ZetaDial: dialing net charge of protein binders at inference time for therapeutic developability

Mohammed Sameer Syed[1] and Tamara Dinneen[2]

[1]*Independent Researcher; E-mail: mohammedsameersyed1@gmail.com*
[2]*Independent Researcher; E-mail: tamaradinneen@gmail.com*

Net charge and isoelectric point are among the strongest sequence-level predictors of whether a protein binder can become a therapeutic: they govern high-concentration viscosity, in-vivo clearance and half-life, nonspecific binding, and aggregation, and the biologics field already uses quantified charge windows as go/no-go developability criteria.[1–4] Yet deep inverse-folding models such as ProteinMPNN,[5] and the de-novo binder pipelines built on them such as BindCraft,[6] expose *no control* over this property: a practitioner cannot request "the same fold, but net-neutral." ProteinMPNN already ships a per-amino-acid logit bias, and recent work uses analogous bias or guidance to shift charge in generative models,[7–9] but these are *open-loop* with respect to a numeric setpoint: they shape a distribution or follow a monotonic knob without measuring the realized charge per design and correcting it. We present ZetaDial, a closed-loop, per-protein charge controller whose adaptive-slope (secant) update re-estimates each protein's charge response online and drives it to a specified net-charge target. On one matched target set under real stochastic sampling, ZetaDial is the best controller: it beats a per-protein fixed-slope loop on both a heterogeneous set of RCSB protein–protein complexes and de-novo Cas13 binder monomers, and beats the best matched global bias decisively on the heterogeneous complexes (mean absolute error 5.2 vs 11.7 charge units) while matching it on the homogeneous monomers (5.6 vs 5.5). This delineates *when* closing the loop helps, a relationship we quantify as a predictive law: the calibration gain grows with a design set's charge-sensitivity heterogeneity (Pearson $r = 0.79$), so a practitioner can tell in advance whether to close the loop. A protein's charge sensitivity is itself predictable from composition and size (cross-validated R^2 of 0.77/0.53), enabling a probe-free warm start, and both controllers crush a best-of-20 unbiased filtering baseline (5-fold lower error). Charge control is essentially free within a characterized foldability band, and transfers into BindCraft across three disease-area targets: against PD-L1 (oncology) the binder is neutralized from net charge -5.8 to -0.5 with interface confidence preserved (best interface pTM $0.85 \rightarrow 0.83$), collapsing into poly-lysine only when driven far past the band, while IL-7Rα (immunology) tolerates a more positive setting, establishing a per-target electrostatic tolerance window with bootstrapped uncertainty. Neutralization is moreover developability-improving: it moves PD-L1 binders from 0% to 21% simultaneously in-window and binding, and the scalar dial demonstrably collapses the largest same-sign surface charge patch, a spatial liability, on both de-novo binders and real antibody Fv domains (the molecule class the windows derive from). We frame this as a *positive capability*, not a repair: on the same data, ProteinMPNN's charge placement is already as good as or better than native. All evidence is computational.

Keywords: Protein design; inverse folding; ProteinMPNN; BindCraft; net charge; developability; closed-loop control; controllable generation; precision medicine.

1. Introduction

Inverse folding, predicting a sequence that folds to a given backbone, has been transformed by deep generative models led by ProteinMPNN,[5] and de-novo binder pipelines such as BindCraft,[6] built on RFdiffusion[10] and AlphaFold2,[11] now compose these foundation models into functional binders in one shot, an instance of generative therapeutic design. Yet they are *unconditioned* on the biophysical properties that decide a molecule's fate downstream, the gap between raw generative output and a clinically viable therapeutic.

Net charge is the prime example. For therapeutic binders, net charge and isoelectric point are a multi-axis developability lever: excess variable-domain charge shortens half-life through non-specific uptake and FcRn-mediated clearance;[4,12,13] surface-charge patches and asymmetry drive self-association and high-concentration viscosity;[1] and both extremes elevate polyspecificity and aggregation.[2,14] The field uses quantified "charge windows" (e.g. an Fv net charge near 0–6)[1] as go/no-go developability criteria,[3,15] yet no inverse-folding or binder-design tool lets a practitioner ask for "the same interface, but net-neutral"; charge is left to whatever the model emits.

We treat charge instead as a first-class, continuously tunable design variable, without retraining, and carry that control into therapeutic binder design. We are explicit about what is new. ProteinMPNN already ships a per-amino-acid logit bias, and applying it to K/R/D/E uses that existing feature; recent work likewise steers net charge or pI in generative models by analytic bias,[7] position-specific bias,[8] or inference-time guidance.[9] What none provides, and ZetaDial contributes, is a way to *aim* the knob at a specified scalar value, per protein, and carry it into a binder pipeline with binding tracked:

(1) **ZetaDial: a closed-loop, per-protein charge controller with an adaptive-slope (secant) update** that tracks the *saturating* charge response a fixed proportional step overshoots. On a matched target set it is the best controller, significantly beating a fixed-slope loop and a global bias on heterogeneous complexes, tying the global bias on homogeneous monomers, and crushing best-of-N filtering.

(2) **A predictive law for when to close the loop, plus predictable sensitivity:** the calibration gain grows with a design set's charge-sensitivity heterogeneity ($r = 0.79$), and sensitivity is itself predictable from composition, enabling a probe-free warm start; a stagewise ablation isolates each loop component.

(3) **A characterization of the charge–foldability trade-off**, both by isolated-chain self-consistency and by a complex-aware AlphaFold2-Multimer refold, locating the band within which control is essentially free and showing its cost is foldability rather than interface disruption.

(4) **Transfer into de-novo binder design (BindCraft) with a per-target tolerance window and a developability payoff:** neutralization moves PD-L1 binders into the developable-and-binding regime, and the scalar dial collapses the largest same-sign surface charge patch on both de-novo binders and real antibody Fv domains.

We do not claim ProteinMPNN is deficient: on held-out data its designs are, if anything, *more* salt-bridged than native structures (a proxy salt-bridge count of 11.3 versus 9.1, $n \approx 170$, $p = 7 \times 10^{-5}$); the value is controllability, not repair, and, newly, its therapeutic relevance.

2. Related Work

Controllable and guided generation. Classifier and classifier-free guidance steer diffusion models toward target properties, and reweighted or biased sampling conditions autoregressive generators without retraining; our steering primitive is an instance of the latter. Several recent methods control composition or net charge: ProtNHF[7] adds analytic bias potentials for "smooth, predictable, approximately monotonic" net-charge control; Feynman–Kac steering[9] resamples diffusion particles toward target charge distributions; and predictor- or classifier-guided samplers over discrete sequence spaces (e.g. ProteinGuide[16]) condition generation on target property values without retraining. These are *open-loop* with respect to a numeric setpoint: they shape a distribution, follow a monotonic knob, or steer toward a property *value*, but do not measure the realized scalar per design and correct it. Our controller does exactly that, removing the protein-dependent residual these methods leave; the matched fixed-β baseline (Sec. 4) is a direct stand-in for such open-loop control, and calibration reduces its error most where it matters, on heterogeneous proteins and at distant targets.

Biased and retrained ProteinMPNN. ProteinMPNN's bias inputs are routinely used to shape outputs: Deng et al.[8] penalize charged surface residues to keep pI in a favorable window inside a binder pipeline; SolubleMPNN[17] retrains the weights to bias surface composition; ProtAlign/MoMPNN[18] fine-tunes ProteinMPNN by preference alignment to improve developability including surface net charge, on binder backbones. These control charge as a *constraint* or through *retraining*; none provides a signed, calibrated dial to an arbitrary net-charge target at inference, and only ProtAlign touches binders, via training, without interface-confidence tracking.

Charge, foldability, and interface electrostatics. Protein supercharging established that surface net charge can be pushed tens of units without abolishing the fold, with a soft ceiling where solubility degrades;[19,20] increased negative surface charge correlates with solubility while positive patches predict insolubility,[21,22] explaining the asymmetric tolerance we observe. In binder design, recent complementarity-based pipelines optimize *shape* complementarity (e.g. HECTOR,[23] which reaches nanomolar binders, including against IL-7Rα), and a large meta-analysis[24] finds interface *electrostatic* complementarity only a weak predictor of success, while the nearest charge-control work imposes interface electrostatics for pH-switching[25] rather than a net-charge setpoint. Net charge as a calibrated, tunable handle on the designed binder is thus largely unexplored; it is what we add to this pipeline.

3. Methods

Steering primitive (a native feature). ProteinMPNN samples autoregressively from per-position logits and exposes a per-amino-acid additive bias; the equivalent input exists in

ColabDesign's MPNN.[26] Our primitive sets that bias with a single signed scalar β, adding β to the lysine and arginine (K, R) logits and subtracting β from the aspartate and glutamate (D, E) logits at every position. Here $\beta > 0$ pushes designs more positive and $\beta = 0$ recovers vanilla ProteinMPNN. We sample at temperature 0.1; because softmax divides logits by the temperature, the effective strength of a given β scales inversely with it, which is why the usable β range differs between standalone runs and BindCraft. Net charge is scored as $(\#\mathrm{K} + \#\mathrm{R}) - (\#\mathrm{D} + \#\mathrm{E})$; developability descriptors (net charge at pH 7.4, isoelectric point) additionally account for histidine and chain termini via a Henderson–Hasselbalch calculation. The two agree closely: across our designs the count and the pH-7.4 net charge differ by 0.21 charge units on average (median 0.17, maximum 0.50), well below the control resolution, so calibrating on the count effectively calibrates the developability-relevant pH-7.4 charge.

Closed-loop calibration (ZetaDial). Because a fixed β produces a protein-dependent charge shift, hitting a *specified* target requires per-protein calibration. For each protein we probe with single samples across a small β grid, estimate the local slope $\mathrm{d}q/\mathrm{d}\beta$ by central difference (floored at a small positive value to avoid overshoot on charge-insensitive proteins), initialise $\beta = (q^\star - q_0)/\text{slope}$, sample, measure the miss, and correct, for a fixed budget of three iterations, keeping the closest sample. The charge response is *saturating*: a fixed slope estimated near $\beta = 0$ over-predicts the effect of large β, so a naive proportional loop overshoots, most damagingly on charge-insensitive monomers. ZetaDial therefore uses an adaptive-slope (secant) update: after each new sample it re-estimates the local slope from the two most recent (β, q) evaluations, guarded to update only when the response is locally monotone and otherwise falling back to the probe slope. This tracks the curvature of the response and converges in the same three-sample budget. All sampling is at temperature 0.1 with recorded seeds, and a single design is drawn per probe and per iteration, so the reported errors already fold in sampling noise rather than averaging it away; averaging two samples per evaluation further widens the secant advantage. As a matched baseline we grant a single global β its best value across all proteins for each target and report the residual it leaves; this is an open-loop global-bias comparator of the same class as analytic monotonic-knob steering.[7] For an internally consistent comparison we run all three controllers, global-β, fixed-slope loop, and ZetaDial's secant loop, on a *single* set of absolute charge-shift targets, scoring each protein only where the target is reachable, so Table 1 is a like-for-like head-to-head under real sampling.

Predicting charge sensitivity. A protein's sensitivity $\mathrm{d}q/\mathrm{d}\beta$ governs how strong a bias a target demands. We predict it from cheap sequence/structure-composition features (length, titratable-residue counts and fractions) with a ridge regressor evaluated by grouped cross-validation within each dataset, and use the prediction as a probe-free warm start for the loop.

Datasets and metrics. We evaluate on (1) charged protein–protein complexes from the RCSB, split by 30% sequence-identity clustering with MMseqs2[27] so that no test protein is a homolog of any training protein, and (2) 96 de-novo Cas13 binder monomers. For the RCSB complexes the controller targets the complex-level net charge summed over all chains,

which is why the reachable span reaches ∼915 units on the largest multi-chain complexes; the Cas13 monomers and the BindCraft binders are single chains, matching the therapeutic scenario in which only the binder is designed. Foldability is measured by ESMFold[28] self-consistency RMSD (scRMSD). BindCraft[6] designs binders by AlphaFold2 back-propagation, ProteinMPNN redesign, and AlphaFold2-multimer[29] re-prediction with Rosetta scoring; we inject the identical charge bias into its MPNN step, attending to that implementation's amino-acid ordering (AlphaFold's restype alphabet, which differs from raw ProteinMPNN's; indexing with the wrong order silently biases the wrong residues and moves no charge). We design against PD-L1 (BindCraft example), IL-7Rα (PDB 3DI2 chain B), and SARS-CoV-2 RBD (PDB 6M0J chain E), each across a per-target β grid. The bias is applied uniformly to all designed positions; because BindCraft holds the interface residues fixed during redesign, it acts on non-interface (largely solvent-exposed) positions by construction, and restricting it to explicitly surface-exposed residues is a straightforward extension. Filters are disabled so every design is recorded; per (target, β) we report achieved net charge, best interface pTM (i_pTM), and a success rate (fraction of designs with i_pTM ≥ 0.5), which are robust to weak designs that mean i_pTM conflates with target difficulty.

4. Results

4.1. *ZetaDial is best controller, and calibration pays off with heterogeneity*

The raw bias moves net charge smoothly and monotonically; closing the loop turns it into a controller that hits specified targets. On one matched set of absolute charge-shift targets, scored per protein wherever reachable and under real stochastic sampling, we compare three controllers head to head (Table 1). Two findings stand out.

First, ZetaDial's secant loop is the best controller in both regimes: it beats the per-protein fixed-slope loop on the heterogeneous RCSB complexes (5.2 vs 6.5 mean absolute error, MAE) and on the homogeneous de-novo Cas13 monomers (5.6 vs 8.2). The adaptive slope is what earns this. Because the charge response saturates, a fixed slope estimated near $\beta = 0$ overshoots at larger β, and on the charge-insensitive Cas13 monomers this makes the fixed-slope loop *worse* than a well-chosen global bias (8.2 vs 5.5); the secant update, which re-reads the slope from the two most recent samples, removes that failure mode.

Second, the value of per-protein calibration over a single global bias depends on target heterogeneity. On the RCSB complexes, whose per-protein sensitivities span an order of magnitude, ZetaDial beats the best matched global bias decisively (5.2 vs 11.7), because no single β can fit proteins with such different responses. On the homogeneous Cas13 monomers a global bias is already near-optimal (5.5) and ZetaDial only matches it (5.6). A cluster bootstrap over proteins confirms the pattern: the ZetaDial-minus-global difference is -6.5 charge units on RCSB (95% CI $[-8.0, -5.0]$, $p < 0.001$) and a statistically indistinguishable $+0.10$ on Cas13 ($[-0.5, +0.7]$, $p = 0.63$). This is the honest boundary of the method, confirmed rather than asserted: closing the loop earns its extra samples precisely when the design set is heterogeneous, and a matched global bias suffices when it is not. The naive alternative of best-of-N filtering (sample many unbiased designs, keep the closest) is far weaker: unbiased designs cluster near the model's natural charge, so even 20 draws miss by 26.7/29.9 charge units on RCSB/Cas13

(Table 1), roughly five times ZetaDial's error, and grow worse as targets move away.

A stagewise ablation (Fig. 1) confirms the mechanism: on the mean charge-response curves, adding the per-protein slope, then corrective iterations, then the adaptive secant update each lowers error, and the same figure exposes the fixed-slope overshoot on Cas13 that the secant step repairs. Across the 114 RCSB complexes the secant advantage over the fixed loop is significant (-1.3, $[-2.4, -0.4]$, $p < 0.001$) but comes from taming the fixed loop's heavy overshoot tail rather than winning every target (the fixed loop is closer on a slight majority of individual cells yet far worse in the mean); on the 96 Cas13 monomers the secant advantage is larger and decisive ($p < 0.001$). Because developability windows span several charge units (e.g. a net charge of -4 to $+4$), the operative requirement is landing *within* such a band rather than hitting a target exactly; achieved charge tracks target charge closely (Pearson $r = 0.95$), with the majority of monomer targets, the single-chain case that matches therapeutic binders, inside the tolerance band.

Table 1. Charge-hit error on one matched set of absolute charge-shift targets, under real ProteinMPNN sampling, scored per protein where reachable (charge units; lower is better; single sample per evaluation). Columns count (protein, target) cells over 114 RCSB complexes and 96 Cas13 monomers. ZetaDial's secant loop is best on both datasets; per-protein calibration beats the matched global bias decisively on the heterogeneous complexes and matches it on the homogeneous monomers, and both crush a best-of-20 unbiased filtering baseline.

Controller	RCSB complexes ($n = 912$)	Cas13 monomers ($n = 768$)
Best-of-20 unbiased filtering	26.7	29.9
Best matched global β (open-loop)	11.7	**5.5**
Per-protein fixed-slope loop	6.5	8.2
ZetaDial (per-protein secant loop)	**5.2**	5.6

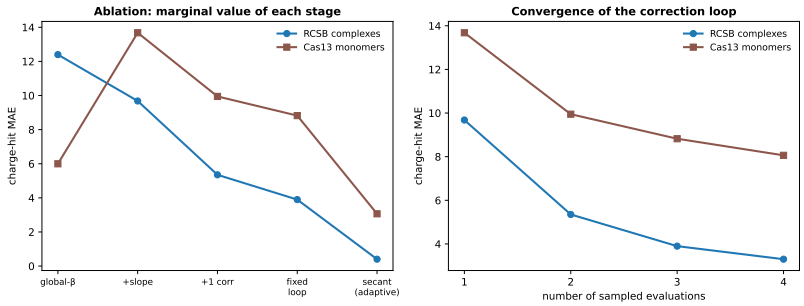

Fig. 1. Marginal value of each controller stage, evaluated on each protein's mean charge-response curve to isolate the algorithm from sampling noise (absolute errors are therefore optimistic relative to the single-sample deployable numbers in Table 1; the ordering is the point). Left: global-β, +per-protein slope, +one correction, the full fixed-slope loop, and the adaptive secant loop. On Cas13 a fixed slope overshoots and is worse than global-β until corrected; the secant update is best throughout. Right: error versus number of sampled evaluations.

4.2. *A predictive law for when to close the loop*

The heterogeneity boundary above is not a two-point anecdote; it is a smooth, quantitative law. Drawing random design subsets pooled across both datasets and measuring, per subset, its charge-sensitivity heterogeneity (the coefficient of variation of $\mathrm{d}q/\mathrm{d}\beta$) against the realized calibration gain (global-β MAE minus ZetaDial MAE), the two are strongly correlated (Pearson $r = 0.79$, Fig. 2). The gain is near zero for homogeneous sets (CV $\lesssim 0.3$, the Cas13 regime) and rises to tens of charge units as heterogeneity grows. This turns the Cas13 tie from an apparent weakness into a *confirmed prediction* and gives a practitioner an a-priori rule: estimate the spread of sensitivities (itself predictable from composition, above), and close the loop when it is large.

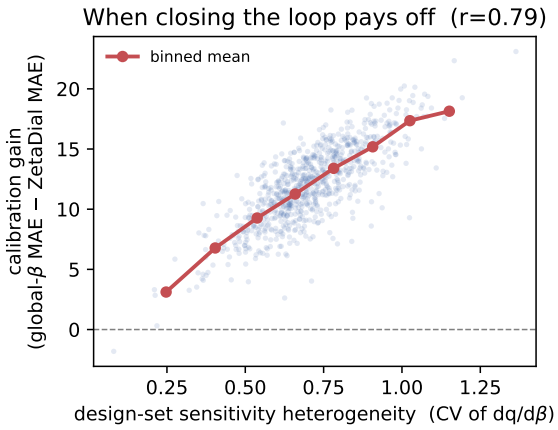

Fig. 2. When closing the loop pays off. Across random design subsets pooled from both datasets, the calibration gain (global-β MAE − ZetaDial MAE) grows with the subset's charge-sensitivity heterogeneity (CV of $\mathrm{d}q/\mathrm{d}\beta$); Pearson $r = 0.79$. Points are subsets; the line is the binned mean. Gain is near zero for homogeneous sets and large for heterogeneous ones.

4.3. *Charge sensitivity is predictable, enabling a probe-free warm start*

The loop spends most of its samples discovering a protein's charge sensitivity $\mathrm{d}q/\mathrm{d}\beta$. That sensitivity is largely set by composition and size and is predictable without any sampling: a ridge regressor on titratable-residue counts and length recovers it with grouped cross-validated R^2 of 0.77 on RCSB complexes and 0.53 on Cas13 monomers (Fig. 3). Because the response saturates, a single prediction is not accurate enough to skip the loop outright, but it warm-starts the loop closer to target and removes the probe phase, and within a homogeneous class most of the sensitivity is shared, so a class-level default already helps.

4.4. *Control is "for free" within a characterized band*

Charge control does not cost foldability until it is pushed hard (Fig. 4). On RCSB complexes, designs at $\beta = 0$ fold well (median scRMSD 1.69 Å, 77% under 5 Å), and a charge shift of

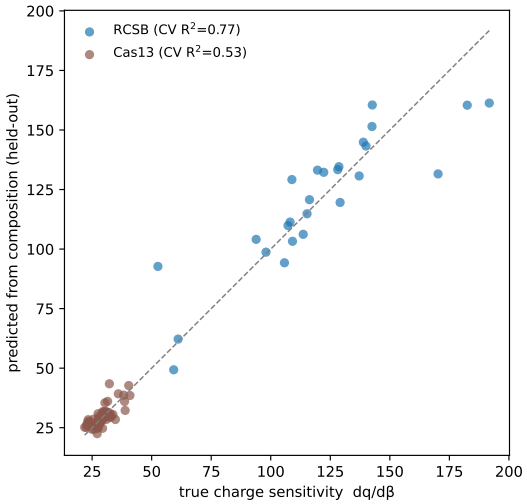

Fig. 3. Charge sensitivity $\mathrm{d}q/\mathrm{d}\beta$ predicted from composition and size (held-out) versus the measured value, per protein. Grouped cross-validated R^2 is 0.77 on RCSB complexes and 0.53 on Cas13 monomers, enabling a probe-free warm start.

$\sim\pm85$ units ($\beta = \pm1.5$) is absorbed at essentially no cost (scRMSD $\sim$2.3 Å); only at the extreme ($\beta = \pm3$) does the fold collapse. The de-novo Cas13 monomers fold near-perfectly at $\beta = 0$ (0.64 Å, 94%) with a narrower band in absolute charge, and these natively negative binders tolerate being pushed *toward* positive better than further negative, mirroring the surface-charge/solubility asymmetry.[21] One caveat: ESMFold self-consistency folds each chain in isolation, so on *complexes* it cannot see the interface where the perturbed charge lives; we therefore take the Cas13 *monomer* results as the cleaner designability evidence.

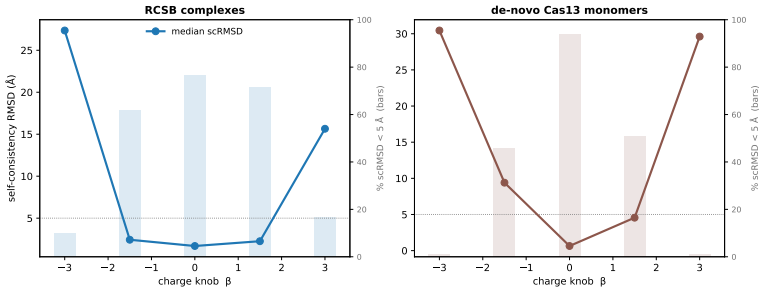

Fig. 4. Charge-vs-foldability trade-off on RCSB complexes (left) and de-novo Cas13 binders (right). Both show a usable band around $\beta = 0$ that widens for larger proteins.

4.5. *Complex-aware validation: the cost is foldability, not the interface*

The self-consistency metric above folds each chain in isolation, so it cannot see the binding interface. To test whether charge control damages the *interface* rather than only isolated-chain foldability, we refold 52 held-out designed complexes across the full charge sweep with AlphaFold2-Multimer (all 52 returned at every β; complex-aware, on GPU with full prove-

nance). Two costs separate cleanly. The dominant effect of charge bias is on *monomer* foldability: the designed-chain pLDDT falls with $|\beta|$ (by 30 points at $\beta = -3$; all 95% CIs exclude zero) and supercharged designs cross into the predicted-disordered regime (Uversky-disordered fraction 15% at $\beta = 0$ rising to 100% at every biased setting). Conditional on the chain still folding (pLDDT within 10 of $\beta = 0$), the interface is instead preserved: the primary endpoint ipSAE (interface Score from Aligned Errors, an interface-focused predicted-aligned-error score) has a change versus $\beta = 0$ whose 95% CI includes zero at every setting from -3 to $+3$ (e.g. -0.02, $[-0.05, +0.01]$ at $\beta = -3$), and interface pTM likewise includes zero at the mild settings. The apparent interface degradation across the full cohort therefore tracks the loss of monomer fold rather than interface disruption; the noisier interface RMSD agrees at mild β (95% CIs include zero at $\beta = \pm 1.5$) and shows residual disruption only at the extremes, the same band edge the monomer analysis locates. Where the calibrated charge lands is also developability-favourable: it accumulates on the surface and is kept out of the buried core (the core absorbs significantly less charge than a structure-blind placement, and this survives matching on residue burial), and it tunes interface electrostatic complementarity monotonically with β. One caveat: the fold-tolerant subset thins at the extremes (12 of 52 complexes at $\beta = -3$), so the conditional interface claim is strongest inside the usable band.

4.6. *Controllable-charge binders and a per-target tolerance window*

The knob operates inside BindCraft on every target: net charge is a smooth, monotonic function of β for PD-L1, IL-7Rα, and RBD alike (the horizontal axis of Fig. 5), spanning -5.8 to $+66.9$ on PD-L1. What the charge *buys*, and at what cost to binding, is target-dependent (Fig. 5, Table 2).

Against PD-L1, the unmodified pipeline ($\beta = 0$) produces binders with mean net charge -5.8 and strong interface confidence (best i_pTM 0.85; 39% of designs ≥ 0.5). A gentle push neutralizes the binder (to -2.1 at $\beta = 0.15$ and -0.5 at $\beta = 0.3$) while binding is *preserved* (best i_pTM 0.83 at both; success rate 39%$\rightarrow$62%$\rightarrow$38%). Only past the band does binding degrade (best i_pTM 0.64 at $\beta = 0.6$, 0.47 at $\beta = 1.0$), collapsing into poly-lysine at $\beta = 2.0$ (net charge $+66.9$, best i_pTM 0.10). This is the developability claim in one target: the binder can be moved into the favorable near-neutral charge window at no detectable cost to binding. Against IL-7Rα, strong binders (best i_pTM 0.79–0.83) are retained at neutral-to-mildly-positive charge ($\beta = 0.2$–0.5), collapsing by $\beta = 1.0$; the usable band thus sits at a *more positive* setting than PD-L1's, establishing that the electrostatic tolerance window is target-specific rather than universal (this target is noisier at two trajectories per point). That an immunology target (IL-7Rα) and an oncology target (PD-L1) admit different charge settings is precisely the target-specific tailoring precision medicine calls for: the same generative pipeline is adapted to the distinct biophysics of each disease target rather than held to one universal rule. Against RBD, charge control transfers cleanly ($-5.1 \rightarrow +17.1$), but BindCraft did not produce strong binders at *any* setting including baseline (best i_pTM ≤ 0.32), so RBD cannot demonstrate a binding-preservation window; we report it as charge-controllability evidence and an honest negative on binder quality. Across targets: charge controllability holds on all three; a binding-preservation tolerance window is clean on PD-L1 and supported on IL-7Rα. Bootstrapped

95% confidence intervals on the success rate confirm the pattern is not sampling noise: on PD-L1 the neutralized setting ($\beta = 0.15$) reaches 62% [42, 79]% success versus 39% [21, 57]% at baseline, and IL-7Rα peaks at 44% [19, 69]% in its positive band while its extremes sit at 0%; a split-half replication over design seeds preserves the target ordering.

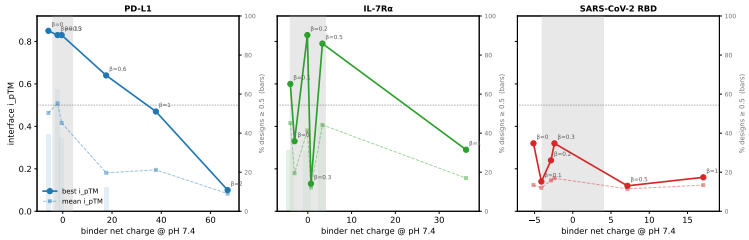

Fig. 5. Per-target electrostatic tolerance window: interface i_pTM (best, solid; mean, dashed) and success rate (% designs ≥ 0.5, bars) versus achieved net charge. Shaded band is literature-motivated favorable developability charge window (-4 to $+4$). PD-L1 retains strong binders through neutralization and collapses far positive; IL-7Rα's band sits more positive; RBD never reaches strong binding.

Table 2. Charge-controlled binder design across three targets in BindCraft (filters off; per-target β grid, key settings shown). "best" is the best interface pTM at that setting; "≥ 0.5" the fraction of designs reaching it.

Target	Setting	n	net charge	best i_pTM	% ≥ 0.5
PD-L1	$\beta = 0$ (vanilla)	28	-5.8	0.85	39%
PD-L1	$\beta = 0.15$	24	-2.1	0.83	62%
PD-L1	$\beta = 0.3$ (neutral)	16	-0.5	0.83	38%
PD-L1	$\beta = 2.0$ (poly-K)	16	$+66.9$	0.10	0%
IL-7Rα	$\beta = 0$ (vanilla)	16	-3.0	0.33	0%
IL-7Rα	$\beta = 0.5$	16	$+3.3$	0.79	44%
IL-7Rα	$\beta = 1.0$	16	$+36.2$	0.29	0%
RBD	$\beta = 0$ (vanilla)	16	-5.1	0.32	0%
RBD	$\beta = 0.5$	16	$+7.2$	0.12	0%

4.7. *Charge control improves computed developability, on binders and antibodies*

The tolerance window shows binding survives neutralization; does neutralization actually *buy* developability? On the PD-L1 binders, the vanilla pipeline (net charge -5.8) produces *no* designs inside the favorable pI/charge developability window, whereas the gently neutralized setting ($\beta = 0.15$, net charge -2.1) places 21% of designs simultaneously in-window and binding

(i_pTM ≥ 0.5), while raising the binding hit-rate from 39% to 62%. The dial thus converts binders into developable-and-binding candidates the unmodified pipeline never yields.

A net-charge scalar, however, is not the whole story: viscosity, polyspecificity, and aggregation track the *spatial* charge distribution. Threading ZetaDial sequences onto fixed backbones and measuring the largest same-sign surface charge patch directly (from Cα geometry, no folding), the patch scales tightly with net-charge magnitude (Fig. 6): neutralizing net charge collapses the dominant charged surface patch from 60-plus residues to about 4, in 100% of proteins, on de-novo Cas13 binders (Pearson $r = 0.81$) and, more strongly, on real antibody variable domains ($r = 0.89$). The scalar dial systematically reshapes the spatial electrostatics in the developability-favourable direction.

Because the charge windows are antibody-derived, we finally test ZetaDial on 25 real antibody Fv variable domains extracted from the PDB (heavy and light variable domains, truncated by their conserved J-region motif). Their native net charges already cluster in the favourable window (17/25 within 0–6, mean +1.5), a validation of the window on real antibodies. ZetaDial controls Fv net charge to a mean absolute error of 3.8 units, halving the 6.4 of a matched global bias (the secant and fixed loops are comparable here, as expected for this homogeneous class and consistent with the heterogeneity law), so charge outliers can be dialed into the window while the same spatial-patch reshaping holds, on the exact molecule class the windows derive from.

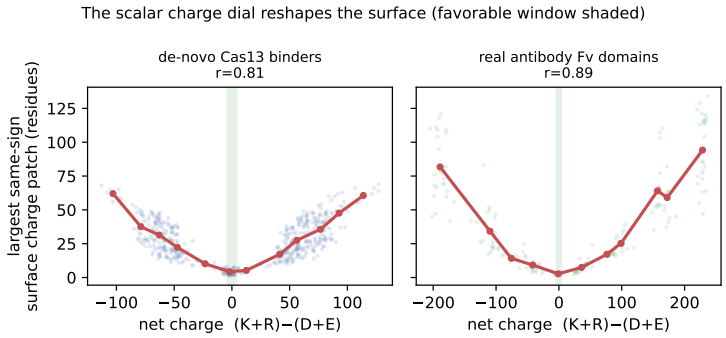

Fig. 6. The scalar charge dial reshapes the surface. Largest same-sign surface charge patch versus net charge for de-novo Cas13 binders (left, $r = 0.81$) and real antibody Fv domains (right, $r = 0.89$); points are designs on fixed backbones, line is the binned mean, shaded band is the favourable window. Neutralizing net charge collapses the dominant charged patch, the spatial liability behind viscosity and polyspecificity.

4.8. *Robustness and a surface-restricted variant*

Two controls show the method is not brittle. First, the controller is not tuned to a single sampling temperature: rerunning the secant loop at temperatures 0.1, 0.2, and 0.3 leaves accuracy essentially unchanged (Cas13 monomer MAE 6.4/5.6/5.6, with no degradation as temperature rises). Second, the uniform K/R/D/E bias can be restricted to solvent-exposed residues, identified by a Cα neighbour-count proxy, through ProteinMPNN's per-residue bias

input, leaving the core untouched. This surface-restricted variant hits the same targets at least as precisely as the uniform bias (RCSB MAE 12.3 vs 12.3; Cas13 4.4 vs 7.9, i.e. better on monomers) while using only about 55% of the reachable range, so charge can be tuned with a smaller and more physical structural footprint at a modest cost in authority.

5. Limitations

We report these results as a capability demonstration. (i) Our primary foldability metric folds each chain in isolation; we address this with a complex-aware AlphaFold2-Multimer refold of 52 designed complexes (Sec. 4.5), which shows charge control's cost is monomer foldability rather than interface disruption. That refold is itself in-silico, and its fold-tolerant subset thins at the charge extremes, so the conditional interface result is strongest within the usable band. (ii) The BindCraft binding evidence is proof-of-concept scale (two-to-three design trajectories per (target, β), in-silico metrics only); IL-7Rα is noisy at this depth. (iii) RBD did not yield strong binders at any setting, so its tolerance window is undetermined; more trajectories and revised hotspots would be needed. (iv) The knob β is an absolute logit shift whose charge effect is size- and temperature-relative, so charge-shift units are the fairer frame than β. (v) The developability windows are antibody-derived; we mitigate this by demonstrating control on real antibody Fv domains (Sec. 4), but the BindCraft binders remain small de-novo proteins for which the exact thresholds are guidance rather than validated cut-offs. (vi) All evidence is computational; wet-lab validation of charge-tuned binders remains the decisive next step. (vii) We show the scalar dial reshapes a key spatial liability (the largest same-sign surface patch) but do not *independently* control spatial descriptors; a calibrated loop that targets patchiness or surface-charge density directly is a natural extension.

Code and data availability

The charge-control and per-protein calibration scripts, the spatial-developability and antibody-Fv analyses, and the charge-controlled BindCraft notebook (including the amino-acid restype-indexing guardrails described in Methods) will be released publicly upon publication.

6. Conclusion

Building on ProteinMPNN's existing bias input, ZetaDial contributes a per-protein closed-loop charge controller whose adaptive secant update makes it the best controller we tested, together with the predictability of charge sensitivity, the charge-versus-foldability characterization, and the transfer into BindCraft, where neutralization is developability-improving on both de-novo binders and real antibody Fv domains. The result turns charge into a precise, training-free design dial, and locates its limits honestly: per-protein calibration earns its samples on heterogeneous design sets and merely matches a good global bias on homogeneous ones. Every number is from held-out data and real inference. By showing that distinct disease-area targets impose distinct electrostatic tolerance windows, ZetaDial makes charge a training-free handle for tailoring binders to each target's biophysics rather than to one universal rule.

References

1. V. K. Sharma *et al.*, In silico selection of therapeutic antibodies for development: Viscosity, clearance, and chemical stability, *Proceedings of the National Academy of Sciences* **111**, 18601 (2014).
2. T. Jain *et al.*, Biophysical properties of the clinical-stage antibody landscape, *Proceedings of the National Academy of Sciences* **114**, 944 (2017).
3. M. I. J. Raybould *et al.*, Five computational developability guidelines for therapeutic antibody profiling, *Proceedings of the National Academy of Sciences* **116**, 4025 (2019).
4. T. Igawa *et al.*, Reduced elimination of IgG antibodies by engineering the variable region, *Protein Engineering, Design and Selection* **23**, 385 (2010).
5. J. Dauparas, I. Anishchenko, N. Bennett, H. Bai, R. J. Ragotte, L. F. Milles, B. I. M. Wicky, A. Courbet, R. J. de Haas, N. Bethel *et al.*, Robust deep learning–based protein sequence design using ProteinMPNN, *Science* **378**, 49 (2022).
6. M. Pacesa, L. Nickel, C. Schellhaas, J. Schmidt, E. Pyatova, L. Kissling, P. Barendse, J. Choudhury, S. Kapoor, A. Alcaraz-Serna *et al.*, One-shot design of functional protein binders with BindCraft, *Nature* **646**, 483 (2025).
7. B. Raghavan and D. M. Rogers, ProtNHF: Neural hamiltonian flows for controllable protein sequence generation, *bioRxiv* (2026).
8. Z. Deng, R. Hou, N. Xie, M. Tyers and M. Koziarski, Discontinuous epitope fragments as sufficient target templates for efficient binder design, *arXiv preprint arXiv:2509.25479* (2025).
9. E. Hartman, J. Wallin, J. Malmström and J. Olsson, Controllable protein design with particle-based feynman–kac steering, *arXiv preprint arXiv:2511.09216* (2025).
10. J. L. Watson, D. Juergens, N. R. Bennett, B. L. Trippe, J. Yim, H. E. Eisenach, W. Ahern, A. J. Borst, R. J. Ragotte, L. F. Milles *et al.*, De novo design of protein structure and function with RFdiffusion, *Nature* **620**, 1089 (2023).
11. J. Jumper, R. Evans, A. Pritzel, T. Green, M. Figurnov, O. Ronneberger, K. Tunyasuvunakool, R. Bates, A. Žídek, A. Potapenko *et al.*, Highly accurate protein structure prediction with AlphaFold, *Nature* **596**, 583 (2021).
12. A. Schoch *et al.*, Charge-mediated influence of the antibody variable domain on FcRn-dependent pharmacokinetics, *Proceedings of the National Academy of Sciences* **112**, 5997 (2015).
13. A. Datta-Mannan *et al.*, Effect of variable domain charge on in vitro and in vivo disposition of monoclonal antibodies, *mAbs* **13**, p. 1993769 (2021).
14. R. L. Kelly *et al.*, High throughput cross-interaction measures for human IgG1 antibodies correlate with clearance rates in mice, *mAbs* **7**, 770 (2015).
15. L. Kalejaye *et al.*, DeepSP: Deep learning-based spatial properties to predict monoclonal antibody stability, *Computational and Structural Biotechnology Journal* **23**, 2220 (2024).
16. J. Xiong, I. Gaur, M. Lukarska, H. Nisonoff, L. M. Oltrogge, D. F. Savage and J. Listgarten, ProteinGuide: On-the-fly property guidance for protein sequence generative models, *arXiv preprint arXiv:2505.04823* (2025).
17. C. A. Goverde, M. Pacesa, N. Goldbach *et al.*, Computational design of soluble and functional membrane protein analogues, *Nature* **631**, 449 (2024).
18. X. Hou, J. Liu, C. Shi, X. Liu, Z. Yang and J. Tang, Property-driven protein inverse folding with multi-objective preference alignment, *arXiv preprint arXiv:2603.06748* (2026), ICLR 2026.
19. M. S. Lawrence, K. J. Phillips and D. R. Liu, Supercharging proteins can impart unusual resilience, *Journal of the American Chemical Society* **129**, 10110 (2007).
20. B. S. Der, M. Machius, M. J. Miley, J. L. Mills, T. Szyperski and B. Kuhlman, Alternative computational protocols for supercharging protein surfaces for reversible unfolding and retention of stability, *PLOS ONE* **8**, p. e64363 (2013).

21. P.-F. Chan, R. A. Curtis and J. Warwicker, Soluble expression of proteins correlates with a lack of positively-charged surface, *Scientific Reports* **3**, p. 3333 (2013).
22. R. M. Kramer, V. R. Shende, N. Motl, C. N. Pace and J. M. Scholtz, Toward a molecular understanding of protein solubility: increased negative surface charge correlates with increased solubility, *Biophysical Journal* **102**, 1907 (2012).
23. K. Maksymenko *et al.*, A complementarity-based approach to de novo binder design, *Advanced Science* (2025), Introduces the HECTOR shape-complementarity algorithm.
24. M. D. Overath, A. Rygaard, P. Sormanni, T. P. Jenkins *et al.*, Predicting experimental success in de novo binder design: A meta-analysis of 3,766 experimentally characterised binders, *bioRxiv* (2025).
25. G. Ahn, B. Coventry, E. Haefner, S. Sadre, J. Hu, M. Van, B. Huang, I. Sappington, A. J. Broerman, M. A. Lichtenstein, M. Glögl, I. Goreshnik, D. Vafeados and D. Baker, Computational design of ph-sensitive binders, *bioRxiv* (2025).
26. S. Ovchinnikov *et al.*, ColabDesign: Making protein design accessible to all via Google Colab `https://github.com/sokrypton/ColabDesign`, (2021).
27. M. Steinegger and J. Söding, MMseqs2 enables sensitive protein sequence searching for the analysis of massive data sets, *Nature Biotechnology* **35**, 1026 (2017).
28. Z. Lin, H. Akin, R. Rao, B. Hie, Z. Zhu, W. Lu, N. Smetanin, R. Verkuil, O. Kabeli, Y. Shmueli *et al.*, Evolutionary-scale prediction of atomic-level protein structure with a language model, *Science* **379**, 1123 (2023).
29. R. Evans, M. O'Neill, A. Pritzel, N. Antropova, A. Senior, T. Green, A. Žídek, R. Bates, S. Blackwell, J. Yim *et al.*, Protein complex prediction with AlphaFold-Multimer, *bioRxiv* (2021).