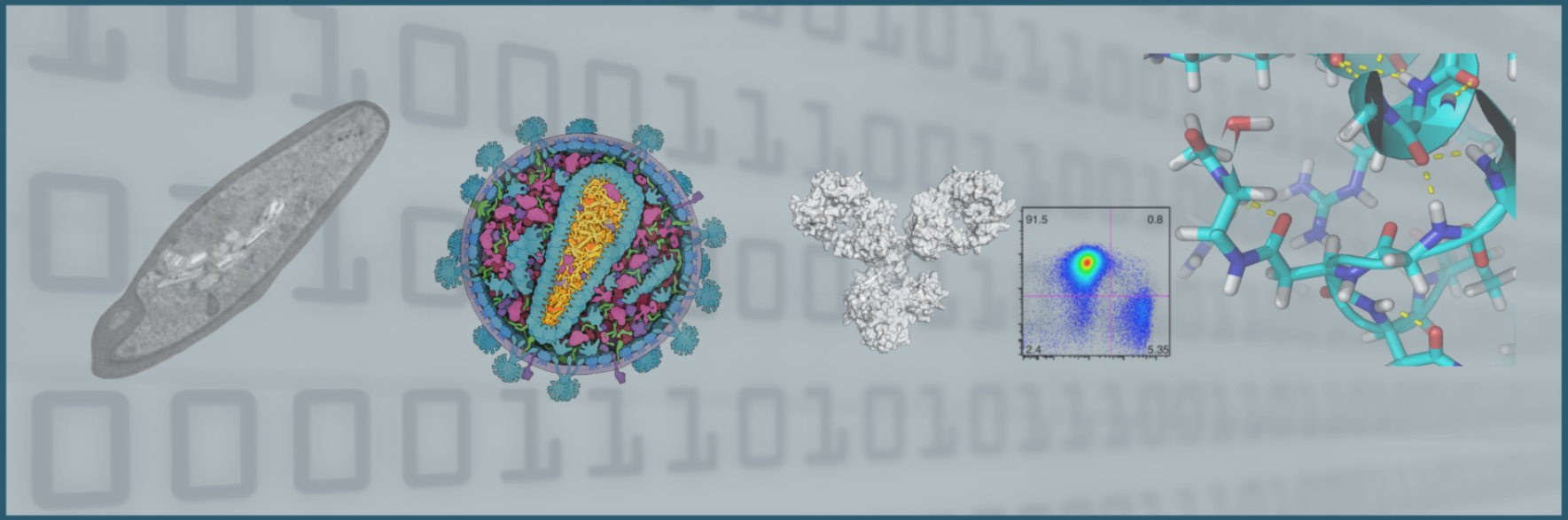
Protein Interface Analysis & Design

Dr. Sebastian Rämisch

Cambrium GmbH & Max Planck Unit for the Science of Pathogens



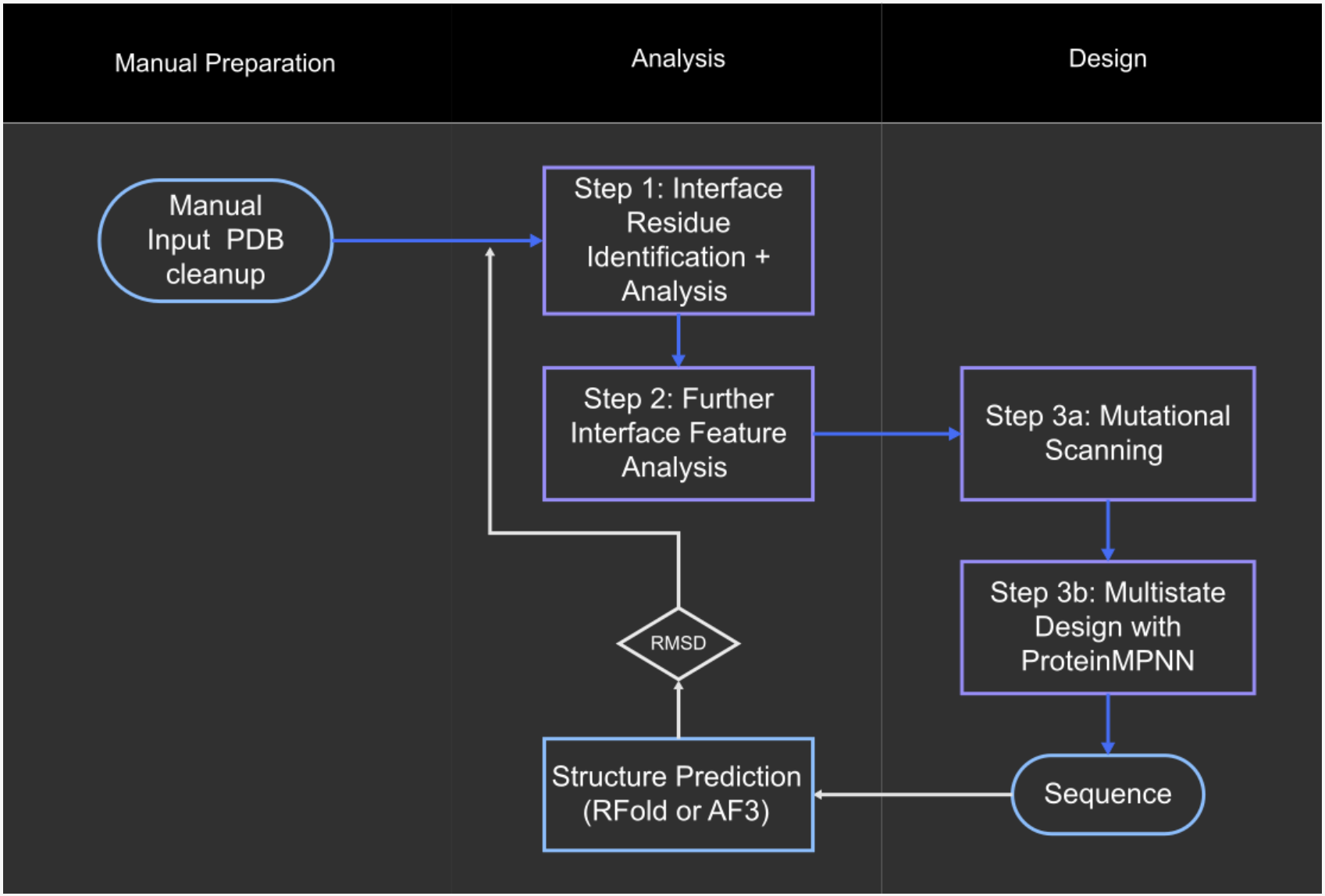
Problem Setting

- **Challenge:** Prot_A binds both Prot_B and Prot_C with overlapping interfaces.
- Use case: **SARS-CoV-2 Spike Protein (Prot_A)**, **Antibody P5A-3C8 (Prot_B)**, and **ACE2 receptor (Prot_C)**.
- The aim is to disrupt the Prot_A–Prot_B interaction while maintaining Prot_A–Prot_C affinity.

Analytical Pipeline Overview

1. Interface Residue Identification + Analysis.
   - Tool: Rosetta **InterfaceAnalyzer**.
   - Output: Interface metrics, Interface residues
2. Further Interface Feature Analysis.
   - Metrics: Hydrophobicity, charge, and shape complementarity.
3. Mutational Scanning.
   - **Identify hot-spot residues** at the interface.
   - Identify candidate point mutations.
4. Computational Design for Abrogation.
   - **ProteinMPNN** (Implemented)
   - Rosetta **Filterscan** (90% implemented)
   - Rosetta Glycosylation site design (not implemented)

Pipeline Visual Representation



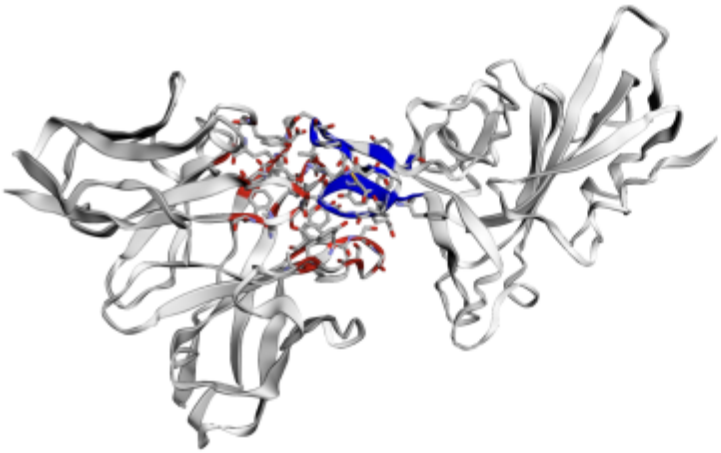
Results from Proof of Concept

	7Z0X	6M0J
dG	-51.52	-44.31
SASA	1162.95	1718.14
hyrophobic SASA	813.9	1280.17
num_res	50	73
unsats	3	3
packstat	0.0 (?)	0.0 (?)
hbond_E	-12.24	-9.55

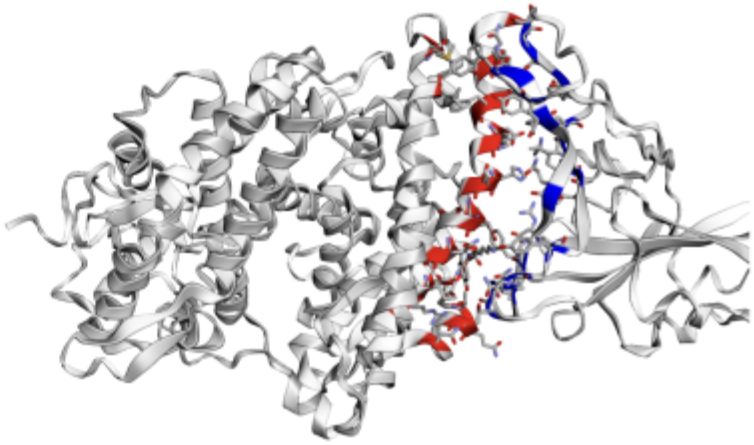
Table 1: Comparison of Protein Interfaces between 7Z0X and 6M0J.

Results from Interface Detection

Detected interface residues



SARS-CoV-2 spike - P5A-3C8



SARS-CoV-2 spike - ACE2

→ Images taken from Jupyter Notebook

Design Experiment: ProteinMPNN

- **Goal:** Disrupt the ACE2 interaction while maintaining or increasing antibody binding.
- **Steps:**
  1. Identify Spike interface residues in both complexes.
  2. Design with ProteinMPNN
     - Spike-Ab-Interface
     - Spike-ACE2-Interface
     - Combined interface on unbound Spike
  3. Get probabilities from each design run
  4. Linearly combine weighted probabilities (negative weight for the ACE2 interface design)
  5. Get the sequence(s) using softmax or argmax
  6. Structure prediction (AF3, RosettaFold2) $\rightarrow$ *not implemented*

Results from Interface design

Multistate design with ProteinMPNN

- ProteinMPNN has been used extensively for interface design
- Current implementation of creating multiple sequences from MPNN probabilities is working as intended.
- No automated processing of filterscan results implemented, yet.

Notes

- An alanine scan can be useful to identify hotspots. The filterscan includes an alanine scan, but extracting those data is not implemented, yet.
- The most straightforward way to abrogate receptor binding is to add glycosylation sites to the interface. It's not hard to do but again, due to time limitations it has not been done here.
- Predicting the complex structure after re-design is an essential step. I tried predicting the wildtype complex structures with ESM and AlphaFold2 (the only automatable tools available to me for free) - both failed to reproduce the structure. Hence, they would be a bad choice as parts of the pipeline.

Questions?

Thank you for your consideration!