DE NOVO ANTIBODY DESIGN HACKATHON

COMPLETE PARTICIPANT HANDBOOK

Scientific Background • Challenge Specifications • Submission Guidelines • Evaluation Criteria

Engineering Next-Generation Immune Checkpoint Therapeutics

PD-1/PD-L1 Pathway Targeting

11-14 December 2025

The Central Question:

"Can we computationally design antibodies that don't just bind, but that look like real drugs?"

Table of Contents

1. Executive Overview

1.1 Competition Philosophy

Traditional antibody development takes 18-36 months from target identification to clinical candidate nomination, with 90%+ attrition rates. This hackathon challenges participants to demonstrate that computational approaches can compress this timeline while simultaneously improving success rates by designing molecules that are not just structurally plausible binders, but genuine drug candidates.

The key insight: *A computationally designed antibody that binds perfectly but aggregates in solution, triggers immunogenic responses, or contains chemical liabilities is worthless for therapeutic development. Real drugs must satisfy multiple constraints simultaneously.*

1.2 Target System: PD-1/PD-L1 Immune Checkpoint

The PD-1/PD-L1 axis represents one of the most validated and clinically successful targets in immuno-oncology. Pembrolizumab (Keytruda®), a humanized IgG4 anti-PD-1 antibody, has achieved over **$25 billion** in annual sales and revolutionized cancer treatment across multiple indications.

Reference Structure: PDB 5GGS (Pembrolizumab Fab in complex with PD-1 extracellular domain)

Target Affinity: Kd ~29 pM (pembrolizumab)

1.3 Competition Structure

Participants compete in two challenges with escalating difficulty:

	Challenge 1: Remix Keytruda	Challenge 2: Invent the Future
Difficulty	★★★☆☆ (Moderate)	★★★★★ (Expert)
Starting Point	PDB 5GGS (Keytruda + PD-1)	Blank canvas
Objective	Redesign heavy-chain CDRs while maintaining binding	Design complete antibody de novo
Key Metric	DockQ + ipSAE	ipSAE (no reference)
Max Score	100 points	100 points

Total Competition Score: Challenge 1 (100) + Challenge 2 (100) + Presentation (50) = **Maximum 250 points**

2. Scientific Background

2.1 The PD-1/PD-L1 Checkpoint Axis

PD-1 (Programmed cell death protein 1) is an inhibitory receptor expressed on activated T cells, B cells, and myeloid cells. Its primary ligand, PD-L1 (B7-H1), is frequently upregulated on tumor cells as an immune evasion mechanism.

Mechanism of Action:

1. Tumor cells upregulate PD-L1 expression in response to inflammatory cytokines (particularly IFN-γ)
2. PD-L1 engages PD-1 on tumor-infiltrating T cells, delivering inhibitory signals
3. T cell exhaustion ensues: reduced proliferation, cytokine production, and cytotoxic function
4. Anti-PD-1 antibodies block this interaction, restoring T cell anti-tumor activity

2.2 Antibody Architecture Fundamentals

Antibodies are Y-shaped proteins composed of two identical heavy chains (~50 kDa each) and two identical light chains (~25 kDa each). The antigen-binding region (Fv) comprises the variable domains of both chains (VH and VL).

Region	Location	Function	Variability
CDR-H1	Heavy chain loop 1	Antigen contact, shape complementarity	Moderate
CDR-H2	Heavy chain loop 2	Antigen contact, interface area	Moderate
CDR-H3	Heavy chain loop 3	Primary specificity determinant	Very High
CDR-L1	Light chain loop 1	Antigen contact	Moderate
CDR-L2	Light chain loop 2	Limited direct contact	Low
CDR-L3	Light chain loop 3	Antigen contact, VH/VL interface	Moderate

Key Insight: CDR-H3 Dominance

CDR-H3 is the most structurally diverse loop in the antibody repertoire. It is generated by VDJ recombination at the junction of three gene segments, creating enormous sequence and length diversity. CDR-H3 typically contributes 30-50% of the paratope surface area and is the primary determinant of antigen specificity.

2.3 Pembrolizumab (Keytruda) Reference

Pembrolizumab is a humanized IgG4κ monoclonal antibody derived from a mouse precursor through CDR grafting and framework optimization. Key structural features:

- **Affinity:** Kd ~29 pM for human PD-1
- **Epitope:** Loop of PD-1, overlapping with PD-L1 binding site
- **Mechanism:** Competitive inhibition of PD-1/PD-L1 interaction
- **Primary Reference:** PDB 5GGS (Fab + PD-1 complex at 2.0 Å resolution)

3. Challenge Specifications

3.1 Challenge 1: Remix Keytruda

Difficulty: ★★★☆☆ (Moderate) | **Max Score:** 100 points

Starting Point: PDB 5GGS — Pembrolizumab Fab bound to human PD-1 extracellular domain

Objective: Redesign the heavy-chain CDR loops (CDR-H1, CDR-H2, CDR-H3) to create a novel antibody that maintains or improves binding to PD-1 while demonstrating significant sequence novelty.

Success Criteria:

- **Sequence Novelty:** CDR-H3 identity to Keytruda must be <95% (minimum, hard cutoff)
- **Binding Preservation:** DockQ ≥ 0.23 vs. reference (higher is better, >0.80 is high quality)
- **Interface Confidence:** ipSAE ≥ 0.60 (hard cutoff)
- **Developability:** NetSolP ≥ 0.50 (hard cutoff)

3.2 Challenge 2: Invent the Future

Difficulty: ★★★★★ (Expert) | **Max Score:** 100 points

Starting Point: Blank canvas — design from scratch

Objective: Design a complete VH/VL antibody de novo that binds PD-1. You may target the same epitope as pembrolizumab, a different epitope, or design multi-epitope approaches.

Success Criteria:

- **Sequence Novelty:** CDR-H3 identity to human germline must be <95%
- **Binding Performance:** ipSAE ≥ 0.60 (hard cutoff, no DockQ reference)
- **Binding Energy:** ΔG ≤ -6 kcal/mol (hard cutoff)
- **Developability:** NetSolP ≥ 0.50 (hard cutoff)

4. Submission Guidelines

CRITICAL: Read This Section Carefully

Submissions that do not follow the exact folder structure and file naming conventions will fail automated validation and may be disqualified. One design per challenge is allowed.

4.1 Folder Structure Overview

Each team submits one master folder named exactly as their TEAM_NAME, compressed as a single ZIP file:

```
TEAM_NAME.zip containing TEAM_NAME/
```

Inside the master folder, create subfolders for each challenge:

```
TEAM_NAME/
├── TEAM_NAME_Challenge1/
├── TEAM_NAME_Challenge2/
└── pitch/
```

4.2 Challenge 1 Folder Structure

```
TEAM_NAME_Challenge1/
├── structures/
│   ├── design_X_complex.pdb
│   └── design_X_pae.json
├── sequences/
│   └── design_X.fasta
├── metrics/          (optional)
└── docs/             (optional)
```

4.2.1 Structure Files (Required)

- **design_X_complex.pdb:** 3D structure of the antibody-antigen complex
- **design_X_pae.json:** AlphaFold-style PAE (Predicted Aligned Error) file for the complex

Note: *The design_X prefix can be any identifier (e.g., design_01, design_PD1), but prefixes must match between files.*

4.2.2 Sequence Files (Required)

The FASTA file must contain exactly three sequences with these specific headers:

```
>Heavy_Chain
[heavy chain amino acid sequence]
>Light_Chain
[light chain amino acid sequence]
>Antigen
[antigen amino acid sequence]
```

Example FASTA format:

>Heavy_Chain
QVQLVQSGVEVKKPGASVKVSCKASGYTFTNYYMYWVRQAPGQGLEWMGGINPSNGGT
NFNEKFKNRVTLTTDSSTTTAYMELKSLQFDDTAVYYCARRDYRFDMGFDYWGQGTTVTV
SSASTKGPSVFPLAPSSKSTSGGTAALGCLVKDYFPEPVTVSWNSGALTSGVHTFPAVLQS
SGLYSLSSVVTVPSSSLGTQTYICNVNHKPSNTKVDKKVEPKSCDKTHHHHHH

>Light_Chain
EIVLTQSPATLSLSPGERATLSCRASKGVSTSGYSYLHWYQQKPGQAPRLLIYLASYLESGV
PARFSGSGSGTDFTLTISSLEPEDFAVYYCQHSRDLPLTFGGGTKVEIKRTVAAPSVFIFPPS
DEQLKSGTASVVCLLNNFYPREAKVQWKVDNALQSGNSQESVTEQDSKDSTYSLSSTLTLS
KADYEKHKVYACEVTHQGLSSPVTKSFNRGEC

>Antigen

DSPDRPWNPPTFSPALLVVTEGDNATFTCSFSNTSESFVLNWYRMSPSNQTDKLAAFPED
RSQPGQDSRFRVTQLPNGRDFHMSVVRARRNDSGTYLCGAISLAPKAQIKESLRAELRVTE RR

Such FASTA must always contain the three headers >Heavy_Chain, >Light_Chain, >Antigen. Remember to consider Chain A as heavy chain , Chain B as light chain and Chain C as antigen

Chain Assignment Convention

Chain A = Heavy Chain, Chain B = Light Chain, Chain C = Antigen. Use these exact header names in your FASTA files.

4.2.2 (Optional) folders for Challenge 1

Inside TEAM_NAME_Challenge1/ you may also include:

- 📁 metrics/
  - Any CSV or tables with your own internal scores.
- 📁 docs/
  - Methods, notes, PDF summaries, code README, etc. These are optional and do not affect automated scoring, but they help the judges understand your pipeline.

4.3 Challenge 2 Folder Structure

4.3.1 Identical structure to Challenge 1: Inside the same master folder create folder named : TEAM_NAME_Challenge2/

This folder also must contain structure and sequence for the design, in the same style.

1. Structure files Create a folder:

📁 TEAM_NAME_Challenge2/structures/ For the design, include:

- 📄 design_X_complex.pdb

3D structure of the antibody and antigen complex for this challenge.

- 📄 design_X_pae.json Matching AlphaFold style PAE file for the same complex.

Again, design_X can be any name, but the prefixes must match.

For example:

- structures/design_01_complex.pdb
- structures/design_01_pae.json

```
TEAM_NAME_Challenge2/
├── structures/
│   ├── design_X_complex.pdb
│   └── design_X_pae.json
├── sequences/
│   └── design_X.fasta
├── metrics/          (optional)
└── docs/             (optional)
```

4.3.2 Sequence files

Create a folder:

📁 TEAM_NAME_Challenge2/sequences/ For the design, include:

- 📄 design_X.fasta The FASTA format is the same as in Challenge 1, with three entries:

>Heavy_Chain [heavy chain sequence]

>Light_Chain [light chain sequence]

>Antigen [antigen sequence]

Each file must contain the three headers and corresponding sequences.

4.3.3 (Optional) folders for Challenge 2

Inside TEAM_NAME_Challenge2/ you may optionally include:

- 📁 metrics/
- 📁 docs/

Same idea as Challenge 1: for your own metrics and documentation

4.4 Pitch Folder (Required)

Submit your presentation explaining your approach and results:

```
TEAM_NAME/pitch/
└── TEAM_NAME_presentation.pptx
```

Presentation Time: 3 minutes maximum
Presentation Score: Up to 50 points

4.5 Final Submission Checklist

- ☐ Master folder name is exactly your TEAM_NAME
- ☐ Inside master folder: TEAM_NAME_Challenge1/ and/or TEAM_NAME_Challenge2/
- ☐ Each challenge folder contains structures/ with design_X_complex.pdb and design_X_pae.json
- ☐ Each challenge folder contains sequences/ with design_X.fasta (3 headers: >Heavy_Chain, >Light_Chain, >Antigen)
- ☐ Pitch folder contains your presentation (5 min, PPT format)
- ☐ Everything compressed into TEAM_NAME.zip
- ☐ One design only per challenge

5. Scoring Rubric & Evaluation Criteria

All submissions are evaluated through an automated pipeline followed by expert review. Raw metrics are converted to 0-10 scores using predefined bands, then weighted to produce final scores.

5.1 Score Composition

Category	Weight	Components
Binding & Structure Quality	**60%**	ipSAE, DockQ*, ΔG, Contacts, Interface pLDDT, CDR SASA
Developability	**20%**	NetSolP solubility score
Novelty	**20%**	CDR-H3 sequence identity

**DockQ is only computed for Challenge 1 (vs. Keytruda reference)*

Final Score Formula

Final Score (0-100) = `[ 0.60 × Binding_struct + 0.20 × Developability + 0.20 × Novelty ] × 10`

5.2 Metric Bands and Thresholds

Each metric is converted to a 0-10 score based on the following bands. Designs failing any minimum requirement are marked non-viable.

Metric	Good (9-10)	Medium (6-8)	Poor (0-5)	Min Required	Challenge
BINDING & INTERFACE METRICS					
ipSAE	**≥ 0.80**	**0.60 – 0.80**	**< 0.60**	**≥ 0.60**	Both
DockQ	**≥ 0.80**	**0.49 – 0.80**	**< 0.49**	**≥ 0.23**	Ch 1 only
ΔG (kcal/mol)	**≤ -12**	**-10 to -12**	**> -10**	**≤ -6**	Both
Contacts	**> 25**	**15 – 25**	**< 15**	**≥ 10**	Both
Interface pLDDT	**> 80**	**70 – 80**	**< 70**	**≥ 65**	Both
CDR SASA (Å²)	**> 600**	**300 – 600**	**< 300**	**> 250**	Both
DEVELOPABILITY METRIC					
NetSolP	**≥ 0.70**	**0.50 – 0.70**	**< 0.50**	**≥ 0.50**	Both
NOVELTY METRIC					
CDR-H3 Identity	**< 70%**	**70 – 90%**	**> 90%**	**< 95%**	Both

6. Detailed Metric Definitions

6.1 Binding & Interface Metrics

6.1.1 ipSAE (interaction prediction Score from Aligned Errors)

What it measures: ipSAE uses the AlphaFold PAE (Predicted Aligned Error) matrix to quantify how self-consistent an interface is. Values closer to 1 mean the interface geometry is well-supported by the predicted errors and more likely to represent a real binding interaction.

Computation: The pipeline runs ipSAE on design_X_pae.json and design_X_complex.pdb, reads rows with Type = max for antibody vs antigen chains, and takes the best ipSAE among them.

Why it matters: ipSAE directly reflects the confidence of the structure prediction at the binding interface. Low ipSAE indicates the model is uncertain about the interface geometry, suggesting the binding pose may be unreliable.

6.1.2 DockQ (Challenge 1 Only)

What it measures: DockQ compares your complex to the reference Keytruda-PD-1 structure (PDB 5GGS) and returns a docking quality score between 0 and 1 that corresponds to CAPRI (Critical Assessment of PRedicted Interactions) quality classes.

CAPRI Classes: High quality (≥0.80), Medium (0.49-0.80), Acceptable (0.23-0.49), Incorrect (<0.23)

Why it matters: For Challenge 1, you're redesigning from a known template. DockQ ensures your redesigned antibody still adopts a similar binding pose to the original, validating that your modifications preserve the binding mode.

6.1.3 Binding Free Energy (ΔG)

What it measures: ΔG is the predicted binding free energy in kcal/mol, related to the dissociation constant Kd. More negative values indicate stronger predicted binding affinity.

Computation: PRODIGY is run on design_X_complex.pdb to predict binding affinity based on interface contacts.

Why it matters: Therapeutic antibodies typically require nanomolar to picomolar affinity. ΔG provides a physics-based estimate of binding strength independent of the structure prediction confidence.

6.1.4 Interface Contacts

What it measures: The number of intermolecular contacts between antibody and antigen atoms at the interface.

Why it matters: More contacts generally correlate with larger buried surface area and stronger binding. Too few contacts suggest a weak or non-specific interaction.

6.1.5 Interface pLDDT

What it measures: Average pLDDT (predicted Local Distance Difference Test) score for residues at the binding interface. pLDDT is AlphaFold's per-residue confidence metric.

Why it matters: High interface pLDDT indicates the structure prediction is confident about the local geometry of the binding site. Low values suggest disorder or uncertainty that may indicate an unrealistic binding pose.

6.1.6 CDR SASA

What it measures: Total Solvent Accessible Surface Area (SASA) of the CDR loops (paratope) on the antibody, measured in Ų.

Computation: CDR residues are identified using IMGT-style numbering, and SASA is computed for those atoms using a standard surface algorithm.

Why it matters: Larger CDR SASA indicates a more exposed and potentially more functional binding surface. Buried or inaccessible CDRs cannot effectively engage the antigen.

6.2 Developability Metric

6.2.1 NetSolP Solubility Score

What it measures: NetSolP estimates the probability that the Fv (VH + VL) region will express and remain soluble in solution, based purely on sequence features.

Computation: NetSolP is run on the Fv sequence extracted from design_X.fasta. Per-chain scores are combined into a single value per design.

Why it matters: Antibodies that aggregate or precipitate during manufacturing are non-viable as therapeutics. Low solubility leads to reduced yields, formulation challenges, and potential immunogenicity from aggregates. This is a primary developability filter in industry.

> **Industry Perspective on Solubility**
>
> In real-world antibody development, solubility issues cause ~30% of candidate failures during CMC (Chemistry, Manufacturing, and Controls) development. NetSolP provides an early computational filter to identify sequences likely to have expression or aggregation problems before committing to expensive experimental validation.

6.3 Novelty Metric

6.3.1 CDR-H3 Sequence Identity

What it measures: How similar your heavy-chain CDR3 sequence is to reference sequences:

- **Challenge 1:** Compared to Keytruda's CDR-H3
- **Challenge 2:** Compared to human germline CDR sequences

Computation: CDR-H3 positions are identified using IMGT-like patterns, aligned to reference sequences, and percent identity is calculated.

Why it matters: The hackathon rewards genuine computational design, not trivial modifications. Lower identity indicates more novel designs that demonstrate the power of de novo methods.

7. Automated Evaluation Pipeline

All submissions pass through a standardized automated evaluation pipeline. Understanding this pipeline will help you optimize your designs and avoid common pitfalls.

7.1 Pipeline Overview

1. **File Validation:** Reads design_X_complex.pdb, design_X_pae.json, and design_X.fasta. Missing files = disqualification.
2. **Binding & Interface Metrics:** Runs PRODIGY (ΔG, contacts), ipSAE (interface confidence, pLDDT), CDR SASA computation, and DockQ (Challenge 1 only).
3. **Developability Assessment:** Runs NetSolP on the VH+VL sequence to estimate solubility probability.
4. **Novelty Calculation:** Extracts CDR-H3 from FASTA, compares to reference sequences (Keytruda for Ch1, germline for Ch2).
5. **Score Mapping:** Maps each raw metric to 0-10 score using predefined bands (Good/Medium/Poor).
6. **Category Aggregation:** Averages metric scores within each category (Binding/Structure, Developability, Novelty).
7. **Final Score:** Applies weights (60/20/20) and scales to 0-100.
8. **Viability Check:** Designs failing ANY minimum threshold are marked non-viable and ranked below all viable designs.

7.2 Hard Cutoffs (Minimum Requirements)

Failing ANY of these thresholds marks your design as non-viable:

Metric	Minimum Value	Consequence of Failure
ipSAE	≥ 0.60	Interface not confident → non-viable
DockQ (Ch 1 only)	≥ 0.23	CAPRI "Incorrect" → non-viable
ΔG	≤ -6 kcal/mol	Weak predicted binding → non-viable
Contacts	≥ 10	Insufficient interface → non-viable
Interface pLDDT	≥ 65	Low confidence structure → non-viable
CDR SASA	> 250 Ų	CDRs not accessible → non-viable
NetSolP	≥ 0.50	Poor solubility → non-viable
CDR-H3 Identity	< 95%	Trivial clone → non-viable

7.3 Final Scoring

Component	Points
Challenge 1 Score	0 – 100
Challenge 2 Score	0 – 100
Presentation Score	0 – 50
Total Hackathon Score	**0 – 250**

8. Recommended Computational Toolkit

The following tools are recommended for the various stages of antibody design. While you're free to use any computational approach, these tools have been validated for the evaluation pipeline.

8.1 Structure Prediction

Tool	Use Case	Access
AlphaFold-Multimer	Complex structure prediction (required for PAE)	ColabFold, LocalColabFold
AlphaFold3	Latest generation structure prediction	AlphaFold Server
IgFold	Antibody-specific folding	GitHub, Hugging Face
ABodyBuilder2	Antibody structure modeling	SAbDab web server
ESMFold	Fast single-chain prediction	ESM Metagenomic Atlas

8.2 Generative Design

Tool	Description
ProteinMPNN	Sequence design for fixed backbone structures. Excellent for redesigning CDRs on existing scaffolds.
LigandMPNN	Interface-aware sequence design. Can condition on binding partners.
RFdiffusion	De novo backbone generation. Can generate antibody-like structures conditioned on target.
ESM-IF	Inverse folding model for sequence design from structure.
AbLang/AntiBERTy	Antibody-specific language models for sequence generation and scoring.

8.3 Binding Evaluation

Tool	Description	Used in Pipeline
ipSAE	Interface confidence from PAE matrix	**Yes**
DockQ	Docking quality vs. reference (CAPRI metric)	**Ch 1 only**
PRODIGY	Binding affinity prediction (ΔG, contacts)	**Yes**
FoldX	ΔΔG mutation scanning for optimization	No (optional)
Rosetta Interface	Physics-based interface analysis	No (optional)

8.4 Developability Assessment

Tool	Description	Used in Pipeline
NetSolP	Solubility prediction from sequence	**Yes**
CamSol	Intrinsic solubility prediction	No (optional)
TANGO/Aggrescan	Aggregation propensity analysis	No (optional)

TAP	Therapeutic Antibody Profiler (comprehensive)	No (optional)
IgBLAST/ANARCI	Germline assignment and humanization scoring	No (optional)

9. Beyond NetSolP: Comprehensive Developability

While NetSolP is the official scoring metric, real-world antibody developability encompasses multiple factors. Understanding these will help you design better candidates.

9.1 The Six Pillars of Antibody Developability

Pillar 1: Solubility

Primary Metric: NetSolP ≥ 0.50 (scored), CamSol > 1.0 (recommended)

Low solubility limits dose delivery, causes manufacturing problems, and creates immunogenic aggregates that can trigger anti-drug antibody responses.

Pillar 2: Aggregation Propensity

Tools: TANGO (< 500 total score), Aggrescan, SAP (spatial aggregation propensity)

Aggregates reduce activity, clog filters during manufacturing, and trigger immune responses. Even soluble proteins can have aggregation-prone regions that cause problems at high concentrations.

Pillar 3: Immunogenicity

Tools: NetMHCIIpan (MHC-II binding), IEDB (T-cell epitopes), Humanization score > 85%

Anti-drug antibodies (ADAs) can neutralize therapeutic efficacy, cause hypersensitivity reactions, and require treatment discontinuation. Higher humanization reduces immunogenicity risk.

Pillar 4: Chemical Stability

Key Liabilities to Avoid:

- **Deamidation:** NG, NS, NT, NH, ND motifs → Consider N→Q mutations
- **Isomerization:** DG, DS, DT, DD, DH motifs → Consider D→E mutations
- **Oxidation:** Exposed M, W, H, C residues → Bury or M→L mutations
- **N-glycosylation:** N-X-S/T (X≠P) in Fv → N→Q or S→A mutations

Pillar 5: Humanization

Tools: IgBLAST (germline assignment), IMGT/DomainGapAlign, T20 score > 80

Preferred Frameworks: IGHV1-69, IGHV3-23, IGHV4-59, IGKV1-39, IGLV1-51

Pillar 6: Manufacturability

Key Metrics: Expression titer > 1 g/L, Thermal stability (Tm > 65°C), pI 6-9, Viscosity < 20 cP at 150 mg/mL

9.2 Quick Developability Checklist

While not all scored, consider these factors during design:

- ☐ NetSolP ≥ 0.50 (REQUIRED for scoring)
- ☐ CamSol > 1.0 (recommended)
- ☐ No high TANGO aggregation regions (< 500)
- ☐ Minimal predicted MHC-II epitopes
- ☐ No NG/DG deamidation/isomerization motifs in CDRs
- ☐ No N-glycosylation sequons (N-X-S/T) in Fv region
- ☐ No exposed methionine residues in CDRs
- ☐ Humanization score > 85% to germline
- ☐ Using preferred germline framework families
- ☐ Predicted pI between 6-9

10. Resources & References

10.1 Essential PDB Structures

PDB ID	Description	Resolution
5GGS	Pembrolizumab Fab + PD-1 (PRIMARY REFERENCE)	2.0 Å
5DK3	Pembrolizumab Fab alone	2.3 Å
5IUS	PD-1/PD-L1 complex (for epitope context)	2.45 Å
5WT9	Nivolumab-PD-1 complex (alternative anti-PD-1)	2.4 Å

10.2 Key Publications

Structure & Mechanism

- Tan et al. (2017) "Structural basis of pembrolizumab recognition of PD-1" Nature Communications
- Na et al. (2017) "Structural basis of PD-1 checkpoint blockade" Cell Research

Computational Methods

- Dauparas et al. (2022) "Robust deep learning–based protein sequence design using ProteinMPNN" Science
- Watson et al. (2023) "De novo design of protein structure and function with RFdiffusion" Nature
- Ruffolo et al. (2023) "Fast, accurate antibody structure prediction from deep learning" Nature

Developability

- Raybould et al. (2019) "Five computational developability guidelines for therapeutic antibody profiling" PNAS
- Jain et al. (2017) "Biophysical properties of the clinical-stage antibody landscape" PNAS

10.3 Online Resources

- **SAbDab:** opig.stats.ox.ac.uk/webapps/sabdab-sabpred
- **IMGT:** imgt.org
- **ColabFold:** github.com/sokrypton/ColabFold
- **ProteinMPNN:** github.com/dauparas/ProteinMPNN
- **RFdiffusion:** github.com/RosettaCommons/RFdiffusion

11. Final Notes

11.1 Code of Conduct

Scientific integrity is paramount. All submissions must represent original work by your team. Plagiarism, fabrication of results, or submission of designs from prior published work will result in immediate disqualification.

11.2 Intellectual Property

Participants retain ownership of their designs. By submitting, you grant organizers a non-exclusive license for evaluation, presentation of results, and publication of competition outcomes.

11.3 Support Channels

- **Please contact : +91- 9910450995, +91-8550077706, +91- 8919049271**
- **Email:** Contact the organizing committee for technical issues or clarifications at
  - yashasdevasurmutt@locksmithbio.org
  - vaibhav@diagnobacs.com
  - ajiteshlunge@locksmithbio.org
- **WhatsApp Channel : Bio-hackathon (IBAB EC × Locksmith Bio)**

11.4 Final Checklist Before Submission

1. Verify all required files are present with correct naming conventions
2. Ensure FASTA headers are exactly: >Heavy_Chain, >Light_Chain, >Antigen
3. Confirm PDB and PAE files have matching prefixes (design_X_complex.pdb ↔ design_X_pae.json)
4. Run your own ipSAE and NetSolP checks to ensure you pass minimum thresholds
5. For Challenge 1: Verify DockQ ≥ 0.23 against PDB 5GGS
6. Prepare your 3-minute presentation in the pitch folder
7. Compress everything into TEAM_NAME.zip
8. Submit before the deadline: **14th December 2025,12pm noon IST.**

Good luck, and may your designs fold well!

— The Organizing Committee