MeloDDE: Melo Drug Design Engine for Multi-Modality Binder Design

Atomelody Team
https://www.atomelody.com

November 14, 2026

Abstract

Computational drug-molecule design must advance experimentally testable molecules across modalities and identify assay-worthy candidates efficiently; public evidence remains concentrated on protein binders, while peptide and single-pipeline multi-modality validation are scarce, and generation / screening efficiency is likewise limiting. We introduce Melo Drug Design Engine (MeloDDE) for head-to-tail macrocycles, bicyclic peptides, linear peptides, nanobodies, and miniproteins: an LLM-based Design Blueprint Engine sets physicochemical constraints, Melo-D Alpha designs candidates, and Melo-1 selects them by all-atom folding and multi-metric screening. In wet-lab studies, 14 target–modality campaigns on 12 antigens yield binding or cell-functional readouts across all five formats (micromolar to nanomolar), including parallel nanobody / miniprotein hits; on independent *in silico* benchmarks, Melo-D Alpha success and sample efficiency in miniprotein and nanobody design, as well as Melo-1 virtual screening, outperform the base model and strong public baselines. Together, MeloDDE integrates constraint setting, generation, and evaluation into end-to-end results for multi-modality validation and efficient triage.

1 Introduction

As AlphaFold 3-like models can now produce plausible structural predictions for many complexes [1], the primary question in computational drug-molecule design is whether experimentally testable molecules can be advanced across multiple therapeutic modalities, not whether reasonable *in silico* poses can be drawn. Recent systems such as RFdiffusion [2], AlphaProteo [3], BoltzGen [4], Latent-X / Latent-X2 [5, 6], Chai-2 [7], and JAM / JAM-2 [8, 9] have advanced design across multiple formats and reported wet-lab hits. Yet the maturity of computational design still differs markedly across modalities: miniproteins are relatively mature, whereas nanobodies and macrocyclic peptides remain more challenging. In the Chai-2 report, for example, miniprotein wet-lab hit rates are about 68% (75/111 designs) while zero-shot antibody and nanobody hit rates are about 16%; these are separate modality-specific studies with distinct denominators [7]. Overall, public wet-lab evidence remains more concentrated on protein-binder pathways such as miniproteins, antibodies, and nanobodies, whereas systematic validation for peptide modalities is still scarce. More critically, multi-modality results that can be compared under one design pipeline remain limited.

Beyond cross-modality validation, the next bottleneck is efficiency. Even when candidates can be generated, synthesis and assay throughput remain limited; what matters is not blindly scaling compute or inflating generation volume, but concentrating limited budgets on molecules that truly merit assay through higher computational efficiency. This requirement has at least two sides: the generator's sample efficiency (reaching usable candidates with fewer designs), and the evaluator's ranking and enrichment ability (whether true binders can be pushed to the front of the queue). Although existing systems already report hits across formats, comparing generation efficiency under a shared protocol and compressing assay queues by confidence-based screening remain a critical bottleneck on the path from "can design" to "can advance at scale."

We therefore introduce Melo Drug Design Engine (MeloDDE) (Figure 1), aiming to advance multi-modality validation and to generate and triage candidates with higher efficiency within one pipeline. An LLM-based Design Blueprint Engine translates biological intent into physicochemical constraints. Under these constraints, Melo-D Alpha designs head-to-tail macrocyclic peptides, bicyclic peptides, linear peptides, nanobodies, and miniproteins, and Melo-1 [10] selects high-quality candidates via all-atom folding and multi-metric screening. MeloDDE also supports hotspot and geometric conditioning, cyclization modeling (head-to-tail, bicycle, and disulfide), and chemical modifications with unnatural amino acids (Figure 1b).

The main contributions are as follows:

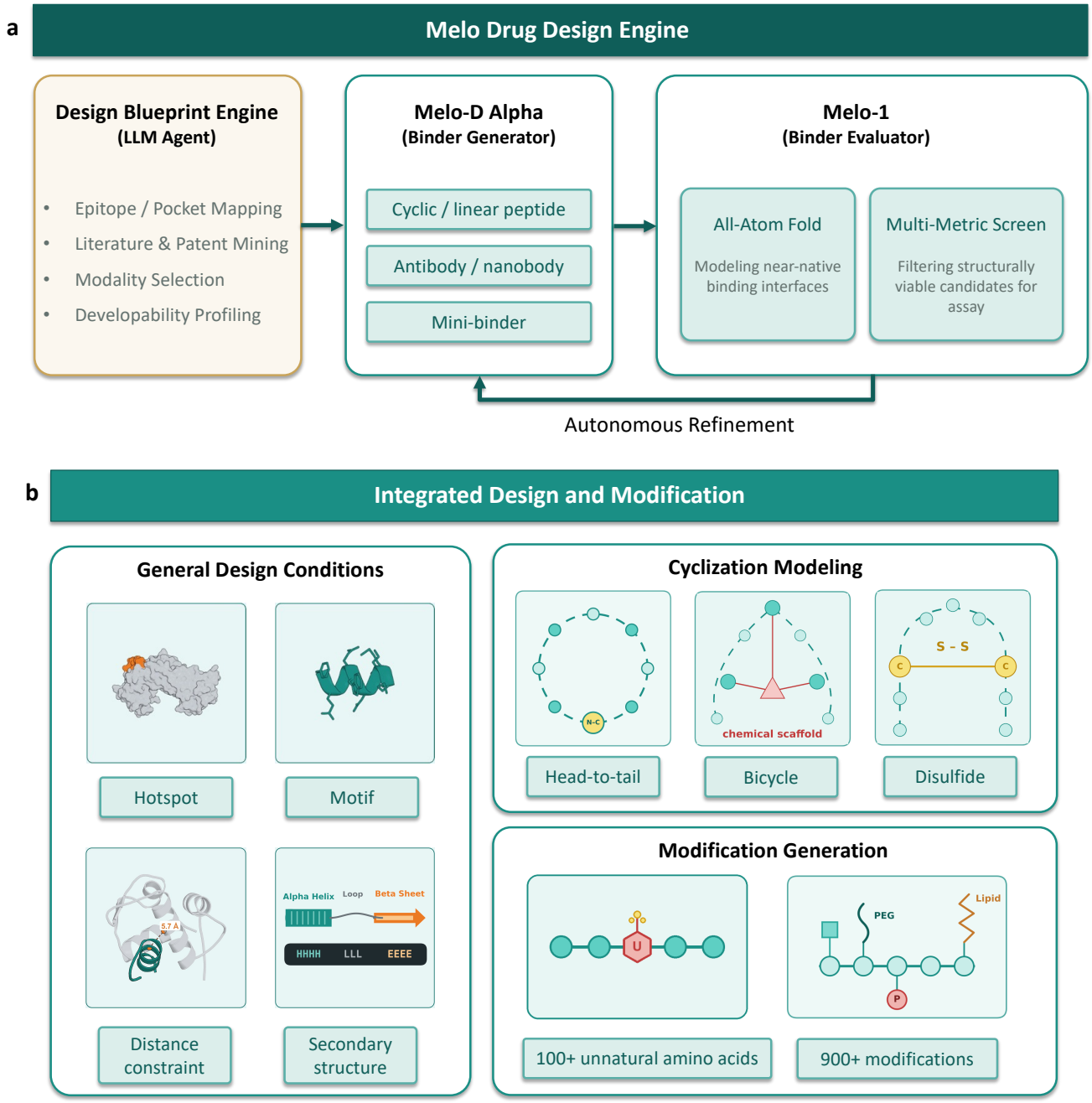

Figure 1: Melo Drug Design Engine (MeloDDE). (a) Pipeline: an LLM-based Design Blueprint Engine sets physicochemical constraints; Melo-D Alpha generates cyclic / linear peptides, nanobodies, and miniproteins; Melo-1 evaluates candidates by all-atom folding and multi-metric screening. Blueprint details are in Section 4. (b) Design conditioning and modifications: general design conditioning (hotspots, motifs, distance constraints, secondary structure); cyclization modeling (head-to-tail, bicycle, disulfide); and modification generation (over 900 modifications and over 100 unnatural amino acids).

1. Demonstrated generalizability across five therapeutic modalities. Fourteen target–modality wet-lab studies across twelve unique antigens cover head-to-tail macrocyclic peptides, bicyclic peptides, linear peptides, nanobodies, and miniproteins, with clear affinity or cell-potency readouts from the micromolar to the nanomolar range, including parallel nanobody and miniprotein hits on the same antigens.
2. Higher computational sample efficiency and virtual screening. On independent *in silico* benchmarks, Melo-D Alpha achieves higher success rates and sample efficiency than its base model and strong public baselines in miniprotein and nanobody design; Melo-1 virtual screening likewise outperforms comparators and enriches candidates with true binding activity.
3. An end-to-end computational design engine that supports the above capabilities. MeloDDE integrates a Design Blueprint Engine, Melo-D Alpha (binder generator), and Melo-1 (binder evaluator) into a closed loop from target insight to assay-ready candidate advancement, enabling multi-modality validation and efficient triage.

2 Broad Generalizability Across Therapeutic Modalities

2.1 Validation Overview

This section reports multi-modality wet-lab validation of Melo Drug Design Engine (MeloDDE) across five binder formats: head-to-tail macrocyclic peptides, bicyclic peptides, linear peptides, nanobodies, and miniproteins. Some wet-lab studies used intermediate design–screening versions that may differ from the Melo-1 / Melo-D Alpha configuration used in the *in silico* benchmark. With fewer than 30 designs tested per target, MeloDDE completed 14 target–modality wet-lab studies (Figure 2; target list in Table A1), yielding clear binding or cell-functional readouts in all five formats, including parallel nanobody and miniprotein hits on VEGF-A and PD-L1. Per-complex affinities / potencies and assay modalities are labeled in Figure 2; methods and fitted constants are in Appendix A. The remainder of this section proceeds along the peptide axis (Section 2.2), nanobodies and miniproteins (Section 2.3), and novelty of hit sequences (Section 2.4).

2.2 Peptides: Head-to-Tail Macrocycles, Bicycles, and Linear Peptides

Peptides are a critical test of MeloDDE's cross-modality reach: relative to protein binders, high-quality public wet-lab evidence for computationally designed peptides remains scarce. This subsection therefore asks both whether head-to-tail macrocyclic and bicyclic peptides can deliver measurable binding on challenging protein interfaces and tumor membrane antigens, and whether linear peptides can drive cell function at metabolism-related GPCRs. The cyclic-peptide work leverages MeloDDE's native cyclization modeling (Figure 1b) to cover head-to-tail and bicyclic scaffolds within a single pipeline.

Head-to-tail macrocyclic peptides MDM2–p53 is a canonical helix-in-groove PPI; KRAS G12D presents a flat, pocket-poor surface on a mutant oncogene. Designs were cyclized as head-to-tail lactams (sequences in Table A5). In SPR, each target was immobilized and the macrocycle injected as analyte; multi-point titrations were fittable (Figure 2a,c). The same cyclization chemistry therefore yields multi-point binding on both a groove-type and a flat interface.

Bicyclic peptides Prostate-specific membrane antigen (PSMA) and fibroblast activation protein (FAP) are tumor-associated membrane targets. Designs were constrained as bicycles; multi-point SPR titrations were fittable (Figure 2a). Bicyclic peptides therefore yield measurable binding on both membrane antigens, providing a second cyclic-peptide evidence path alongside head-to-tail macrocycles.

Linear peptides In parallel with the cyclic-peptide binding readouts above, the linear-peptide studies ask for membrane-receptor function. MC3R / MC4R participate in energy homeostasis and related metabolic regulation; MC4R is also an established target in obesity-related indications. On cells overexpressing the corresponding receptor, the linear peptides were assayed by FLIPR fluorescence agonism (α-MSH as control), giving pico- to nanomolar cell EC_{50} values (Figure 2a) and showing that the designed products can activate the receptor pathway.

Taken together, the peptide axis supplies three distinguishable lines of evidence: head-to-tail macrocycles and bicycles show measurable binding on hard-to-drug PPIs and tumor membrane antigens, respectively, while linear peptides show GPCR cell function. Affinities and potencies are labeled in Figure 2a.

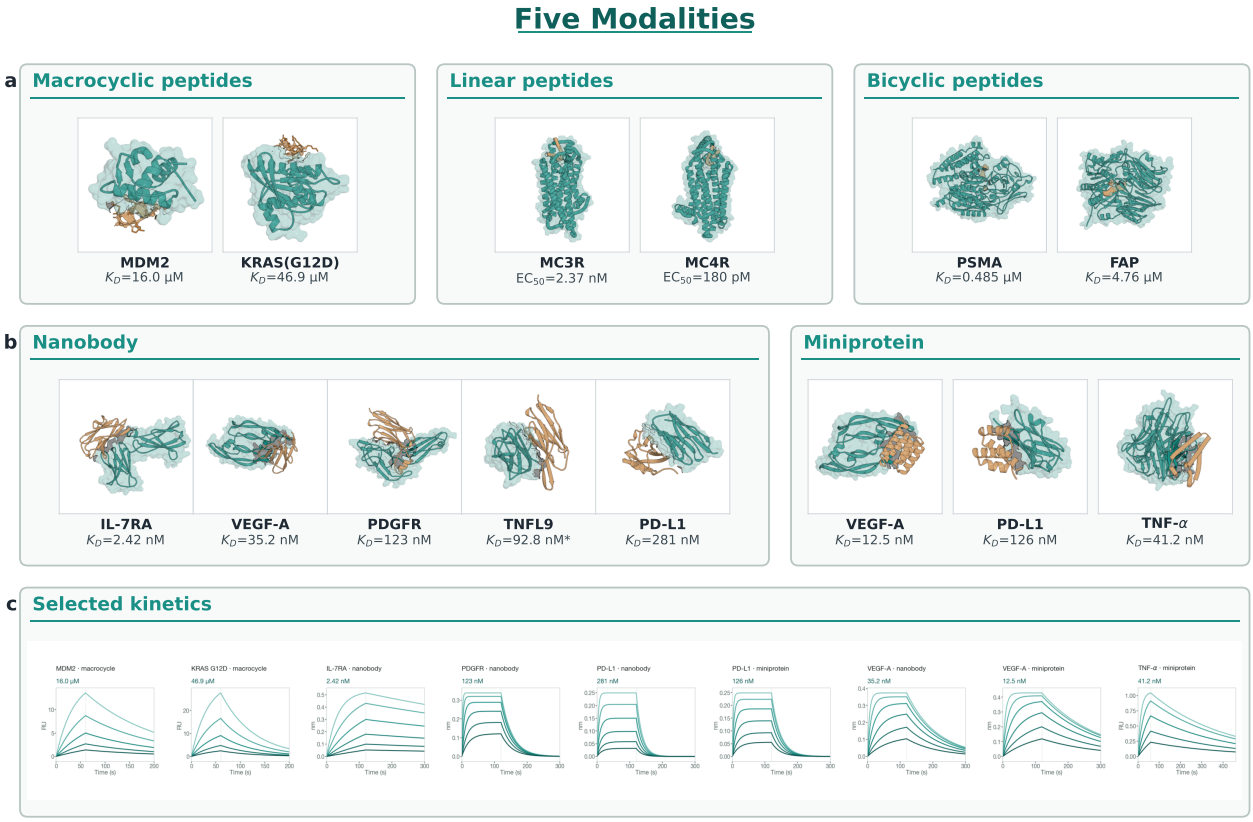

Figure 2: Broad generalizability of MeloDDE across five therapeutic modalities. Wet-lab validation overview and selected kinetics spanning head-to-tail macrocyclic peptides, bicyclic peptides, linear peptides, nanobodies, and miniproteins (fewer than 30 designs tested per target). Targets are depicted in teal and binders in orange; affinity / potency labels are annotated on each complex panel. (a) Head-to-tail macrocyclic peptides (MDM2 / KRAS G12D; multi-point SPR K_D), linear peptides (MC3R / MC4R; FLIPR cell EC_{50}), and bicyclic peptides (PSMA / FAP; multi-point SPR K_D). (b) Left: nanobodies (IL-7RA, VEGF-A, PDGFR, TNFL9, PD-L1; multi-point BLI K_D; TNFL9*: avidity); right: miniproteins (VEGF-A / PD-L1; multi-point BLI K_D; TNF-α; multi-point SPR K_D). (c) Selected multi-point traces matching panels a–b (SPR in RU for MDM2 / KRAS macrocycles and the TNF-α miniprotein; BLI in nm for IL-7RA, PDGFR, PD-L1, and VEGF-A nanobody / miniprotein pairs). Detailed protocols and fitted constants are provided in Appendix A and Appendix B.

2.3 Nanobodies and Miniproteins

Turning to protein binders. Relative to peptides, public wet-lab evidence is more concentrated on nanobodies and miniproteins; this subsection asks whether, within one MeloDDE pipeline, the two paths can each deliver measurable binding and whether they can succeed in parallel on the same antigen.

Nanobodies The nanobody panel spans five target classes: a cytokine receptor (IL-7RA), an angiogenic growth factor (VEGF-A), a receptor tyrosine kinase (PDGFR), an immune-checkpoint ligand (PD-L1), and a TNF-family costimulatory ligand (TNFL9 / 4-1BBL). Designs were framework-conditioned VHHs; binding was measured by multi-point BLI (Figure 2b). All five targets yield nanomolar-range binding: IL-7RA reaches the low-nanomolar range, and VEGF-A, PDGFR, and PD-L1 likewise fall in the nanomolar band. The TNFL9 readout against multivalent TNFL9–hFc is avidity and is not interchangeable with a monovalent K_D. Fitted constants and co-assayed public controls are in Table A6.

Miniproteins On the miniprotein side, the key point is parallelism with nanobodies. On VEGF-A and PD-L1, both paths yield clear binding (Figure 2b,c), so the same target can be advanced in two formats rather than only one. Both pairs were

read out by multi-point BLI and fall in a similar nanomolar band to the matched nanobodies.

In addition, TNF-α assembles as a trimer with a receptor-facing groove between subunits; a MeloDDE miniprotein designed toward that groove gives a fittable multi-point SPR binding signal (Figure 2b,c). This target was advanced only in the miniprotein format, providing a third miniprotein evidence path complementary to the parallel antigens.

Taken together, the protein-binder side supplies two distinguishable lines of evidence: nanobodies give nanomolar binding across five target classes, and miniproteins succeed in parallel with nanobodies on VEGF-A / PD-L1 while also yielding an SPR readout on TNF-α. On some targets, the same expression / assay batches also included several public reference sequences, listed to complete the assay record rather than as a ranked method comparison. Per-complex affinities are labeled in Figure 2; disclosed sequences, co-assayed references, and assay details are in Appendix A and Appendix B.

2.4 Novelty of Hit Sequences

For the nanobody and miniprotein hits above, we assessed sequence similarity against the full Protein Data Bank (PDB) (Tables 1–2). Peptide hits are omitted because their sequences are short and are not suited to the same criteria.

Antibody frameworks are highly conserved, so nanobodies were scored by full variable-domain global identity, taking the highest BLAST score among references with framework-region differences FR ≤ 1 (Appendix B). The five hits show 74.5–80.6% global identity with 21–36 CDR amino acid differences (Table 1): framework proximity to known VHHs is expected, while the CDR differences indicate that these candidates are not local rewrites of known antibodies. For PD-L1, no reference satisfied FR ≤ 1; the tabulated entry is the top hit after relaxing that filter.

Miniproteins were evaluated by unbiased whole-PDB BLAST: none of the three hits has a significant homolog at $E \leq 10^{-5}$ (Table 2). Taken together, the nanobodies retain clear CDR divergence beyond framework neighbors, and the miniproteins show no significant PDB homology; search parameters are in Appendix B.

Table 1: Sequence similarity between MeloDDE nanobodies and known proteins in the PDB. Each row presents the highest-scoring match from the exhaustive database search. Identity is reported as full variable-domain global identity; amino acid differences in CDR and FR regions are calculated based on the IMGT numbering scheme [11]. Detailed computational parameters are provided in Appendix B. Top-scoring PDB matches are distinct from the experimental antigens. The PD-L1 row is the top hit after relaxing an initial FR ≤ 1 filter.

Target	Top PDB ID	Identity (%)	CDR1	CDR2	CDR3	FR	CDR total
PD-L1	8z8v	80.6	3	8	10	5	21
IL-7RA	7xl0	79.8	5	8	14	0	27
VEGF-A	8coh	78.1	8	7	17	0	32
TNFL9	8coh	76.3	8	7	17	0	32
PDGFR	7a6o	74.5	9	5	22	0	36

Table 2: Sequence similarity between MeloDDE miniproteins and known proteins in the PDB. Each row presents the highest-scoring match. Identity and coverage are based on BLAST local alignments [12]. None of the hits meets the $E \leq 10^{-5}$ significance threshold, and none is the experimental antigen.

Target	Top PDB ID	Identity (%)	Coverage (%)	E
VEGF-A	8ch0	34.2	85	0.045
PD-L1	6n5u	40.7	42	1.1
TNF-α	3lkv	30.4	88	1.4

3 High-Efficiency Generation and Virtual Screening

This section evaluates MeloDDE on the computational side in two parts (Figures 3 and 4): Section 3.1 compares Melo-D Alpha, Melo-D Base, and public generators on success rate and sample efficiency under a shared *in silico* filter; Section 3.2 assesses Melo-1 alone on confidence-based ranking and enrichment of experimental binders. Melo-1 is the all-atom complex evaluator [10]; Melo-D Base is trained within the same framework, initialized from Melo-1's structure-prediction pretrained weights, and Melo-D Alpha is obtained from Melo-D Base by self-evolution (Appendix C).

3.1 Melo-D Alpha *In Silico* Design Success and Sample Efficiency

This subsection compares Melo-D Alpha, Melo-D Base, and public generators on two tasks: minibinder design and framework-conditioned nanobody CDR design. Both tasks use the same 11 targets (Table A7), with 800 designs planned per method–target cell (8,800 evaluable designs per method for minibinder and likewise for nanobody). Each generator configuration is held fixed across targets. Success rate and sample efficiency are both computed under that filter (definitions below). Minibinder design is compared across eight generators: classical RFdiffusion [2], RFdiffusion3 [13], RFD4-Proteina [14], BoltzGen [4], PXDesign [15], Proteina-Complexa [16], Melo-D Base, and Melo-D Alpha. The nanobody task requires a generator to accept a target structure, framework information, and specified hotspots; preserve the framework with constrained internal geometry while redesigning designated CDRs; and allow the nanobody pose relative to the target to vary. The public methods that meet this full conditioning setup are BoltzGen and RFD4-Proteina, so the nanobody comparison comprises these two together with Melo-D Base and Melo-D Alpha.

All generator outputs were refolded with an external model, ESMFold2-Fast (a fast ESMFold-family refolder [17, 18]), and then scored under a shared *in silico* success filter (Melo-1 is not used, to avoid evaluator bias in the generator comparison). Under fixed sampling settings, a design counts as successful only if min ipAE < 1.5 Å (interface confidence), binder pTM > 0.8 (fold confidence), and complex backbone RMSD < 2.5 Å (complex geometric agreement) hold simultaneously, following the AlphaProteo AF3-based in silico filter [3] (metric definitions as in AlphaFold 3 [1]). The success rate is $n_{\text{pass}}/n_{\text{evaluable}}$; uncertainty is reported as Wilson score 95% confidence intervals for the binomial proportion [19, 20]. Sample efficiency is $N = K/p$, where $K = 30$ is a shared reference yield for method comparison (Figure 3c,d), not a wet-lab test budget. Results are in Figure 3.

On minibinder design, Melo-D Alpha reaches 36.25% pooled *in silico* success, outperforming RFD4-Proteina (29.00%) and Melo-D Base (16.57%), and highest among the eight compared generators (Figure 3a). With $K = 30$, the expected design counts derived from the pooled rates are about 83, 103, and 181 (Figure 3c). Per-target rates vary substantially (Figure 3e): Melo-D Alpha ranks first on six targets, RFD4-Proteina and others lead on the remaining targets, and H1 stays at zero for all methods.

On the nanobody task, Melo-D Alpha reaches 19.45% pooled success, outperforming Melo-D Base (5.68%), RFD4-Proteina (4.86%), and BoltzGen (1.98%) (Figure 3b). The corresponding $K = 30$ expected design counts are about 154, 528, 617, and 1,517 (Figure 3d). Melo-D Alpha ranks first on every target with any successes among the four compared methods; H1, IL-17A, and TNF-α remain at zero for all four (Figure 3f).

3.2 Melo-1 Molecular Screening

This subsection evaluates the molecular screening performance of Melo-1 [10]: whether confidence-based ranking enriches experimental binders toward the front of a candidate pool. Unlike Section 3.1, which scores generators with external ESMFold2-Fast, this subsection examines Melo-1's own ranking performance.

The data cover Anthropic Binder Design [21] (1,230 candidates across 14 targets) and Overath [22] (3,669 candidates across 15 targets), shown per target beyond the target-macro means in the Melo-1 technical report (Figure 4). Melo-1, Protenix-v2 [23], OpenDDE [24], and ESMFold2 [18] are evaluated on the same candidates and labels within each dataset, with one prediction per design. Candidates are ranked separately within each target by minimum ipSAE, an interface-focused confidence score [25]. We report enrichment factor among the top 10% of candidates (EF@10%, relative to random sampling) and average precision (AP) over the full ranked list. The final row of each panel averages targets equally, matching the endpoint in the Melo-1 report; Appendix C reports binder–target ipTM as a complementary ranking.

Under minimum-ipSAE ranking, Melo-1 reaches target-macro AP / EF@10% of 0.558 / 2.695 on Anthropic Binder Design and 0.564 / 5.212 on Overath, the highest macro point estimates among Melo-1 and the three baselines on both datasets and both metrics (Figure 4). Per-target heterogeneity remains substantial, and high AP need not coincide with high EF@10%.

4 Autonomous Intent Translation via the Blueprint Engine

As the capabilities of generative models (like Melo-D Alpha) and evaluators (like Melo-1) improve, structure generation is no longer the only hard problem; equally important is how to translate high-level biological intent (e.g., "blocking a specific pathway") into precise, machine-executable constraints (e.g., "targeting specific residues with a specific modality"). The Design Blueprint Engine is the cognitive hub built for this upstream step. It elevates multi-modality binder design from a manual, expert-dependent task into an end-to-end, data-driven systems engineering process.

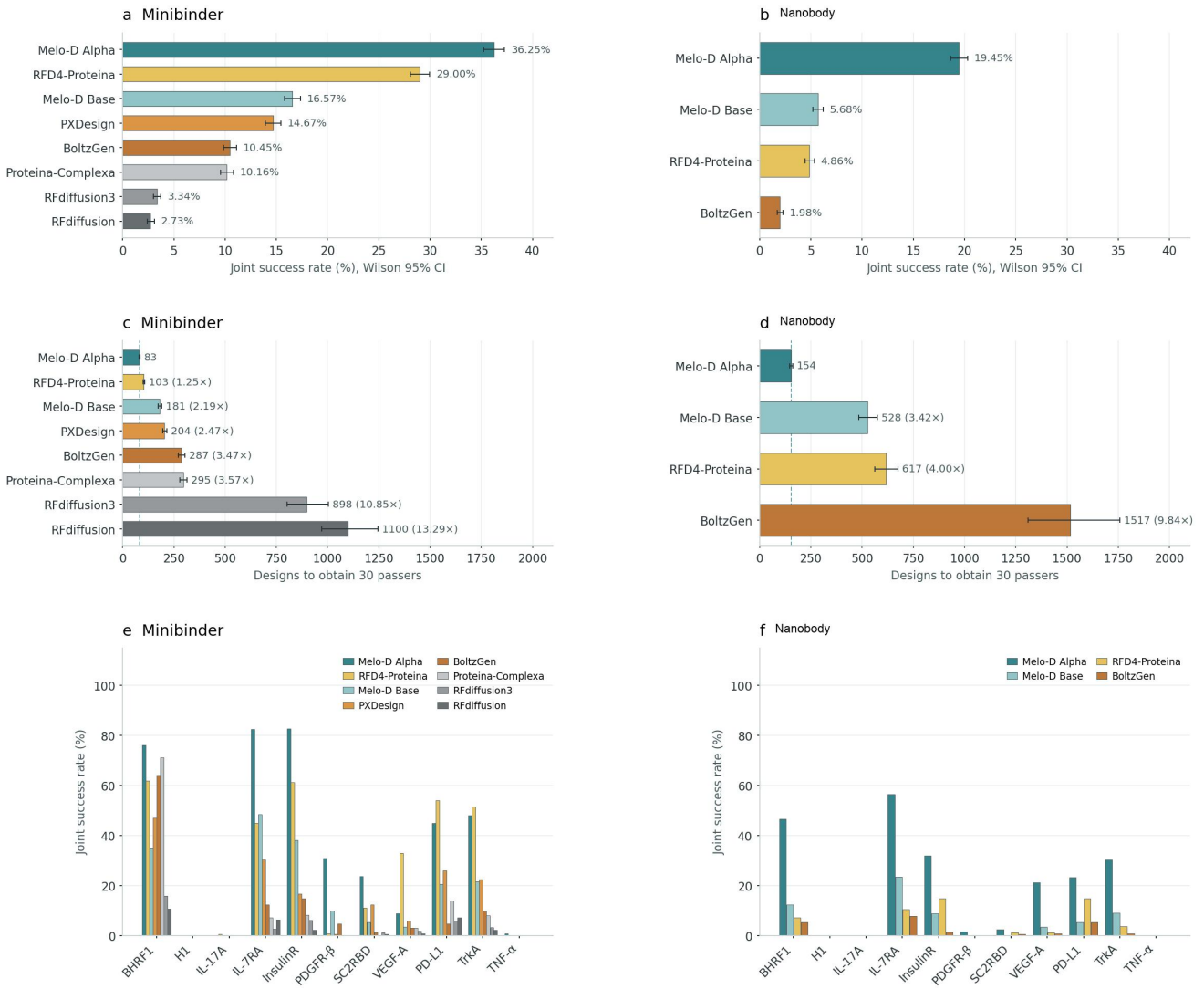

Figure 3: Minibinder (left) and framework-conditioned nanobody CDR design (right) under the shared *in silico* success filter. (a,b) Pooled success rates with Wilson score 95% CI. (c,d) Expected number of designs required to reach a successfully screened yield of $K = 30$ candidates ($N = K/p$). (e,f) Per-target success rates. Nanobody public baselines are restricted to BoltzGen and RFD4-Proteina under the inclusion criteria in the text.

4.1 Architecture and workflow

The core function of the Design Blueprint Engine is to reduce ambiguous biological goals into precise computational blueprints. Acting as the LLM agent in MeloDDE (Figure 5a), its workflow relies on two key mechanisms:

1. Knowledge-grounded constraint anchoring: The engine is contextually grounded by an Asset Repository (data sources, tools including Melo-1 and Melo-D Alpha, and design knowledge). Before sampling, it automatically carries out four decisions: epitope / pocket mapping, literature & patent mining, modality selection, and developability profiling. These decisions translate macroscopic biological and patent evidence into residue-mappable site hypotheses and lock the design configuration (e.g., modality type, hotspot / mask, geometric constraints).

2. Reflection-based dynamic closed loop: The engine outputs a design blueprint to an AI compute cluster, invoking Melo-D Alpha generation and Melo-1 evaluation. Subsequently, a reflection mechanism feeds the generated candidates, metrics, and knowledge back to the engine to dynamically revise sites, modality, or filters. This closed-loop mechanism upgrades the design process from open-loop, one-pass execution into self-correcting hypothesis iteration.

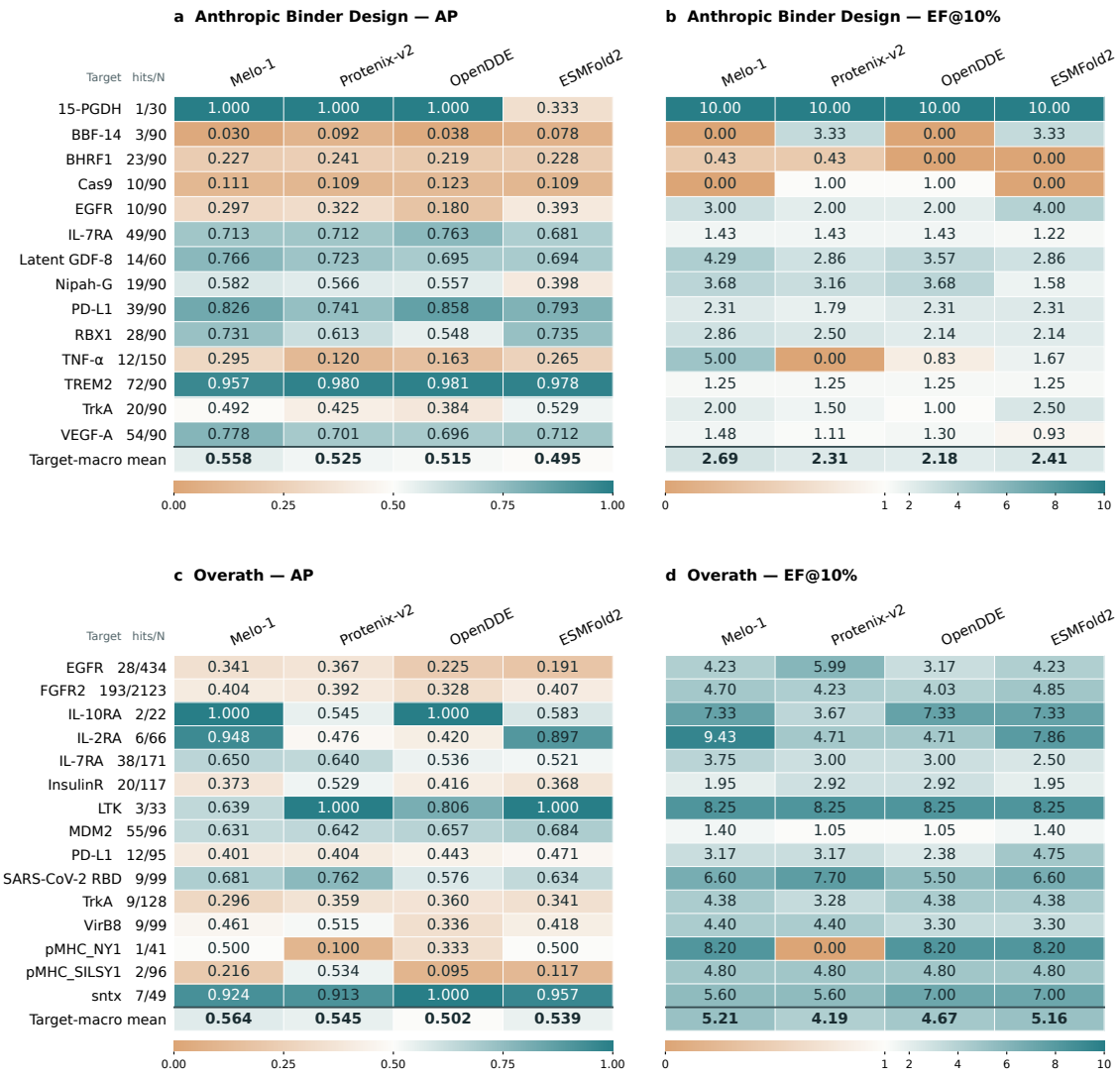

Figure 4: Per-target virtual screening ranked by minimum ipSAE. Within each target, candidates are ranked by descending minimum ipSAE; AP and EF@10% are then computed and compared across Melo-1, Protenix-v2, OpenDDE, and ESMFold2 under matched candidates and experimental binding labels. (a) AP on Anthropic Binder Design; (b) EF@10% on the same dataset; (c) AP on Overath; (d) EF@10% on the same dataset. In each heatmap, rows are targets (with hits/N) and columns are methods; the final row is the unweighted target-macro mean. EF@10% is the enrichment of experimental binders in the top 10% relative to random sampling; AP is average precision over the full ranked list.

4.2 Case study

To illustrate how the Design Blueprint Engine automates this "constraint anchoring" and guides downstream generation, we present a multi-target, multi-site nanobody design case (Figure 5b).

In this case, the engine autonomously partitioned leptin (LEP), interleukin-20 (IL20), and oncostatin M (ONCM) into five non-overlapping surface sites each, ensuring at least one blocking-oriented site at every native receptor interface. For each site, the engine issued explicit blueprint constraints (locking the framework, varying CDRs). Under these precise local constraints, Melo-D Alpha and Melo-1 successfully designed and screened about six high-confidence candidates per site.

This case shows that the Design Blueprint Engine can autonomously execute complex site planning and constraint issuance, so subsequent generation and evaluation compute is precisely focused on high-value local hypotheses rather than unconstrained global blind search.

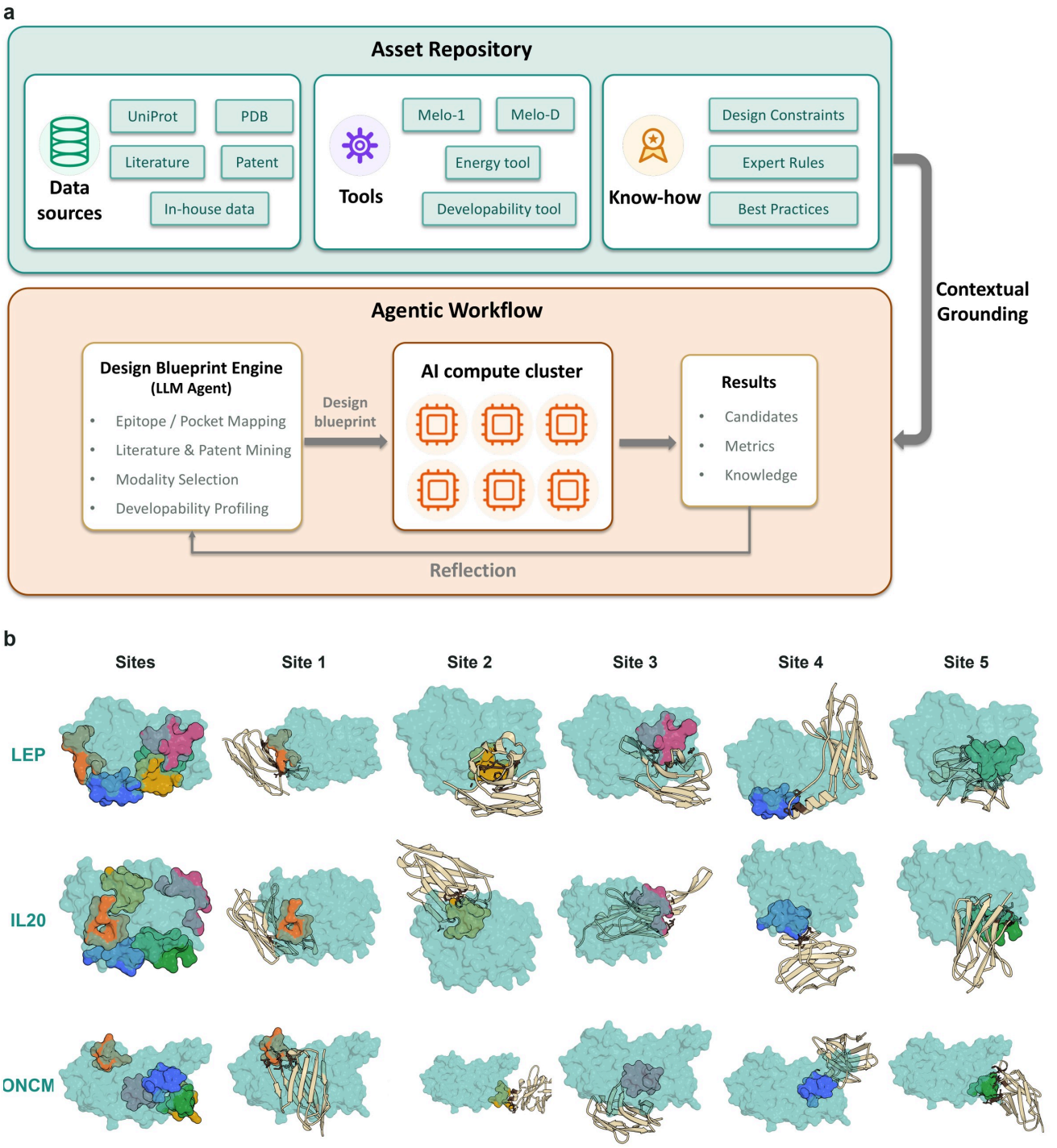

Figure 5: Design Blueprint Engine. (a) Architecture: an Asset Repository (data sources, tools including Melo-1 and Melo-D Alpha, and design knowledge) grounds an LLM agent that performs epitope / pocket mapping, literature & patent mining, modality selection, and developability profiling; outputs a design blueprint to an AI compute cluster; and reflects over candidates, metrics, and knowledge. (b) Case study: multi-target, multi-site nanobody design on LEP, IL20, and ONCM. The leftmost *Sites* column displays the target surface with five colored epitopes; Site 1–5 panels use the corresponding colors and a fixed target viewpoint within each row. Each Site panel shows a representative Melo-1 predicted complex for that epitope, selected among designs retained by an interface-contact filter as the candidate with the largest predicted contact set. Structures are Melo-1 complex models from the design set reported here.

5 Discussion

Sections 2–3 show that MeloDDE can deliver measurable binding or cell readouts across multiple therapeutic formats, and can raise generator success and screening enrichment under shared protocols. At the same time, wet-lab affinities span a wide range (micromolar to picomolar), and *in silico* readouts mainly reflect foldability and interface confidence rather than high affinity or developability. The remainder of this section discusses the next steps along two axes: affinity maturation and front-loading developability assessment.

From measurable binding to higher affinity. The wet-lab readouts in Section 2 span multiple affinity bands: many cyclic peptides (head-to-tail macrocycles and bicycles) start in the micromolar range, nanobodies and miniproteins fall in the nanomolar band, and linear GPCR peptides show pico- to nanomolar cell potencies. The next priority is to mature and locally refine weak-but-real binders, especially micromolar cyclic peptides.

Front-loading developability assessment. Even when affinity reaches a usable range, candidates can still stall on modality-specific downstream criteria: for instance, expression levels and aggregation propensity for nanobodies, or membrane permeability and intracellular accessibility for cyclic peptides. These developability thresholds are often no less challenging than affinity optimization; deferring them until after binding is confirmed more readily wastes design and assay effort. A more robust path is to front-load developability assessment: as described in Section 4, the Design Blueprint Engine already integrates developability profiling with epitope mapping and modality selection; the next step is to bring permeability and other developability filters earlier into Blueprint constraint generation and Melo-1 screening. Affinity maturation and front-loaded developability thus form the two main next lines for MeloDDE.

6 Conclusion

The central challenge in computational binder design has evolved beyond merely generating plausible conformations; it now lies in bridging the gap between *in silico* design and the physical world to reliably deliver candidates with true therapeutic potential. The Melo Drug Design Engine (MeloDDE) was developed as a systematic solution to this industry-wide bottleneck. By deeply integrating the Design Blueprint Engine for intent translation, Melo-D Alpha for efficient generation, and Melo-1 for rigorous evaluation, MeloDDE establishes a closed-loop pipeline from target insight to molecule delivery.

Across both computational benchmarks and multi-modality wet-lab validations, MeloDDE has demonstrated exceptional generalizability and design efficiency. It not only outperforms current state-of-the-art generative models but also breaks through the limitations of single-modality design. The engine successfully unifies the generation and advancement of five distinct therapeutic formats: head-to-tail macrocycles, bicycles, linear peptides, nanobodies, and miniproteins. These results validate that front-loading high-accuracy structure prediction and multi-dimensional physicochemical constraints fundamentally improves the translation success rate from *in silico* screening to *in vitro* validation.

Looking ahead, computational protein design must inevitably expand toward more complex developability frontiers. The next phase of MeloDDE will focus on overcoming the affinity maturation bottleneck for initial micromolar hits, while further embedding critical developability metrics (such as expression yield, aggregation propensity, and membrane permeability) directly into the generation constraints. As its evaluation and feedback mechanisms continue to evolve, MeloDDE aims to provide a highly predictable and scalable foundational infrastructure for AI-driven drug discovery.

7 Availability

To facilitate community research and application translation in computational protein design, the core evaluation capabilities of Melo-1 have been made publicly accessible through Melo Playground (`https://platform.atomelody.com`). Melo Playground currently provides two primary functionalities:

- High-Accuracy Structure Prediction: Users can submit monomeric or complex sequences to generate all-atom 3D conformations rapidly. Melo Playground provides interactive structural visualization alongside detailed residue-level confidence (pLDDT) and predicted aligned error (PAE) matrices, allowing users to intuitively assess folding reliability.
- High-Throughput Virtual Screening: Given a target and a library of candidate binders, Melo-1 performs batch folding and binding interface evaluation. The system automatically scores and ranks candidate molecules based on interface confidence (e.g., iPTM), binding affinity proxies, and conformational stability, thereby pinpointing high-probability candidates for experimental validation.

Users can interact with these features directly via the Melo Playground web interface, or automate task submission and result parsing using the REST API (API reference: https://www.atomelody.com/docs/api/task) and MCP-capable (Model Context Protocol) AI agents. For a detailed walkthrough of a completed screening report on Melo Playground, please refer to Appendix D.

Appendix

A Wet-Lab Methods

This appendix records structural design inputs and assay protocols for the wet-lab studies in Figure 2. Some studies used earlier MeloDDE configurations that may differ from the Melo-1 / Melo-D Alpha setup in the *in silico* benchmark. Across those studies, fewer than 30 designs per target were advanced to wet-lab testing. Unless noted otherwise, reported K_D values are from multi-point kinetic fits; single-point screens were used only to select candidates for the multi-point panel. Sequences and fitted constants are in Appendix B (Tables A5–A6); recombinant antigens are listed in Table A2. Tables A3–A4 summarize PDB-conditioned design inputs for macrocycles, nanobodies, and miniproteins; MC3R / MC4R linear peptides were advanced as cell-based FLIPR agonists and are not listed there (Appendix B). Assay details follow by modality: constrained peptides (SPR; Section A.1), nanobody / miniprotein binders (BLI; Section A.2), and the TNF-α miniprotein (SPR; Section A.3).

Table A1: Wet-lab targets corresponding to Figure 2. Twelve unique antigens across 14 target–modality wet-lab studies.

Target	Brief description	Modality
MDM2	E3 ubiquitin ligase that binds p53; oncology PPI target	Head-to-tail macrocycle
KRAS G12D	Oncogenic GTPase mutant common in solid tumors	Head-to-tail macrocycle
MC3R	GPCR in energy homeostasis and metabolic regulation	Linear peptide
MC4R	GPCR linked to appetite control and obesity	Linear peptide
PSMA	Membrane peptidase enriched on prostate tumor cells	Bicyclic peptide
FAP	Membrane protease expressed in tumor stroma	Bicyclic peptide
IL-7RA	Cytokine receptor for lymphoid IL-7 signaling	Nanobody
VEGF-A	Angiogenic growth factor acting on VEGF receptors	Nanobody; miniprotein
PDGFR	Receptor tyrosine kinase for PDGF ligands	Nanobody
TNFL9	4-1BBL / TNFSF9; TNF-family costimulatory ligand for 4-1BB	Nanobody
PD-L1	Immune-checkpoint ligand that engages PD-1	Nanobody; miniprotein
TNF-α	Canonical proinflammatory TNF-family cytokine	Miniprotein

Table A2: Recombinant target antigens and catalog numbers (Sino Biological).

Target	Catalog ID	Construct (vendor description)
MDM2	M45-31BH	Human MDM2, His tag
KRAS G12D	R06-32BH	Human KRAS(G12D), His tag
PD-L1	10084-H02H	Human PD-L1/B7-H1 ECD, hFc tag
PDGFR	10514-H02H1	Human PDGFRB, Fc tag
IL-7RA/CD127	10975-H02H	Human CD127/IL-7RA, Fc tag
VEGF-A (VEGF165)	11066-HNAH	Human/cynomolgus VEGF165, HPLC-verified
TNFL9	15693-H01H	Human TNFL9 ECD, hFc tag
TNF-α	10602-H49H-B	Recombinant Human TNF-alpha Protein

Table A3: Macrocycle design inputs (MDM2 and KRAS G12D).

	MDM2	KRAS G12D
PDB ID	4HFZ	6WGN
Target chain and residue numbers	A26–108	A1–34, A38–169
Binder length	12–20 aa	12–20 aa
Cyclization	Head-to-tail lactam	Head-to-tail lactam
Target hotspot residues	A54, A58, A61	A9, A58, A60, A62, A69, A72, A73, A96, A99, A102, A103

Table A4: Nanobody and miniprotein design inputs for wet-lab studies.

Modality	Target	PDB ID	Target chain and residue numbers	Binder length	Target hotspot residues
Miniprotein	VEGF-A	1BJ1	V14–107, W14–107	50–140	W81, W83, W91
Miniprotein	PD-L1	5O45	A17–132	50–120	A56, A115, A123
Miniprotein	TNF-α	1TNF	A6–157, B6–157, C6–157	50–100	A75, A77–79, A86–92, A95, A97, A135–137; B20–21, B23, B31–33, B65, B67, B110–113, B115, B143–147, B149
Nanobody	IL-7RA	3DI3	B17–209	—	B58, B80, B139
Nanobody	VEGF-A	1BJ1	V14–107, W14–107	—	W81, W83, W91
Nanobody	PDGFR	3MJG	X93–280	—	X265, X264, X210
Nanobody	TNFL9	6A3V	G88–243	—	G100, G101, G150–153, G227, G230
Nanobody	PD-L1	5O45	A17–132	—	A56, A115, A123

A.1 Constrained Peptides (SPR)

Bicyclic PSMA and FAP designs were measured by multi-point SPR (Section 2).

For head-to-tail macrocycles (design inputs in Table A3), quantitative affinities were obtained from multi-concentration SPR on a CM5 chip. His-tagged KRAS G12D (Table A2) was prepared at 20 μg ml^{-1} (pH 4.0) and immobilized to approximately 4660 RU; His-tagged MDM2 was prepared at 30 μg ml^{-1} (pH 4.0) and immobilized to approximately 4900 RU. Peptide analytes were injected as a twofold dilution series (50, 25, 12.5, 6.25, 3.125, and 0 μM) in 1× HBS-EP+ at 25°C and 30 μl min^{-1} (60 s association / 140 s dissociation), followed by a running-buffer regeneration step (~145 s). Sensorgrams were fit to a 1:1 binding model; fitted constants are in Table A5.

A.2 Nanobody and Miniprotein Binders (BLI)

Protein binders carried a C-terminal 10His tag. IL-7RA, PDGFRβ, and PD-L1 binders were expressed in CHO cells; VEGF-A and TNFL9 (TNFSF9) binders were produced by cell-free expression. Multi-point BLI was performed on a Sartorius Octet RH16 in 1× PBST (120 s association / 180 s dissociation) and analyzed with Data Analysis 12.0 under Kinetics 1:1 Binding. Antigens are listed in Table A2; design inputs in Table A4; fitted constants in Table A6. Target-specific sensor configurations were as follows.

IL-7RA. IL-7RA–Fc was captured on Pro A sensors from 10 μg ml^{-1} antigen to approximately 2.5 nm loading; binders were analytes in serial twofold dilutions (typically 62.5–1.95 nM; one miniprotein series started at 31.3 nM).

PDGFRβ. PDGFRB–Fc was captured on AHC2 sensors from 20 μg ml^{-1} antigen (target loading ~1.0–2.0 nm); binder analytes were tested at 2000, 1000, 500, 250, 125, and 62.5 nM. Comparator designs lacking specific binding or dominated by non-specific responses were excluded from quantitative affinity reporting.

PD-L1. Nanobody assays captured PD-L1–hFc on AHC2 sensors from 10 μg ml^{-1} antigen (loading ~1.8–2.0 nm), with analytes at 1000–31.3 nM. Miniprotein assays captured the same antigen on Pro A sensors from 20 μg ml^{-1} antigen, with analytes at 500–15.6 nM or 250–7.81 nM.

VEGF-A. The reported nanobody affinity used HIS1K with His-tagged binder immobilized and VEGF165 as analyte (2500–78.15 nM). Reported miniprotein affinities used AR2G with VEGF165 immobilized from 10 μg ml^{-1} antigen for 300 s and miniprotein as analyte (500–15.6 nM or 250–7.81 nM).

TNFL9 (TNFSF9). His-tagged nanobodies were immobilized on HIS1K sensors from 10 μg ml^{-1} binder for 300 s, and multivalent TNFSF9–hFc was applied as analyte (2000–62.5 nM). Because the antigen is a multivalent Fc fusion, the starred affinities in Table A6 are avidity-related and are not interchangeable with monovalent K_D.

A.3 TNF-α Miniprotein (SPR)

The TNF-α miniprotein was produced in CHO and assayed by multi-point SPR. Recombinant TNF-α (Table A2) was immobilized on an SA chip to ~220 RU; binder analytes were injected as a twofold dilution series in 1× HBS-EP+ at 25°C (60 s association / 400 s dissociation; regeneration with 3 M $MgCl_2$). Sensorgrams were fit to a 1:1 Langmuir binding model (K_D = 41.2 nM).

B Binder Sequences, Affinities, and PDB Similarity

B.1 Sequences and Affinities

Table A5 lists the macrocyclic peptide designs reported for MDM2 and KRAS G12D (multi-point SPR). Sequences are written N→C in one-letter code; both designs were cyclized by head-to-tail lactam chemistry (amide bond between the termini), with the linear string shown for readability. Table A6 lists amino-acid sequences and measured affinities for nanobody and miniprotein entries that are disclosed here (multi-point BLI unless noted). Melo-D Alpha rows are designs from this work; other source labels in Table A6 are co-assayed public or control sequences from the same studies (same expression / assay panels where applicable), listed for completeness of the assay record rather than as a ranked comparison: AlphaProteo [3], PXDesign [15], Latent-X / Latent-X2 [5, 6], BoltzGen [4], PPIFlow [26], and Germinal [27]. Sequences for the bicyclic PSMA / FAP peptides, the linear MC3R / MC4R peptides, and the TNF-α miniprotein are omitted here as these studies are part of ongoing programs; however, the corresponding multi-point affinities and cell potencies are fully reported in Section 2. SPR and BLI assay parameters are in Appendix A.1–A.3.

Table A5: Macrocyclic peptide sequences with multi-point SPR affinities.

Target	Modality	K_D	Sequence
MDM2	Macrocycle	16.0 μM	EFEGEWNKMMGLPPSSG
KRAS G12D	Macrocycle	46.9 μM	TSSSLWKKYDPR

Table A6: Binder sequences with measured affinities (multi-point BLI). Melo-D Alpha = designs from this work; other sources = co-assayed panel entries (citations in the appendix text).

Target	Source	Modality	K_D	Sequence
IL-7RA	Melo-D Alpha	Nanobody	2.42 nM	EVQLVESGGG LVQPGGSLRL SCAASGDVTI YTMAWYRQAP GKGRELVAGM SRDADDATSY ADSVKGRFTI SRDNAKNTLY LQMNSLRPED TAVYYCAVLA EDHSSKLGAN LYWGQGTLVT VSS
	AlphaProteo	Miniprotein	5.9 nM	AVEPVLSKEE VGEIARIYAK EIGKDYGIEL SDEEIDLAAE LARELYGKSP EEAKEFLEEV YKKLSKELSK ETLKIIIAAA VGALEAAELA GRLAEEYRAG VIDADELREE LSKFLPDELV DRVLARAEA
	PXDesign	Miniprotein	6.82 nM	MEEFLKEIEE AGKEAKKIGK EIAKEFIEEN KNDLEKIKEE LERAKGLAEL FKEVADEDLA KFKANPEGLT EEEKEDIKYD YAQAEAAEAF YKELKEYLEK K
	Latent-X	Miniprotein	9.86 nM	SAAAQINTLK NASNLINEGK YGEAVTLLTN LLPTVADNAV KTEINNAIDN AKEADLYAKK ADLAEQYLAE NPGLSAADVD KVKYLIDAYK DISADSATAA KTSINNAISL LETNS
PDGFR	Melo-D Alpha	Nanobody	123 nM	EVQLVESGGG LVQPGGSLRL SCAASGSFSW NQIVFMGWFR QAPGKGRELV AAINYSGSYT YYPDSVEGRF TISRDNAKRM VYLQMNSLRA EDTAVYYCVA LTVTTSAILE AKLKGKAVKP DTYWGQGTQV TVSS
	PPIFlow	Nanobody	—†	EVQLVESGGG LVQPGGSLRL SCAASGTTQS IVMMAWYRQA PGKGRELVAG ISNTGRYISY ADSVKGRFTI SRDNAKNTLY LQMNSLRPED TAVYYCAVWQ ISYSGSTTWT VYWGQGTLVT VSS
	PPIFlow	Nanobody	—	EVQLVESGGG LVQPGGSLRL SCAASGGTFI ETFMGWFRQA PGKGRELVAA ISSSGSRTYY PDSVEGRFTI SRDNAKRMVY LQMNSLRAED TAVYYCAAYY LDRSSKTLKV SWGQGTQVTV SS
	PPIFlow	Nanobody	—	EVQLVESGGG LVQPGGSLRL SCAASGSSSS SYFMAWFRQA PGQEREFVAG INWSGNRTLY ADSVRGRFTN SRDNSKNTLY LQMNSLRAED TAVYYCTTLD MRNRSVSYWG QGTLVTVSS
	PPIFlow	Nanobody	—	EVQLVESGGG LVQPGGSLRL SCAASGASLH SGFMGWFRQA PGKGRELVAA INNRNGDTYY PDSVEGRFTI SRDNAKRMVY LQMNSLRAED TAVYYCAVMG YYTSYSSKTT NKAYWGQGTQ VTVSS
VEGF-A	Melo-D Alpha	Nanobody	35.2 nM	EVQLVESGGG LVQPGGSLRL SCAASGANTS DMVMMAWFRQ APGQEREFVA GIEGYNNDET LYADSVRGRF TNSRDNSKNT LYLQMNSLRA EDTAVYYCAV ITSFAHGSTL TWGQGTLVTV SS

Continued on next page

Table A6 — *continued*

Target	Source	Modality	K_D	Sequence
	Melo-D Alpha	Miniprotein	12.5 nM	ADALIKAAEA LLEGKVEEAI KNLEEIKKKY SDNEGVVHYI NKAIEALKEG RLAEAAVNLL DVVLQIKDDM SEEQYKKLLE LIEKALEKWG KK
	AlphaProteo	Miniprotein	8.34 nM	AEKKEKIIKA LELLAEAAKK LEEAAEDPSL KEALKELKEK LKEIKEKLKK GEISLEDAAN QIGALGAMII DFADGMLAMG KIDEAEEVLK LVKEAAKALI EGGGEAGRAG RSISAKIASL EKRIAAAK
	PPIFlow	Miniprotein	26.5 nM	MSGSEKERIR ALFHRLLMLL VKDERTTDEA RALAIAMEIY FTSTKDLEIV DINVNGMSQE EVEEKISKGE VILGKIEITF ELDGKKHTIY IEIYEEDGKI KAEISADPND KVSQELVKKI KELAEMILAA EKA
	BoltzGen	Miniprotein	—	SGATLAALLA AVTAQVLYTK HGKSPEEALA EGLKVAERLA KEAGWDLEEA LAAAKEVVEK VLELLEAGKY EEAKALLDEL EEEVRE
PD-L1	Melo-D Alpha	Nanobody	281 nM	EVQLVESGGG LVQPGNSLRL SCAASGFTFS DYWMSWVRQA PGKGLEWVSS IGSKSYSYAT LYADSVKGRF TISRDNAKTT LYLQMNSLRP EDTAVYYCTV WGTADWRKVF EGKATGTLVT VSS
	Melo-D Alpha	Miniprotein	126 nM	MKSSDEIEEE YSKLSKEELI ERIKKILDEL NFPEAAKNFY LSNLPKMSKD ELVKLATLLE KIKQS
	Germinal	Nanobody	50.7 nM	QVQLVESGGG LVQPGGSLRL SCAASPPSSS EYELKYLGWF RQAPGQGLEA VAAIAGRQER RYYADSVKGR FTISRDNSKN TLYLQMNSLR AEDTAVYYCY SVWMPKDWQK DVPLVHKWGQ GTLVTVSSRG R
	AlphaProteo	Miniprotein	24 nM	SAEEKILANL EAMKAKALAA KTEEEKLFYA KALLAVAISY AIRGDYELAR RAAELAVEVI KSLSKEEQKK VMDFLINIIK NITDPEDREK AIELAIAIAE RLDEEVREEA LKKIEELKKE
TNFL9	Melo-D Alpha	Nanobody	92.8 nM*	EVQLVESGGG LVQPGGSLRL SCAASSLPSG YSYVMAWFRQ APGQEREFVA GITSGRYQQT LYADSVRGRF TNSRDNSKNT LYLQMNSLRA EDTAVYYCAV LALIKLHPDG SETQTPVDYW GQGTLVTVSS
	Latent-X2	Nanobody	34.4 nM*	EVQLVESGGG LVQPGGSLRL SCAASLGSSY LAGMGWFRQA PGKGRELVAA IDWATGDTYY PDSVEGRFTI SRDNAKRMVY LQMNSLRAED TAVYYCGAWT TTAAPGTLDP AAVNGFAWGQ GTQVTVSS

Legend. K_D values are multi-point BLI fits (nM) unless marked. * avidity (TNFL9: multivalent Fc-fused antigen); not interchangeable with monovalent affinity. — no measurable binding under the study assay. —† no expression (no soluble product for an affinity readout).

B.2 Sequence Similarity to the PDB

Tables 1 and 2 evaluate the sequence novelty of the experimentally validated hits. To rigorously quantify sequence divergence, we expanded the search scope to an exhaustive, unbiased whole-database query against all canonical `polypeptide(L)` entities in the Protein Data Bank (514,768 distinct entity sequences, non-deduplicated). To eliminate search blind spots arising from antigen preference or restricted antibody sub-libraries, queries were not restricted to the experimental target; SAbDab [28] was employed exclusively as a post hoc annotation tool rather than a search database.

To guarantee the deterministic capture of the nearest phylogenetic neighbors, we intentionally relaxed the BLASTP 2.12.0+ [12] reporting thresholds (e.g., set to 10 for miniproteins) and removed target count limits (`max_target_seqs=1000000`). This forced the algorithm to report the closest available match trajectories, even when they entirely lacked statistical significance. The operational significance cutoff for homology remained strict at $E \leq 10^{-5}$.

We decoupled our scoring strategies to account for the architectural differences between modalities. Because antibodies share highly conserved folding frameworks, standard heuristic local alignments yield misleadingly high false-positive identities. Consequently, nanobodies were strictly subjected to full-domain global alignments (Biopython 1.88, BLOSUM62, gap open/extend $-10/-0.5$) based on ANARCI 2.0.8 [29] IMGT numbering [11]. This cleanly isolates binding-specific CDR divergence from generic framework conservation. For each nanobody, the tabulated top PDB match was selected among references with framework-region differences FR ≤ 1 (substitutions, insertions, and deletions under IMGT numbering), taking the highest BLAST score within that set; CDR1/2/3 and FR differences are reported relative to that same reference. The PD-L1 hit had no FR ≤ 1 reference across the full entity set; its Table 1 entry is therefore the highest-scoring match after relaxing that filter (global identity 80.6%; FR = 5). Conversely, miniproteins, lacking predefined regional architectures, were subjected to the most stringent unconstrained local BLAST evaluation to capture any potentially homologous short motifs.

This differentiated methodology ensures a rigorous and unbiased assessment of sequence novelty. The sequence-level

divergence established here serves as an orthogonal complement to the physical receptor-binding validated in our wet-lab experiments, jointly forming a complete chain of evidence for designed binders.

C *In Silico* Design and Screening Details

This appendix collects computational details that support Section 3: Melo-D training, the shared generator benchmark targets, and Melo-1 virtual-screening ranking beyond the main-text minimum-ipSAE analysis.

C.1 Melo-D Training

The Melo-D Base model is trained within the Melo-1 framework, initialized from the Melo-1 structure-prediction pretrained weights. Melo-D Alpha subsequently builds upon Melo-D Base through a self-evolution process; throughout this report, Melo-D Alpha refers to the complete end-to-end design system. The training corpus comprises experimental PDB structures deposited on or before September 30, 2021, augmented with curated antibody structures and high-confidence monomer distillation data. The training objectives are formulated to jointly optimize whole-entity design, local-fragment design, and structural folding across multiple conditioning regimes.

C.2 Shared Benchmark Targets

Table A7 details the eleven targets utilized for the minibinder and framework-conditioned nanobody CDR benchmarking presented in Section 3. Note that this computational target set is distinct from the wet-lab validation panel detailed in Table A1, though they share several key targets (IL-7RA, VEGF-A, PD-L1, the PDGFR family, and TNF-α).

Table A7: Minibinder and nanobody *in silico* benchmark targets used in Section 3.

Target	Brief description
BHRF1	Epstein–Barr virus Bcl-2 homolog; anti-apoptotic viral protein
H1	Viral envelope glycoprotein mediating host-cell entry
IL-17A	Proinflammatory cytokine in the IL-17 axis
IL-7RA	Cytokine receptor for lymphoid IL-7 signaling
InsulinR	Receptor tyrosine kinase for insulin / IGF signaling
PDGFR-β	Receptor tyrosine kinase for PDGF ligands
SC2RBD	Receptor-binding domain of the coronavirus spike protein
VEGF-A	Angiogenic growth factor acting on VEGF receptors
PD-L1	Immune-checkpoint ligand that engages PD-1
TrkA	Neurotrophin receptor for nerve growth factor (NGF)
TNF-α	Canonical proinflammatory TNF-family cytokine

C.3 Complementary Ranking by Binder–Target ipTM

While Figure 4 in Section 3.2 ranks the Anthropic Binder Design and Overath cohorts using minimum ipSAE, Figure A1 provides a complementary analysis using binder–target ipTM as the primary ranking metric. All candidate sets, ground-truth labels, and structural predictions remain identical to the main-text analysis. The binder–target ipTM is calculated by averaging the chain-pair ipTM scores between the binder and all target protein chains. Evaluated under this metric, Melo-1 achieves a target-macro Average Precision (AP) and Enrichment Factor at 10% (EF@10%) of 0.538 and 2.477 on the Anthropic Binder Design dataset, and 0.522 and 4.896 on the Overath dataset, respectively.

For computational reproducibility: The Anthropic Binder Design evaluation excludes inconclusive MatureGDF8 measurements, designs lacking a final binding label, and MBP (utilizing the `binder_final` labels). The Overath evaluation excludes seven incomplete records (utilizing the `binder` labels). During inference, target-side Multiple Sequence Alignments (MSAs) and structural templates were enabled, whereas binder-side MSAs and templates were explicitly disabled. The minimum ipSAE calculation [25] employs a 10 Å PAE cutoff without applying an additional coordinate-distance filter.

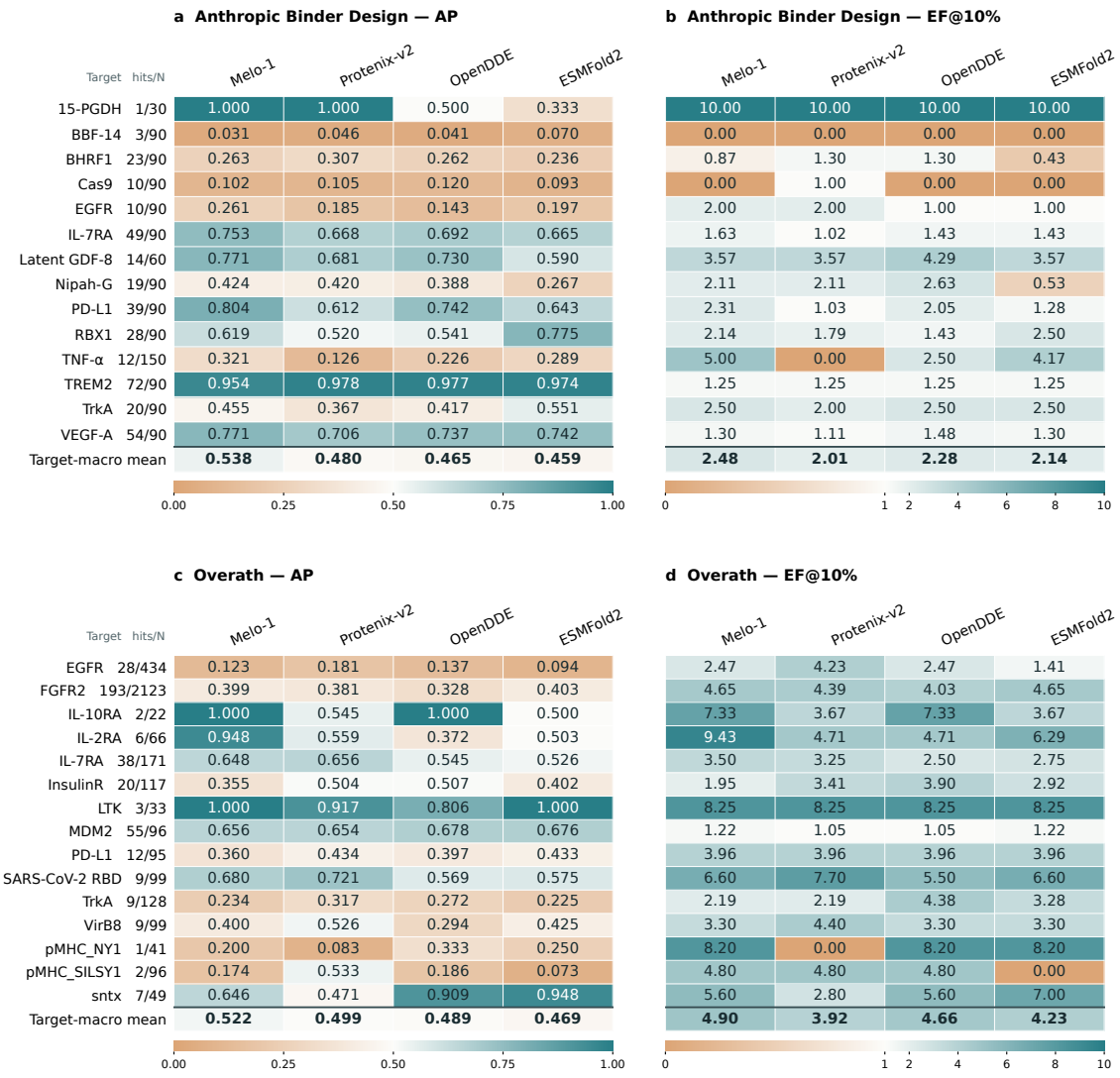

Figure A1: Complementary virtual screening performance ranked by binder–target ipTM. This analysis evaluates the same Anthropic Binder Design and Overath cohorts presented in Figure 4, substituting minimum ipSAE with binder–target ipTM as the ranking metric. The overall layout and metric definitions correspond to those in the main-text figure, though the resulting numerical values differ.

D Melo Playground and Melo-1 API

This appendix details the user interfaces and interactive reporting features of Melo Playground, the unified deployment environment for the Melo-1 model. Melo Playground provides comprehensive online access to Melo-1 for structure prediction and batch candidate screening across diverse molecular systems, including proteins, DNA, RNA, small-molecule ligands, and ions.

Users can programmatically submit and manage computational tasks via a REST API or through AI agents utilizing the Model Context Protocol (MCP). Alternatively, the Melo Playground web interface provides a graphical environment to monitor task status, inspect submitted inputs, explore structural predictions, and download results. While Section 4 focuses on the autonomous multi-site design planning executed by the Design Blueprint Engine, this appendix highlights the shared Melo-1 evaluation surface used to analyze the resulting candidates.

Figures A2–A5 illustrate a representative virtual-screening workflow for a single protein target against ten candidate proteins, generating five structure samples per candidate. The platform's interactive reports seamlessly integrate candidate ranking, structure-sample selection, and three-dimensional molecular visualization. Furthermore, they provide linked confidence

diagnostics, enabling users to navigate residue-level metrics and perform detailed PAE, pLDDT, and inter-chain confidence analyses.

It is important to note that while prediction confidence and computational ranking scores are strongly correlated with success rates, they do not inherently guarantee experimentally validated binding or biological activity.

API reference: https://www.atomelody.com/docs/api/task; Melo Playground: https://platform.atomelody.com.

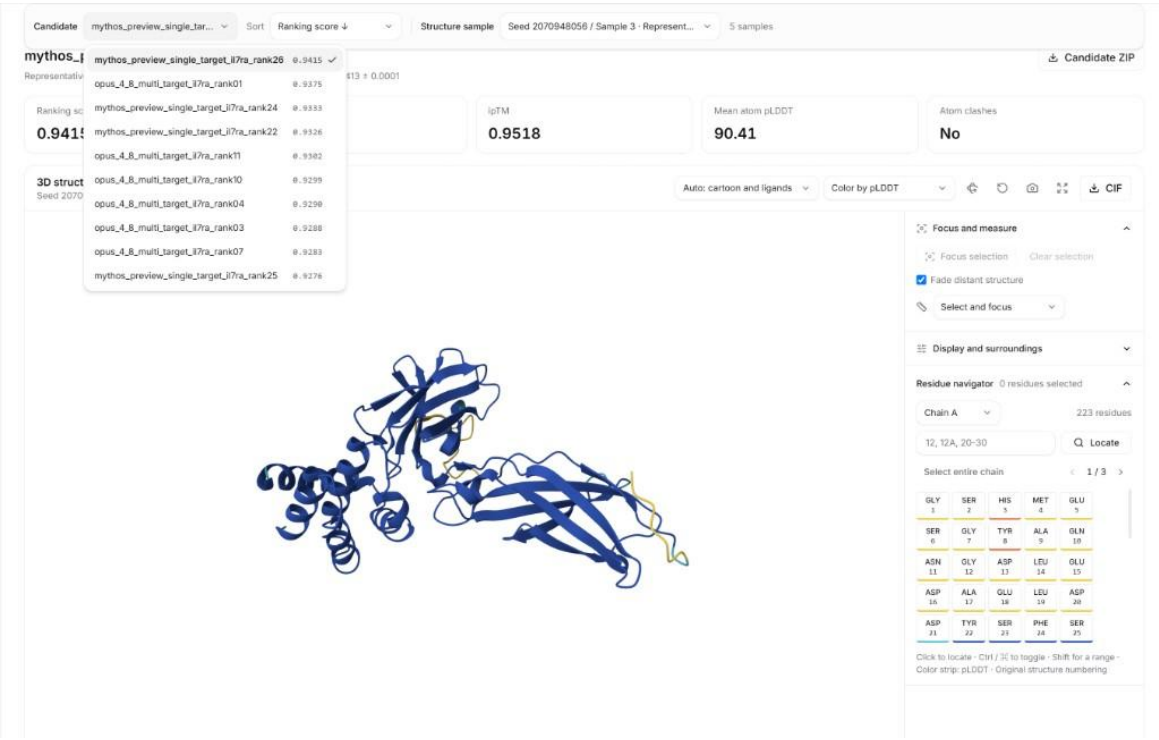

Figure A2: Interactive inspection of candidate complex structures. The view displays candidate and sample selection, prediction scores, and a 3D structural model colored by pLDDT for the top-ranked representative (ipTM 0.9518; mean atomic pLDDT 90.41). The workflow generates five structural samples per candidate.

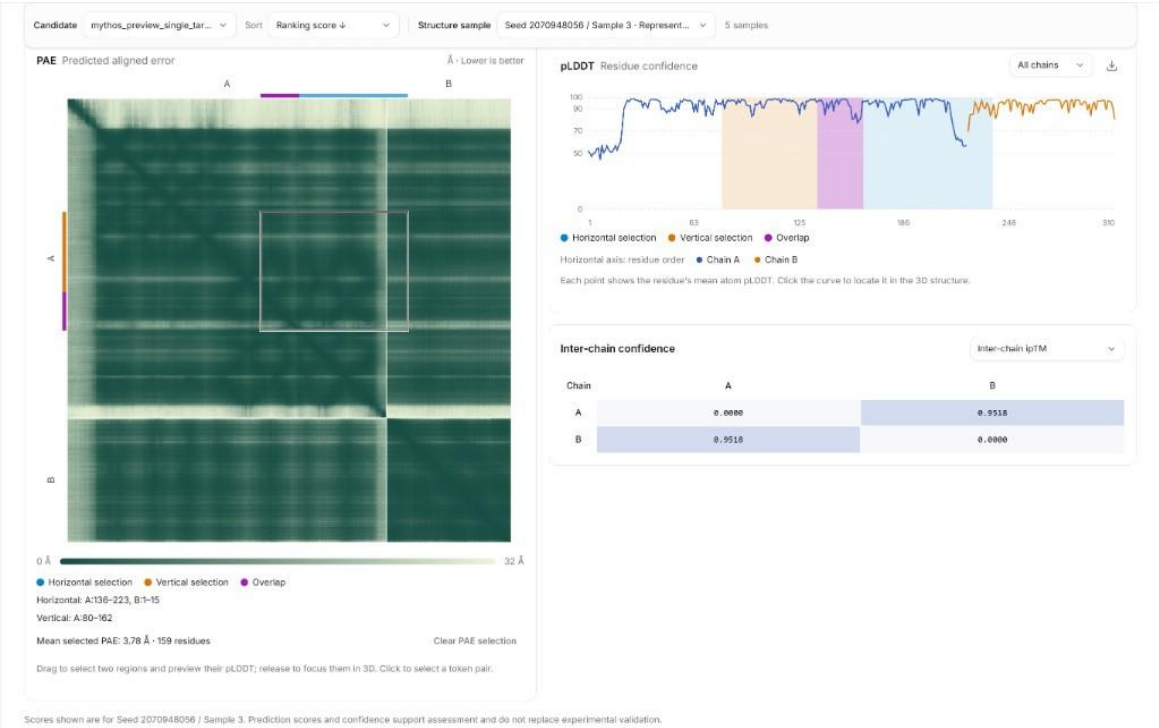

Figure A3: Linked analysis of structural confidence. The diagnostic panel displays the Predicted Aligned Error (PAE) matrix, residue-level pLDDT profile, and chain-pair ipTM scores. Selecting a region within the PAE matrix (mean PAE 3.65 Å) highlights the corresponding residues in the 3D viewer (blue, orange, and purple indicate horizontal, vertical, and overlapping selections, respectively).

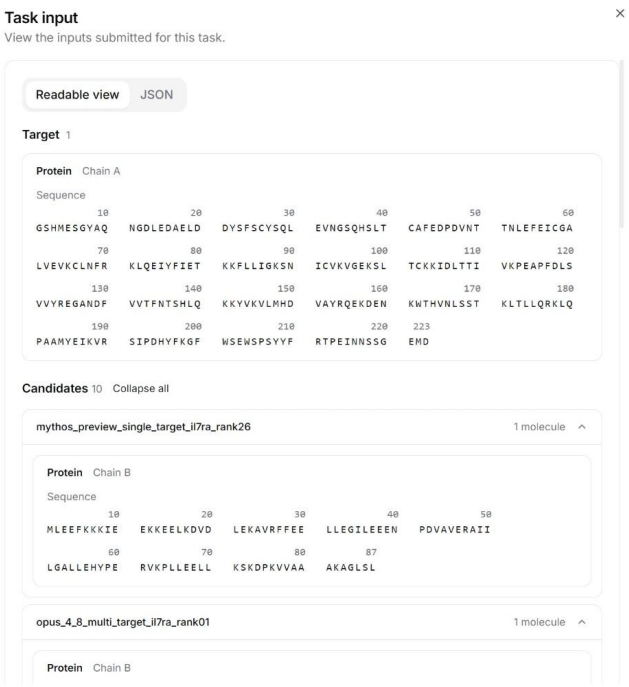

Figure A4: Inspection of submitted molecular inputs. The Melo Playground interface provides human-readable and JSON-formatted views of the input parameters, detailing molecule types, chain IDs, and sequences. This example displays a shared target (chain A, 223 amino acids) screened against ten distinct candidates (with the first candidate, chain B, expanded to show its 87-amino-acid sequence).

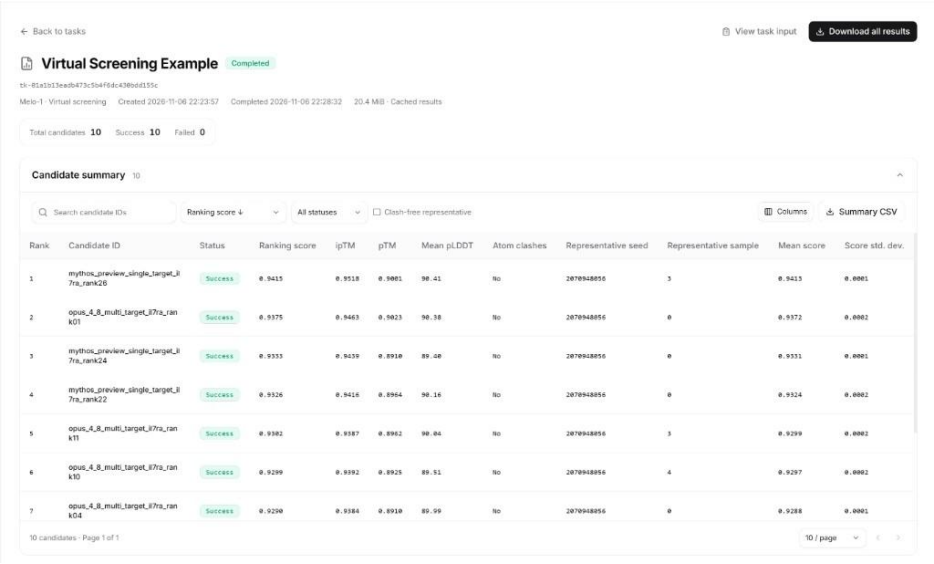

Figure A5: Summary and comparison of batch screening results. The ranked table aggregates task status, platform ranking score, ipTM, pTM, mean atomic pLDDT, clash metrics, and sample-score statistics across all ten evaluated candidates. A "success" status indicates successful computational completion rather than confirmed experimental binding.

References

[1] Josh Abramson, Jonas Adler, Jack Dunger, et al. Accurate structure prediction of biomolecular interactions with AlphaFold 3. *Nature*, 630:493–500, 2024. doi:10.1038/s41586-024-07487-w.

[2] Joseph L. Watson, David Juergens, Nathaniel R. Bennett, et al. De novo design of protein structure and function with RFdiffusion. *Nature*, 620:1089–1100, 2023. doi:10.1038/s41586-023-06415-8.

[3] Vinicius Zambaldi, David La, Alexander E. Chu, et al. De novo design of high-affinity protein binders with AlphaProteo, 2024. arXiv:2409.08022.

[4] Hannes Stark, Felix Faltings, MinGyu Choi, et al. BoltzGen: Toward universal binder design. *bioRxiv*, 2025. doi:10.1101/2025.11.20.689494.

[5] Latent Labs Team et al. Latent-X: An atom-level frontier model for de novo protein binder design, 2025. arXiv:2507.19375.

[6] Latent Labs Team et al. Drug-like antibodies with low immunogenicity in human panels designed with Latent-X2, 2025. arXiv:2512.20263.

[7] Chai Discovery Team et al. Zero-shot antibody design in a 24-well plate. *bioRxiv*, 2025. doi:10.1101/2025.07.05.663018.

[8] Nabla Bio and Surojit Biswas. De novo design of epitope-specific antibodies against soluble and multipass membrane proteins with high specificity, developability, and function. *bioRxiv*, 2025. doi:10.1101/2025.01.21.633066.

[9] Nabla Bio. JAM-2: Fully computational design of drug-like antibodies with high success rates. Technical report, Nabla Bio, 2025. Technical report.

[10] Atomelody Team. Melo-1: Accurate all-atom complex prediction with near-native interfaces and geometric validity. Technical report, Atomelody, 2026. https://www.atomelody.com/research/melo-1-report.pdf.

[11] Marie-Paule Lefranc, Christelle Pommié, Manuel Ruiz, et al. IMGT unique numbering for immunoglobulin and T cell receptor variable domains and Ig superfamily V-like domains. *Developmental & Comparative Immunology*, 27:55–77, 2003. doi:10.1016/S0145-305X(02)00039-3.

[12] Christiam Camacho, George Coulouris, Vahram Avagyan, et al. BLAST+: architecture and applications. *BMC Bioinformatics*, 10:421, 2009. doi:10.1186/1471-2105-10-421.

[13] Jasper Butcher, Rohith Krishna, Raktim Mitra, et al. De novo design of all-atom biomolecular interactions with RFdiffusion3. *bioRxiv*, 2025. doi:10.1101/2025.09.18.676967.

[14] Nathaniel Corley, Kieran Didi, Maxwell W. Kazman, Rafael I. Brent, et al. Biomolecular design through latent reasoning with RFD4-Proteina. 2026. Preprint; https://rfd4-proteina.org/assets/paper/RFD4-Proteina.pdf.

[15] Protenix Team et al. PXDesign: Fast, modular, and accurate de novo design of protein binders. *bioRxiv*, 2025. doi:10.1101/2025.08.15.670450.

[16] Kieran Didi, Zuobai Zhang, Guoqing Zhou, et al. Scaling atomistic protein binder design with generative pretraining and test-time compute, 2026. Proteina-Complexa; arXiv:2603.27950.

[17] Zeming Lin, Halil Akin, Roshan Rao, et al. Evolutionary-scale prediction of atomic-level protein structure with a language model. *Science*, 379:1123–1130, 2023. doi:10.1126/science.ade2574.

[18] EvolutionaryScale / Biohub. Esmfold2 protein structure prediction. https://biohub.ai/models/esmfold2, 2026.

[19] Edwin B. Wilson. Probable inference, the law of succession, and statistical inference. *Journal of the American Statistical Association*, 22:209–212, 1927.

[20] Lawrence D. Brown, T. Tony Cai, and Anirban DasGupta. Interval estimation for a binomial proportion. *Statistical Science*, 16:101–133, 2001.

[21] Anthropic. Claude protein binder design dataset. https://huggingface.co/datasets/Anthropic/claude-protein-binder-design, 2026.

[22] Max D. Overath et al. Predicting experimental success in de novo binder design: A meta-analysis of 3,766 experimentally characterised binders. *bioRxiv*, 2025.

[23] ByteDance Seed. Protenix-v2: Broadening the reach of structure prediction and biomolecular design. https://github.com/bytedance/Protenix/blob/main/docs/PX2.pdf, 2026.

[24] Aureka Research. Folding, reasoning, and scaling with open-source drug discovery engine, 2026.

[25] Jr. Dunbrack, Roland L. Rēs ipSAE loquuntur: What's wrong with AlphaFold's ipTM score and how to fix it. *bioRxiv*, 2025. doi:10.1101/2025.02.10.637595.

[26] Qilin Yu, Liangyue Guo, Xiayan Qin, et al. High-affinity protein binder design via flow matching and in silico maturation. *bioRxiv*, 2026. doi:10.64898/2026.01.19.700484.

[27] Luis S. Mille-Fragoso, Claudia L. Driscoll, John N. Wang, et al. Efficient generation of epitope-targeted antibodies with Germinal. *Nature Biotechnology*, 2026. doi:10.1038/s41587-026-03187-0.

[28] James Dunbar, Konrad Krawczyk, Jinwoo Leem, et al. SAbDab: the structural antibody database. *Nucleic Acids Research*, 42:D1140–D1146, 2014. doi:10.1093/nar/gkt1043.

[29] James Dunbar and Charlotte M. Deane. ANARCI: antigen receptor numbering and receptor classification. *Bioinformatics*, 32:298–300, 2016. doi:10.1093/bioinformatics/btv552.