SPECTRAL-PATH FLOW MATCHING FOR ANTIBODY GENERATION

Anonymous authors
Paper under double-blind review

ABSTRACT

Therapeutic antibody design requires generating complementarity-determining region (CDR) loops whose sequences and three-dimensional conformations support antigen recognition. CDR-H3 remains particularly challenging because of its high sequence and structural diversity. Recent diffusion and flow-matching models can generate CDR sequence and structure, but their positional generative paths typically evolve all spatial components under a shared notion of progress, even though native CDR deformation is dominated by low-frequency components corresponding to overall loop shape. We introduce SP-FMAG, a spectral-path flow-matching framework that decomposes framework-anchored CDR deformation into ordered spatial modes and assigns each mode its own generative clock. This produces a coarse-to-fine positional path in which broad loop geometry is established before finer structural variation. On the standard SAbDab benchmark with 19 held-out antibody–antigen complexes, SP-FMAG reduces CDR-H3 backbone RMSD from 3.313 Å to 2.958 Å and increases Rosetta binding-energy improvement rate from 23.8% to 35.7% relative to linear positional flow matching. The structural advantage persists under reduced sampling budgets. Reversing the modal ordering to fine-to-coarse worsens H3 RMSD to 3.499 Å, indicating that the benefit depends on the coarse-to-fine direction. These results demonstrate that coarse-to-fine spatial ordering in the probability path can improve antibody loop generation without changing the underlying neural architecture.

1 INTRODUCTION

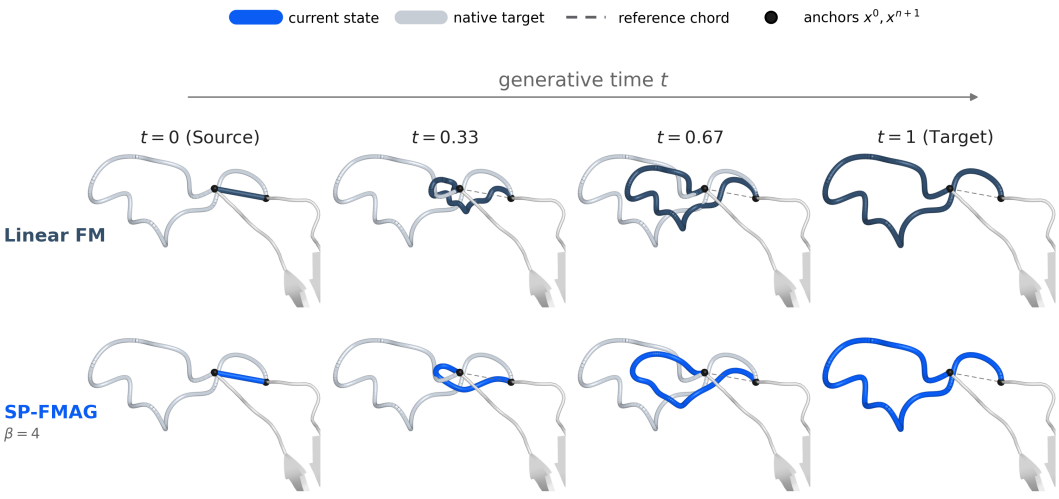


Figure 1: **Linear and spectral positional probability paths.** The CDR loop is generated between fixed framework anchors x^0 and x^{n+1}, where n is the number of CDR residues. Both paths connect the same source chord to the same native CDR-H3 target. Linear flow matching advances all modes uniformly, whereas SP-FMAG delays higher-frequency modes, yielding smoother intermediate states and a coarse-to-fine progression.

Reviewers: please read the Reviewer Guidelines (iclr.cc/Conferences/2027/ReviewerGuidelines) and the AI Policy for Reviewers (iclr.cc/Conferences/2027/AIPolicyForReviewers). If you used AI to expand, edit, or polish your review, please provide the input text to the LLM. Better still, consider skipping the LLM and submitting your original text: we, and the authors, are much more interested in your unedited thoughts than in what an LLM has to say. AI-assisted or not, you are putting your name and reputation behind your review: LLM-generated falsehoods, hallucinations or misrepresentations are subject to disciplinary action, which may include desk-rejecting all papers you have authored.

Therapeutic antibodies recognize antigens through six complementarity-determining region (CDR) loops that together form the principal antigen-binding interface (Sela-Culang et al., 2013). Their binding properties depend strongly on the sequence and three-dimensional geometry of these loops, making CDR design a central problem in computational antibody engineering (Zhao et al., 2018). Among the six loops, CDR-H3 is particularly challenging because it exhibits the greatest sequence and structural diversity and is less amenable to canonical loop classification (Shirai et al., 1999), while also playing a major role in antigen recognition and binding specificity (Xu & Davis, 2000; Marks & Deane, 2017). Experimental screening of the large antibody sequence–structure space is difficult to perform exhaustively, motivating computational approaches that can prioritize or generate promising candidates (Szataniak & Packi, 2026).

Recent deep generative models increasingly address this problem by jointly designing CDR sequence and structure. DiffAb formulates antigen-conditioned CDR design as diffusion over residue identities, positions, and orientations (Luo et al., 2022), while subsequent methods have introduced equivariant architectures, physically motivated guidance, and alternative generative objectives (Kong et al., 2023a;b; Zhu et al., 2024; Zhou et al., 2024). FlowDesign more recently applies flow matching to antibody sequence–structure co-design (Wu et al., 2025). Despite these differences, diffusion- and flow-based approaches generally evolve the generated CDR coordinates under a shared positional schedule. Consequently, fine structural variation is resolved over the same interval as broad loop geometry, rather than after it, as illustrated in Figure 1.

CDR geometry gives specific reason to question this shared treatment. A CDR is attached to the surrounding antibody framework at both ends, providing two fixed spatial reference points. Relative to the straight chord connecting these anchors, its deformation can be decomposed into fixed-end spatial modes ordered from smooth, broad variation to progressively finer geometric detail. Native CDR structures use this spectrum highly unevenly. For CDR-H3, the lowest 25% of the normalized mode spectrum captures a median 95.1% of native deformation energy across the 19 held-out complexes. Across all six CDR types, the corresponding median ranges from 78.6% to 98.1%, showing that CDR geometry is strongly concentrated toward low-frequency components. Full spectral-content analysis is provided in Appendix B. This concentration motivates a coarse-to-fine generative path in which broad loop geometry is established before progressively finer variation.

There is precedent for treating spatial scales differently in generative modelling. Frequency-structured image generation and cascading-resolution protein diffusion both exploit differences across spatial scales (Rissanen et al., 2023; Hoogeboom & Salimans, 2023; Jing et al., 2023). Related protein modelling work also associates low-frequency modes with collective conformational motion (Tirion, 1996). These results motivate scale-aware generation, but do not address fixed-boundary CDR generation with flow matching.

We therefore introduce SP-FMAG, a spectral-path flow-matching framework in which positional transport is explicitly ordered across spatial scales. CDR deformation relative to its framework-anchor chord is represented in a fixed-end sine basis, and each mode is assigned a generative clock according to its spectral frequency. Higher-frequency modes are progressively delayed, producing a coarse-to-fine probability path in which broad geometric components progress toward the target before finer variation. The ordering is introduced entirely through the probability path rather than through a spectral neural architecture or additional guidance.

On the standard SAbDab benchmark (Dunbar et al., 2014) with 19 held-out antibody–antigen complexes, SP-FMAG reduces CDR-H3 backbone root-mean-square deviation (RMSD) from 3.313 Å to 2.958 Å relative to the matched linear-flow model and increases Rosetta binding-energy improvement rate (IMP) from 23.8% to 35.7%. The structural advantage persists under substantially reduced sampling budgets. Reversing the same range of mode-dependent clocks instead increases H3 RMSD to 3.499 Å, performing worse than the linear baseline and showing that the benefit depends on the direction of the spectral ordering rather than heterogeneous timing alone. Beyond de novo generation, the same trained path also supports controlled antibody optimization, providing a continuous trade-off between preserving an existing CDR and permitting greater sequence–structure redesign.

Our contributions are

- We introduce a fixed-end spectral representation and mode-dependent flow-matching path that orders positional transport from broad loop deformation to finer geometric variation without changing the underlying neural architecture or adding external guidance.

- Our construction recovers ordinary linear positional flow matching when the mode-dependent ordering is removed, enabling a direct comparison that isolates the effect of spectral ordering.
- We demonstrate improved CDR-H3 structural generation, robustness under reduced sampling budgets, sensitivity to the direction of the spectral schedule, and controlled antibody optimization using the same probability path.

2 Related Work

Generative antibody design. Recent antibody-design methods increasingly model CDR sequence and structure jointly. DiffAb formulates antigen-conditioned CDR generation as diffusion over residue identities, backbone positions, and orientations (Luo et al., 2022), while dyMEAN uses an end-to-end E(3)-equivariant framework for joint sequence–structure design and antibody–antigen docking (Kong et al., 2023b). Subsequent work has explored physical guidance, multi-scale geometric representations, and more expressive reverse transitions. DIFFFORCE incorporates force-based guidance during diffusion (Kulytė et al., 2024), AbMEGD combines atomic- and residue-level representations in an E(3)-equivariant diffusion model (Chen et al., 2025), and MDMAD introduces multimodal positional and rotational reverse transitions (Durvasula et al., 2026). FlowDesign extends antibody co-design to flow matching (Wu et al., 2025), while TiDE-Ab uses conditional SE(3) flow matching with time-dependent classifier-free guidance for epitope-specific generation (Kim & Baek, 2026). Collectively, these methods explore architectural, guidance-based, and generative-objective improvements for antibody sequence–structure design.

Frequency-structured generative processes. Generative processes that distinguish spatial frequencies have been studied outside antibody design. In images, inverse heat dissipation constructs a coarse-to-fine trajectory by preferentially removing high-frequency information, while blurring diffusion expresses a related process through frequency-dependent Gaussian noise (Rissanen et al., 2023; Hoogeboom & Salimans, 2023). In protein generation, EigenFold uses harmonic diffusion over protein eigenmodes to obtain a cascading-resolution reverse process (Jing et al., 2023). These works establish that spatial frequencies can be treated differently during generation, but they do not address fixed-boundary CDR generation with flow matching.

Spectral structure in protein modelling. Spectral information can also enter protein models without modifying the generative path. Classical normal-mode and elastic-network approaches associate low-frequency modes with collective protein motion (Tirion, 1996). More recently, PI-Mamba incorporates a Rouse-polymer spectrum into its state-space architecture within a flow-matching framework (Wu & Zhu, 2026). In these approaches, spectral structure enters through the physical representation or model architecture. SP-FMAG instead places the spectral hierarchy directly in the probability path, with the fixed CDR boundaries defining the basis and the neural network receiving only the shared base time.

3 Background and Setup

3.1 Flow matching

Flow matching. Flow matching learns a continuous transport from a source distribution ρ_0 to a target distribution ρ_1 (Lipman et al., 2023). The transport is described by a time-dependent flow map ψ_t, where a source sample $x_0 \sim \rho_0$ follows the trajectory $x_t = \psi_t(x_0)$ with $\psi_0(x_0) = x_0$. The trajectory is generated by a time-dependent velocity field u_t, satisfying

$$\frac{\mathrm{d}}{\mathrm{d}t}\psi_t(x_0) = u_t(\psi_t(x_0))\,. \tag{1}$$

As t increases from 0 to 1, the flow transports samples from ρ_0 toward ρ_1. A neural network $v_\theta(x_t, t)$ is trained to approximate the velocity field, and during generation its predicted velocity is used to advance the current state along the trajectory.

Conditional flow matching. Directly constructing the marginal velocity field $u_t(x)$ can be difficult. Conditional flow matching instead defines tractable trajectories between sampled source and target states (Lipman et al., 2023). Let $x_0 \sim \rho_0$ and $x_1 \sim \rho_1$ be sampled independently. Conditioned on x_1, we define the flow map

$$x_t = \psi_t(x_0 \mid x_1) = [1 - \alpha(t)]\, x_0 + \alpha(t) x_1, \tag{2}$$

where $\alpha(0) = 0$ and $\alpha(1) = 1$. Differentiating gives

$$\frac{\mathrm{d}}{\mathrm{d}t}\psi_t(x_0 \mid x_1) = \alpha'(t)(x_1 - x_0),$$

and since $x_1 - x_t = [1 - \alpha(t)](x_1 - x_0)$, the conditional velocity can be written in endpoint form as

$$u_t(x_t \mid x_1) = q(t)(x_1 - x_t), \qquad q(t) = \frac{\alpha'(t)}{1 - \alpha(t)}. \tag{3}$$

The network is trained against these conditional velocities using

$$\mathcal{L}_{\mathrm{CFM}}(\theta) = \mathbb{E}_{t,x_0,x_1}\left[\|v_\theta(x_t, t) - u_t(x_t \mid x_1)\|_2^2\right]. \tag{4}$$

Although the targets are defined conditionally, $\mathcal{L}_{\mathrm{CFM}}$ has the same gradients with respect to θ as the marginal flow-matching objective (Lipman et al., 2023, Theorem 2). The network can therefore predict the clean endpoint $\hat{x}_1(x_t, t)$, while the corresponding velocity is recovered analytically.

3.2 Antibody CDR co-design

We consider conditional sequence–structure co-design of an antibody complementarity-determining region while keeping the surrounding antibody framework and antigen fixed. Following prior antibody generative models (Luo et al., 2022), each generated residue i is represented by an amino-acid identity $s^i \in \{1, \ldots, 20\}$, a C_α coordinate $x^i \in \mathbb{R}^3$, and a backbone orientation $o^i \in \mathrm{SO}(3)$.

For a CDR containing n residues, we stack its coordinates as $X = [x^1; \ldots; x^n] \in \mathbb{R}^{n\times 3}$. Let $\mathcal{M}$ denote these generated residues and let $\mathcal{M}^c$ contain the fixed framework and antigen context. We model

$$p_\theta\left(\{s^i, x^i, o^i\}_{i\in\mathcal{M}} \mid \mathcal{M}^c\right). \tag{5}$$

A CDR differs from a free protein chain in an important way. It is a contiguous segment attached at both ends to the surrounding antibody framework. We denote the fixed C_α coordinates immediately preceding and following the generated segment by x^0 and x^{n+1}. These boundary conditions provide the spatial structure used to construct our positional probability path.

4 Spectral-Path Flow Matching

Standard positional flow matching evolves all CDR coordinates according to the same temporal schedule. SP-FMAG instead decomposes loop deformation into ordered spatial modes and assigns different progression rates across those modes. The resulting coarse-to-fine structure is introduced entirely through the probability path, while the underlying geometric network remains unchanged.

4.1 Fixed-end spectral representation

Using the fixed framework anchors x^0 and x^{n+1} defined above, we separate the anchor-to-anchor placement of the CDR from its loop deformation. The straight line between the two anchors defines the reference chord

$$\bar{X}_j = \left(1 - \frac{j}{n+1}\right) x^0 + \frac{j}{n+1} x^{n+1}, \qquad j = 1, \ldots, n. \tag{6}$$

The deformation of the loop relative to this reference is $D = X - \bar{X}$. Because the reference chord meets the two fixed framework anchors, the extended deformation satisfies $D_0 = D_{n+1} = 0$.

To describe this deformation at different spatial scales, we use a spectral representation based on the fixed-end discrete Laplacian along the residue index,

$$(KD)_j = 2D_j - D_{j-1} - D_{j+1}, \qquad K = \text{tridiag}(-1, 2, -1). \tag{7}$$

This operator measures how rapidly the deformation changes between neighboring residues. Its eigenvectors therefore provide an ordered set of spatial deformation patterns. Under the fixed-end conditions $D_0 = D_{n+1} = 0$, these eigenvectors are the discrete sine modes

$$U_{jk} = \sqrt{\frac{2}{n+1}} \sin\left(\frac{\pi jk}{n+1}\right), \qquad j, k = 1, \ldots, n, \tag{8}$$

with corresponding eigenvalues

$$\lambda_k = 2\left[1 - \cos\left(\frac{\pi k}{n+1}\right)\right]. \tag{9}$$

The eigenvalues increase with the mode index k. Low-index modes therefore vary slowly along the CDR and represent broad loop deformation, whereas higher-index modes vary more rapidly and represent progressively finer spatial detail.

Because U is orthonormal, the deformation can be projected into modal coordinates as

$$C = U^\top (X - \bar{X}), \qquad X = \bar{X} + UC. \tag{10}$$

Writing $C = [c_1; \ldots; c_n]$ with $c_k \in \mathbb{R}^3$, each row c_k gives the three-dimensional coefficient of spatial mode k. Thus the same CDR deformation is represented not by individual residue displacements, but by the amplitudes of deformation modes ordered from broad to fine spatial variation. The basis depends only on the CDR length and its fixed boundaries and introduces no learned parameters. Detailed derivations of the eigenvectors and eigenvalues are provided in Appendix A.1.

4.2 Mode-dependent probability path

The spectral representation provides an ordering of loop deformation from broad to fine spatial scales. We use this ordering to assign different progression rates to the modes along the conditional flow-matching path. Because CDRs may contain different numbers of residues, their Laplacian eigenvalues occupy different numerical ranges. We therefore normalize the eigenvalue of mode k as

$$r_k = \frac{\lambda_k - \lambda_1}{\lambda_n - \lambda_1}, \qquad r_k \in [0, 1]. \tag{11}$$

Here $r_1 = 0$ corresponds to the lowest-frequency, broadest mode and $r_n = 1$ to the highest-frequency, finest mode. We then assign each mode its own interpolation clock

$$\gamma_k = 1 + \beta r_k, \qquad A_k(t) = t^{\gamma_k}, \qquad \beta \geq 0, \tag{12}$$

where β controls the separation between spatial scales. For $0 < t < 1$, t^γ decreases as the exponent γ increases. Since γ_k increases with spectral frequency, higher-frequency modes have smaller $A_k(t)$ at the same base time and are therefore delayed relative to lower-frequency modes. The lowest-frequency mode follows the ordinary linear clock $A_1(t) = t$.

For source and target modal coordinates $c_{0,k}$ and $c_{1,k}$, the conditional positional trajectory is

$$c_{t,k} = [1 - A_k(t)]\, c_{0,k} + A_k(t) c_{1,k}. \tag{13}$$

The scalar $A_k(t)$ determines how far mode k has progressed from its source value toward its target value at base time t. Because lower-frequency modes have larger $A_k(t)$ than higher-frequency

modes, broad loop deformation approaches its target earlier along the path, while finer spatial variation is introduced progressively later.

The mode-dependent clocks change only the trajectory between source and target and leave both endpoint distributions unchanged. When $\beta = 0$, $\gamma_k = 1$ and $A_k(t) = t$ for every mode, giving $C_t = (1-t)C_0 + tC_1$. Since the spectral transform is linear, this implies $X_t = (1-t)X_0 + tX_1$, so the construction exactly recovers ordinary linear positional flow matching when mode-dependent ordering is removed.

4.3 Joint sequence–structure learning

The positional source is sampled independently in modal space as $c_{0,k} \sim \mathcal{N}(0, I_3)$ and mapped back through $X_0 = \bar{X} + UC_0$. This source is independent of β, so changing spectral strength modifies only the probability path rather than the initial distribution.

Let $s_t = \{s_t^i\}_{i\in\mathcal{M}}$ and $O_t = \{o_t^i\}_{i\in\mathcal{M}}$ denote the current sequence and orientation states of the generated residues. SP-FMAG uses a shared geometric encoder to jointly predict CDR sequence, position, and orientation. At base time t, the network receives the current sequence states s_t, positions X_t, orientations O_t, the framework–antigen context, and a single time embedding. Although positions follow the mode-dependent clocks $A_k(t)$, the network itself receives only the shared base time t. The spectral ordering is therefore imposed by the probability path rather than explicitly provided as a frequency-dependent conditioning signal.

The network predicts the clean positional endpoint $\hat{X}_1$, amino-acid logits, and a clean-orientation update. Sequence follows an absorbing `MASK` process on the ordinary base time, while orientations follow geodesic interpolation on $\mathrm{SO}(3)$. Their full parameterization is given in Appendix A.4.

For positions, the predicted Cartesian endpoint $\hat{X}_1$ is transformed into modal coordinates as $\hat{C}_1 = U^\top(\hat{X}_1 - \bar{X})$. Differentiating Eq. 13 gives $\frac{\mathrm{d}c_{t,k}}{\mathrm{d}t} = A_k'(t)(c_{1,k} - c_{0,k})$. Since $c_{1,k} - c_{t,k} = [1 - A_k(t)](c_{1,k} - c_{0,k})$, the modal velocity can be written in endpoint form as

$$\frac{\mathrm{d}c_{t,k}}{\mathrm{d}t} = q_k(t)\,(c_{1,k} - c_{t,k})\,, \qquad q_k(t) = \frac{A_k'(t)}{1 - A_k(t)}. \tag{14}$$

Because the network predicts the clean endpoint rather than the velocity directly, velocity matching is equivalent to weighting the endpoint error for mode k by $q_k(t)^2$. We therefore normalize these weights within each CDR,

$$w_k(t) = \frac{q_k(t)^2}{\frac{1}{n}\sum_{m=1}^{n} q_m(t)^2}, \qquad \frac{1}{n}\sum_{k=1}^{n} w_k(t) = 1. \tag{15}$$

The positional objective is then

$$\mathcal{L}^x = \frac{1}{n}\sum_{k=1}^{n} w_k(t)\,\|\hat{c}_{1,k} - c_{1,k}\|_2^2\,. \tag{16}$$

The normalization preserves the relative emphasis induced by the spectral path without changing the average positional-loss scale. At $\beta = 0$, $w_k(t) = 1$ for all modes, recovering ordinary clean-coordinate regression.

The full training objective is $\mathcal{L} = \lambda_s\mathcal{L}^s + \lambda_x\mathcal{L}^x + \lambda_o\mathcal{L}^o$, where $\mathcal{L}^s$ and $\mathcal{L}^o$ denote the sequence cross-entropy and orientation losses described in Appendix A.4. We use $\lambda_s = \lambda_x = \lambda_o = 1$.

4.4 Sampling and antibody optimization

Generation starts from the source state at $t_0 = 0$ and advances over T uniformly spaced base times. At each step, the network is evaluated once to predict the clean sequence, orientation, and positional endpoints.

Because the positional path is analytic, each mode can be advanced directly between consecutive times. Holding the current endpoint estimate $\hat{c}_{1,k}$ fixed over one step,

$$c_{t_{j+1},k} = \hat{c}_{1,k} + \frac{1 - A_k(t_{j+1})}{1 - A_k(t_j)} \left(c_{t_j,k} - \hat{c}_{1,k} \right). \tag{17}$$

The Cartesian loop is then reconstructed as $X_{t_{j+1}} = \bar{X} + UC_{t_{j+1}}$. Sequence and orientation advance over the same base-time interval using their corresponding state-space paths. Each step therefore requires one network evaluation, so the number of function evaluations is exactly T. The same probability path also enables local optimization of an existing antibody without retraining. Instead of starting at $t = 0$, we partially perturb a native CDR to an intermediate point $t_{\text{start}} = 1 - \frac{m}{T}$, where $m \in \{1, \ldots, T\}$ denotes the number of generative steps remaining. We then resume generation from t_{start} to 1 using the same sequence, position, and orientation paths employed during training. Small m retains most of the starting antibody, whereas larger m permits progressively greater redesign. At $m = T$, the procedure reduces to ordinary de novo generation.

5 Experiments

Dataset and evaluation. We use the standard Structural Antibody Database (SAbDab) benchmark released with DiffAb (Luo et al., 2022). Following the original split, complexes are clustered at 50% CDR-H3 sequence identity and 19 antibody–antigen complexes from five clusters are held out for evaluation. We compare against RAbD (Adolf-Bryfogle et al., 2018), DiffAb (Luo et al., 2022), DIFFFORCE (Kulytė et al., 2024), AbMEGD (Chen et al., 2025), and MDMAD (Durvasula et al., 2026). Baseline numbers are taken from MDMAD and SP-FMAG is trained and evaluated under the same data processing, relaxation, and metric-computation pipeline.

We report amino-acid recovery (AAR), C_α RMSD, and Rosetta binding-energy improvement rate (IMP). AAR is the fraction of generated CDR residues matching the native sequence, RMSD measures deviation of generated CDR C_α coordinates from the native structure, and IMP is the fraction of generated designs with improved Rosetta binding energy, $\Delta\Delta G = \Delta G_{\text{design}} - \Delta G_{\text{native}} < 0$. Generated structures are processed using the same OpenMM (Eastman et al., 2017) and Rosetta relaxation protocol (Alford et al., 2017) before evaluation. Unless stated otherwise, all results use 100 generated samples per antibody–antigen complex. Full dataset, implementation, relaxation, and metric-computation details are provided in the Appendix D.

5.1 Comparison with baseline methods

We report the primary benchmark comparison on CDR-H3, the most structurally diverse and challenging CDR and the setting in which the proposed spectral ordering is expected to be most consequential. Table 1 compares SP-FMAG with existing sequence–structure antibody design methods using 100 generated samples per antibody–antigen complex. Complete results across all six CDRs are reported in Appendix B.1.

Table 1: CDR-H3 generation on SAbDab using 100 samples per complex. Best results are **bold**; second-best are underlined.

Method	AAR↑	IMP↑	RMSD↓
RAbD	22.14	23.25	**2.900**
DiffAb	26.79	19.90	3.616
DiffAb + analytic posterior	<u>40.78</u>	24.90	3.360
DIFFFORCE	29.52	30.22	3.612
AbMEGD	25.06	23.20	3.914
MDMAD	**42.85**	<u>32.40</u>	3.381
SP-FMAG, $\beta = 4$	36.78	**35.70**	<u>2.958</u>

SP-FMAG obtains an H3 RMSD of 2.958 Å, lower than all learned generative baselines considered here and close to the physics-based RAbD result of 2.900 Å. It also achieves the highest H3 IMP at 35.7%, while its AAR remains below MDMAD and DiffAb with the analytic posterior. The gain is therefore concentrated in structural recovery and binding-related quality rather than native sequence reconstruction.

Results on the remaining CDRs are more mixed. Their matched linear-flow RMSDs are already substantially lower than H3 at approximately 1.1–1.6 Å, leaving less headroom for structural improvement. Reversing the ordering does not degrade H1 or H2 but clearly degrades H3 (Appendix B.6). Within H3, the spectral benefit tends to increase with loop length (Figure 2). SP-FMAG remains

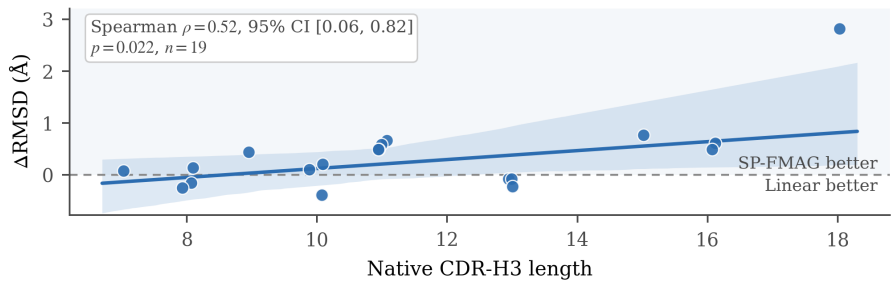
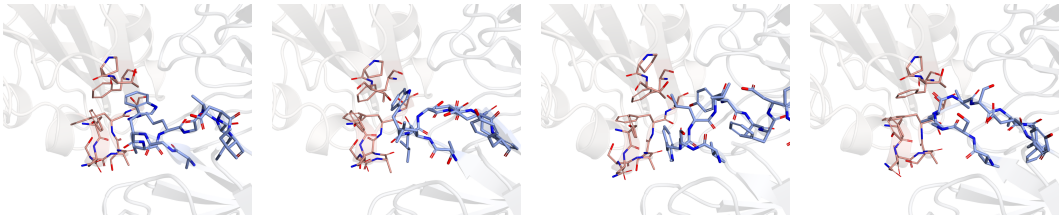


Figure 2: **Structural gain tends to increase with CDR-H3 length.** Per-complex RMSD improvement $\Delta\text{RMSD} = \text{RMSD}_{\beta=0} - \text{RMSD}_{\beta=4}$ versus native H3 length across the 19 held-out complexes; positive values favour SP-FMAG. The line shows a Theil–Sen fit with a pointwise 95% bootstrap band. Spearman $\rho = 0.52$, $p = 0.022$. Length-group stratification and per-complex values are provided in Appendix B.5.



Reference

Sample 1
RMSD = 1.998 Å
$\Delta\Delta G = -7.84$

Sample 2
RMSD = 3.072 Å
$\Delta\Delta G = -18.23$

Sample 3
RMSD = 2.568 Å
$\Delta\Delta G = -13.95$

Figure 3: **Representative CDR-H3 samples generated by SP-FMAG.** Native and generated structures are shown for the same held-out complex. CDR-H3 is shown in blue and the local antigen interface in salmon. RMSD is measured to the native CDR-H3 backbone, and $\Delta\Delta G < 0$ indicates improved Rosetta binding energy.

competitive on the other CDRs, including the highest IMP on L2 and lowest RMSD on L3. We therefore focus the controlled analyses on H3.

Figure 3 shows representative CDR-H3 samples from the $\beta = 4$ model, with all three examples achieving improved Rosetta binding energy ($\Delta\Delta G < 0$).

5.2 Effect of spectral separation

Table 2 isolates the effect of the probability path by varying only β while keeping the architecture, source distribution, training procedure, and evaluation pipeline fixed. The $\beta = 0$ model exactly recovers ordinary linear positional flow matching.

Moderate spectral separation improves H3 structural recovery, with $\beta = 4$ reducing RMSD from 3.313 Å to 2.958 Å and increasing IMP from 23.8% to 35.7%. Stronger separation at $\beta = 8$ partially reverses these gains, indicating that increasingly delaying fine-scale modes is not uniformly beneficial. A paired bootstrap analysis of the 19 held-out complexes gives a mean RMSD reduction of 0.355 Å with a 95% confidence interval of 0.098–0.683 Å. Full paired statistics are reported in Appendix B.4.

Table 2: CDR-H3 performance as spectral separation is varied. $\beta = 0$ is the matched linear-flow baseline.

β	AAR↑	IMP↑	RMSD↓
0	**38.97**	23.80	3.313
2	36.96	26.30	3.079
4	36.78	**35.70**	**2.958**
8	34.22	30.90	3.143

5.3 Sampling efficiency

Because the sampling budget is chosen only at inference, the same checkpoint can be evaluated at different numbers of function evaluations. Table 3 compares the matched $\beta = 0$ and $\beta = 4$ models.

The H3 structural advantage persists under aggressive reductions in sampling budget. SP-FMAG yields lower RMSD than the matched linear model at both $T = 5$ and $T = 1$. AAR and IMP are less consistent, so we interpret this result specifically as robustness of the structural benefit rather than a uniform improvement across all metrics.

Table 3: CDR-H3 performance under reduced sampling budgets.

	$T = 1$		$T = 5$		$T = 100$	
Metric	$\beta = 0$	$\beta = 4$	$\beta = 0$	$\beta = 4$	$\beta = 0$	$\beta = 4$
AAR↑	39.15	37.37	38.77	37.57	38.97	36.78
IMP↑	22.30	26.40	26.30	21.00	23.80	35.70
RMSD↓	3.180	**3.005**	3.281	**2.928**	3.313	**2.958**

5.4 Does the ordering matter?

The β ablation establishes that heterogeneous modal timing can improve H3 generation, but not whether its direction matters. We therefore reverse the frequency assignment while preserving the same range of clock exponents. The forward schedule uses $\gamma_k = 1 + \beta r_k$, whereas the reversed schedule uses $\gamma_k^{\text{rev}} = 1 + \beta(1 - r_k)$.

As shown in Table 4, reversing the ordering removes the structural advantage. RMSD increases from 2.958 Å to 3.499 Å and IMP falls from 35.7% to 21.1%, while AAR remains nearly unchanged. Because the two schedules use the same degree of temporal heterogeneity and differ only in its assignment across frequencies, the result supports the specific coarse-to-fine ordering rather than mode-dependent timing alone.

Table 4: Effect of spectral ordering on CDR-H3. Both schedules use $\beta = 4$.

Schedule	AAR↑	IMP↑	RMSD↓
Forward	36.78	**35.70**	**2.958**
Reverse	37.01	21.10	3.499

5.5 Antibody optimization

We finally test whether the same trained path can refine an existing CDR without retraining. Starting from an intermediate point on the probability path, increasing the optimization strength m progressively increases the amount of sequence and structural redesign.

As shown in Table 5, increasing m produces the expected preservation–redesign trade-off. At $m = 1$, generated CDRs retain 99.46% sequence identity with an RMSD of 1.632 Å, while 25.1% improve predicted binding energy. At $m = 16$, IMP reaches 27.5% while retaining 89.35% sequence identity. Thus the same pretrained model supports a continuum from local refinement toward increasingly unconstrained redesign.

Table 5: CDR-H3 optimization with $\beta = 4$. RMSD and SeqID are relative to the starting CDR.

m	IMP↑	RMSD (Å)↓	SeqID↑
1	25.1	**1.632**	**99.46**
2	15.7	1.714	98.72
4	11.1	1.829	97.63
8	17.0	1.894	95.05
16	**27.5**	2.037	89.35
32	26.5	2.215	80.26
64	17.1	2.632	60.09

6 Discussion and Limitations

SP-FMAG introduces multiscale structure directly into the flow-matching probability path rather than the neural architecture. By delaying finer spatial modes, it encourages broad loop geometry to approach the target before finer variation. The strongest gains occur on CDR-H3 and persist under reduced sampling budgets. Reversing the ordering degrades both RMSD and IMP, supporting the importance of the coarse-to-fine direction rather than heterogeneous timing alone.

Evaluation is limited to the standard DiffAb benchmark of 19 held-out complexes, and binding improvement is assessed using Rosetta energy rather than experimental affinity. The fixed-end sine basis is deliberately simple, encodes no explicit residue-specific interactions, and only positional transport is mode-dependent. Future work could explore learned spectral bases, extend mode-dependent paths to sequence and orientation, and evaluate the same principle in broader molecular conformational generation and under additional antibody-design objectives.

AI use statement

In this work, we used generative AI tools to improve the clarity, readability, organization, and presentation of the manuscript. We have not used generative AI tools for other tasks requiring disclo-

sure, and the remaining required disclosure tasks are not applicable to this work. We reviewed all AI-assisted edits and verified that the final text accurately reflects the authors' scientific content and intended meaning. We take responsibility for the final content of this work, including text, claims, or artifacts produced with the aid of generative AI.

Reproducibility statement

The spectral construction, mode-dependent probability path, and path-consistent sampler are derived in Appendix A, including exact recovery of linear positional flow matching at $\beta = 0$. Dataset construction, the DiffAb split, model architecture, training hyperparameters, sampling configuration, and the OpenMM and PyRosetta relaxation and evaluation protocol are described in Appendix D. All reported ablations use the same architecture, source distribution, training procedure, and evaluation pipeline, with only β or the spectral ordering varied where stated. Random seeds are also reported in Appendix D.

References

Jared Adolf-Bryfogle, Oleks Kalyuzhniy, Michael Kubitz, Brian D Weitzner, Xiaozhen Hu, Yumiko Adachi, William R Schief, and Roland L Dunbrack Jr. Rosettaantibodydesign (rabd): A general framework for computational antibody design. *PLoS Computational Biology*, 14(4):e1006112, 2018. doi: 10.1371/journal.pcbi.1006112.

Rebecca F Alford, Andrew Leaver-Fay, Jeliazko R Jeliazkov, Matthew J O'Meara, Frank P DiMaio, Hahnbeom Park, Maxim V Shapovalov, P Douglas Renfrew, Vikram K Mulligan, Kalli Kappel, et al. The rosetta all-atom energy function for macromolecular modeling and design. *Journal of chemical theory and computation*, 13(6):3031–3048, 2017.

Jacob Austin, Daniel D. Johnson, Jonathan Ho, Daniel Tarlow, and Rianne van den Berg. Structured denoising diffusion models in discrete state-spaces. In *Advances in Neural Information Processing Systems*, volume 34, pp. 17981–17993. Curran Associates, Inc., 2021. URL https://proceedings.neurips.cc/paper_files/paper/2021/file/958c530554f78bcd8e97125b70e6973d-Paper.pdf.

Sidhartha Chaudhury, Sergey Lyskov, and Jeffrey J Gray. Pyrosetta: a script-based interface for implementing molecular modeling algorithms using rosetta. *Bioinformatics*, 26(5):689–691, 2010.

Jiameng Chen, Xiantao Cai, Jia Wu, and Wenbin Hu. Antibody design and optimization with multi-scale equivariant graph diffusion models for accurate complex antigen binding. In *Proceedings of the Thirty-Fourth International Joint Conference on Artificial Intelligence, IJCAI-25*, pp. 2722–2730. International Joint Conferences on Artificial Intelligence Organization, 2025. doi: 10.24963/ijcai.2025/303. URL https://doi.org/10.24963/ijcai.2025/303.

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

James Dunbar, Konrad Krawczyk, Jinwoo Leem, Terry Baker, Angelika Fuchs, Guy Georges, Jiye Shi, and Charlotte M Deane. Sabdab: the structural antibody database. *Nucleic acids research*, 42(D1):D1140–D1146, 2014.

Vasanth Durvasula, Tiara Natasha Binte Sayuti, and Jagath C Rajapakse. Mixture diffusion model for multimodal antibody design. *Briefings in Bioinformatics*, 27(4):bbag409, 2026. doi: 10.1093/bib/bbag409. URL https://doi.org/10.1093/bib/bbag409.

Peter Eastman, Jason Swails, John D Chodera, Robert T McGibbon, Yutong Zhao, Kyle A Beauchamp, Lee-Ping Wang, Andrew C Simmonett, Matthew P Harrigan, Chaya D Stern, et al. Openmm 7: Rapid development of high performance algorithms for molecular dynamics. *PLoS computational biology*, 13(7):e1005659, 2017.

Emiel Hoogeboom and Tim Salimans. Blurring diffusion models. In *The Eleventh International Conference on Learning Representations*, 2023. URL https://openreview.net/forum?id=OjDkC57x5sz.

Bowen Jing, Ezra Erives, Peter Pao-Huang, Gabriele Corso, Bonnie Berger, and Tommi S. Jaakkola. Eigenfold: Generative protein structure prediction with diffusion models. In *ICLR 2023 - Machine Learning for Drug Discovery workshop*, 2023. URL https://openreview.net/forum?id=BgbRVzfQqFp.

Yunji Kim and Minkyung Baek. De novo epitope-specific antibody design via time-dependent guidance. *Bioinformatics*, 42(Supplement 2):btag455, 2026. doi: 10.1093/bioinformatics/btag455. URL https://doi.org/10.1093/bioinformatics/btag455.

Xiangzhe Kong, Wenbing Huang, and Yang Liu. Conditional antibody design as 3d equivariant graph translation. In *The Eleventh International Conference on Learning Representations*, 2023a. URL https://openreview.net/forum?id=LFHFQbjxIiP.

Xiangzhe Kong, Wenbing Huang, and Yang Liu. End-to-end full-atom antibody design. In *Proceedings of the 40th International Conference on Machine Learning*, volume 202 of *Proceedings of Machine Learning Research*, pp. 17409–17429. PMLR, 2023b. URL https://proceedings.mlr.press/v202/kong23c.html.

Paulina Kulytė, Francisco Vargas, Simon Valentin Mathis, Yu Guang Wang, José Miguel Hernández-Lobato, and Pietro Liò. Improving antibody design with force-guided sampling in diffusion models. *arXiv preprint arXiv:2406.05832*, 2024.

Yaron Lipman, Ricky T. Q. Chen, Heli Ben-Hamu, Maximilian Nickel, and Matthew Le. Flow matching for generative modeling. In *The Eleventh International Conference on Learning Representations*, 2023. URL https://openreview.net/forum?id=PqvMRDCJT9t.

Shitong Luo, Yufeng Su, Xingang Peng, Sheng Wang, Jian Peng, and Jianzhu Ma. Antigen-specific antibody design and optimization with diffusion-based generative models for protein structures. In *Advances in Neural Information Processing Systems*, volume 35, pp. 9754–9767. Curran Associates, Inc., 2022. doi: 10.52202/068431-0709. URL https://proceedings.neurips.cc/paper_files/paper/2022/file/3fa7d76a0dc1179f1e98d1bc62403756-Paper-Conference.pdf.

Claire Marks and Charlotte M Deane. Antibody h3 structure prediction. *Computational and Structural Biotechnology Journal*, 15:222–231, 2017. doi: 10.1016/j.csbj.2017.01.010.

Severi Rissanen, Markus Heinonen, and Arno Solin. Generative modelling with inverse heat dissipation. In *The Eleventh International Conference on Learning Representations*, 2023. URL https://openreview.net/forum?id=4PJUBT9f2Ol.

Inbal Sela-Culang, Vered Kunik, and Yanay Ofran. The structural basis of antibody-antigen recognition. *Frontiers in Immunology*, 4:302, 2013. doi: 10.3389/fimmu.2013.00302.

Hiroki Shirai, Akinori Kidera, and Haruki Nakamura. H3-rules: identification of cdr-h3 structures in antibodies. *FEBS Letters*, 455(1-2):188–197, 1999. doi: 10.1016/S0014-5793(99)00821-2.

Martin Steinegger and Johannes Söding. Mmseqs2 enables sensitive protein sequence searching for the analysis of massive data sets. *Nature biotechnology*, 35(11):1026–1028, 2017.

Ida Szataniak and Kacper Packi. The evolution of artificial intelligence in antibody design: From structure-based engineering to generative models. *Antibodies*, 15(5):81, 2026. doi: 10.3390/antib15050081. URL https://doi.org/10.3390/antib15050081.

Monique M. Tirion. Large amplitude elastic motions in proteins from a single-parameter, atomic analysis. *Phys. Rev. Lett.*, 77:1905–1908, 1996. doi: 10.1103/PhysRevLett.77.1905. URL https://link.aps.org/doi/10.1103/PhysRevLett.77.1905.

Jun Wu, Xiangzhe Kong, Ningguan Sun, Jing Wei, Sisi Shan, Fuli Feng, Feng Wu, Jian Peng, Linqi Zhang, Yang Liu, et al. Flowdesign: Improved design of antibody cdrs through flow matching and better prior distributions. *Cell Systems*, 16(6), 2025.

Tianyu Wu and Lin Zhu. Pi-mamba: linear-time protein backbone generation via spectrally initialized flow matching. *Bioinformatics*, 42(Supplement 2):btag370, 2026. doi: 10.1093/bioinformatics/btag370. URL https://doi.org/10.1093/bioinformatics/btag370.

John L Xu and Mark M Davis. Diversity in the cdr3 region of vh is sufficient for most antibody specificities. *Immunity*, 13(1):37–45, 2000.

Jun Zhao, Ruth Nussinov, Wen-Jin Wu, and Buyong Ma. In silico methods in antibody design. *Antibodies*, 7(3):22, 2018. doi: 10.3390/antib7030022.

Xiangxin Zhou, Dongyu Xue, Ruizhe Chen, Zaixiang Zheng, Liang Wang, and Quanquan Gu. Antigen-specific antibody design via direct energy-based preference optimization. In *Advances in Neural Information Processing Systems*, volume 37, pp. 120861–120891. Curran Associates, Inc., 2024. doi: 10.52202/079017-3840. URL https://proceedings.neurips.cc/paper_files/paper/2024/file/daef77101ba5711084a57442c8cf2709-Paper-Conference.pdf.

Tian Zhu, Milong Ren, and Haicang Zhang. Antibody design using a score-based diffusion model guided by evolutionary, physical and geometric constraints. In *Forty-first International Conference on Machine Learning*, 2024. URL https://openreview.net/forum?id=1YsQI04KaN.

A Mathematical Details of Spectral-Path Flow Matching

This section provides the derivations underlying the fixed-end spectral representation, the mode-dependent probability path, and the path-consistent sampler used in the main text.

A.1 Fixed-End Discrete Laplacian

Let $D = X - \bar{X} \in \mathbb{R}^{n\times 3}$ denote the CDR deformation relative to the straight chord between the two framework anchors. Since the generated segment is bounded by fixed framework residues, its extended deformation satisfies $D_0 = D_{n+1} = 0$.

We use the negative discrete second-difference operator along the residue index,

$$(KD)_j = 2D_j - D_{j-1} - D_{j+1}, \qquad j = 1, \dots, n, \tag{18}$$

with $K = \mathrm{tridiag}(-1, 2, -1)$. The same operator is applied independently to each Cartesian component. Under $D_0 = D_{n+1} = 0$,

$$\mathrm{tr}(D^\top KD) = \sum_{j=0}^{n} \|D_{j+1} - D_j\|_2^2. \tag{19}$$

Thus the quadratic form measures variation between neighboring residues. Modes with smaller eigenvalues vary more smoothly along the residue index, whereas modes with larger eigenvalues exhibit progressively finer spatial variation.

A.2 Eigenvectors and Eigenvalues of the Fixed-End chain

For $k = 1, \dots, n$, let $u_j^{(k)} = \sin(j\theta_k)$ with $\theta_k = \frac{\pi k}{n+1}$. The fixed-end conditions are automatically satisfied since $u_0^{(k)} = u_{n+1}^{(k)} = 0$. Applying the discrete Laplacian gives

$$(Ku^{(k)})_j = 2\sin(j\theta_k) - \sin((j-1)\theta_k) - \sin((j+1)\theta_k) \tag{20}$$

$$= 2(1 - \cos\theta_k)\sin(j\theta_k), \tag{21}$$

where the second equality follows from $\sin((j-1)\theta) + \sin((j+1)\theta) = 2\sin(j\theta)\cos\theta$. Thus $u^{(k)}$ is an eigenvector of K with eigenvalue $\lambda_k = 2\left[1 - \cos\left(\frac{\pi k}{n+1}\right)\right]$.

The sine vectors satisfy the discrete orthogonality relation $\sum_{j=1}^{n} \sin\left(\frac{\pi jk}{n+1}\right)\sin\left(\frac{\pi j\ell}{n+1}\right) = \frac{n+1}{2}\delta_{k\ell}$. and the normalized eigenvectors are therefore $U_{jk} = \sqrt{\frac{2}{n+1}}\sin\left(\frac{\pi jk}{n+1}\right)$, giving $U^\top U = I$ and $KU = U\Lambda$. Since $\frac{\pi k}{n+1} \in (0, \pi)$ and $1-\cos\theta$ is strictly increasing on this interval, $\lambda_1 < \lambda_2 < \cdots < \lambda_n$. The fixed-end spectrum therefore orders the deformation modes from smooth, low-frequency variation to progressively finer spatial variation.

A.3 Properties of the Mode-Dependent Probability Path

Recall that the normalized spectral coordinate is $r_k = \frac{\lambda_k - \lambda_1}{\lambda_n - \lambda_1}$, with $r_k \in [0, 1]$, and that SP-FMAG defines $\gamma_k = 1 + \beta r_k$ and $A_k(t) = t^{\gamma_k}$.

For $\beta \geq 0$, each clock satisfies $A_k(0) = 0$, $A_k(1) = 1$ and $A_k'(t) = \gamma_k t^{\gamma_k - 1} \geq 0$, so every mode moves monotonically between the same source and target endpoints. Since λ_k increases with k, so do r_k and γ_k. For $0 < t < 1$, the map $\gamma \mapsto t^\gamma$ is strictly decreasing. Hence, for $\beta > 0$,

$$k < \ell \quad \Longrightarrow \quad A_k(t) > A_\ell(t). \tag{22}$$

Lower-frequency modes therefore contain a larger fraction of their target contribution at the same base time, while higher-frequency modes are delayed. The modal conditional path is

$$c_{t,k} = A_k(t)c_{1,k} + [1 - A_k(t)]\, c_{0,k}, \tag{23}$$

Differentiating $c_{t,k}$ with respect to time gives $\frac{\mathrm{d}c_{t,k}}{\mathrm{d}t} = A'_k(t)(c_{1,k} - c_{0,k})$, Since $c_{1,k} - c_{t,k} = [1 - A_k(t)](c_{1,k} - c_{0,k})$, the velocity can be written in endpoint form as

$$\frac{\mathrm{d}c_{t,k}}{\mathrm{d}t} = q_k(t)\,(c_{1,k} - c_{t,k})\,, \qquad q_k(t) = \frac{A'_k(t)}{1 - A_k(t)}. \tag{24}$$

For the power-law clock used in SP-FMAG, this becomes $q_k(t) = \frac{\gamma_k t^{\gamma_k - 1}}{1 - t^{\gamma_k}}$.

Recovery of linear flow matching When $\beta = 0$, every $\gamma_k = 1$ and $A_k(t) = t$, so $C_t = tC_1 + (1-t)C_0$. Using $X = \bar{X} + UC$,

$$\begin{aligned} X_t &= \bar{X} + U\,[tC_1 + (1-t)C_0] \\ &= t(\bar{X} + UC_1) + (1-t)(\bar{X} + UC_0) \\ &= tX_1 + (1-t)X_0. \end{aligned} \tag{25}$$

Thus the $\beta = 0$ construction exactly recovers ordinary linear positional flow matching in Cartesian coordinates.

A.4 Flow Matching over Sequence and Rotational Spaces

Only positional coordinates use the mode-dependent spectral clocks described in Section 4.2. Sequence and orientation retain state-space-appropriate paths on the shared base time $t \in [0, 1]$, with $t = 0$ denoting the source state and $t = 1$ the native target.

Sequence path. We use an absorbing masking process over amino-acid identities, following the absorbing-state construction used in discrete diffusion models (Austin et al., 2021). Let $s_1^i \in \{1, \ldots, 20\}$ denote the native amino-acid identity of generated residue i, and let `MASK` denote an additional source state. At time t, the sequence state is

$$s_t^i = \begin{cases} s_1^i, & \text{with probability } t, \\ \texttt{MASK}, & \text{with probability } 1 - t. \end{cases} \tag{26}$$

Thus all generated residues are masked at $t = 0$ and fully revealed at $t = 1$. The network predicts amino-acid logits from the current joint sequence–structure state. Cross-entropy is evaluated only on generated residues that remain masked, since the identities of revealed residues are already present in the network input and including them in the loss would reward trivial copying. The sequence objective is

$$\mathcal{L}^s = -\frac{1}{|\mathcal{M}_t|} \sum_{i \in \mathcal{M}_t} \log p_\theta \left(s_1^i \mid s_t, X_t, O_t, t, \mathcal{M}^c\right), \tag{27}$$

where $\mathcal{M}_t \subseteq \mathcal{M}$ contains the generated residues whose sequence identities remain masked at time t.

During sampling, let the current and next base times be t_j and t_{j+1}. Each residue that remains masked is revealed with conditional probability

$$\eta_j = \frac{t_{j+1} - t_j}{1 - t_j}. \tag{28}$$

When revealed, its identity is sampled from the network's predicted categorical distribution and remains fixed thereafter. This update preserves the forward marginal exactly. Writing S_j for the probability that a residue remains masked at t_j,

$$S_{j+1} = S_j(1 - \eta_j) = S_j \frac{1 - t_{j+1}}{1 - t_j}, \tag{29}$$

which telescopes from $S_0 = 1$ to $S_j = 1 - t_j$, matching Eq. 26.

Rotational path. Let $O_1^i \in \mathrm{SO}(3)$ denote the native backbone orientation of residue i. For generated residues, the source orientation O_0^i is sampled uniformly from $\mathrm{SO}(3)$. We interpolate from source to target along the geodesic

$$O_t^i = O_0^i \exp\left(t \log\left((O_0^i)^\top O_1^i\right)\right). \tag{30}$$

Orientations of fixed framework and antigen residues remain unchanged.

Given the current orientation O_t^i, the network predicts a three-dimensional rotational update ϵ_θ^i, which is mapped to a valid rotation matrix $Q(\epsilon_\theta^i) \in \mathrm{SO}(3)$ through a normalized quaternion. The predicted clean orientation is

$$\hat{O}_1^i = O_t^i Q(\epsilon_\theta^i). \tag{31}$$

Following DiffAb (Luo et al., 2022), orientation prediction is trained with a cosine loss over corresponding columns of the predicted and native rotation matrices,

$$\mathcal{L}^o = \frac{1}{|\mathcal{M}|} \sum_{i \in \mathcal{M}} \sum_{d=1}^{3} \left[1 - \langle \hat{o}_{1,d}^i, o_{1,d}^i \rangle\right], \tag{32}$$

where $\hat{o}_{1,d}^i$ and $o_{1,d}^i$ denote column d of $\hat{O}_1^i$ and O_1^i, respectively.

At sampling time, the network produces a clean-orientation estimate $\hat{O}_1^i$ at each step. The current orientation is advanced along the geodesic toward this estimate using the same base-time conditional fraction η_j,

$$O_{t_{j+1}}^i = O_{t_j}^i \exp\left[\eta_j \log\left((O_{t_j}^i)^\top \hat{O}_1^i\right)\right], \qquad \eta_j = \frac{t_{j+1} - t_j}{1 - t_j}. \tag{33}$$

Sequence and orientation therefore remain on the ordinary base-time path, while positions alone use the mode-dependent clocks $A_k(t)$. Consequently, β changes only positional transport and leaves the sequence and rotational probability paths unchanged.

A.5 Path-Consistent Sampling

At sampling step j, let $\hat{c}_{1,k}$ denote the predicted clean endpoint for mode k. Holding this prediction fixed over $[t_j, t_{j+1}]$, the current state can be expressed as

$$c_{t_j,k} = A_k(t_j)\hat{c}_{1,k} + [1 - A_k(t_j)]\,\tilde{c}_{0,k}, \tag{34}$$

where $\tilde{c}_{0,k}$ is the source coefficient implied by the current state and endpoint prediction. Rearranging gives $\tilde{c}_{0,k} = \frac{c_{t_j,k} - A_k(t_j)\hat{c}_{1,k}}{1 - A_k(t_j)}$.

Evaluating the same conditional path at t_{j+1} and substituting this expression yields

$$\begin{aligned} c_{t_{j+1},k} &= A_k(t_{j+1})\hat{c}_{1,k} + [1 - A_k(t_{j+1})]\,\tilde{c}_{0,k} \\ &= \hat{c}_{1,k} + \frac{1 - A_k(t_{j+1})}{1 - A_k(t_j)}\left(c_{t_j,k} - \hat{c}_{1,k}\right). \end{aligned} \tag{35}$$

This recovers the update used in Eq. 17. If $\hat{c}_{1,k} = c_{1,k}$ and the current state lies on the true conditional path, then $\tilde{c}_{0,k} = c_{0,k}$ and Eq. 35 gives

$$c_{t_{j+1},k} = A_k(t_{j+1})c_{1,k} + [1 - A_k(t_{j+1})]\,c_{0,k}. \tag{36}$$

Hence an exact endpoint prediction advances the state exactly along the prescribed modal path rather than approximating the trajectory through Euler integration.

A.6 Relation Between Velocity Matching and Endpoint Regression

Let $\hat{c}_{1,k}$ denote the modal coefficient obtained from the network's predicted clean structure. From Eq. 24, the predicted and exact conditional velocities are

$$v_{\theta,k} = q_k(t)\,(\hat{c}_{1,k} - c_{t,k})\,, \qquad u_{t,k} = q_k(t)\,(c_{1,k} - c_{t,k})\,. \tag{37}$$

Their difference is $v_{\theta,k} - u_{t,k} = q_k(t)(\hat{c}_{1,k} - c_{1,k})$, so

$$\|v_{\theta,k} - u_{t,k}\|_2^2 = q_k(t)^2\,\|\hat{c}_{1,k} - c_{1,k}\|_2^2\,. \tag{38}$$

Velocity matching under the spectral path is therefore equivalent to clean-endpoint regression with a mode-dependent factor $q_k(t)^2$. To avoid changing the overall positional-loss scale with t and β, we normalize these factors as

$$w_k(t) = \frac{q_k(t)^2}{\frac{1}{n}\sum_{m=1}^{n} q_m(t)^2}, \qquad \frac{1}{n}\sum_{k=1}^{n} w_k(t) = 1. \tag{39}$$

The resulting positional objective is

$$\mathcal{L}^x = \frac{1}{n}\sum_{k=1}^{n} w_k(t)\,\|\hat{c}_{1,k} - c_{1,k}\|_2^2\,. \tag{40}$$

When $\beta = 0$, all modes share the same clock and therefore the same $q_k(t)$, giving $w_k(t) = 1$ for every k. The objective reduces to

$$\mathcal{L}^x = \frac{1}{n}\sum_{k=1}^{n} \|\hat{c}_{1,k} - c_{1,k}\|_2^2\,. \tag{41}$$

Since U is orthonormal, Parseval's identity gives

$$\sum_{k=1}^{n} \|\hat{c}_{1,k} - c_{1,k}\|_2^2 = \left\|\hat{X}_1 - X_1\right\|_F^2\,. \tag{42}$$

Thus the $\beta = 0$ objective also recovers ordinary clean-coordinate regression in Cartesian space, up to the same normalization by the number of generated residues.

B Additional Empirical Analyses

B.1 Complete per-CDR benchmark results

Table 6 reports complete benchmark results across all six CDRs using 100 generated samples per antibody–antigen complex. We compare SP-FMAG with RAbD, DiffAb, DiffAb with the analytic categorical posterior, DIFFFORCE, AbMEGD, and MDMAD using amino-acid recovery (AAR), binding-energy improvement rate (IMP), and backbone RMSD.

Across the non-H3 CDRs, SP-FMAG remains competitive but does not consistently dominate every metric. It achieves the highest IMP on L2 and the lowest RMSD on L3, while other methods remain stronger on several H1, H2, and L1 metrics. The main text therefore focuses on H3, where the spectral path yields its clearest structural advantage. The relationship between CDR length and the expected benefit of spectral separation is examined in Appendix B.5.

Table 6: Complete per-CDR benchmark results using 100 samples per complex. SP-FMAG uses $\beta = 4$. Best results are shown in **bold** and second-best results are underlined.

Heavy-chain CDRs

Method	H1			H2			H3		
	AAR↑	IMP↑	RMSD↓	AAR↑	IMP↑	RMSD↓	AAR↑	IMP↑	RMSD↓
RAbD	22.85	43.88	2.261	25.50	**53.50**	1.641	22.14	23.25	**2.900**
DiffAb	65.75	36.40	1.200	49.31	38.40	**1.058**	26.79	19.90	3.616
DiffAb + analytic posterior	79.72	**61.40**	**1.058**	68.72	41.10	1.125	40.78	24.90	3.360
DIFFFORCE	60.78	49.45	1.561	53.51	36.81	1.401	29.52	30.22	3.612
AbMEGD	61.78	43.30	1.218	41.23	38.70	1.084	25.06	23.20	3.914
MDMAD	**83.73**	56.40	1.734	**78.73**	35.00	1.691	**42.85**	32.40	3.381
SP-FMAG, $\beta = 4$	79.06	46.60	1.302	70.20	42.60	1.272	36.78	**35.70**	2.958

Light-chain CDRs

Method	L1			L2			L3		
	AAR↑	IMP↑	RMSD↓	AAR↑	IMP↑	RMSD↓	AAR↑	IMP↑	RMSD↓
RAbD	34.27	**46.81**	**1.204**	26.30	56.94	1.767	20.73	**55.63**	1.624
DiffAb	56.67	28.40	1.349	59.32	49.60	1.424	46.47	42.40	1.665
DiffAb + analytic posterior	78.36	31.10	1.368	88.65	52.10	**1.368**	76.06	54.90	1.569
AbMEGD	52.27	39.80	1.388	54.33	43.90	1.606	41.72	42.90	1.713
MDMAD	**83.61**	28.50	2.021	**89.39**	45.30	1.907	75.65	44.80	1.971
SP-FMAG, $\beta = 4$	83.19	38.20	1.427	88.75	**58.90**	1.426	**79.30**	45.40	**1.535**

B.2 Spectral Content of Native CDR Loops

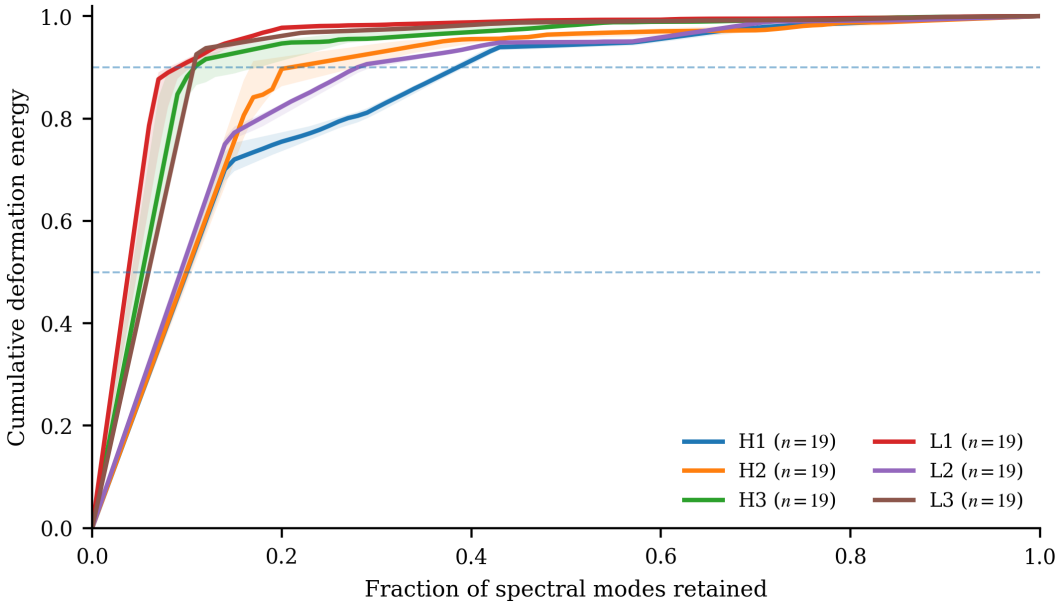


Figure 4: Cumulative deformation energy across the fixed-end spectral modes for all six CDR types. Curves show the median cumulative energy, with shaded interquartile ranges across held-out complexes.

To quantify how native CDR deformation is distributed across the fixed-end spectrum (Figure 4), we define the spectral deformation energy of mode k as $E_k = \|c_k\|_2^2$, where $c_k \in \mathbb{R}^3$ is the corresponding modal coefficient. We normalize these values within each loop as $\tilde{E}_k = E_k / \sum_m E_m$ and

compute the cumulative spectral content

$$F(K) = \sum_{k=1}^{K} \tilde{E}_k.$$

Across all six CDR types, native deformation is strongly concentrated toward the low-frequency portion of the fixed-end spectrum. At 25% of the normalized mode spectrum, the median cumulative deformation energy is 78.6% for H1, 91.3% for H2, 95.1% for H3, 98.1% for L1, 87.1% for L2, and 97.0% for L3. Thus, a relatively small fraction of low-frequency modes captures most of the deformation energy, while progressively higher-frequency modes account for the remaining finer geometric variation.

This concentration supports the fixed-end basis as an ordered multiscale representation of CDR geometry. The ordering experiment in Section 5.4 separately tests whether this hierarchy should be traversed coarse-to-fine during generation.

B.3 Does the sampler follow the spectral path?

The preceding results show that the spectral schedule improves H3 structural recovery, but they do not establish whether the learned sampler remains close to the prescribed mode-dependent trajectory. We therefore measure path tracking directly, with the resulting residuals shown in Figure 5.

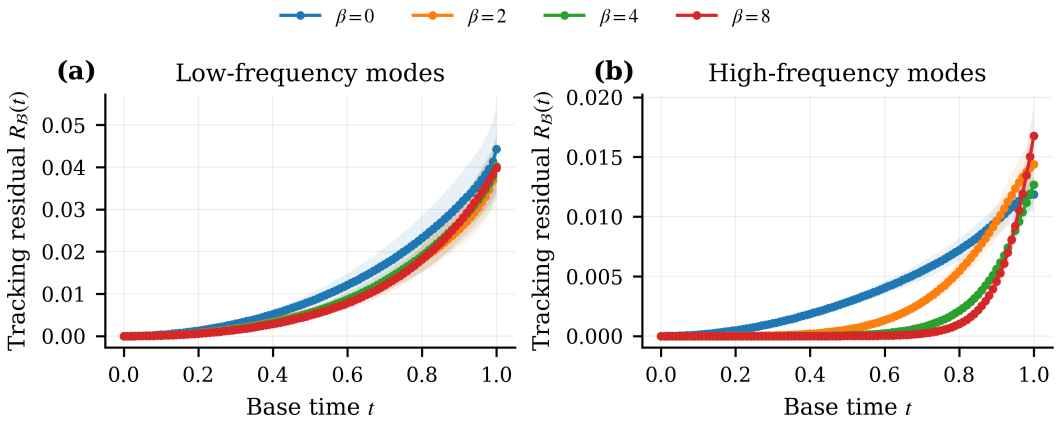


Figure 5: **Tracking residual along the sampling trajectory.** Normalized deviation from the prescribed conditional path for the lower and upper halves of the spectrum under different spectral strengths β. Low-frequency tracking is largely insensitive to spectral strength, whereas high-frequency trajectories differ substantially across spectral schedules. Shaded regions denote the standard error across the 19 held-out H3 complexes. At $t = 1$, the reference state equals the native endpoint, so the residual reduces to normalized terminal generation error.

For a sampled source state and its native target, the reference state of mode k at time t is $c_{t,k}^{\mathrm{ref}} = A_k(t)c_{1,k} + (1 - A_k(t))\, c_{\mathrm{src},k}$. where $c_{\mathrm{src},k}$ is the source modal coordinate and $c_{1,k}$ is the native modal coordinate. For a frequency band $\mathcal{B}$, we measure the normalized deviation from this trajectory as

$$R_{\mathcal{B}}(t) = \frac{\sum_{k\in\mathcal{B}} \|c_{t,k} - c_{t,k}^{\mathrm{ref}}\|_2^2}{\sum_{k\in\mathcal{B}} \|c_{\mathrm{src},k} - c_{1,k}\|_2^2}. \tag{43}$$

We compute $R_{\mathcal{B}}(t)$ separately over the lower and upper halves of the spectrum. Across all schedules, the residual remains small over most of the trajectory, indicating that the learned sampler stays close to the conditional path prescribed during training. Tracking in the low-frequency band is largely insensitive to β, consistent with the comparatively small change in the corresponding clocks. In contrast, the high-frequency trajectories differ more strongly across spectral schedules while remaining close to their respective prescribed paths. Near $t = 1$, the residual becomes normalized endpoint generation error, so its ordering across β should be interpreted as terminal accuracy rather than intermediate path tracking.

B.4 Statistical Analysis of CDR-H3 Improvements

To quantify uncertainty in the CDR-H3 structural improvement, we perform paired bootstrap resampling over the 19 held-out test complexes. For each of 10,000 bootstrap replicates, we sample 19 complexes with replacement from the original test set, so that some complexes may appear multiple times while others may be absent, and recompute the mean paired RMSD difference between the $\beta = 4$ spectral model and the matched $\beta = 0$ linear baseline. Repeating this procedure produces an empirical distribution of the main improvement, from which we obtain the 95% confidence interval. Resampling is performed at the complex level rather than over individual generated designs, preserving the paired comparison between the two models.

Table 7: Paired bootstrap analysis of CDR-H3 RMSD over the 19 held-out complexes. Improvement is defined as $\mathrm{RMSD}_{\beta=0} - \mathrm{RMSD}_{\beta=4}$, so positive values favor the spectral path.

Linear $\beta = 0$	Spectral $\beta = 4$	Paired improvement
3.313 Å	2.958 Å	**0.355** Å 95% CI [0.098, 0.683] Å

As reported in Table 7, the spectral path reduces mean CDR-H3 RMSD by 0.355 Å relative to the matched linear baseline. The 95% paired bootstrap confidence interval spans 0.098–0.683 Å and remains entirely above zero. Lower RMSD is also observed on 13 of the 19 held-out complexes, indicating that the aggregate improvement is not driven solely by a small subset of test cases.

B.5 Where Does the Spectral Path Help?

To examine where the spectral path provides the largest structural benefit, we compare the paired CDR-H3 RMSD improvement across held-out complexes as a function of native H3 length. We define

$$\Delta\mathrm{RMSD} = \mathrm{RMSD}_{\beta=0} - \mathrm{RMSD}_{\beta=4}$$

so that positive values indicate improved structural recovery under the spectral path.

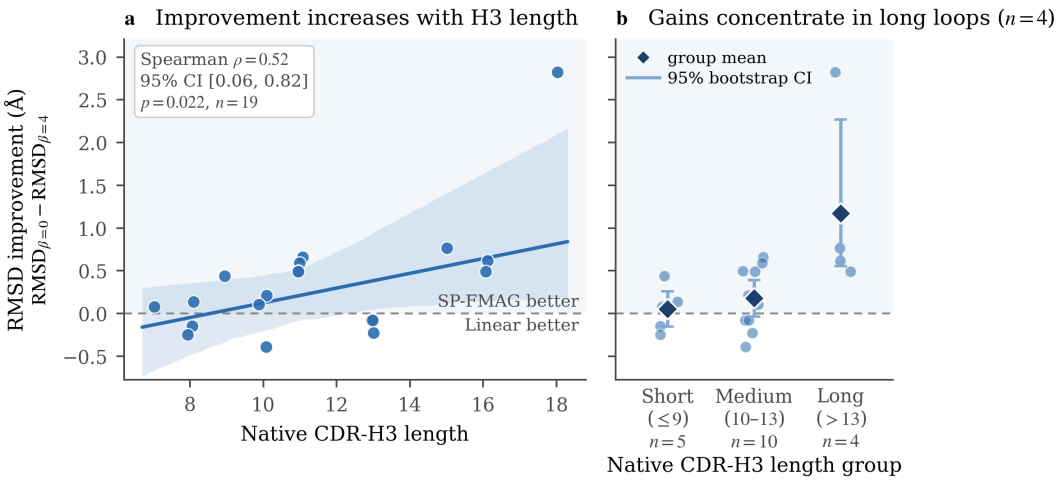


Figure 6: CDR-H3 structural improvement versus loop length. (a) Per-complex RMSD improvement, defined as $\mathrm{RMSD}_{\beta=0} - \mathrm{RMSD}_{\beta=4}$, against native CDR-H3 length ($n = 19$ complexes); positive values favour SP-FMAG. The line is a Theil–Sen fit and the band its pointwise 95% bootstrap interval (2000 resamples of complexes). Spearman $\rho = 0.52$ (95% bootstrap CI [0.06, 0.82]), $p = 0.022$. Points are jittered horizontally by at most ± 0.12 residues for legibility. (b) The same values stratified by length group (short (≤ 9), medium (10–13), long (> 13) residues). Diamonds are group means and whiskers are 95% percentile bootstrap CIs (10000 resamples); Dashed lines mark zero improvement.

Across the 19 held-out complexes (Figure 6), larger improvements tend to occur for longer H3 loops. This suggests that the multiscale spectral path is particularly useful in structurally more challenging H3 cases.

B.6 Reversed Spectral Ordering

Table 8 reports the complete comparison between the forward and reversed spectral schedules across all six CDRs. Both models use $\beta = 4$ and differ only in the assignment of clock exponents across spectral frequencies.

Table 8: Per-CDR comparison of forward and reversed spectral ordering at $\beta = 4$. Forward assigns earlier progression to lower-frequency modes, whereas Reverse applies the opposite ordering. We report amino-acid recovery (AAR), improvement rate (IMP), and backbone RMSD. Higher AAR and IMP are better, while lower RMSD is better.

Heavy-chain CDRs

	H1			H2			H3		
Schedule	AAR↑	IMP↑	RMSD↓	AAR↑	IMP↑	RMSD↓	AAR↑	IMP↑	RMSD↓
Forward	79.06	**46.6**	1.302	70.20	42.6	1.272	36.78	**35.7**	**2.958**
Reverse	**82.86**	41.2	**1.187**	**74.44**	**44.5**	**1.184**	**37.01**	21.1	3.499

Light-chain CDRs

	L1			L2			L3		
Schedule	AAR↑	IMP↑	RMSD↓	AAR↑	IMP↑	RMSD↓	AAR↑	IMP↑	RMSD↓
Forward	83.19	38.2	**1.427**	88.75	**58.9**	**1.426**	**79.30**	45.4	**1.535**
Reverse	**84.15**	**43.7**	1.459	**89.39**	58.2	1.659	72.90	**60.6**	1.790

As shown in Table 8, the effect of spectral ordering varies across CDRs. The clearest difference is observed for CDR-H3, where reversing the ordering increases RMSD from 2.958 Å to 3.499 Å and reduces IMP from 35.7% to 21.1%, while AAR remains nearly unchanged. Forward ordering also yields lower RMSD for L1, L2, and L3, whereas the reversed schedule performs better on H1 and H2. These results indicate that the advantage of coarse-to-fine ordering is most pronounced for the structurally challenging H3 loop rather than being uniform across all CDRs.

C Extended Benchmark Results

C.1 Spectral-Strength Ablation Across All CDRs

As shown in Table 9, the effect of spectral strength varies across CDRs. The clearest benefit is observed for CDR-H3, where moderate spectral separation improves structural recovery and binding-related performance relative to the matched linear path, with the best RMSD and IMP obtained at $\beta = 4$. Increasing the strength further to $\beta = 8$ reduces this gain, consistent with the main-text observation that excessive separation between spatial scales is detrimental.

Across the remaining CDRs, the effect is more heterogeneous. Several shorter CDRs perform best under the linear $\beta = 0$ path for RMSD or AAR while individual spectral settings improve selected metrics such as IMP on L2 and structural recovery on L3.

Table 9: Ablation of the spectral path strength β across all six CDRs (100 samples per complex). $\beta = 0$ recovers the standard flow-matching positional path, while increasing β progressively delays higher-frequency modes relative to low-frequency modes. We report amino-acid recovery (AAR), improvement rate (IMP), and backbone RMSD. Best results among the tested spectral schedules are shown in **bold**.

Heavy-chain CDRs

	H1			H2			H3		
β	AAR↑	IMP↑	RMSD↓	AAR↑	IMP↑	RMSD↓	AAR↑	IMP↑	RMSD↓
0	**82.82**	**57.30**	1.286	**76.22**	39.40	**1.105**	**38.97**	23.80	3.313
2	78.83	44.00	**1.218**	74.19	**42.90**	1.166	36.96	26.30	3.079
4	79.06	46.60	1.302	70.20	42.60	1.272	36.78	**35.70**	**2.958**
8	79.00	44.60	1.363	68.60	45.70	1.279	34.22	30.90	3.143

Light-chain CDRs

	L1			L2			L3		
β	AAR↑	IMP↑	RMSD↓	AAR↑	IMP↑	RMSD↓	AAR↑	IMP↑	RMSD↓
0	81.69	**40.90**	**1.321**	**90.68**	53.60	**1.366**	73.98	**49.90**	1.582
2	84.47	31.10	1.478	89.77	53.20	1.612	76.64	40.50	1.724
4	83.19	38.20	1.427	88.75	**58.90**	1.426	**79.30**	45.40	**1.535**
8	81.85	29.20	1.510	85.78	48.20	1.433	79.38	50.70	1.589

C.2 Low-NFE Results Across All CDRs

Table 10: Per-CDR performance with only $T = 5$ sampling steps. We compare the standard flow-matching path ($\beta = 0$) with the spectral path ($\beta = 4$) under the same low-NFE sampling budget. We report amino-acid recovery (AAR), improvement rate (IMP), and backbone RMSD. Higher AAR and IMP are better, while lower RMSD is better.

Heavy-chain CDRs

	H1			H2			H3		
β	AAR↑	IMP↑	RMSD↓	AAR↑	IMP↑	RMSD↓	AAR↑	IMP↑	RMSD↓
0	82.75	52.10	1.313	76.02	41.60	1.165	38.77	26.30	3.281
4	79.04	42.80	1.341	70.36	37.50	1.260	37.57	21.00	2.928

Light-chain CDRs

	L1			L2			L3		
β	AAR↑	IMP↑	RMSD↓	AAR↑	IMP↑	RMSD↓	AAR↑	IMP↑	RMSD↓
0	81.43	40.80	1.399	90.90	50.70	1.418	74.21	50.20	1.562
4	83.30	44.20	1.479	88.89	54.00	1.503	79.22	52.30	1.572

Tables 10 and 11 report per-CDR performance at $T = 5$ and $T = 1$. Across the remaining CDRs, the effect of the spectral path under low-NFE sampling is mixed. The clearest structural advantage is retained for CDR-H3, where $\beta = 4$ achieves lower RMSD than the matched $\beta = 0$ model at both $T = 5$ and $T = 1$. For the other CDRs, differences are smaller and vary across metrics, with no consistent advantage for either schedule. The results therefore support the main-text conclusion that the low NFE benefit is concentrated primarily on the structurally challenging H3 loop rather than being uniform across all CDRs.

Table 11: Per-CDR performance with only $T = 1$ sampling step. We compare the standard flow-matching path ($\beta = 0$) with the spectral path ($\beta = 4$) under the same one-step sampling budget. We report amino-acid recovery (AAR), improvement rate (IMP), and backbone RMSD. Higher AAR and IMP are better, while lower RMSD is better.

Heavy-chain CDRs

	H1			H2			H3		
β	AAR↑	IMP↑	RMSD↓	AAR↑	IMP↑	RMSD↓	AAR↑	IMP↑	RMSD↓
0	81.53	53.40	1.554	74.80	41.20	1.311	39.15	22.30	3.180
4	79.08	43.80	1.568	70.50	39.10	1.530	37.37	26.40	3.005

Light-chain CDRs

	L1			L2			L3		
β	AAR↑	IMP↑	RMSD↓	AAR↑	IMP↑	RMSD↓	AAR↑	IMP↑	RMSD↓
0	80.31	36.30	1.625	90.70	46.90	1.631	74.50	59.10	1.652
4	83.37	38.50	1.748	89.29	56.20	1.789	79.20	40.50	1.774

D Experimental and Implementation Details

D.1 Dataset and preprocessing

We use the SAbDab split released with DiffAb (Luo et al., 2022). Structures are restricted to protein antigens with resolution at most 4.0 Å, and antibody chains are processed using Chothia numbering (Chothia & Lesk, 1987). CDR3 loops longer than 30 residues are removed. Complexes are clustered using MMseqs2 (Steinegger & Söding, 2017) with 50% minimum sequence identity and 80% coverage. Test complexes are drawn from the five antigen groups used by DiffAb, while 20 non-test complexes are reserved for validation and the remainder are used for training.

Training examples mask one CDR at a time, merge antibody and antigen chains, and extract a local patch around the CDR anchors. Native sequence and structural information of the generated CDR are masked before residue and pair feature construction to prevent leakage.

D.2 Model and training

Residue and pair feature dimensions are 128 and 64, respectively. We use a six-layer geometry-aware attention encoder with 12 heads, query/key and value dimensions of 32, and eight query and value points per head. Separate MLP heads predict amino-acid identity, clean orientation, and clean position. The sequence state uses 20 amino-acid tokens plus an explicit `MASK` token.

All models are trained for 200,000 optimization steps with batch size 4 and gradient accumulation over four batches. We use Adam with learning rate 5×10^{-5}, $(\beta_1, \beta_2) = (0.9, 0.999)$, no weight decay, gradient clipping at 10, and seed 2022. Validation is performed every 1,000 steps. A plateau scheduler uses factor 0.8, patience 10, and minimum learning rate 5×10^{-6}. EMA decay is 0.995. Sequence, position, and orientation losses are equally weighted.

The reported models use the same isotropic positional source $c_{0,k} \sim \mathcal{N}(0, I_3)$ and differ only in the spectral clock. We evaluate $\beta \in \{0, 2, 4, 8\}$, with $\beta = 4$ as the main model and $\beta = 0$ as the matched linear-flow baseline.

D.3 Sampling and optimization

Unless otherwise stated, sampling uses $T = 100$ steps and generates 100 samples per complex and CDR with seed 2022. Reduced-budget experiments use the same checkpoint with $T \in \{1, 5, 100\}$. Native CDR sequence and structure are removed before generation, and predicted residue frames are reconstructed into backbone coordinates afterward.

Antibody optimization uses the pretrained $\beta = 4$ model without retraining. For optimization strength m, sampling starts from $t_{\text{start}} = 1 - m/T$ after partially perturbing the native CDR

using the same sequence, rotation, and positional paths used during training. We evaluate $m \in \{1, 2, 4, 8, 16, 32, 64\}$.

D.4 Relaxation and evaluation

Generated structures are first minimized using OpenMM (Eastman et al., 2017) after PDBFixer preprocessing. We use the Amber99SB force field with hydrogen-bond constraints and restrain non-generated non-hydrogen atoms with harmonic stiffness 10. Minimization uses tolerance $10\,\mathrm{kJ\,mol^{-1}\,nm^{-1}}$.

Structures are then relaxed using PyRosetta (Chaudhury et al., 2010) FastRelax with the `ref2015` score function (Alford et al., 2017), two repeats, and at most 1,000 iterations. Sequence design is disabled. Backbone motion is allowed for the generated CDR, while side chains are repacked within the generated region and its neighborhood. We report amino-acid recovery, C_α RMSD, and binding-energy improvement rate. Binding energy is computed using PyRosetta `InterfaceAnalyzerMover` with separated-state repacking. For each sample, $\Delta\Delta G = \Delta G_{\mathrm{design}} - \Delta G_{\mathrm{native}}$, and a design is counted as improved when $\Delta\Delta G < 0$. IMP denotes the percentage of improved samples.

D.5 Controls

All SP-FMAG ablations use the same architecture, source distribution, training procedure, and evaluation pipeline. The $\beta = 0$ model exactly recovers linear positional flow matching. The reverse-ordering control preserves the same range of clock exponents and changes only their assignment across spatial frequencies.