ISSN: 2476-2377

Research Article

International Journal of Cancer Research & Therapy

Computational Design of a De Novo Binder for Claudin-6 (CLDN6)

Jingxing Chen*

Independent Researcher, USA

*Corresponding Author
Jingxing Chen, Independent Researcher, USA

Submitted: 2025, Oct 16; **Accepted**: 2025, Nov 17; **Published**: 2025, Dec 05

Citation: Chen, J. (2025). Computational Design of a De Novo Binder for Claudin-6 (CLDN6). *Int J Cancer Res Ther, 10*(2), 01-08.

Abstract

Claudin-6 (CLDN6) is a tight-junction membrane protein normally expressed in embryonic tissues but abnormally activated in several cancers, where it contributes to tumor growth and metastasis. Its restricted oncofetal expression makes CLDN6 an attractive target for selective cancer therapies, yet designing molecules that bind its small, flexible extracellular loops has proven challenging. Here, we report the first de novo computational design of a protein binder to our knowledge that specifically engages the extracellular domain of human CLDN6.

Using RFdiffusion, we guided backbone generation around a defined CLDN6 surface epitope, performing multiple independent diffusion trajectories to explore diverse binding geometries. Candidate backbones were filtered through key metrics and visual inspection. The best-scoring backbone was then optimized using ProteinMPNN to create sequences that stabilize both the binder fold and its interface with CLDN6.

The final binder was validated in silico with AlphaFold3, which predicted a well-structured CLDN6–binder complex (interface TM-score ≈ 0.73). Rosetta InterfaceAnalyzer predicted a binding free energy of –60 REU and a buried surface area of ~1200 Å², values comparable to those of antibody–antigen complexes. Computational alanine scanning identified a cluster of key interface residues (B9–B12) that contributed most to binding. We performed in silico saturation mutagenesis and discovered that a tryptophan substitution at B12 (B12W) enhanced predicted binding energy by ~7.4 REU, increased buried surface area by +129 Å², and strengthened hydrogen bonding by +0.85 bonds ($p < 0.01$).

These results show us that our designed binder is able to bind tightly and form an energetically favorable interface with CLDN6. Our binder may block or occupy these regions, thereby inhibiting the oncogenic functions of CLDN6. Beyond this specific application, the study demonstrates that a workflow of integrating backbone generation, sequence design, and energetic validation can create high-affinity protein binders for membrane protein targets.

Keywords: De Novo Protein Binder, Claudin-6, Rfdiffusion, Protein Engineering, Computational Design

1. Introduction

Claudin-6 (CLDN6) is a tetraspan transmembrane protein and member of the tight junction claudin family. It contains four transmembrane helices with two extracellular loops (ECL1 and ECL2) that mediate cell–cell contacts [1]. Unlike most other claudins, CLDN6 is essentially absent from healthy adult tissues, being expressed only during embryonic development. This cancer-specific expression makes CLDN6 an appealing therapeutic target, as interventions against CLDN6 are less likely to damage normal adult tissues [1]. Indeed, CLDN6 is currently under investigation as a target in cancer immunotherapy. However, obtaining high-affinity binders for CLDN6 has been challenging due to its membrane-bound, multispan architecture.

Traditional methods to discover protein binders involve immunization or large library screening, which can be labor-intensive and offer limited control over binder properties [2]. By contrast, computational de novo design offers a faster, more customizable route to create binders targeting specific epitopes. Recent advances in deep learning–guided protein design have markedly improved success rates in designing high-affinity binders from a target's structure alone [2]. In particular, generative models like RFdiffusion can "sculpt" a de novo protein backbone that wraps around a chosen target site, and network-based sequence

design tools like ProteinMPNN can then optimize the amino acid sequence for stability and binding [3]. By combining these tools, researchers have created novel proteins that bind targets with affinities approaching those of natural antibodies.

2. Materials and Methods

2.1. Computational Binder Design

To design a protein that could recognize CLDN6 with high specificity, we used RFdiffusion to generate new protein backbones guided by a target structure. We provided the extracellular domain of CLDN6, which includes its ECL1 and ECL2 loops, to RFdiffusion as the target input for diffusion modeling. We ran multiple independent diffusion trajectories, gradually extending the binder by two residues at a time to explore different binders and interface geometries.

From roughly one hundred generated backbones, we identified one candidate that showed the most promising interface confidence scores, which is a high iPTM, low iPAE, and favorable interface RMSD, and is supported by visual inspection in PyMOL. This backbone was then selected as the foundation for sequence design and refinement.

2.2. Sequence Optimization

We next used ProteinMPNN to design amino acid sequences for this backbone, with the goal of stabilizing both the binder's fold and its interface with CLDN6. Unlike traditional Rosetta-based sequence design, ProteinMPNN leverages neural networks trained on vast structural data to predict sequences that naturally fit a given 3D fold. This approach is faster and often produces better-packed and more realistic interfaces [2]. We generated a wide range of sequence variants and evaluated each for predicted stability, energy, and interfacial complementarity. After ranking and inspection, we selected one sequence and carried it forward for in silico validation.

2.3. Structure Prediction and Validation

To predict whether the designed binder would fold properly and bind CLDN6 as intended, we used AlphaFold3 (Multimer mode) to model the CLDN6–binder complex, and named it our wildtype. AlphaFold's predicted interface TM-score (iPTM) for the binder–CLDN6 complex was 0.73, indicating moderate confidence that the two chains interact in the predicted orientation (Figure 1) [4]. The overall predicted TM-score (pTM) for the complex was 0.83 (with the CLDN6 chain pTM ≈0.72). Notably, the binder chain alone received a very low pTM (~0.2), suggesting that the binder may be intrinsically disordered in isolation and only folds into a stable structure upon binding to CLDN6. Nonetheless, the AlphaFold3 prediction provided confidence that the designed binder can indeed engage CLDN6 at the intended epitope, even if the binder relies on the target for its full structure. We designate CLDN6 as chain A and the binder as chain B in all subsequent analyses.

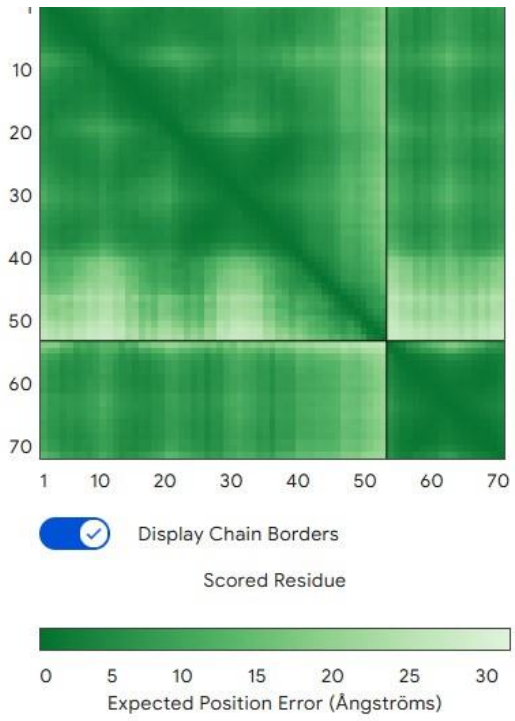


Figure 1: Predicted CLDN6–Binder Complex Structure. The AlphaFold3-Predicted Model of the CLDN6 (Chain A) and Binder (Chain B) Complex is Shown

2.4. Interface Analysis

We evaluated the CLDN6–binder interface using the Rosetta protein modeling suite to calculate key interface metrics. These metrics included the binding energy (Rosetta ΔG_separated, where a more negative value indicates stronger predicted binding), the buried interface surface area (dSASA_int), shape complementarity (SC), and the number of interfacial hydrogen bonds. We also performed computational alanine scanning on the complex to identify "hotspot" residues. In these scans, individual residues at the interface were mutated to alanine, and the change in binding free energy (ΔΔG_bind) relative to the wild-type was computed using Rosetta's ddG protocol. Residues yielding large positive ΔΔG upon alanine mutation are considered binding hot spots, which contribute significantly to interface affinity [5].

2.5. In Silico Mutagenesis

To probe structure–function relationships and potentially improve affinity, we carried out rational in silico mutagenesis on the binder. Guided by the alanine-scanning results (see Results), we focused on binder positions contacting the identified hotspot region of the epitope. Two variants that were designed and evaluated are B10K and B12W. For each variant, we used Rosetta to relax the complex and compute binding energies and interface metrics as above. Additionally, heavy-atom contacts within 4.5 Å were tallied to generate residue–residue contact maps for the wild-type and mutant interfaces, facilitating visualization of any interface remodeling. All structural models and contact analyses were examined using PyMOL and ChimeraX for visualization.

2.6. Computational Setup

All computational design and analysis steps were performed on a Linux platform using standard builds of RFdiffusion, ProteinMPNN, AlphaFold3 (multimer), and Rosetta. No experimental (wet-lab) data were generated in this study; all results are predictive and require experimental validation.

3. Results

3.1. Structure Prediction and Initial Interface Properties

We visualized the protein-binder complex in PyMOL, and the model confirmed that the binder is capable of wrapping around CLDN6's extracellular region, forming the intended interface (Figure 2).

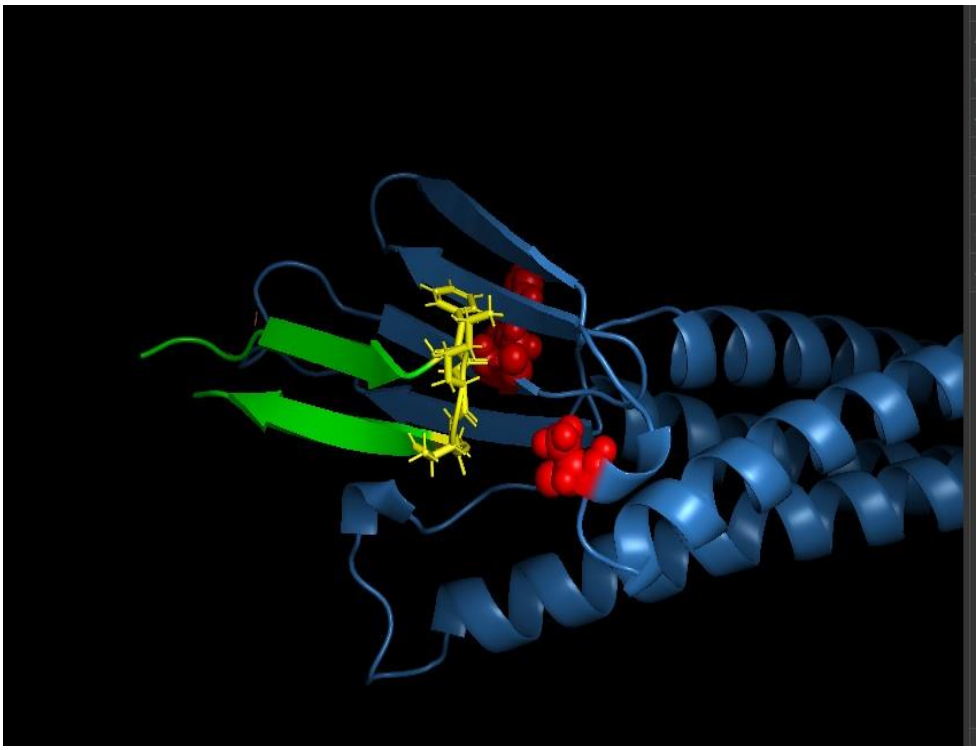


Figure 2: Design Model of the CLDN6 Binder Interface

A close-up view of the CLDN6–binder interface is shown (visualized in PyMOL). CLDN6 (blue surface) and the binder (green cartoon) interact primarily through the ECL1 region and contacts to ECL2. Key binder side chains forming the interface are shown in stick representation. The designed binder conforms snugly to the CLDN6 surface, with shape complementarity evident from the interlocking surfaces.

We next analyzed the predicted CLDN6–binder interface using Rosetta energy calculations. Across an ensemble of 40 relaxed complex models, the average Rosetta interface binding energy (ΔG_separated) was –60.05 ± 0.57 REU, placing it comfortably in the range of high-affinity protein complexes (approximately –40 to –80 REU for many antibody–antigen interfaces). This corresponds

to roughly –4.86 ± 0.08 REU per 100 Å² of buried interface area (mean buried surface area ≈1236 ± 19 Å²). About 61% of the buried surface was hydrophobic (≈755 ± 22 Å²). Notably, the total buried area (~1200 Å²) and hydrophobic fraction align with typical values for antibody–antigen interfaces of moderate to high affinity (often on the order of 1200–2000 Å² buried). The Rosetta energy decomposition is consistent with a tightly packed interface: a large favorable van der Waals term (fa_atr ≈ –1332 REU) is partly offset by modest steric repulsion (fa_rep ≈ +145 REU) and desolvation penalties (fa_sol ≈ +751 REU), while electrostatic interactions contribute favorably (fa_elec ≈ –434 REU). The complex also features an extensive hydrogen-bonding network, with ~9–11 interfacial hydrogen bonds (average ~9.6) contributing roughly –239 REU in total hbond energy. On a per-residue basis, interfacial residues (averaging ~55 residues total on both sides) contribute about –3.15 REU each to the binding energy [6]. These metrics paint a picture of a well-optimized interface dominated by shape complementarity and dispersion forces, with a supportive but not extreme electrostatic component. In summary, the designed binder forms a densely packed, desolvated interface with CLDN6, comparable in quality to natural protein complexes.

3.2. Interface Contacts and Hotspot Identification

We analyzed the interface in detail to identify which contacts are most critical for binding. Figure 3 shows a residue–residue contact map for the wild-type complex, indicating all pairwise contacts within 4.5 Å between CLDN6 (vertical axis, chain A residues) and the binder (horizontal axis, chain B residues). One main cluster is approximately A31–A80 (covering much of ECL1's N-terminal region), and the other involves CLDN6 residues ~A150–A160 (encompassing ECL2 and adjacent segments). This pattern implies that the binder simultaneously engages CLDN6's first and second extracellular loops. Such bipartite binding, wherein both ECL1 and ECL2 are contacted, likely enhances affinity by increasing the total buried surface area and providing two anchor points on the target. The contact map also reveals a contiguous stretch of binder residues (B8–B13) that contact ECL1, hinting at a focused paratope region on the binder.

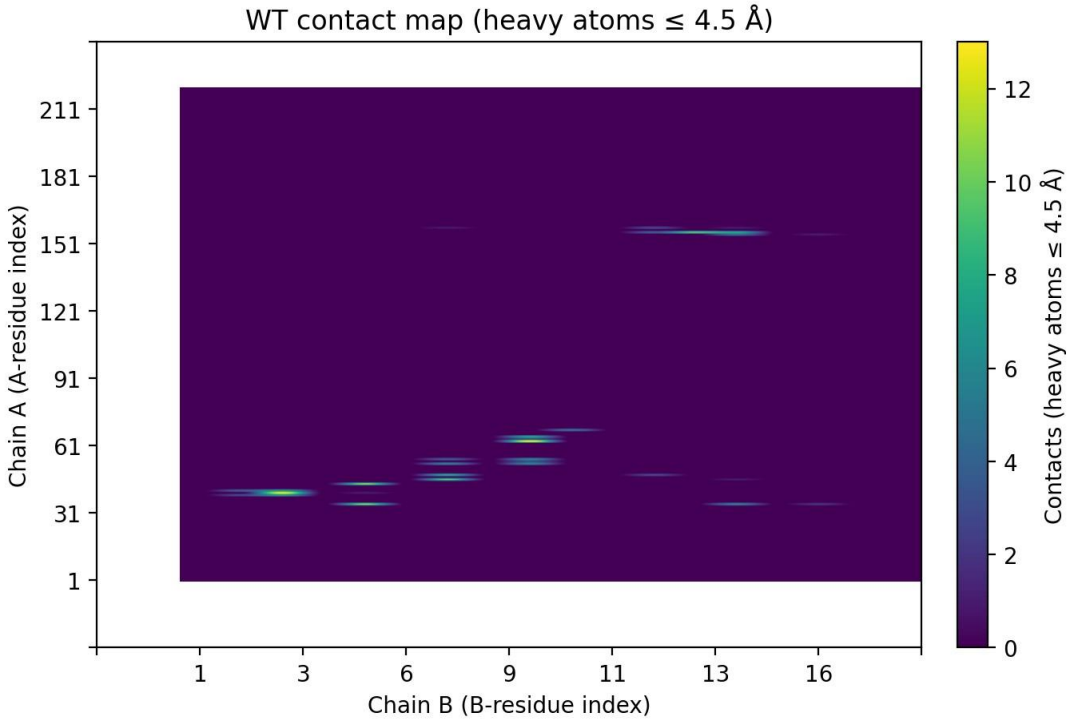


Figure 3: Contact Map of the CLDN6–Binder Interface

The heatmap depicts contacts between specific CLDN6 residues (vertical axis, chain A) and binder residues (horizontal axis, chain B) within 4.5 Å in the wild-type complex. Each filled dot indicates a contact between a particular residue pair. Two clusters of interactions are evident: (i) the binder contacts a region of CLDN6 ECL1 roughly spanning A31–A80, and (ii) the binder contacts part of ECL2 around A150–A160. This demonstrates that the de novo binder concurrently engages both extracellular loops of CLDN6. Bipartite engagement of ECL1 and ECL2 could contribute to higher affinity and specificity by enlarging the binding interface.

To quantify the contribution of individual interface residues to binding, we performed computational alanine scanning on the complex. This analysis was done through scanning the CLDN6

protein(epitope) and the binder side (paratope). For the target side, each CLDN6 residue at the interface was mutated in silico to alanine, and the change in binding energy (ΔΔG_bind) was calculated (Figure 4a). Several CLDN6 mutations significantly weakened binding. Notably, substituting Ala at CLDN6 positions A34, A63, and A76 (all in ECL1) produced ΔΔG_bind values of +0.832, +0.654, and +0.798 REU, respectively [6]. This hotspot overlaps with known functional sites on CLDN6 ECL1 that have been implicated in cancer cell signaling [7].

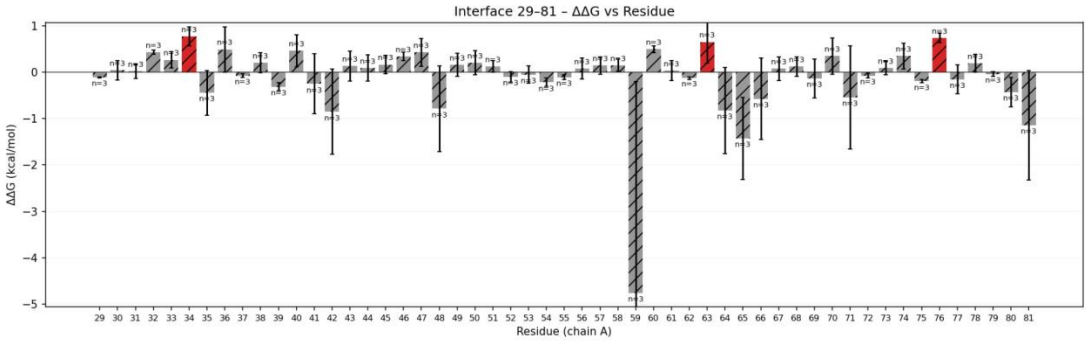


Figure 4: (a). Alanine-Scanning Analysis of Binding Hotspots. (A) Epitope Side (CLDN6): Bar graph of ΔΔG_bind upon mutating key CLDN6 ECL1 residues to alanine. Larger positive values indicate greater loss of binding affinity. CLDN6 residues A34, A63, and A76 stand out as hotspot positions; mutating any of these to Ala significantly destabilizes the complex, pinpointing a critical surface patch on ECL1.

We next scanned key binder-side residues to see which contribute most to affinity. When each binder interface residue was individually mutated to Ala, four positions showed significant drops in predicted binding energy. Mutating binder residues B9, B10, B11, or B12 to alanine increased the binding free energy by +5.062, +2.678, +1.048, and +0.729 REU, respectively. By contrast, an Ala mutation at the neighboring position B8 had only a negligible effect (+0.045 REU) [6]. These data indicate that the paratope cluster B9–B12 accounts for the majority of the binder's binding energy, essentially anchoring the binder to the CLDN6 epitope. This contiguous energetic hotspot on the binder side mirrors the contact cluster seen in the contact map (B9–B12 contacting CLDN6 ECL1) and demonstrates that our design captured a highly focused interaction surface rather than relying on diffuse contacts. Figure 4(b) illustrates the results of the alanine scan, highlighting the energetic importance of the B9–B12 region on the binder. Knowledge of these binder hotspots is valuable for subsequent affinity optimization efforts.

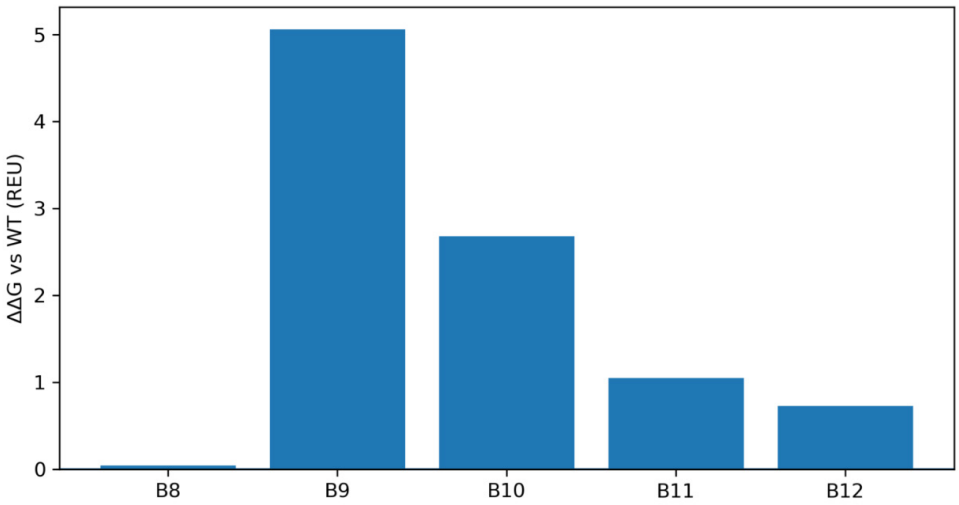


Figure 4: (b) Paratope Side (Binder): Bar graph of ΔΔG_bind upon mutating binder interface residues to alanine. Binder positions B9–B12 show the largest increases in ΔΔG from Ala substitution, which means they are the core contributors to binding energy. These results guided subsequent binder optimization efforts.

3.3. Binder Variant Design and Affinity Evaluation

Our wild-type binder showed strong predicted affinity in silico, but we wanted to see if we could push it further through rational mutation design. Based on the hotspot analysis above, we focused on residues B10 and B12. Both sit next to the key hotspot B9 and are positioned near the edge of the interface, making them good candidates for fine-tuning.

3.3.1. Variant B10K

Residue B10 lies beside hotspot B9, which interacts with an acidic patch on the CLDN6 ECL1 loop. Since this region contains several negatively charged residues (Asp and Glu), we reasoned that introducing a positively charged residue at B10 could strengthen electrostatic interactions or form new hydrogen bonds. We substituted lysine at this position, creating variant B10K. Lysine's long, flexible side chain seemed well-suited to reach into the interface and form potential salt bridges.

When we relaxed the B10K complex in Rosetta, the predicted binding energy improved slightly by about 0.2–0.3 REU, but this difference falls within the typical uncertainty of Rosetta scoring. Consistent with this, contact map analysis showed that the overall interfacial contacts in B10K were nearly identical to those of the wild-type (Figure 5A), revealing that B10K doesn't have better binding interfaces compared to the wild-type.

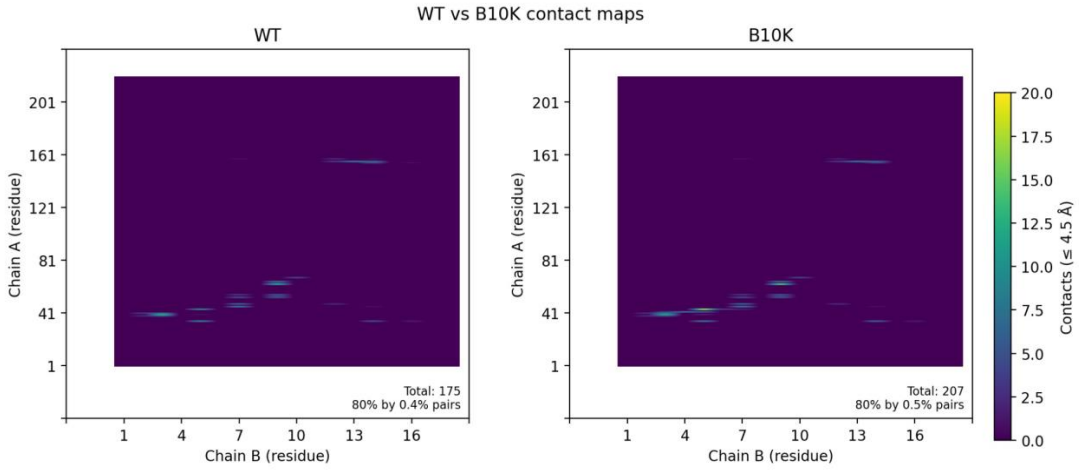


Figure 5: Effects of Binder Mutations on Interface Contacts. (A) Wild-type vs. B10K contact maps: The residue contact matrix for variant B10K (right) is virtually identical to the wild-type binder (left), indicating that introducing a Lys at binder position 10 does not substantially alter the interface contacts with CLDN6. No new salt bridge or electrostatic interaction was observed; the variant maintains the original binding mode.

3.3.2. Variant B12W Packing Optimization

Binder residue B12 is located at the periphery of the binding interface, next to the critical B9–B10 loop. In the wild-type design, B12 is only partially buried upon complex formation, where its side chain is only partially buried. This means that part of the residue remains solvent-exposed rather than fully shielded within the hydrophobic core. This type of partial burial often signals suboptimal packing, as it leaves small cavities or unpaired hydrophobic surfaces that neither contribute to binding affinity nor structural stability.

To improve this, we conducted an in silico saturation mutagenesis at position B12, evaluating all 20 amino acid substitutions. The screen revealed that aromatic or bulky hydrophobic residues consistently improved predicted binding energies, whereas polar or charged substitutions were often destabilizing. Among all variants, tryptophan (B12W) emerged as the most favorable, suggesting that its large, planar indole ring could effectively occupy a shallow pocket on the CLDN6 surface. The Trp side chain can reinforce local hydrophobic clustering and establish additional π–π or edge-to-face contacts, improving van der Waals complementarity.

The B12W variant improved Rosetta binding energetics relative to wild type. Its average binding free energy (dG_separated) was –67.46 ± 3.31 REU, compared with –60.05 ± 0.57 REU for WT, an improvement of ~–7.41 REU. The interface also grew: dSASA_int = 1383.79 ± 74.58 Å² for B12W vs 1235.85 ± 18.64 Å² for WT (+148 Å²). Both hydrophobic and polar buried surface areas increased (~+130 Å² hydrophobic; ~+18 Å² polar), consistent with Trp12 forming additional contacts. Interfacial hydrogen bonds rose from 9.60 ± 0.55 to 10.45 ± 0.82. Although the larger Trp side chain raised fa_rep (+1.48), this was balanced by a stronger fa_atr (~6.72 REU more favorable). Statistical testing confirmed the directional improvements for B12W over WT as dG_separated decreased by –7.41 REU (Welch one-sided $p \approx 1.6\times10^{-17}$, 95% CI [–8.34, –6.30]); dSASA_int increased by +148 Å² ($p \approx 5.9\times10^{-16}$, 95% CI [+123, +170]); hydrophobic and polar contributions rose

by ~+130 Å² and ~+18 Å² ($p \leq 3.0\times10^{-5}$ each); and interfacial H-bonds increased by +0.85 ($p \approx 3.3\times10^{-7}$, 95% CI [+0.55, +1.15]). Although fa_rep increased (+1.48, $p \approx 9.1\times10^{-4}$), this was offset by a more favorable fa_atr (–6.72 REU, $p \approx 1.6\times10^{-8}$). Overall, B12W creates a tighter, more extensive interface with CLDN6, with gains driven primarily by increased contact area and van der Waals attraction.

Contact map analysis further supported these findings (Figure 6). Introduction of Trp12 expanded the interfacial footprint of the binder, increasing the number of residue–residue contacts, particularly between CLDN6 residues 47–72 and binder residue 12. Collectively, these data demonstrate that a single-point substitution at the edge of the binding interface can significantly remodel packing interactions, yielding a tighter, more extensive, and energetically favorable complex.

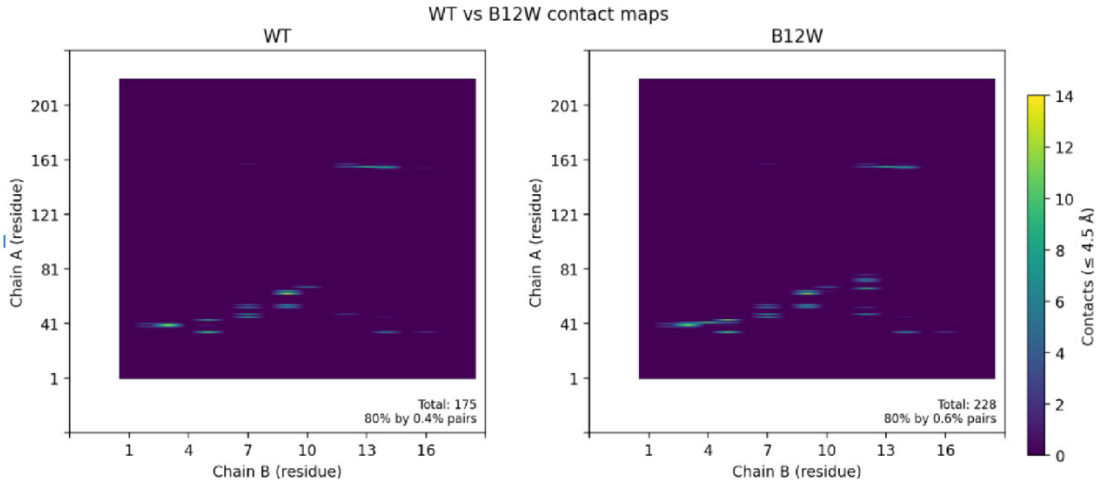
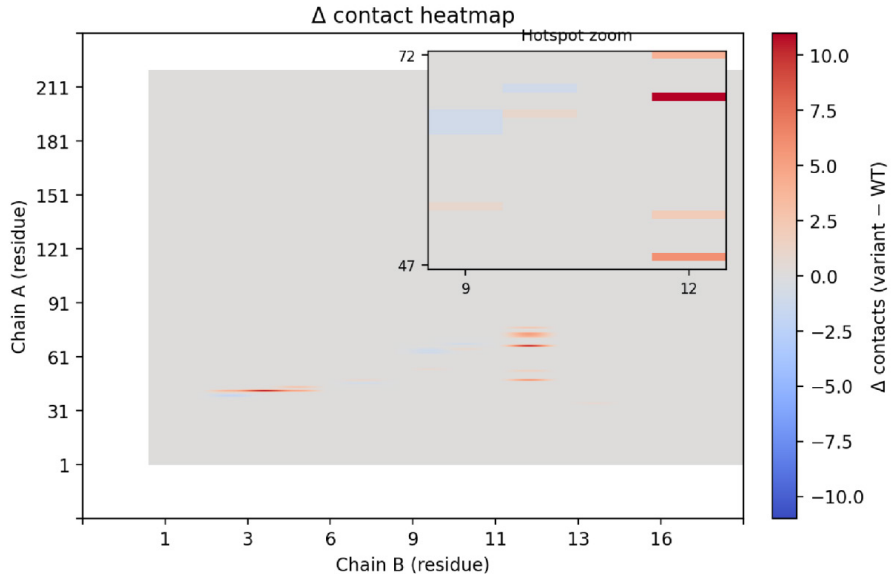


Figure 6: (A)Contact Maps Highlighting the Differences between the Wild-Type Binder and the B12W Variant. The B12W mutation introduces additional contacts with CLDN6, indicated by the greater density of contact dots. Specifically, Trp12 fills a hydrophobic crevice on CLDN6's ECL1, increasing the binder's interfacial complementarity. The broadened contact pattern for B12W correlates with its ~7.4 REU binding energy improvement over the wild-type binder.



(B) Contact map contrasting contact points between WT and B12W

Figure 6: (B) Contact Map Contrasting Contact Points between WT and B12W

4. Discussion

The CLDN6 binder could have several potential applications in oncology if wet-lab experiments validate the function of this binder. It could function as a steric blocker, attaching directly to CLDN6's extracellular loops and preventing their participation in tumor-promoting signaling or cell–cell adhesion. In particular, CLDN6's ECL2 region has been implicated in epithelial–mesenchymal transition (EMT) and metastasis blocking this surface could help inhibit cancer cell invasiveness and reduce metastatic potential [1]. Such an approach may parallel strategies used for CLDN18.2- and HER2-targeted biologics, which achieve tumor growth inhibition through ligand occlusion rather than cytotoxic payload delivery [5].

Beyond direct inhibition, the binder could serve as a targeting module across a range of therapeutic and diagnostic platforms. When fused to an Fc domain, it could engage immune effector mechanisms such as antibody-dependent cell cytotoxicity (ADCC) or complement-dependent cytotoxicity (CDC), providing antibody-like activity with a smaller molecular footprint. In an antibody–drug conjugate (ADC) configuration, the binder could guide cytotoxic payloads such as MMAE or DM1 specifically to CLDN6-positive tumor cells, minimizing off-target toxicity. This strategy aligns with the growing success of anti-CLDN6 ADCs like CLDN6–23-ADC, which have shown potent preclinical activity in xenograft models and are progressing through early clinical development [7].

The binder also holds promise in cellular immunotherapy. By incorporating its sequence into a chimeric antigen receptor (CAR) construct, engineered T cells could be directed to recognize and eliminate CLDN6-expressing cancer cells. Early-stage trials of CLDN6 CAR-T therapy (e.g., BNT211) have reported encouraging safety profiles and measurable antitumor responses [8]. Alternatively, conjugating the binder to imaging agents or nanoparticle carriers could enable precise visualization and targeted delivery of therapeutics to CLDN6-positive tumors, aiding both diagnostic imaging and surgical navigation.

For any of these applications, further optimization and validation will be required. Our design already takes advantage of sequence differences unique to CLDN6's extracellular loops to enhance selectivity, but experimental studies will be critical for confirming this. Because AlphaFold3 suggests that the binder may rely on its target for structural stability, biochemical handling may require stabilizing mutations or fusion to a carrier protein to maintain a folded state in isolation.

By combining deep generative modeling (RFdiffusion), AI-based sequence design (ProteinMPNN), and physics-based validation (Rosetta), this workflow demonstrates that it is now feasible to rationally design high-affinity, target-specific binders from scratch. The CLDN6 binder, therefore serves as an example illustrating how AI-driven protein design can accelerate the creation of new molecular tools for cancer therapy and diagnostics [9-12].

References

1. Qu, H., Jin, Q., & Quan, C. (2021). CLDN6: from traditional barrier function to emerging roles in cancers. *International journal of molecular sciences, 22*(24), 13416.
2. Bakerlab. (2024). Designing binders with the highest affinity ever reported.
3. Wang, H., Liu, C., & Deng, L. (2018). Enhanced prediction of hot spots at protein-protein interfaces using extreme gradient boosting. *Scientific reports, 8*(1), 14285.
4. Watson, J. L., Juergens, D., Bennett, N. R., Trippe, B. L., Yim, J., Eisenach, H. E., ... & Baker, D. (2023). De novo design of protein structure and function with RFdiffusion. *Nature, 620*(7976), 1089-1100.
5. Du, H., Yang, X., Fan, J., & Du, X. (2021). Claudin 6: Therapeutic prospects for tumours, and mechanisms of expression and regulation (Review). *Molecular Medicine Reports, 24*(3).
6. Fleishman, S. J., Leaver-Fay, A., Corn, J. E., Strauch, E. M., Khare, S. D., Koga, N., ... & Baker, D. (2011). RosettaScripts: a scripting language interface to the Rosetta macromolecular modeling suite. *PloS one, 6*(6), e20161.
7. McDermott, M. S., O'Brien, N. A., Hoffstrom, B., Gong, K., Lu, M., Zhang, J., ... & Slamon, D. J. (2023). Preclinical efficacy of the antibody–drug conjugate CLDN6–23-ADC for the treatment of CLDN6-positive solid tumors. *Clinical Cancer Research, 29*(11), 2131-2143. https://doi.org/10.1158/1078-0432.ccr-22-2981.
8. Mackensen, A., Haanen, J. B., Koenecke, C., Alsdorf, W., Wagner-Drouet, E., Borchmann, P., ... & Şahin, U. (2023). CLDN6-specific CAR-T cells plus amplifying RNA vaccine in relapsed or refractory solid tumors: the phase 1 BNT211-01 trial. *Nature medicine, 29*(11), 2844-2853.
9. Abramson, J., Adler, J., Dunger, J., Evans, R., Green, T., Pritzel, A., ... & Jumper, J. M. (2024). Accurate structure prediction of biomolecular interactions with AlphaFold 3. *Nature, 630*(8016), 493-500.
10. Tyka, M. D., Keedy, D. A., André, I., DiMaio, F., Song, Y., Richardson, D. C., ... & Baker, D. (2011). Alternate states of proteins revealed by detailed energy landscape mapping. *Journal of molecular biology, 405*(2), 607-618.
11. Khatib, F., Cooper, S., Tyka, M. D., Xu, K., Makedon, I., Popović, Z., ... & Players, F. (2011). Algorithm discovery by protein folding game players. *Proceedings of the National Academy of Sciences, 108*(47), 18949-18953.
12. Simon, A. G., Lyu, S. I., Laible, M., Wöll, S., Türeci, Ö., Şahin, U., ... & Quaas, A. (2023). The tight junction protein claudin 6 is a potential target for patient-individualized treatment in esophageal and gastric adenocarcinoma and is associated with poor prognosis. *Journal of Translational Medicine, 21*(1), 552.