ConTact: Contact-First Antibody CDR Design via Explicit Interface Reasoning

Mansoor Ahmed [1 2] **Spencer VonBank** [3] **Nadeem Taj** [4] **Sujin Lee** [2] **Naila Jan** [5] **Murray Patterson** [1]

Abstract

Computational antibody CDR design methods condition on antigen structure to generate binding loops. Yet, the existing architectures conflate two fundamentally distinct sub-problems: identifying which CDR positions will contact the antigen, and selecting amino acids at those positions. This forces models to learn contact reasoning implicitly through uniform message passing, diluting antigen signal across all positions equally. We introduce CONTACT, a contact-then-act architecture that explicitly decomposes CDR design into three cascaded stages: learning surface complementarity fingerprints, predicting CDR-antigen contacts, and injecting contact-gated antigen features into the prediction head. A distance-biased cross-attention module encodes geometric priors favoring spatial neighbors, while a contact-weighted cross-entropy loss concentrates gradient signal on binding-critical positions. On the CHIMERA-BENCH dataset, CONTACT achieves the lowest backbone RMSD on every split (a 5 to 6% improvement over the best baseline) and the best fraction of native contacts, interface RMSD, and epitope F1 on the antigen-fold and temporal splits, while remaining competitive on the harder epitope-group split. The source code is available at: https://github.com/mansoor181/ConTact.git

1. Introduction

Antibodies bind antigens through their complementarity-determining regions (CDRs), six hypervariable loops whose sequence and structure determine binding specificity (Chothia & Lesk, 1987). Computational CDR design methods condition on antigen structure to generate sequences and backbone conformations for these loops (Luo et al., 2022; Kong et al., 2023a;b; Wu et al., 2025b). Yet a growing body of evidence shows that existing methods largely fail to leverage antigen information. The predictions remain nearly unchanged when the antigen is removed (Li et al., 2025), and BLOSUM substitution matrices explain model outputs as well as learned likelihoods (Uçar & Sormanni, 2025; Chinery et al., 2024).

We argue that a fundamental cause is architectural, where the current methods conflate two distinct sub-problems into a single prediction head. The first sub-problem is *where* the CDR will contact the antigen, i.e., which CDR positions form binding interactions. The second is *what* amino acids to place at those positions, given the local chemistry of the binding partner. Equivariant GNNs such as MEAN (Kong et al., 2023a) and RAAD (Wu et al., 2025b) propagate antigen information through uniform message passing that treats all antigen residues equivalently. Diffusion-based methods like DiffAb (Luo et al., 2022) concatenate antibody and antigen residues into a flat graph with only a fragment-type embedding to distinguish them. Even dyMEAN (Kong et al., 2023b), which uses a shadow paratope mechanism and edge distance prediction for contact-aware graph construction, does not use predicted contacts to modulate sequence and structure prediction. In all cases, the model must simultaneously discover which positions are binding-relevant and what residues belong there, with a uniform cross-entropy loss that allocates equal learning capacity to every position.

The CDR-antigen interface is inherently sparse. A CDR-H3 of length 10–25 typically forms only 5–15 contacts with the antigen, and the amino acid identity at contact positions is directly constrained by the chemistry of the binding partner. For example, hydrophobic pockets select for complementary hydrophobic CDR residues, while charged patches favor oppositely charged side chains (Chothia & Lesk, 1987; Li et al., 2025). The non-contact positions are primarily constrained by backbone geometry and loop stability. Therefore, treating these two classes of positions equally wastes learning capacity on the less informative non-contact positions.

We propose CONTACT, an encoder-decoder architecture that decomposes CDR design into three explicit stages, addressing the *where* before the *what*. First, the model learns surface complementarity fingerprints that characterize the

[1]Georgia State University, Atlanta, USA [2]Georgia Institute of Technology, Atlanta, USA [3]DePauw University, Indiana, USA [4]University of Engineering and Technology, Lahore, Pakistan [5]Accenture, Chesterbrook, USA. Correspondence to: Mansoor Ahmed <mahmed76@student.gsu.edu>.

Accepted at the 2026 Workshop on Generative and Agentic AI for Biology (ICML 2026)

local binding environment at each CDR position, inspired by molecular surface fingerprints (Gainza et al., 2020; 2023). Second, it predicts which CDR positions will contact the antigen using a supervised contact predictor. Third, it selectively injects local antigen features into the CDR representation, gated by the predicted contact confidence, so that antigen information flows preferentially to binding-critical positions. A distance-biased cross-attention module provides geometric inductive bias by favoring spatial neighbors, and a contact-weighted cross-entropy loss concentrates gradient signal on positions the model identifies as contacts.

Our main contributions are:

1. We identify the conflation of contact identification and sequence prediction as a structural limitation of existing CDR design architectures, and propose the *contact-first* design paradigm that decomposes these sub-problems into an explicit three-stage cascade.
2. We introduce a contact-gated injection mechanism that selectively routes antigen information to binding-relevant CDR positions, preventing noise from distant antigen residues at non-contact positions.
3. We demonstrate on CHIMERA-BENCH that CONTACT achieves the lowest backbone RMSD on every split and the best fnat, interface RMSD, and epitope F1 on the antigen-fold and temporal splits.

2. Related Work

Equivariant GNN methods. MEAN (Kong et al., 2023a) introduced multi-channel equivariant attention with alternating intra-chain and inter-chain segment layers for CDR design. dyMEAN (Kong et al., 2023b) extended this with a shadow paratope mechanism that predicts inter-chain edge distances for dynamic graph construction, making it the closest existing work to contact-aware CDR design. RAAD (Wu et al., 2025b) uses a multi-relational graph representation and a contrastive specificity objective for binding specificity. CONTACT differs from all three in that it uses predicted contacts to directly modulate the sequence, structure, and binding prediction through gated injection and position-specific loss weighting.

Diffusion and flow methods. DiffAb (Luo et al., 2022) models CDR generation as a joint diffusion process over coordinates, orientations, and amino acid types. AbFlowNet (Abir et al., 2025) extends this with flow matching and trajectory balance loss. AbMEGD (Chen et al., 2025) and RADAb (Wang et al., 2024) add retrieval-augmented and multi-expert components. dyAb (Tan et al., 2025) applies flow matching with structure relaxation. FlowDesign (Wu et al., 2025a) identifies that standard Gaussian priors are poorly suited for CDR generation and replaces them with data-driven prior distributions. All these methods treat antigen conditioning as a flat concatenation of antibody and antigen residues with fragment-type embeddings, applying uniform attention without distinguishing contact from non-contact positions. CONTACT addresses this as a complementary limitation of the conditioning mechanism itself, and not the prior distribution.

Antigen conditioning failures. Multiple studies have documented that existing CDR design methods fail to effectively use antigen information. Li et al. (2025) showed that predictions remain nearly unchanged when the antigen is removed. Uçar & Sormanni (2025) demonstrated that BLOSUM substitution matrices explain model outputs as well as learned likelihoods. Chinery et al. (2024) found that simple computational methods can outperform deep learning in generating diverse, binder-enriched antibody libraries. RefineGNN (Jin et al., 2022b), which receives no antigen input, achieves the second-best binding metrics on CHIMERA-BENCH, further corroborating this failure. The contact-first decomposition in CONTACT directly targets this problem by providing an explicit, supervised pathway for antigen information to reach the sequence and structure prediction heads.

Predict-then-design paradigms. The idea of predicting binding-relevant features before designing sequences has precedent in broader protein design. MaSIF-seed (Gainza et al., 2023) predicts favorable interaction sites on molecular surfaces using learned surface fingerprints, then designs binders targeting those sites. RFdiffusion (Watson et al., 2023) generates protein backbones first, then designs sequences with ProteinMPNN (Dauparas et al., 2022). CONTACT applies a similar strategy at the residue-contact level, stated as: *predict which CDR positions will contact the antigen, then condition sequence and structure design on those predictions*. Unlike MaSIF-seed, which operates on molecular surfaces in a separate pipeline, CONTACT performs contact prediction and sequence design end-to-end within a single differentiable architecture.

3. Preliminaries

3.1. Task Definition

We adopt the formulation from CHIMERA-BENCH (Ahmed et al., 2026). Given an antigen structure $A = \{(s_j, \mathbf{x}_j) \mid j \in V_A\}$, an epitope specification $E \subseteq V_A$, and an antibody framework $F = \{(s_i, \mathbf{x}_i) \mid i \in V_{\text{FR}}\}$, the task is to design CDR residues $R = \{(s_k, \mathbf{x}_k) \mid k \in V_{\text{CDR}}\}$ that maximize the conditional likelihood subject to epitope contact

constraints:

$$R^* = \arg\max_R \; p_\theta\big(R \mid A, E, F\big), \quad \text{s.t. } \mathcal{C}(R, A) \neq \emptyset, \tag{1}$$

where each residue has amino acid type $s_k \in \{1, \ldots, 20\}$ and Cα coordinate $\mathbf{x}_k \in \mathbb{R}^3$. We denote by $\mathcal{C}(R, A) = \{j \in V_A \mid \exists\, k \in V_{\text{CDR}} : \|\mathbf{x}_k - \mathbf{x}_j\| < d_c\}$ the set of antigen residues contacted within cutoff d_c. Specifically, the goal is to design CDRs (H1, H2, H3, L1, L2, L3), where the most variable loop and primary determinant of antigen specificity is CDR-H3 (Chothia & Lesk, 1987).

3.2. Graph Construction

We represent the antibody-antigen complex as a heterogeneous graph $\mathcal{G} = (V, \mathcal{E})$. The node set $V = V_{\text{HC}} \cup V_{\text{LC}} \cup V_A \cup V_{\text{glob}} \cup V_{\text{vn}}$ contains residue nodes from the heavy chain (V_{HC}), light chain (V_{LC}), and antigen (V_A), three global delimiter tokens ($V_{\text{glob}} = \{\text{BOH}, \text{BOL}, \text{BOA}\}$), and N_{vn} virtual nodes (Sestak et al., 2026). Each residue node i carries amino acid type $s_i \in \{1, \ldots, 20\}$ and four backbone atom coordinates $\mathbf{X}_i = [\mathbf{x}_i^{\text{N}}, \mathbf{x}_i^{\text{C}\alpha}, \mathbf{x}_i^{\text{C}}, \mathbf{x}_i^{\text{O}}] \in \mathbb{R}^{4\times 3}$.

The edge set $\mathcal{E}$ is partitioned into 10 typed subsets that capture different structural relationships between amino acids. Within each chain, we construct *radial edges* connecting all pairs within a Cα distance cutoff, *sequential edges* linking residues separated by one or two positions in primary sequence, and *KNN edges* connecting each residue to its nearest spatial neighbors. Across chains, we add *inter-chain radial edges* and *inter-chain KNN edges* that enable direct communication between antibody and antigen residues. Moreover, three *global-to-chain edges* connect the delimiter tokens to their respective chains. Two *virtual node edge types* connect each virtual node bidirectionally to all epitope and all CDR residues. This creates a two-hop shortcut between epitope and CDR, directly addressing the over-squashing problem (Alon & Yahav, 2021) where information from distant epitope residues dilutes through many layers of sequential message passing.

Each edge (i, j) carries a feature vector $\mathbf{e}_{ij}$ encoding edge type (one-hot), relative position in local coordinate frames, pairwise distance RBFs between backbone atom pairs, a quaternion encoding of relative backbone orientation, and local frame direction features. Virtual node edges use learnable feature vectors rather than geometric features.

3.3. Contact Definition

We define a CDR residue k as contacting the antigen if its Cα atom lies within 8 Å of any antigen Cα atom (Xue et al., 2015; Ovchinnikov et al., 2014):

$$c_k = \mathbb{1}\left[\min_{j \in V_A} \|\mathbf{x}_k - \mathbf{x}_j\| < 8\ \text{Å}\right] \tag{2}$$

The binary labels $c_k \in \{0, 1\}$ serve as supervision for the contact prediction stage and as weights in the contact-weighted sequence loss.

4. Method

ConTact consists of three components: (i) an EGNN encoder that performs E(3)-equivariant message passing over the heterogeneous graph, (ii) a distance-biased cross-attention module that combines CDR and antigen representations with spatial priors, and (iii) a three-stage contact-first decoder that cascades complementarity fingerprinting, contact prediction, and contact-guided sequence generation. Figure 1 illustrates the full architecture.

4.1. Feature Encoding

Each residue i in the antibody-antigen complex is represented by a feature vector $\mathbf{f}_i$ composed of five groups. **Amino acid identity**: a one-hot encoding of the residue type, masked to the zero vector for all CDR positions during training to prevent trivial teacher-forcing. **Backbone distance RBFs**: intraresidue bond lengths (N–Cα, Cα–C, C–O) each expanded into Gaussian basis functions:

$$\phi_{\text{rbf}}(d)_m = \exp\left(-\frac{(d - \mu_m)^2}{2\varsigma^2}\right), \quad m = 1, \ldots, M, \tag{3}$$

where μ_m are uniformly spaced centers and ς is the basis width. **Backbone angles**: bond angles and dihedral angles (ϕ, ψ, ω), each encoded as sine-cosine pairs. **Local frame directions**: unit vectors along the three local coordinate axes defined by the N-Cα-C backbone triangle. **Sinusoidal position embedding**: encoding of the residue index within its chain at multiple frequency scales.

A segment type indicator distinguishes heavy chain, light chain, and antigen residues. A dual-path MLP processes geometric and chemical features through separate pathways with SiLU activations, fuses the outputs, and projects to embedding dimension d:

$$\mathbf{h}_i^{(0)} = \text{MLP}_{\text{fuse}}\big([\text{MLP}_{\text{geom}}(\mathbf{f}_i^{\text{geom}}),\ \text{MLP}_{\text{chem}}(\mathbf{f}_i^{\text{chem}})]\big) \tag{4}$$

Epitope residues (those in E) receive an additional learnable embedding $\mathbf{e}_{\text{epi}}$ added to their representation, providing an explicit signal that these residues are part of the designated binding site.

4.2. EGNN Encoder

The encoder applies multiple relation-aware E(3)-equivariant GNN layers (Satorras et al., 2021) on graph $\mathcal{G}$. The virtual nodes (VN) with learnable feature vectors and learnable coordinates participate in message passing through the VN-to-epitope and VN-to-CDR edge types.

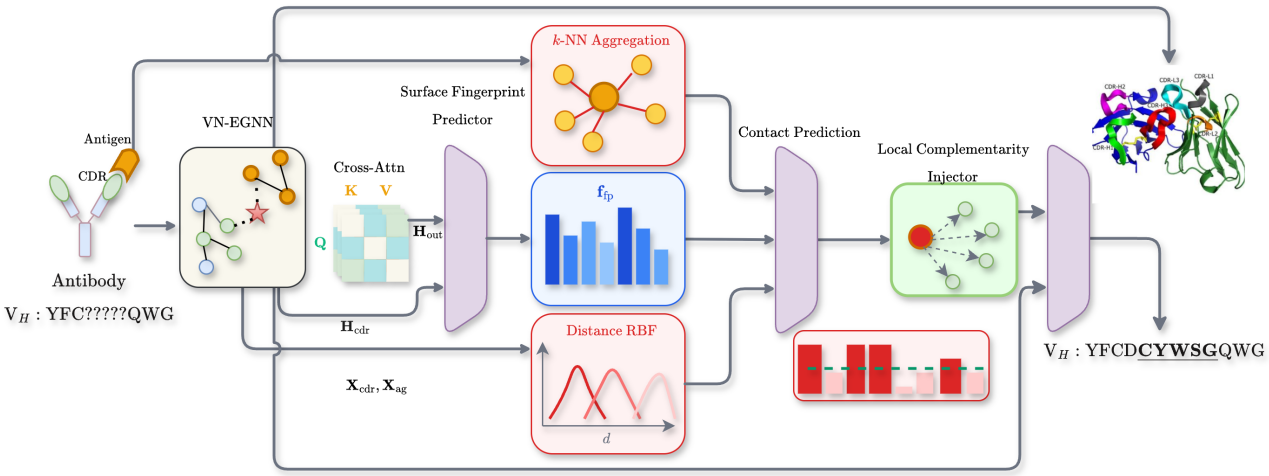

Figure 1. **ConTact architecture.** The encoder maps residue features through a GNN to produce per-residue embeddings and updated coordinates. CDR and antigen embeddings are combined via distance-biased cross-attention. The three-stage decoder cascades complementarity fingerprinting, contact prediction, and contact-guided local complementarity injection.

Because each virtual node connects to all epitope and CDR residues, information flows from any epitope residue to any CDR residue in exactly two message-passing steps. Without virtual nodes, this information must traverse the graph via sequential edges, suffering from over-squashing (Alon & Yahav, 2021) at bottleneck residues.

Each layer l updates node features and coordinates simultaneously. For edge (i, j) of type t, the message function takes the concatenation of sender and receiver embeddings, an outer product geometry term, and the edge features:

$$\mathbf{m}_{ij}^{(l)} = \text{MLP}_{\text{msg}}^{(l)}\Big([\mathbf{h}_i^{(l)},\ \mathbf{h}_j^{(l)},\ \text{vec}\big(\Delta\mathbf{x}_{ij}(\Delta\mathbf{x}_{ij})^\top\big),\ \mathbf{e}_{ij}]\Big), \tag{5}$$

where $\Delta\mathbf{x}_{ij} = \mathbf{x}_i^{(l)} - \mathbf{x}_j^{(l)}$ and $\text{vec}(\cdot)$ flattens the 3×3 outer product matrix into a 9-dimensional vector. The entries of $\Delta\mathbf{x}_{ij}(\Delta\mathbf{x}_{ij})^\top$ are dot products of displacement components, which are invariant to rotations, translations, and reflections.

The model aggregates messages from all edge types with type-specific linear projections and updates node features via a residual connection:

$$\mathbf{h}_i^{(l+1)} = \mathbf{h}_i^{(l)} + \text{MLP}_{\text{node}}^{(l)}\left(\left[\mathbf{h}_i^{(l)},\ \sum_t \mathbf{W}_t^{(l)} \sum_{j\in\mathcal{N}_t(i)} \mathbf{m}_{ij}^{(l)}\right]\right), \tag{6}$$

where $\mathbf{W}_t^{(l)}$ is a type-specific projection matrix and $\mathcal{N}_t(i)$ denotes the neighbors of node i under edge type t. Coordinates are updated equivariantly by adding a weighted sum of displacement vectors:

$$\mathbf{x}_i^{(l+1)} = \mathbf{x}_i^{(l)} + \sum_t \frac{1}{|\mathcal{N}_t(i)|} \sum_{j\in\mathcal{N}_t(i)} \Delta\mathbf{x}_{ij}\cdot\text{MLP}_t^{\text{coord},(l)}(\mathbf{m}_{ij}^{(l)}), \tag{7}$$

where $\text{MLP}_t^{\text{coord},(l)}$ produces a scalar weight. The product of a displacement vector and a scalar computed from invariant inputs is equivariant by construction. After all layers, the encoder produces residue embeddings $\mathbf{h} \in \mathbb{R}^{N\times D}$ and updated backbone coordinates $\hat{\mathbf{X}}$.

4.3. Distance-Biased Cross-Attention

After encoding, we extract CDR node embeddings $\mathbf{H}_{\text{cdr}} \in \mathbb{R}^{L\times D}$ and antigen node embeddings $\mathbf{H}_{\text{ag}} \in \mathbb{R}^{M\times D}$. We note that standard cross-attention computes alignment scores purely from learned feature similarity, making no distinction between an antigen residue 5 Å from the CDR and one 30 Å away. ConTact adds a Gaussian spatial bias that encodes a geometric inductive bias, since binding contacts are necessarily spatial neighbors.

We project queries $\mathbf{Q} = \mathbf{H}_{\text{cdr}}\mathbf{W}_Q$ and keys $\mathbf{K} = \mathbf{H}_{\text{ag}}\mathbf{W}_K$ into H attention heads. The attention score between CDR position i and antigen position j is:

$$\alpha_{ij}^{(h)} = \text{softmax}_j\left(\frac{(\mathbf{q}_i^{(h)})^\top \mathbf{k}_j^{(h)}}{\sqrt{D_h}} + \beta_{ij}\right), \tag{8}$$

where the distance bias β_{ij} decays as a Gaussian function of the Cα–Cα distance between the updated coordinates from the encoder:

$$\beta_{ij} = \exp\left(-\frac{d_{ij}^2}{2\sigma^2}\right), \quad d_{ij} = \|\hat{\mathbf{x}}_i^{\text{C}\alpha} - \hat{\mathbf{x}}_j^{\text{C}\alpha}\|_2, \tag{9}$$

with bandwidth σ. The Gaussian decay is near 1.0 for residues within van der Waals contact distance, drops to $e^{-2} \approx 0.14$ at 2σ (approximately the contact threshold), and becomes negligible beyond 3σ. The bias is shared

across heads to provide a consistent spatial prior, while the learned query-key projections specialize to different aspects of the binding interaction.

The output for each CDR position concatenates the multi-head weighted sum with the original CDR embedding:

$$\mathbf{o}_i = \Big[\mathbf{h}_i^{\text{cdr}},\ \Big\|_{h=1}^{H} \sum_{j=1}^{M} \alpha_{ij}^{(h)} \mathbf{v}_j^{(h)}\Big], \tag{10}$$

where $\|$ denotes concatenation and $\mathbf{v}_j^{(h)} = \mathbf{h}_j^{\text{ag}} \mathbf{W}_V^{(h)}$ are value projections. The skip connection preserves the CDR embedding for downstream stages that primarily need structural context.

4.4. Decoder

STAGE 1: COMPLEMENTARITY FINGERPRINTING

The first decoder stage compresses the cross-attention output into a compact representation of the local surface complementarity at each CDR position. The binding interactions follow chemical complementarity patterns (hydrophobic-hydrophobic, charge-charge, donor-acceptor) that can be captured in a low-dimensional fingerprint, analogous to molecular fingerprints in cheminformatics. Given the cross-attention output $\mathbf{o}_i$ for CDR position i, an MLP produces a fingerprint vector:

$$\mathbf{f}_i = \text{MLP}_{\text{fp}}(\mathbf{o}_i) \tag{11}$$

We train the fingerprint predictor with a contrastive loss that aligns CDR fingerprints with their local antigen environment. The antigen embeddings are projected to the fingerprint space via a separate MLP. For each contacting CDR position i, its nearest antigen neighbor serves as the positive, and distant antigen residues (beyond $1.5\times$ the contact cutoff) as negatives. The loss combines an InfoNCE (Chen et al., 2020) term with a margin penalty on non-contact positions:

$$p_{ij^*} = \frac{\exp(\mathbf{f}_i^\top \mathbf{g}_{j^*}/\tau_{\text{fp}})}{\exp(\mathbf{f}_i^\top \mathbf{g}_{j^*}/\tau_{\text{fp}}) + \sum_{k\in\mathcal{N}_i} \exp(\mathbf{f}_i^\top \mathbf{g}_k/\tau_{\text{fp}})}, \tag{12}$$

$$\mathcal{L}_{\text{fp}} = -\frac{1}{|\mathcal{C}|} \sum_{i\in\mathcal{C}} \log p_{ij^*} + \frac{1}{2}\mathcal{L}_{\text{margin}}, \tag{13}$$

where $\mathcal{C}$ is the set of contacting CDR positions, $j^* = \arg\min_j d_{ij}$ is the nearest antigen neighbor, $\mathbf{g}_j$ are the projected antigen fingerprints, $\mathcal{N}_i$ are the distant negatives, and $\mathcal{L}_{\text{margin}}$ penalizes non-contact positions whose maximum antigen similarity exceeds a margin threshold. The contrastive objective ensures that the fingerprint captures *what kind of binding environment* a CDR position faces, conditioning the subsequent contact prediction stage.

STAGE 2: CONTACT PREDICTION

The second stage predicts which CDR positions will form contacts with the antigen. This is the central component of the contact-first decomposition. Existing methods leave contact identification as an implicit byproduct of message passing (Kong et al., 2023a; Wu et al., 2025b) or graph construction (Kong et al., 2023b). CONTACT supervises contact prediction explicitly and uses the predictions to gate downstream information flow.

For each CDR position i, we aggregate features from its K nearest antigen neighbors (by Cα distance) using softmax-weighted attention over distances:

$$\mathbf{a}_i = \sum_{j\in\text{KNN}_K(i)} w_{ij}\, \mathbf{h}_j^{\text{ag}}, \quad w_{ij} = \text{softmax}_j(-d_{ij}), \tag{14}$$

where the softmax weights assign higher importance to closer neighbors. The contact predictor takes a concatenation of four inputs: the CDR embedding, the KNN-aggregated antigen features, an RBF encoding of the minimum distance to any antigen residue, and the complementarity fingerprint from Stage 1:

$$\hat{c}_i = \sigma\big(\text{MLP}_{\text{ct}}\big([\mathbf{h}_i^{\text{cdr}},\ \mathbf{a}_i,\ \phi_{\text{rbf}}(d_i^{\text{min}}),\ \mathbf{f}_i]\big)\big), \tag{15}$$

where $d_i^{\text{min}} = \min_{j\in V_A} \|\hat{\mathbf{x}}_i - \hat{\mathbf{x}}_j\|$ is the minimum Cα distance to any antigen residue and σ denotes the sigmoid function. Including the fingerprint $\mathbf{f}_i$ from Stage 1 creates a cascaded dependency, so that the quality of contact prediction depends on the learned complementarity representation.

The contact predictor outputs a soft probability $\hat{c}_i \in [0, 1]$ rather than hard binary predictions. The sigmoid output already provides smooth gradients, and hard thresholding introduced training instability in preliminary experiments. We supervise the contact predictor with a focal binary cross-entropy loss (Lin et al., 2017) that addresses the inherent class imbalance, where non-contact positions typically outnumber contacts by 3–5$\times$:

$$\mathcal{L}_{\text{contact}} = -\frac{1}{L}\sum_{i=1}^{L} (1-\hat{p}_i)^\gamma \left[c_i \log \hat{c}_i + (1-c_i)\log(1-\hat{c}_i)\right], \tag{16}$$

where $c_i \in \{0, 1\}$ is the ground-truth contact label (Equation (2)), $\hat{p}_i$ denotes the predicted probability of the correct class, and γ is the focusing parameter. The factor $(1-\hat{p}_i)^\gamma$ down-weights well-classified examples, concentrating the learning signal on hard, ambiguous positions near the contact boundary.

STAGE 3: CONTACT-GUIDED LOCAL COMPLEMENTARITY INJECTION

The third stage uses the predicted contact confidence $\hat{c}_i$ from Stage 2 to selectively inject local antigen information into

the CDR embeddings. Antigen features should influence CDR representations primarily at positions the model predicts will form contacts, while non-contact positions should rely mainly on their backbone geometry context.

For each CDR position i, we aggregate features from K-nearest antigen neighbors using learned attention weights:

$$\mathbf{h}_i^{\text{local}} = \sum_{j \in \text{KNN}_K(i)} \text{softmax}_j\left(\mathbf{w}_a^\top \mathbf{h}_j^{\text{ag}}\right) \mathbf{h}_j^{\text{ag}} \tag{17}$$

The contact confidence from Stage 2 directly gates the injection. The enriched CDR embedding combines the original representation with the contact-gated antigen information:

$$\mathbf{h}_i^{\text{enriched}} = \mathbf{h}_i^{\text{cdr}} + \hat{c}_i \cdot \text{MLP}_{\text{proj}}(\mathbf{h}_i^{\text{local}}) \tag{18}$$

At non-contact positions ($\hat{c}_i \approx 0$), the antigen contribution is effectively zeroed out, preventing noise from distant antigen residues. At contact positions ($\hat{c}_i \approx 1$), the full projected local complementarity is injected. This single-gate design uses the contact predictor's output directly as the information bottleneck, avoiding the redundancy of additional learned gates.

The final representation for the sequence head concatenates the enriched embedding with a soft-masked cross-attention output:

$$\mathbf{z}_i = [\mathbf{h}_i^{\text{enriched}},\ m_i \cdot \mathbf{o}_i^{\text{attn}}], \quad m_i = \epsilon + (1-\epsilon)\,\hat{c}_i \quad , \tag{19}$$

where $\mathbf{o}_i^{\text{attn}}$ is the attention output from Section 4.3 and $\epsilon = 0.15$ is a minimum weight floor. Rather than fully zeroing attention at non-contact positions, the floor preserves a residual antigen signal that allows the model to recover from imperfect contact predictions.

Contact-Weighted Sequence Head

The sequence head maps the final representation $\mathbf{z}_i$ to amino acid logits $\boldsymbol{\ell}_i \in \mathbb{R}^{|\mathcal{V}|}$ via an MLP, where $\mathcal{V}$ includes the 20 standard amino acids plus special tokens (masked during inference). We apply label smoothing ($\epsilon_{\text{ls}} = 0.1$) and a contact-weighted variant of cross-entropy that allocates more learning capacity to positions predicted to form binding contacts:

$$\mathcal{L}_{\text{seq}} = -\frac{1}{L}\sum_{i=1}^{L} w_i\, \text{CE}_{\epsilon_{\text{ls}}}(\boldsymbol{\ell}_i, y_i), \tag{20}$$

where y_i is the ground-truth amino acid at position i and the position-specific weight is:

$$w_i = 1 + \alpha \cdot \hat{c}_i \tag{21}$$

The hyperparameter α controls the relative up-weighting of contact positions. This reweighting follows from the observation that standard cross-entropy distributes learning capacity uniformly, treating a non-contact glycine at the loop apex the same as a contact-forming tryptophan buried in an antigen pocket. By up-weighting contacts, the model receives stronger gradient signal at precisely the positions where amino acid identity is most constrained by the antigen.

At inference, the predicted amino acid at each position is $\hat{s}_i = \arg\max_a \ell_i^a$. The contact predictions $\hat{c}_i$ can also be inspected to verify which positions the model believes form contacts.

4.5. Training Objective

The full training objective combines seven loss terms:

$$\begin{aligned}\mathcal{L} = \mathcal{L}_{\text{seq}} + \lambda_{\text{coord}}\mathcal{L}_{\text{coord}} + \lambda_{\text{contact}}\mathcal{L}_{\text{contact}} + \lambda_{\text{fp}}\mathcal{L}_{\text{fp}} \\ + \lambda_{\text{pair}}\mathcal{L}_{\text{pair}} + \lambda_{\text{dock}}\mathcal{L}_{\text{dock}} + \lambda_{\text{aux}}\mathcal{L}_{\text{aux}}\end{aligned} \tag{22}$$

The coordinate loss $\mathcal{L}_{\text{coord}}$ is a smooth-ℓ_1 (Huber) loss on predicted versus true Cα coordinates for CDR positions:

$$\mathcal{L}_{\text{coord}} = \frac{1}{L}\sum_{k \in V_{\text{CDR}}} \text{smooth}_{\ell_1}\left(\hat{\mathbf{x}}_k^{\text{C}\alpha} - \mathbf{x}_k^{\text{C}\alpha,\text{true}}\right) \tag{23}$$

The pairing loss $\mathcal{L}_{\text{pair}}$ is an InfoNCE contrastive loss that matches mean-pooled CDR and antigen embeddings within the batch, treating cognate pairs as positives:

$$\mathcal{L}_{\text{pair}} = -\frac{1}{B}\sum_{i=1}^{B} \log \frac{\exp(\bar{\mathbf{h}}_i^{\text{cdr}} \cdot \bar{\mathbf{h}}_i^{\text{ag}}/\tau_p)}{\sum_{k=1}^{B}\exp(\bar{\mathbf{h}}_i^{\text{cdr}} \cdot \bar{\mathbf{h}}_k^{\text{ag}}/\tau_p)} \tag{24}$$

The docking loss $\mathcal{L}_{\text{dock}}$ penalizes the minimum predicted Cα distance from each CDR residue to epitope atoms when it exceeds a cutoff, encouraging the predicted backbone to dock near the epitope. The auxiliary loss $\mathcal{L}_{\text{aux}}$ is a CDR feature reconstruction regularizer that prevents representation collapse. All loss weights λ are determined by hyperparameter sweeps using Weights & Biases (W&B).

4.6. Germline-Prior Fusion

CDR-H3 sequences inherit a positional signature from V(D)J recombination, particularly at the conserved J-anchor (IMGT positions 113–117), where a few germline-encoded residues dominate (Lefranc et al., 2003). ConTact captures this signature with a log-prior fused additively into the sequence head. We tabulate a categorical distribution $p_{\text{germline}}(a \mid J, \text{pos})$ over amino acids conditioned on the ANARCI-assigned J-gene and IMGT position, estimated from training-set counts and a backoff to the position marginal when a cell has fewer than three counts. The fused logits are

$$\tilde{\boldsymbol{\ell}}_i = \boldsymbol{\ell}_i + \lambda \log p_{\text{germline}}(\cdot \mid J_i, \text{pos}_i), \tag{25}$$

Table 1. CDR-H3 design on CHIMERA-BENCH, antigen-fold split. Best in **bold**, second-best underlined. Perplexity (PPL) is shown only where available; for diffusion and flow models it is undefined.

Method	AAR↑	CAAR↑	PPL↓	RMSD↓	fnat↑	iRMSD↓	DockQ↑	epiF1↑
MEAN (Kong et al., 2023a)	0.40	0.20	**2.79**	1.66	0.62	1.42	0.70	0.78
RAAD (Wu et al., 2025b)	0.38	0.20	3.06	1.65	0.62	1.42	0.70	0.78
dyMEAN (Kong et al., 2023b)	0.37	0.20	3.35	2.06	0.61	1.79	0.66	0.75
RefineGNN (Jin et al., 2022b)	0.26	0.12	8.17	2.59	0.60	1.35	0.68	0.80
DiffAb (Luo et al., 2022)	0.23	0.10	–	2.24	0.52	2.10	0.58	0.65
AbFlowNet (Abir et al., 2025)	0.23	0.13	–	2.70	0.56	2.46	0.60	0.65
AbMEGD (Chen et al., 2025)	0.24	0.13	–	2.25	0.55	2.13	0.59	0.65
dyAb (Tan et al., 2025)	0.28	0.17	–	2.40	0.56	1.81	0.65	0.72
RADAb (Wang et al., 2024)	0.25	0.12	–	7.79	0.51	4.78	0.57	0.65
AbODE (Verma et al., 2023)	0.27	0.18	–	13.17	0.16	4.00	0.39	0.40
AbDockGen (Jin et al., 2022a)	0.24	0.10	7.89	3.65	0.46	2.41	0.57	0.75
CONTACT (ours)	**0.41**	**0.21**	3.25	**1.58**	**0.69**	**1.31**	**0.73**	**0.82**

where λ is a single learnable scalar shared across positions. The prior is built from training data alone, so it introduces no test-time leakage, and it converges to $\lambda \approx 0.88$ on all three splits. Because the prior modifies only the sequence logits, it leaves the predicted backbone unchanged, and its effect is concentrated at the conserved positions the antigen does not constrain.

5. Experiments

5.1. Setup

Dataset and metrics. We evaluate on CHIMERA-BENCH (Ahmed et al., 2026), comprising 2,922 antibody-antigen complexes, across three pre-defined splits. The antigen-fold split holds out antigen structural folds, the epitope-group split clusters complexes by epitope so that test epitopes are unseen during training, and the temporal split trains on older structures and tests on newer ones. Each split uses roughly 2,338 training and 292 test complexes.

We report the results for all heavy chain CDRs (H1, H2, H3) across different metrics. Sequence quality is measured by amino acid recovery (AAR), contact AAR (CAAR, restricted to positions within 8 Å of the antigen), and perplexity (PPL). Structural quality is measured by Cα RMSD. Binding quality is measured by fraction of native contacts (fnat), interface RMSD (iRMSD), DockQ (Basu & Wallner, 2016), and epitope F1. All interface metrics use symmetric Cα–Cα contacts at 8 Å restricted to CDR residues.

Baselines. We compare against eleven baselines spanning different architectural families: equivariant GNNs (RAAD (Wu et al., 2025b), MEAN (Kong et al., 2023a), dyMEAN (Kong et al., 2023b)), diffusion and flow models (DiffAb (Luo et al., 2022), AbFlowNet (Abir et al., 2025), AbMEGD (Chen et al., 2025), dyAb (Tan et al., 2025), RADAb (Wang et al., 2024)), conjoined ODEs (AbODE (Verma et al., 2023)), and autoregressive models (RefineGNN (Jin et al., 2022b), AbDockGen (Jin et al., 2022a)). All models are retrained on CHIMERA-BENCH dataset with their original hyperparameters.

Implementation details. CONTACT has 9.68M trainable parameters. Specifically, the feature encoder projects 108D input features to 32D embeddings. The EGNN encoder uses 4 layers with 256D hidden features and 3 virtual nodes. The cross-attention module uses 2 heads with bandwidth $\sigma = 4$ Å. The complementarity fingerprint is 32-dimensional, while the contact predictor MLP has hidden layers of 256 and 128 units with LayerNorm. The loss weights are $\lambda_{\text{coord}} = 0.598$, $\lambda_{\text{contact}} = 1.763$, $\lambda_{\text{fp}} = 0.020$, $\lambda_{\text{pair}} = 0.103$, $\lambda_{\text{dock}} = 0.233$, $\lambda_{\text{aux}} = 0.200$, determined by Weights & Biases sweep. Furthermore, the contact weight $\alpha = 4.47$ and focal loss $\gamma = 2$. We train with Adam (lr $= 6.31 \times 10^{-4}$, exponential decay $\gamma_{\text{lr}} = 0.944$ per epoch), gradient clipping at 0.5, batch size 8, and dropout 0.1. It takes approximately 1.6 hours on a single NVIDIA H100 80GB GPU for one hyperparameter sweeping experiment.

5.2. Results and Discussion

Tables 1 to 3 present the performance comparison of CONTACT with the baselines across the three splits. CONTACT leads on structure and binding on the antigen-fold and temporal splits, is competitive on the epitope-group split, and maintains a lead on AAR throughout all splits. The full results for all heavy-chain CDRs for all three splits are provided in Table 4.

Structural quality. CONTACT produces the most accurate backbones, especially on the harder and most variable CDR-H3. It achieves the lowest RMSD on every split (1.58 Å on antigen-fold, 1.72 Å on temporal, 1.90 Å on epitope-group), improving over the best GNN baseline by 5% on antigen-fold (1.58 vs MEAN 1.66 Å) and by 6% on temporal (1.72 vs RAAD 1.81 Å). Coordinate supervi-

Table 2. CDR-H3 design on CHIMERA-BENCH, epitope-group split. Best in **bold**, second-best underlined.

Method	AAR↑	CAAR↑	PPL↓	RMSD↓	fnat↑	iRMSD↓	DockQ↑	epiF1↑
MEAN	**0.42**	0.21	**3.00**	2.01	0.48	1.66	0.63	0.67
RAAD	0.38	0.21	3.31	1.95	0.49	1.64	0.64	0.67
RefineGNN	0.23	0.17	7.80	3.07	**0.62**	**1.55**	**0.71**	**0.76**
DiffAb	0.22	0.11	–	2.64	0.48	2.27	0.58	0.56
AbFlowNet	0.22	0.13	–	2.70	0.49	2.37	0.57	0.55
AbMEGD	0.22	0.11	–	2.76	0.50	2.41	0.57	0.56
dyAb	0.27	0.12	–	3.31	0.35	2.27	0.54	0.50
RADAb	0.22	0.09	–	12.28	0.47	5.82	0.57	0.57
AbODE	0.31	**0.22**	–	16.40	0.08	5.12	0.34	0.20
AbDockGen	0.25	0.10	7.41	3.97	0.35	2.72	0.52	0.63
CONTACT (ours)	0.41	0.19	3.25	**1.92**	0.50	1.62	0.64	0.68

Table 3. CDR-H3 design on CHIMERA-BENCH, temporal split. Best in **bold**, second-best underlined.

Method	AAR↑	CAAR↑	PPL↓	RMSD↓	fnat↑	iRMSD↓	DockQ↑	epiF1↑
MEAN	0.39	**0.22**	3.31	1.84	0.60	1.44	0.71	0.73
RAAD	0.38	0.20	3.24	1.81	0.61	1.46	0.71	0.74
dyMEAN	0.35	0.20	3.28	2.32	0.50	2.00	0.63	0.64
RefineGNN	0.23	0.13	7.91	2.85	0.65	**1.38**	**0.73**	**0.78**
DiffAb	0.23	0.12	–	2.46	0.51	2.18	0.61	0.61
AbFlowNet	0.23	0.13	–	2.67	0.52	2.33	0.61	0.61
dyAb	0.29	0.15	–	2.89	0.51	1.92	0.64	0.65
RADAb	0.22	0.10	–	11.57	0.48	7.91	0.57	0.57
AbODE	0.37	0.19	–	15.63	0.07	4.65	0.35	0.19
AbDockGen	0.23	0.09	7.55	3.89	0.37	2.52	0.55	0.68
CONTACT (ours)	**0.40**	**0.22**	**3.09**	**1.72**	**0.66**	**1.38**	**0.73**	0.77

sion over the equivariant encoder anchors the backbone, and removing it collapses every structural and interface metric (Table 5). Contact guidance then refines the geometry toward the binding interface, and supplying ground-truth contacts raises structural and binding quality further, which shows that contact localization is the operative lever for interface geometry.

Binding quality. On antigen-fold and temporal, CONTACT achieves the best fnat (0.69 and 0.69), iRMSD (1.31 and 1.39 Å), and DockQ (0.73 and 0.74). On the epitope-group split the binding lead narrows. RefineGNN attains the best fnat (0.62) and DockQ (0.71) there despite receiving no antigen input, confirming that backbone geometry alone carries substantial information about interface contacts (Li et al., 2025), and CONTACT is second on those metrics (fnat 0.58, DockQ 0.67). CONTACT matches or surpasses RefineGNN on binding while additionally conditioning on the antigen, as reflected in its higher epitope F1 on the structured splits.

Epitope awareness. CONTACT achieves the best epitope F1 on antigen-fold (0.82, surpassing MEAN at 0.78 by 5%) and ties RefineGNN for the best on temporal (0.78). On epitope-group, where test epitopes are held out by cluster, RefineGNN leads (0.76) and CONTACT is second (0.69). The contact-gated injection and the contact-weighted sequence loss are the components that move epitope F1 in the ablations, so the epitope-awareness gains track the contact-first mechanism rather than the antigen branch alone.

Sequence recovery. A position-wise marginal that predicts the most frequent amino acid per IMGT position from the training set already reaches an AAR of 0.40 (Table 2), so sequence recovery on CDR-H3 is bounded by a strong positional prior that all current methods operate close to. CONTACT reaches an AAR of 0.40 to 0.41, matching MEAN, by fusing the germline prior explicitly into the sequence head. Contact AAR (CAAR) stays at 0.19 to 0.22 across splits and tracks the baselines, and the germline prior leaves CAAR essentially unchanged. The gap between AAR and CAAR shows that the binding-relevant positions are precisely the ones the positional prior cannot resolve. The contact-weighted loss concentrates gradient on these positions and produces the best binding metrics, while recovering the exact antigen-specific residue from Cα-level geometry remains the open bottleneck.

5.3. Ablation Study

Table 5 isolates the contribution of each contact-first component on the epitope-group split. The full row here tracks the

Table 4. Per-CDR heavy-chain results on the **antigen-fold** test split. Best values are in **bold**, second-best are underlined.

Method	AAR↑			RMSD↓			iRMSD↓			fnat↑			EpiF1↑		
	H1	H2	H3	H1	H2	H3	H1	H2	H3	H1	H2	H3	H1	H2	H3
RAAD	0.69	**0.60**	0.38	0.54	0.49	1.65	0.32	**0.33**	1.42	0.63	0.59	0.62	0.92	**0.93**	0.78
MEAN	0.68	0.54	0.40	0.77	1.01	1.66	0.48	0.61	1.42	0.57	0.52	0.62	0.87	0.86	0.78
dyMEAN	0.67	0.53	0.37	0.99	1.37	2.06	0.64	0.99	1.79	0.55	0.40	0.61	0.84	0.78	0.75
DiffAb	0.59	0.26	0.23	0.84	0.90	2.24	0.50	0.70	2.10	0.47	0.50	0.52	0.81	0.86	0.65
AbFlowNet	0.57	0.28	0.23	0.80	0.82	2.70	0.52	0.79	2.46	0.43	0.43	0.56	0.84	0.83	0.65
AbMEGD	0.61	0.28	0.24	0.82	0.71	2.25	0.51	0.72	2.13	0.45	0.49	0.55	0.84	0.84	0.65
RADAb	0.65	0.36	0.25	3.18	0.86	7.79	1.31	0.71	4.78	0.44	0.50	0.51	0.82	0.84	0.65
dyAb	0.52	0.41	0.28	1.70	1.72	2.40	0.79	1.09	1.81	0.51	0.34	0.56	0.83	0.71	0.72
RefineGNN	0.64	0.37	0.26	2.58	2.16	2.59	1.40	0.92	1.35	0.29	0.38	0.60	0.80	0.84	0.80
AbODE	0.45	0.47	0.27	4.82	7.27	13.17	1.95	3.43	4.00	0.39	0.20	0.16	0.45	0.57	0.40
CONTACT (ours)	**0.70**	0.59	**0.41**	**0.50**	**0.47**	**1.58**	**0.30**	0.37	**1.31**	**0.64**	**0.59**	**0.69**	**0.93**	0.88	**0.82**

Table 5. Component and oracle ablations on CHIMERA-BENCH (CDR-H3, epitope-group split). Each row removes one component from the full model or supplies an oracle signal. Best in **bold**.

Variant	AAR	CAAR	RMSD	fnat	DockQ	epiF1
CONTACT (full)	0.41	0.19	1.91	0.57	0.66	0.70
− antigen	0.41	0.18	1.91	0.52	0.65	0.68
− cascade	0.41	0.18	1.91	0.54	0.65	0.69
− contact loss	0.41	0.20	1.91	0.53	0.65	0.70
− contact weighting	0.41	0.19	1.93	0.52	0.64	0.68
− coordinate loss	0.41	0.20	6.26	0.05	0.37	0.12
Oracle contacts	0.41	0.19	**1.84**	**0.57**	**0.67**	**0.72**

converged CONTACT model in Table 2 (AAR 0.41, RMSD 1.92), so the absolute scale is comparable while the relative effects are the object of study.

Antigen conditioning and the cascade. Removing the antigen branch lowers epitope F1 from 0.70 to 0.68 and fnat from 0.57 to 0.52, and removing the three-stage cascade lowers them to 0.69 and 0.54. The antigen pathway and its staged routing contribute specifically to interface quality.

Contact supervision and weighting. Two mechanisms carry the contact signal into the design. Removing the auxiliary contact-prediction loss leaves epitope F1 unchanged and lowers fnat slightly to 0.53, while setting the contact weight to zero, so the sequence loss treats every position equally, lowers fnat to 0.52 and epitope F1 to 0.68. Contact weighting of the sequence objective is the more important of the two, and it is the single change inside the sequence head that improves binding, which supports the contact-first decomposition.

Coordinate supervision. Removing the coordinate loss collapses every structural and interface metric, with RMSD rising from 1.9 to 6.3 Å and fnat falling to 0.05, while AAR is untouched. Structure and sequence are decoupled in the model, and the coordinate loss holds the entire interface subsystem up.

Contact oracle. Supplying ground-truth contacts in place of the predicted ones gives the best RMSD, fnat, DockQ, and epitope F1 in the study. This establishes the ceiling of the contact-first design and shows that the remaining headroom on binding is bounded by contact-prediction accuracy rather than by the conditioning mechanism.

Limitations. Some limitations of this work include not having a fully accurate contact predictor and struggling to translate structural and binding improvements into better sequence predictions. Future work should therefore target backbone geometry directly through richer antigen representations such as side-chain geometry and surface electrostatics, and should couple the predicted backbone into the sequence head so that improved geometry can translate into antigen-specific residue choices. The contact prediction accuracy could be improved with pre-training by employing a multi-step training scheme.

6. Conclusion

In this paper, we propose CONTACT, which decomposes antibody CDR design into explicit contact prediction followed by contact-guided sequence generation. The three-stage cascade (complementarity fingerprinting, contact prediction, contact-gated injection) provides a supervised pathway for antigen information to reach the sequence head at binding-relevant positions. Experiments on CHIMERA-BENCH across three splits demonstrate that this contact-first decomposition achieves the lowest backbone RMSD on every split and the best interface quality on the antigen-fold and temporal splits, along with consistent performance on the sequence recovery.

Impact Statement

This paper presents work whose goal is to advance computational antibody design. Designed sequences require extensive experimental validation before any therapeutic application. We see no specific negative societal consequences that must be highlighted.

Acknowledgments

The work used the Jetstream2 Supercomputer at Indiana University through allocation CIS251200 from the Advanced Cyberinfrastructure Coordination Ecosystem: Services & Support (ACCESS) program (Boerner et al., 2023), which is supported by National Science Foundation grants #2138259, #2138286, #2138307, #2137603, and #2138296.

References

Abir, A. R., Shahgir, H. S., Ratul, M. R. Z., Tahmid, M. T., Steeg, G. V., and Dong, Y. AbFlowNet: Optimizing antibody-antigen binding energy via Diffusion-GFlowNet fusion. *arXiv preprint arXiv:2505.12358*, 2025.

Ahmed, M., Taj, N., Khan, I. U., Venkateswara, H., and Patterson, M. CHIMERA-bench: A benchmark dataset for epitope-specific antibody design. In *ICLR 2026 Workshop on Generative and Experimental Perspectives for Biomolecular Design*, 2026.

Alon, U. and Yahav, E. On the bottleneck of graph neural networks and its practical implications. In *International Conference on Learning Representations*, 2021.

Basu, S. and Wallner, B. DockQ: a quality measure for protein-protein docking models. *PloS one*, 11(8): e0161879, 2016.

Boerner, T. J., Deems, S., Furlani, T. R., Knuth, S. L., and Towns, J. Access: Advancing innovation: NSF's advanced cyberinfrastructure coordination ecosystem: Services & support. In *Practice and experience in advanced research computing 2023: Computing for the common good*, pp. 173–176. 2023.

Chen, J., Cai, X., Wu, J., and Hu, W. Antibody design and optimization with multi-scale equivariant graph diffusion models for accurate complex antigen binding. *arXiv preprint arXiv:2506.20957*, 2025.

Chen, T., Kornblith, S., Norouzi, M., and Hinton, G. A simple framework for contrastive learning of visual representations. In *International conference on machine learning*, pp. 1597–1607. PmLR, 2020.

Chinery, L., Hummer, A. M., Mehta, B. B., Akbar, R., Rawat, P., Slabodkin, A., Le Quy, K., Lund-Johansen, F., Greiff, V., Jeliazkov, J. R., and Deane, C. M. Simple computational methods can outperform deep learning in designing diverse, binder-enriched antibody libraries. *bioRxiv*, 2024. doi: 10.1101/2024.03.26.586756.

Chothia, C. and Lesk, A. M. Canonical structures for the hypervariable regions of immunoglobulins. *Journal of molecular biology*, 196(4):901–917, 1987.

Dauparas, J., Anishchenko, I., Bennett, N., Bai, H., Ragotte, R. J., Milles, L. F., Wicky, B. I., Courbet, A., de Haas, R. J., Bethel, N., et al. Robust deep learning–based protein sequence design using proteinmpnn. *Science*, 378 (6615):49–56, 2022.

Gainza, P., Sverrisson, F., Monti, F., Rodola, E., Boscaini, D., Bronstein, M., and Correia, B. Deciphering interaction fingerprints from protein molecular surfaces using geometric deep learning. *Nature Methods*, 17(2):184–192, 2020.

Gainza, P., Wehrle, S., Van Hall-Beauvais, A., Marchand, A., Scheck, A., Harteveld, Z., Buckley, S., Ni, D., Tan, S., Sverrisson, F., et al. De novo design of protein interactions with learned surface fingerprints. *Nature*, 617 (7959):176–184, 2023.

Jin, W., Barzilay, R., and Jaakkola, T. Antibody-antigen docking and design via hierarchical structure refinement. In *International Conference on Machine Learning*, pp. 10217–10227. PMLR, 2022a.

Jin, W., Wohlwend, J., Barzilay, R., and Jaakkola, T. Iterative refinement graph neural network for antibody sequence-structure co-design. In *International Conference on Learning Representations*, 2022b.

Kong, X., Huang, W., and Liu, Y. Conditional antibody design as 3D equivariant graph translation. In *International Conference on Learning Representations*, 2023a.

Kong, X., Huang, W., and Liu, Y. End-to-end full-atom antibody design. In *International Conference on Machine Learning*, pp. 17409–17429. PMLR, 2023b.

Lefranc, M.-P., Pommié, C., Ruiz, M., Giudicelli, V., Foulquier, E., Truong, L., Thouvenin-Contet, V., and Lefranc, G. IMGT unique numbering for immunoglobulin and T cell receptor variable domains and Ig superfamily V-like domains. *Developmental & Comparative Immunology*, 27(1):55–77, 2003.

Li, Y., Lang, Y., Xu, C., Zhou, Y., Pang, Z., and Greisen, P. J. Benchmarking inverse folding models for antibody CDR sequence design. *PLOS ONE*, 20(6):e0324566, 2025. doi: 10.1371/journal.pone.0324566.

Lin, T.-Y., Goyal, P., Girshick, R., He, K., and Dollár, P. Focal loss for dense object detection. In *IEEE International Conference on Computer Vision*, pp. 2980–2988, 2017.

Luo, S., Su, Y., Peng, X., Wang, S., Peng, J., and Ma, J. Antigen-specific antibody design and optimization with diffusion-based generative models for protein structures. *Advances in Neural Information Processing Systems*, 35: 9754–9767, 2022.

Ovchinnikov, S., Kamisetty, H., and Baker, D. Robust and accurate prediction of residue-residue interactions across protein interfaces using evolutionary information. *eLife*, 3:e02030, 2014.

Satorras, V. G., Hoogeboom, E., and Welling, M. E(n) equivariant graph neural networks. In *International conference on machine learning*, pp. 9323–9332. PMLR, 2021.

Sestak, F., Schneckenreiter, L., Brandstetter, J., Hochreiter, S., Mayr, A., and Klambauer, G. VN-EGNN: E(3)-equivariant graph neural networks with virtual nodes enhance protein binding site identification. *Journal of Cheminformatics*, 18:11, 2026. doi: 10.1186/s13321-025-01127-9.

Tan, C., Zhang, Y., Gao, Z., Huang, Y., Lin, H., Wu, L., Wu, F., Blanchette, M., and Li, S. Z. dyab: Flow matching for flexible antibody design with Alphafold-driven prebinding antigen. In *Proceedings of the AAAI Conference on Artificial Intelligence*, volume 39, pp. 782–790, 2025.

Uçar, T. and Sormanni, P. BLOSUM is all you learn—generative antibody models reflect evolutionary priors. *bioRxiv*, pp. 2025–10, 2025.

Verma, Y., Heinonen, M., and Garg, V. AbODE: Ab initio antibody design using conjoined ODEs. In *International Conference on Machine Learning*, pp. 35037–35050. PMLR, 2023.

Wang, Z., Ji, Y., Tian, J., and Zheng, S. Retrieval augmented diffusion model for structure-informed antibody design and optimization. *arXiv preprint arXiv:2410.15040*, 2024.

Watson, J. L., Juergens, D., Bennett, N. R., Trippe, B. L., Yim, J., Eisenach, H. E., Ahern, W., Borber, A. J., Ragotte, R. J., et al. De novo design of protein structure and function with RFdiffusion. *Nature*, 620(7976): 1089–1100, 2023.

Wu, J., Kong, X., Sun, N., Wei, J., Shan, S., Feng, F., Wu, F., Peng, J., Zhang, L., Liu, Y., and Ma, J. FlowDesign: Improved design of antibody CDRs through flow matching and better prior distributions. *Cell Systems*, 2025a. doi: 10.1016/j.cels.2025.101270.

Wu, L., Lin, H., Huang, Y., Gao, Z., Tan, C., Liu, Y., Wu, T., and Li, S. Z. Relation-aware equivariant graph networks for epitope-unknown antibody design and specificity optimization. In *Proceedings of the AAAI Conference on Artificial Intelligence*, volume 39, pp. 895–904, 2025b.

Xue, L. C., Dobbs, D., Bonvin, A. M., and Honavar, V. Computational prediction of protein interfaces: A review of data driven methods. *FEBS Letters*, 589(23):3516–3526, 2015.