Seeing the Molecules Behind the Density

Anonymous authors
Paper under double-blind review

Abstract

Biomolecular structure predictors can generate plausible conformations, but these may not correspond to the state observed in a particular sample. Cryo-electron microscopy (cryo-EM) density provides spatial constraints from the target experiment; using them to guide both global structural relationships and local atomic placement remains a central challenge. We introduce EviBoltz, which uses global density interactions to update the single and pair representations of the Boltz-2 Pairformer, and atom-centered, multiscale density queries continually guide diffusion-based coordinate generation. Among evaluable cases in an independent benchmark, EviBoltz raises acceptable protein–protein interface success from 27.8% to 70.4% and joint ligand-pocket success from 7.0% to 34.9% relative to Boltz-2. The protein–peptide and antibody–antigen tasks show the same direction of change, with success rates rising from 0.0% to 75.0% and from 0.0% to 66.7%, respectively. In addition, long-loop analysis shows improved backbone agreement with the input density. In a controlled, supervised experiment on two states of the same URAT1 sequence, representations induced by different target maps guide a frozen generator toward the corresponding states. These results connect improved interaction recovery and local geometry with the ability of map-conditioned representations to carry information about the observed state.

1 Introduction

The three-dimensional structures of biomolecules underpin our understanding of molecular mechanisms, recognition, and functional regulation. Since AlphaFold 2 substantially advanced accurate protein structure prediction in 2021, subsequent models have extended prediction to more complex biomolecular compositions and interactions, providing powerful priors for experimental structure determination (Jumper et al., 2021; Baek et al., 2021; Lin et al., 2023; Krishna et al., 2024; Abramson et al., 2024; Chai Discovery et al., 2024; Wohlwend et al., 2024; Passaro et al., 2025). Yet generating a plausible structural candidate and explaining the structure present in a given experiment are distinct tasks. Even when the sequence, interaction partners, and ligands are known, a model's preferred conformation may differ from the state in the sample; identifying that state requires information from the target experiment itself (Abramson et al., 2024; Raghu et al., 2025; Xu et al., 2026). Antibody–antigen complexes with little similarity to known training examples illustrate the difficulty of recovering the correct binding pose from molecular inputs alone (Fromm et al., 2026).

Single-particle cryo-electron microscopy (cryo-EM) provides a direct route to combining structural priors with observations of the sample. Its reconstructed three-dimensional potential map (hereafter, density map) records spatial information about the particle population under analysis (Nakane et al., 2020). Building an all-atom model from such a map requires coordinating chain tracing, sequence assignment, complex assembly, and local refinement (Jamali et al., 2024; Chen et al., 2024). Automated methods have advanced this process substantially (Terwilliger et al., 2018; Pfab et al., 2021; Terashi et al., 2024; Jamali et al., 2024; Wang et al., 2025; Su et al., 2026). Complete modeling across molecular components remains challenging: some primarily produce protein structures and cannot simultaneously recover small-molecule ligand poses, while others cover only a subset of target residues (Jamali et al., 2024; Wang et al., 2025; Su et al., 2026).

Density-guided structure prediction faces difficulties at two scales. Globally, correctly folded individual domains do not establish their relative orientation, channel opening state, or subunit interface in the sample; even when multiple conformations are sampled, the experimentally observed one may be difficult to identify without constraints from the target map (Abramson et al., 2024; Raghu

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.

et al., 2025). Locally, density is not equally interpretable everywhere: some loops lack a continuous backbone signal, whereas other regions locate the backbone but do not resolve side-chain orientations (Kucukelbir et al., 2014; Pintilie et al., 2020). Local motion, partial occupancy, and conformational heterogeneity can all produce weak density and thereby limit the guidance available from the experiment (Beton et al., 2024; Lawson et al., 2021).

We ask how target density can shape both the global relationships and the local atomic coordinates a predictor generates. Existing approaches use density at different stages. Some retain much of the pretrained Pairformer prior and introduce density during inference or target-specific optimization; others construct models directly from maps or retrain structure generators to use density. Their molecular scope, structural coverage, and training requirements differ, as discussed in Section 2.

The Pairformer integrates sequence, evolutionary, and chemical information into single-token and pairwise representations that condition coordinate generation (Abramson et al., 2024; Passaro et al., 2025). Learned geometric relationships can also carry experimental constraints into structure generation (Liu et al., 2023; Yan et al., 2026; Stahl et al., 2023). A molecule can adopt different conformations; changing evolutionary input or adding experimental distances can also alter a model's preferred state or interface (Morcos et al., 2013; Wayment-Steele et al., 2024; Piomponi et al., 2025; Stahl et al., 2024). Pairformer representations therefore contain useful structural priors, but they have yet to respond to the target map. We seek to retain these priors while allowing density to update structural relationships.

We introduce EviBoltz, a density-conditioned cofolding model built on pretrained Boltz-2. Joint fine-tuning of the density encoder and Pairformer lets map tokens update both single representations s and pair representations z. During diffusion, the model also queries local density around moving atoms. These two routes bring experimental information into structural relationships and coordinate generation across PPI, PL, protein–peptide, and AbAg systems.

We compare EviBoltz with Boltz-2 and five existing methods on an independent test set of 120 cases spanning four interaction classes, and further examine local conformations and the transfer of state information. Relative to Boltz-2, EviBoltz improves global structural accuracy and prediction success while fully covering the selected target. In 113 cases with evaluable long loops, the mean RMSD of loop shape after independent alignment decreases by 1.653 Å, showing that the gains extend to local geometry. A supervised, single-system probe using two states of the same URAT1 sequence further indicates that the joint single/pair representations induced by different target maps can convey the difference between the observed states to a frozen generator. Together, these results link changes in global structure, local shape, and internal representations, illustrating how experimental density can help a structure model recover the conformation of the target sample.

2 Related Work

Sequence-driven structure prediction Multicomponent predictors from RoseTTAFold All-Atom and AlphaFold 3 to open models such as Chai-1, Protenix, and Boltz-2 provide strong priors for modeling molecular assemblies (Krishna et al., 2024; Abramson et al., 2024; Chai Discovery et al., 2024; ByteDance AML AI4Science Team et al., 2025; Passaro et al., 2025). These models do not take the target map as experimental evidence. Boltz-2's experimental-method field records how a structure was determined but carries no spatial observation from the sample. A plausible conformation generated from evolutionary information and learned structural priors may therefore differ from the observed apo or holo state. Target-specific evidence reflects the sample's state and can guide conformational selection (Abramson et al., 2024; Xu et al., 2026). This distinction matters even if every molecular component is specified correctly: a sequence and ligand list do not identify which of several accessible states was populated in the measured specimen. AlphaFold 3 reports disagreements with corresponding apo and holo observations, while cross-category evaluation identifies allosteric sites as a difficult setting (Abramson et al., 2024; Xu et al., 2026). Density conditioning addresses this source of ambiguity by supplying spatial evidence from the particular sample, rather than merely enlarging the set of structures that can be generated.

Geometric representations and denoising Three-dimensional pretraining and diffusion models show how learned geometric representations can guide molecular coordinates and ligand poses (Liu et al., 2022; Zhou et al., 2023; Zaidi et al., 2023; Corso et al., 2023). In EviBoltz, these two roles meet:

map-conditioned relationships guide generation, while local density is read again as the coordinates change.

Experimental restraints Experimental restraints provide a way to use target observations in structure prediction. This is especially relevant for antibody–antigen modeling because interchain coevolutionary signals are often unavailable. RosettaHDX uses exchange-derived epitope information for docking, and GRASP combines restraints from experiments such as deep mutational scanning and covalent labeling (Tran et al., 2025; Xie et al., 2025). For conformational selection, AlphaLink and AlphaLink2 incorporate cross-link-derived distances, while DEERFold uses measured distance distributions (Stahl et al., 2023; 2024; Wu et al., 2025). These approaches establish that experimental evidence can change a predicted interface or state. A density map offers a different form of evidence: where the map is resolved, it can constrain the placement of many atoms and their surrounding geometry rather than only specified residue pairs.

Density-driven modeling Direct tracing, refinement of an existing model, and density-guided generation also make different assumptions about their starting structures. Earlier map-building tools and subsequent systems such as DeepTracer, DeepMainmast, and DiffModeler span these alternatives (Frenz et al., 2017; Pfab et al., 2021; Terashi et al., 2024; Wang et al., 2024). ModelAngelo, E3-CryoFold, and CryoAtom construct structures from density and sequence with different input priors and molecular scopes (Jamali et al., 2024; Wang et al., 2025; Su et al., 2026). Their ability to produce a backbone does not imply recovery of every selected component of a multicomponent target, which motivates reporting both accuracy and output coverage.

Among approaches using a pretrained generator, CryoBoltz applies global and local density constraints during sampling, and EmbedOpt optimizes conditioning embeddings for individual targets (Raghu et al., 2025; Li et al., 2026a). CryoACE fuses density and sequence representations after the Pairformer, then repeatedly queries local density at atomic positions for structural feedback (Li et al., 2026b). CoCoFold adapts AlphaFold with cryo-EM particle observations, CryoSampler uses a target map during test-time supervision of Boltz-2, and experimentally guided AlphaFold 3 incorporates observations during sampling (Liao et al., 2026; Shenoy et al., 2026; Maddipatla et al., 2026). These studies differ in where density or other evidence enters the model, whether adaptation is target-specific, and which molecular outputs are supported. EviBoltz instead jointly fine-tunes its density encoder, Pairformer, and denoiser so that the target map can update relational representations and be queried around moving atoms. Appendix A surveys the broader literature; the next section specifies these two conditioning pathways.

3 The EviBoltz Method

EviBoltz builds two density-conditioning pathways, one global and one local, on a pretrained Boltz-2 backbone, incorporating experimental observations into both structural relationship modeling and all-atom coordinate generation. As shown in Fig. 1, a density encoder converts the experimental map into global map tokens and multiscale spatial features. The global pathway uses map tokens to update the Pairformer's single and pair representations layer by layer, while also providing the diffusion network with global density context. The local pathway samples a 13-point neighborhood around each current noisy atomic position and combines these samples with the noise level to form local conditions. These conditions enter denoising through Density Modulation and atom-level feature residuals. As coordinates change, the local queries move with them, continually combining global relationships with density around each atom.

3.1 Density encoding and global structural relationships

We use Boltz-2 as the all-atom generation backbone because it handles proteins, peptides, small molecules, and their complexes (Passaro et al., 2025). Sequence and molecular features first yield single and pair representations (Abramson et al., 2024; Passaro et al., 2025). A three-dimensional density encoder produces fine-, medium-, and coarse-scale spatial fields; pooled coarse features form global map tokens, while the fields remain available for local queries. Before each Pairformer block, cross-attention to the map tokens produces context h_i for token i. It then updates single tokens and

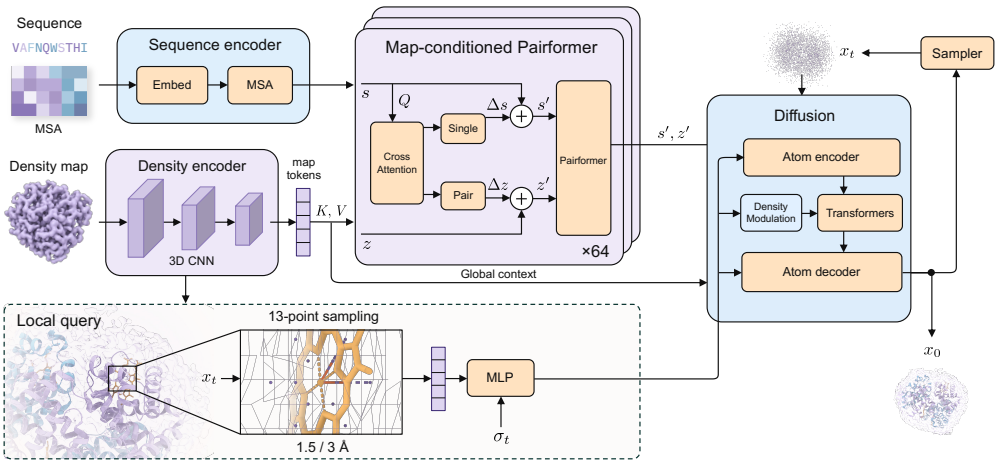

Figure 1: Overview of the EviBoltz architecture. EviBoltz uses density-map guidance during both training and inference. Map tokens guide updates to Pairformer representations. The associated global context and additional local queries then guide internal diffusion latents. See section 3 for details.

their pairwise relationships:

$$\begin{aligned} h_i &= \text{CrossAttn}(\text{LN}(s_i), T_\rho, T_\rho), \\ \widetilde{s}_i &= s_i + W_s h_i, \\ \widetilde{z}_{ij} &= z_{ij} + W_z\,\text{SiLU}(v_i + v_j + v_i \odot v_j), \quad v_i = W_e h_i. \end{aligned} \tag{1}$$

Repeating this interaction through the Pairformer yields density-conditioned relationships (s_ρ, z_ρ). A separate token-specific global context C_ρ queries the map tokens and directly conditions diffusion. The layer-wise formulation appears in Appendix B.

3.2 Atom-centered density guidance

Global map tokens summarize the density but do not retain every local detail. At each denoising step and noise level σ_t, EviBoltz maps every current noisy atomic position into the physical coordinate system of the density and queries the multiscale feature fields there. Each query reads a fixed 13-point neighborhood: the atomic position and points 1.5 Å and 3.0 Å away in both directions along each coordinate axis. An MLP combines these sampled features with a descriptor at the center, a noise-level embedding, and the parent token context to form an atom-level condition $q_{t,a}$. Because the queries are centered on the current coordinates, they move as the sampler updates the structure.

The atom encoder incorporates density through feature residuals. Pooling atom conditions by parent token then provides a residual and the condition $c_{t,i}$ for Density Modulation, to gate and offset Transformer features:

$$\widetilde{u}_i = u_i + W_o\,\text{SiLU}(\text{sigmoid}(\gamma_{t,i}) \odot \text{LN}(u_i) + \beta_{t,i})\,, \quad (\gamma_{t,i}, \beta_{t,i}) = f_{\text{mod}}(c_{t,i}). \tag{2}$$

Pairformer single representations and pair attention biases continue to encode relationships between tokens. The atom decoder combines updated token features with retained atomic features and local density residuals to predict coordinate updates. Appendix B gives the full sampling and layer-wise modulation rules.

3.3 Training and iterative generation

We jointly fine-tune the density encoder, both conditioning pathways, and the Boltz-2 Pairformer and diffusion backbone using paired maps and reference structures. Training adds Gaussian noise to reference coordinates X_0 and asks the denoiser to recover them from $X_t = X_0 + \sigma_t \epsilon$. The objective combines an aligned all-atom loss, a local-distance loss, and a density-patch loss:

$$\mathcal{L} = \mathbb{E}_{t,\epsilon}[w(\sigma_t)\ell_{\text{atom}} + \ell_{\text{sLDDT}} + \lambda_\rho \ell_{\text{patch}}]\,. \tag{3}$$

The moving 13-point queries construct the denoiser input, whereas the patch loss supplies separate training supervision. Reference coordinates generate noisy examples and losses; they are never an independent structural condition to the denoiser.

At inference, the model starts from Gaussian coordinate noise without access to a target reference. Each step queries density around the current atoms, predicts a noise-free structure under the global conditions, and updates coordinates on the sampler's noise schedule. The new positions anchor the next local queries; Appendix B gives the complete objective, denoiser, and sampling equations.

4 Experiments

Our evaluation addresses three questions: whether density conditioning improves recovery of complete structures and interactions; whether these gains extend to local conformations; and whether internal representations can transmit information about the experimentally observed state.

4.1 Data, baselines, and evaluation protocol

Benchmark and training data Our independent cryo-EM benchmark contains 120 PPI, PL, protein–peptide, and AbAg cases, testing subunit arrangement, ligand placement, flexible-peptide binding, and antibody interfaces. Together, these tasks span global structural relationships and local geometry. Source selection for the benchmark and training data, map resolution and input-size distributions, and the target-chain and ligand specifications are detailed in Appendix C and Fig. 7. The cohort contains 57 PPI, 43 PL, 13 protein–peptide, and 7 AbAg cases. We selected structures released after June 1, 2023, required maps with nominal resolution at or better than 4.0 Å, and kept selected protein-chain lengths between 100 and 2,000 residues. Our training and test sets share no PDB or EMDB identifier. These rules limit direct source overlap and set a common input scale, although they cannot rule out related structures in other training data or guarantee uniformly resolved local density. The same deposited source, map, target chains, and relevant ligand instances were specified for every method, so differences in reported performance do not arise from choosing different target components. The paired training set contains 10,714 structure–map sources after benchmark and preparation exclusions; further construction details and the resolution and input-size distributions appear in Appendix C and Fig. 7.

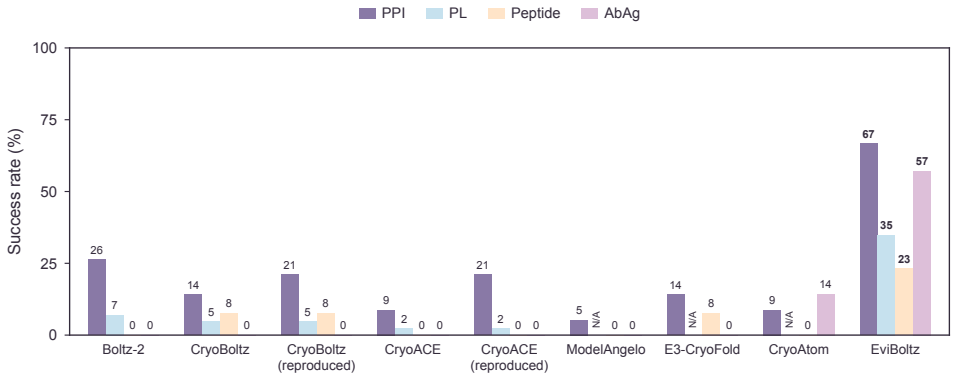

Figure 2: Prediction success across four tasks. Colored bars give task successes divided by the full task cohort. Interface tasks use DockQ > 0.23; PL uses pocket-aligned RMSD < 2 Å and LDDT-PLI > 0.8. N/A marks methods that do not produce ligand coordinates.

Baselines and metrics We compare EviBoltz with Boltz-2, CryoBoltz, CryoACE, ModelAngelo, E3-CryoFold, and CryoAtom (Passaro et al., 2025; Raghu et al., 2025; Li et al., 2026b; Jamali et al., 2024; Wang et al., 2025; Su et al., 2026). The original CryoBoltz and CryoACE runs completed 80 and 61 cases, respectively; the remaining cases include out-of-memory and input-processing failures as well as pending runs. Engineering-adapted reproduced versions each completed 118 cases. ModelAngelo, E3-CryoFold, and CryoAtom do not output ligand coordinates. We use DockQ for protein interfaces, pocket-aligned ligand RMSD and LDDT-PLI for ligand cases, and TM-score for global structure (Basu & Wallner, 2016; Mirabello & Wallner, 2024; Ma et al., 2025; Zhang & Skolnick, 2004). The table also reports protein and target-heavy-atom coverage; DockQ_0 counts output-coverage failures as zero within a fixed evaluable set. Task thresholds, alignment, denominators, and uncertainty estimates appear in Appendix C.

Table 1: Benchmark results on four interaction tasks. Task subscripts give total cases. PL ligand RMSD uses pocket alignment. * and metric subscripts identify restricted evaluation sets; – means not applicable or unavailable. **Bold** and underline identify the best and second-best eligible results within each task. Methods missing predictions for more than three task cases are excluded from ranking.

Task type	Method	Success rate		Structure quality			Coverage	
		DockQ > 0.23	DockQ > 0.49	DockQ ↑	DockQ$_0$ ↑	TM-score ↑	C_α ↑	Heavy atom ↑
$Peptide_{13}$	Boltz2	0.0%	0.0%	0.0224^*	0.0224	0.3762	100.0%	100.0%
	CryoBoltz	$50.0\%_2$	$0.0\%_2$	0.1668^*_2	0.1668^*_2	0.2317^*_6	$100.0\%^*_6$	$100.0\%^*_6$
	∟ *Reproduced*	25.0%	0.0%	0.0872^*	0.0872	0.3207	100.0%	100.0%
	CryoACE	0.0%	0.0%	$\underline{0.0981}^*$	0.0981	0.3015^*_{11}	$100.0\%^*_{11}$	$100.0\%^*_{11}$
	∟ *Reproduced*	0.0%	0.0%	0.0981^*	0.0981	0.3160	100.0%	100.0%
	ModelAngelo	0.0%	0.0%	–	0.0000	$\underline{0.8733}$	87.7%	86.9%
	E3-CryoFold (updated)	$\underline{25.0\%}$	**25.0%**	0.7208^*_1	$\underline{0.1802}$	**0.9178**	96.8%	97.2%
	CryoAtom	0.0%	0.0%	–	0.0000	0.8572	86.2%	89.0%
	EviBoltz	**75.0%**	**25.0%**	$\mathbf{0.3216}^*$	**0.3216**	0.8618	100.0%	100.0%
$AbAg_7$	Boltz2	0.0%	0.0%	$\underline{0.0282}^*$	0.0282	0.3872	100.0%	100.0%
	CryoBoltz	0.0%	0.0%	0.0071^*	0.0071	0.5029	100.0%	100.0%
	∟ *Reproduced*	0.0%	0.0%	0.0071^*	0.0071	0.5029	100.0%	100.0%
	CryoACE	$0.0\%_4$	$0.0\%_4$	0.0081^*_4	0.0081^*_4	0.4983^*_5	$100.0\%^*_5$	$100.0\%^*_5$
	∟ *Reproduced*	0.0%	0.0%	0.0076^*	0.0076	0.5274	100.0%	100.0%
	ModelAngelo	0.0%	0.0%	–	0.0000	0.7791	78.3%	78.5%
	E3-CryoFold (updated)	0.0%	0.0%	–	0.0000	0.5567	84.3%	85.3%
	CryoAtom	$\underline{16.7\%}$	$\underline{16.7\%}$	0.9150^*_1	$\underline{0.1525}$	$\underline{0.7949}$	79.9%	80.7%
	EviBoltz	**66.7%**	**33.3%**	$\mathbf{0.3425}^*$	**0.3425**	**0.8473**	100.0%	100.0%
PPI_{57}	Boltz2	$\underline{27.8\%}$	$\underline{22.2\%}$	$\underline{0.1954}^*$	$\underline{0.1954}$	0.4929	100.0%	100.0%
	CryoBoltz	$24.2\%_{33}$	$9.1\%_{33}$	0.1663^*_{33}	0.1663^*_{33}	0.4808^*_{34}	$100.0\%^*_{34}$	$100.0\%^*_{34}$
	∟ *Reproduced*	$23.1\%_{52}$	$9.6\%_{52}$	0.1602^*_{52}	0.1602^*_{52}	0.4643^*_{55}	$100.0\%^*_{55}$	$100.0\%^*_{55}$
	CryoACE	$27.8\%_{18}$	$22.2\%_{18}$	0.2273^*_{18}	0.2273^*_{18}	0.4995^*_{18}	$100.0\%^*_{18}$	$100.0\%^*_{18}$
	∟ *Reproduced*	$23.1\%_{52}$	$15.4\%_{52}$	0.1778^*_{52}	0.1778^*_{52}	0.4764^*_{55}	$100.0\%^*_{55}$	$100.0\%^*_{55}$
	ModelAngelo	5.6%	5.6%	0.9646^*_3	0.0536	$\underline{0.7948}$	80.5%	82.2%
	E3-CryoFold (updated)	14.8%	14.8%	0.6942^*_9	0.1157	0.6676	88.5%	90.6%
	CryoAtom	9.3%	9.3%	0.9469^*_9	0.0877	**0.8217**	83.7%	86.0%
	EviBoltz	**70.4%**	**57.4%**	$\mathbf{0.5160}^*$	**0.5160**	0.7852	100.0%	100.0%
		Success Rate		Ligand RMSD ↓	LDDT-PLI ↑	TM-score ↑	C_α ↑	Heavy atom ↑
PL_{43}	Boltz2	$\underline{7.0\%}$		$\underline{10.6177}$	$\underline{0.3794}$	0.6316	100.0%	100.0%
	CryoBoltz	$6.1\%_{33}$		11.0734^*_{33}	0.2447^*_{33}	0.4901^*_{33}	$100.0\%^*_{33}$	$100.0\%^*_{33}$
	∟ *Reproduced*	4.7%		12.1666	0.2686	0.5320	100.0%	100.0%
	CryoACE	$3.7\%_{27}$		18.6143^*_{27}	0.1366^*_{27}	0.5045^*_{27}	$100.0\%^*_{27}$	$100.0\%^*_{27}$
	∟ *Reproduced*	2.3%		15.6024	0.2207	0.5292	100.0%	100.0%
	ModelAngelo	0.0%		–	–	0.8038	80.8%	82.6%
	E3-CryoFold (updated)	0.0%		–	–	$\underline{0.8089}$	94.7%	93.8%
	CryoAtom	0.0%		–	–	0.7780	78.4%	80.0%
	EviBoltz	**34.9%**		**4.0532**	**0.6127**	**0.8989**	100.0%	100.0%

4.2 Results

Figure 2 compares recovery across tasks; we then examine local conformations and probe map use through density-input interventions and representation experiments.

Table 1 summarizes all four tasks. Relative to Boltz-2, EviBoltz improves global structure and partner placement: acceptable PPI interfaces become more than twice as common, and joint ligand-pocket success rises markedly. The reproduced CryoBoltz and CryoACE runs extend coverage of their original evaluations, but their task-specific results remain below those of EviBoltz.

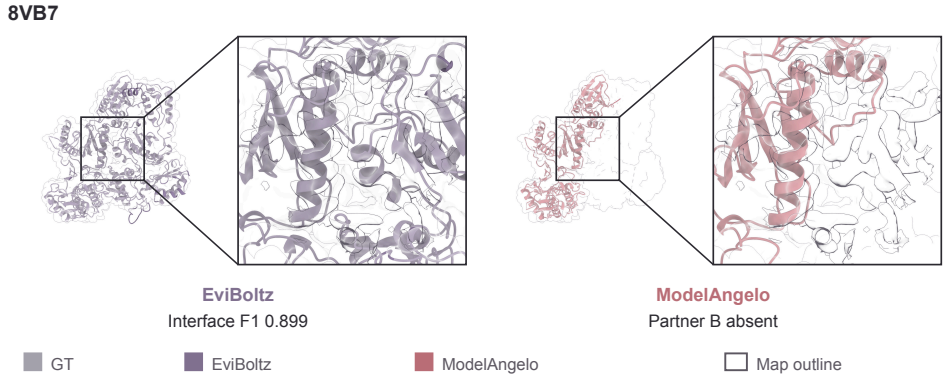

Figure 3: Incomplete partner recovery in 8VB7. The experimental reference (gray) is overlaid with EviBoltz (purple) or ModelAngelo (red) at matched scale.

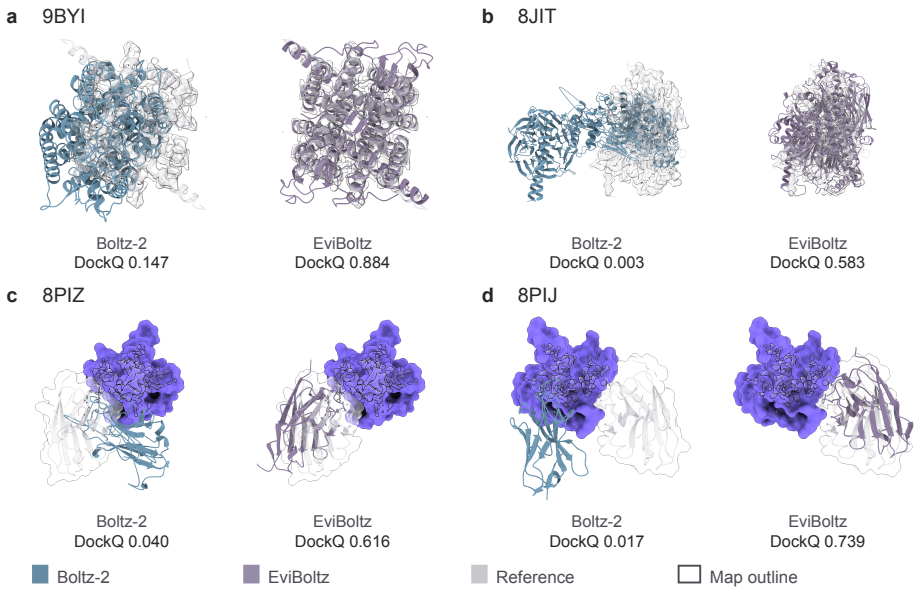

Figure 4: Representative interface comparisons. a,b, interface recovery in PPI complex 9BYI and protein–peptide complex 8JIT; **c,d**, antibody placement relative to antigen in 8PIZ and 8PIJ. Each pair shows Boltz-2 on the left and EviBoltz on the right, overlaid with the reference.

For the automated builders, TM-score must be read alongside coverage and interface quality: a model can score well on the residues it produces while omitting target residues or misplacing a partner. For instance, in 8VB7, ModelAngelo builds part of the complex but leaves the designated interaction partner B unmodeled (Fig. 3); this example illustrates why partner coverage must be reported alongside interface accuracy.

On all 120 sources, the paired EviBoltz-minus-Boltz-2 mean gains are 0.245 in LDDT (95% source-bootstrap CI, 0.219–0.271) and 0.314 in TM-score (0.265–0.361). The paired comparison uses the same source for both models and weights each source equally; it therefore complements the task-level success rates, whose denominators depend on the output and metric being evaluated. The intervals describe variation across benchmark sources, not variation across separately trained checkpoints. Appendix D reports the corresponding paired summaries for the other methods.

Interface recovery On interfaces calculable for both models, DockQ rises from 0.17 to 0.49. In the PPI task, 38 of 54 evaluable interfaces pass the acceptable-quality threshold, compared with 15 for Boltz-2. Protein–peptide global structure improves as well, with task counts and threshold results reported in Appendix D.

The representative cases in Fig. 4**a–d** show how improved partner placement appears across interaction types. EviBoltz corrects the subunit arrangement in PPI complex 9BYI, recovers the fold and interface in protein–peptide complex 8JIT, and places antibodies relative to antigens in AbAg complexes 8PIZ and 8PIJ.

Protein–ligand pocket placement and interactions For PL systems, ligand placement must be accurate both within the pocket and at its local contacts. Mean pocket-aligned RMSD falls from 10.6 Å to 4.1 Å, while mean LDDT-PLI rises from 0.38 to 0.61; together, these changes raise joint success from 3 to 15 of 43 cases. The NADPH complex 8XCO illustrates both aspects: the predicted ligand moves into the reference pocket and makes better local contacts (Fig. 5**e**).

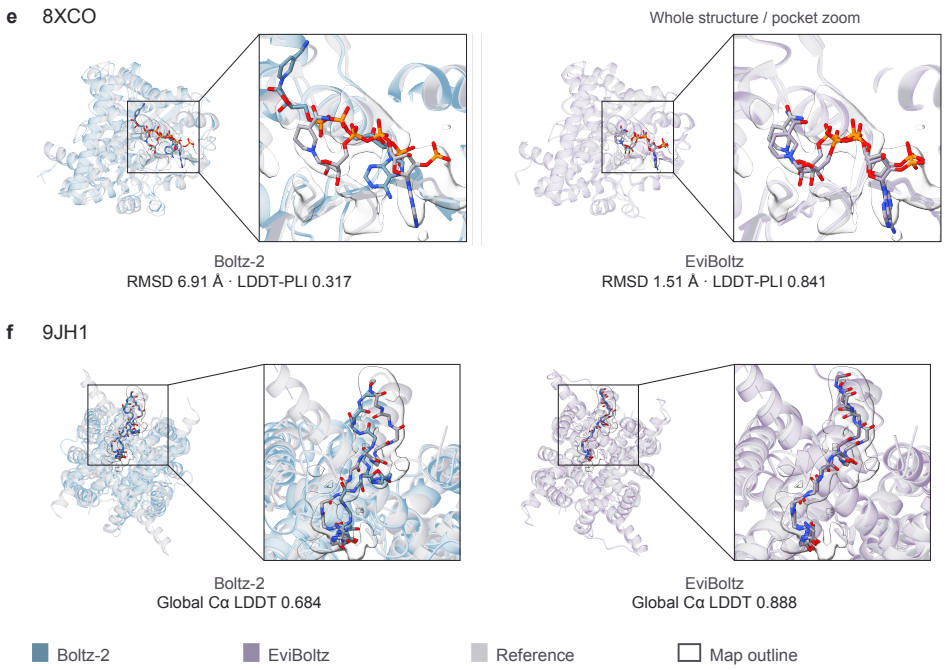

Figure 5: Representative ligand and loop comparisons. **e**, global structure and NADPH pocket in 8XCO; **f**, the 9JH1 long loop and its local enlargement. Boltz-2 is on the left and EviBoltz on the right in each case; light gray shows the reference structure.

Local conformations To test whether gains extend beyond global placement, we analyzed long reference loops annotated using P-SEA (Labesse et al., 1997). We measured three aspects separately: loop position after alignment on the structural core; loop shape after independent alignment of each segment; and backbone support from density for the core-aligned model. Across 113 evaluable cases, mean local-shape RMSD improves by 1.65 Å (95% CI, 1.38–1.92 Å), and mean density support increases by 0.29 (95% CI, 0.24–0.35). Because the independent loop alignment removes rigid-body placement, the shape gain supports a local conformational improvement. Figure 5**f** illustrates one loop; full group results appear in Appendix F.

5 Ablation Studies

5.1 Pairformer fine-tuning and density-map input

We first compare the original Boltz-2 model with EviBoltz run without density input, then compare EviBoltz with and without the target map at the same checkpoint. These comparisons distinguish the effect associated with joint fine-tuning from the additional effect of supplying density at inference. We also probe the two routes by which the encoded map guides prediction, removing either the Pairformer density path or diffusion-block guidance while retaining the other route.

Without a map at inference, EviBoltz improves on Boltz-2 in TM-score (0.635 versus 0.530) and DockQ (0.243 versus 0.172). Supplying the target map to the same checkpoint raises these values further to 0.838 and 0.488, respectively, while reducing ligand RMSD from 9.01 Å to 4.05 Å (Table 2). The matched-checkpoint comparison shows a substantial additional gain from inference-time density.

Table 2: Density input and guidance-pathway ablations. The Pairformer-path ablation retains diffusion guidance; the diffusion-guidance ablation retains the Pairformer density path.

Model	Map	TM-score	DockQ	RMSD (Å)
EviBoltz	Yes	0.838	0.488	4.053
∟ *w/o Pairformer density path*	Yes	0.642	0.296	6.132
∟ *w/o diffusion-block guidance*	Yes	0.750	0.379	7.715
EviBoltz	No	0.635	0.243	9.013
Boltz-2	No	0.530	0.172	11.721

The pathway ablations show distinct metric patterns. Removing the Pairformer density

path yields a TM-score of 0.642 and DockQ of 0.296; retaining that path while removing diffusion-block guidance yields 0.750 and 0.379, respectively. In contrast, the latter condition has higher ligand RMSD (7.715 Å versus 6.132 Å). This pattern is consistent with the design in which the two paths influence structural relationships and local coordinate placement, while the full model outperforms either ablation on all three metrics.

5.2 Map-guided conformational-state selection

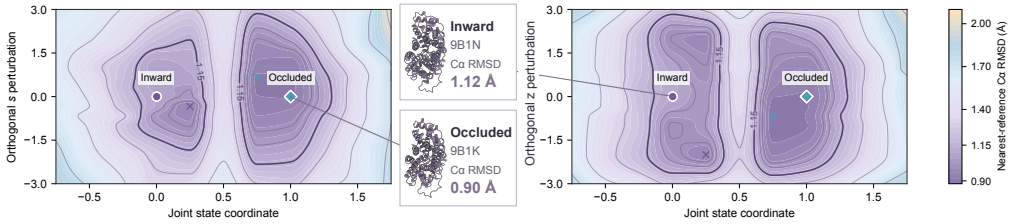

Figure 6: URAT1 joint-representation intervention and endpoint structures. The horizontal coordinate x jointly interpolates single and pair representations: $x = 0$ corresponds to the inward-map state and $x = 1$ to the occluded-map state. The vertical axis applies an equal-norm orthogonal perturbation in either space. Insets show endpoint structures with C_α RMSD to their matching references.

Following the pathway ablations, we ask whether the density encoder can influence conformational selection through Pairformer representations. The same URAT1 sequence has experimentally observed inward-facing (9B1N) and occluded (9B1K) states. We place the two maps in a shared coordinate frame and grid and hold the sequence, MSA, and all other non-density inputs fixed. The two maps yield (s_I, z_I) and (s_O, z_O), respectively. During generation, we disable direct map input to the diffusion module to test whether the joint representations can guide the frozen generator toward the matching conformation.

Yet structures generated from each map-conditioned representation remain closer to the matching experimental state without direct map input during generation. The inward-map representation yields C_α RMSD of 1.10 Å to the inward reference versus 2.55 Å to the occluded reference; the occluded-map representation yields 0.91 Å to the occluded reference versus 2.35 Å to the inward reference. Figure 6 shows joint-representation interpolation and orthogonal perturbations; sampling details appear in Appendix E.

The measured endpoints lie closer to their corresponding reference states: the two maps yield different joint representations that guide the frozen generator toward the observed conformations. The response surfaces show how structural error changes along these representation directions. The horizontal intervention changes s and z together, so the response surface does not isolate the contribution of either representation. Its color reports error to the nearer of two measured references, not a conformational probability or a free-energy landscape. The endpoint comparison is the direct test of state transfer in this supervised URAT1 system; whether the same behavior generalizes to other proteins or map pairs remains to be evaluated.

6 Conclusion and Future Work

EviBoltz incorporates experimental density into structural relationships and queries the map around moving atoms during denoising, combining molecular priors with observations of the target sample. Across 120 independent test cases, it improves global conformations, designated-partner interfaces, and ligand recovery while covering the selected reference target atoms. Gains in local loop shape and the two-state URAT1 experiment show how density can influence local geometry and conformational selection. Target density thus conditions both relationship modeling and coordinate generation.

Future work will address two directions. Noise, artifacts, and variable map quality complicate training the density encoder; augmentation suited to the current pipeline could expand usable training data and improve tolerance to input quality. We also plan to assess performance as dataset size and training scale increase.

AI Use Statement

We used AI tools to assist with polishing the manuscript's English text, preparing figures and LaTeX, setting up computational environments, and adapting, debugging, and running baseline reproduction workflows, including the engineering-adapted CryoACE and CryoBoltz reproductions. The reported predictions and metrics come from the stated computational and evaluation workflows. The authors reviewed the AI-assisted text, code, and outputs and take responsibility for the final manuscript and results.

References

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

Pavel V. Afonine, Billy K. Poon, Randy J. Read, Oleg V. Sobolev, Thomas C. Terwilliger, Alexandre Urzhumtsev, and Paul D. Adams. Real-space refinement in PHENIX for cryo-EM and crystallography. *Acta Crystallographica Section D: Structural Biology*, 74:531–544, 2018. doi: 10.1107/S2059798318006551.

Minkyung Baek, Frank DiMaio, Ivan Anishchenko, et al. Accurate prediction of protein structures and interactions using a three-track neural network. *Science*, 373:871–876, 2021. doi: 10.1126/science.abj8754.

Sankar Basu and Björn Wallner. DockQ: A quality measure for protein-protein docking models. *PLOS ONE*, 11(8):e0161879, 2016. doi: 10.1371/journal.pone.0161879.

Helen M. Berman, John Westbrook, Zukang Feng, Gary Gilliland, T. N. Bhat, Helge Weissig, Ilya N. Shindyalov, and Philip E. Bourne. The protein data bank. *Nucleic Acids Research*, 28(1):235–242, 2000. doi: 10.1093/nar/28.1.235.

Joseph G. Beton, Thomas Mulvaney, Tristan Cragnolini, et al. Cryo-EM structure and B-factor refinement with ensemble representation. *Nature Communications*, 15:444, 2024. doi: 10.1038/s41467-023-44593-1.

ByteDance AML AI4Science Team, Xinshi Chen, Yuxuan Zhang, Chan Lu, Wenzhi Ma, Jiaqi Guan, Chengyue Gong, Jincai Yang, Hanyu Zhang, Ke Zhang, Shenghao Wu, Kuangqi Zhou, Yanping Yang, Zhenyu Liu, Lan Wang, Bo Shi, Shaochen Shi, and Wenzhi Xiao. Protenix - advancing structure prediction through a comprehensive AlphaFold3 reproduction. *bioRxiv*, 2025. doi: 10.1101/2025.01.08.631967. Preprint.

Chai Discovery, Jacques Boitreaud, Jack Dent, Matthew McPartlon, Joshua Meier, Vinicius Reis, Alex Rogozhnikov, and Kevin Wu. Chai-1: Decoding the molecular interactions of life. *bioRxiv*, 2024. doi: 10.1101/2024.10.10.615955. Preprint.

Sheng Chen, Sen Zhang, Xiaoyu Fang, et al. Protein complex structure modeling by cross-modal alignment between cryo-EM maps and protein sequences. *Nature Communications*, 15:8808, 2024. doi: 10.1038/s41467-024-53116-5.

Gabriele Corso, Hannes Stärk, Bowen Jing, Regina Barzilay, and Tommi S. Jaakkola. DiffDock: Diffusion steps, twists, and turns for molecular docking. In *International Conference on Learning Representations*, 2023. URL https://mlanthology.org/iclr/2023/corso2023iclr-diffdock/.

Bradley Efron. Bootstrap methods: Another look at the jackknife. *The Annals of Statistics*, 7(1):1–26, 1979. doi: 10.1214/aos/1176344552.

Brandon Frenz, Alexandra C. Walls, Edward H. Egelman, David Veesler, and Frank DiMaio. RosettaES: a sampling strategy enabling automated interpretation of difficult cryo-EM maps. *Nature Methods*, 14:797–800, 2017. doi: 10.1038/nmeth.4340.

Samuel Fromm, Marko Ludaic, and Arne Elofsson. Evaluating deep learning based structure prediction methods on antibody-antigen complexes. *Bioinformatics*, 42(4):btag136, 2026. doi: 10.1093/bioinformatics/btag136.

Nabin Giri and Jianlin Cheng. De novo atomic protein structure modeling for cryoem density maps using 3d transformer and HMM. *Nature Communications*, 15:5511, 2024. doi: 10.1038/s41467-024-49647-6.

Rajan Gyawali, Ashwin Dhakal, and Jianlin Cheng. Multimodal deep learning integration of cryo-EM and AlphaFold3 for high-accuracy protein structure determination. *Communications Chemistry*, 8:320, 2025. doi: 10.1038/s42004-025-01718-5.

Jiahua He and Sheng-You Huang. Full-length de novo protein structure determination from cryo-EM maps using deep learning. *Bioinformatics*, 37(20):3480–3490, 2021. doi: 10.1093/bioinformatics/btab357.

Jiahua He, Peicong Lin, Ji Chen, Hong Cao, and Sheng-You Huang. Model building of protein complexes from intermediate-resolution cryo-EM maps with deep learning-guided automatic assembly. *Nature Communications*, 13:4066, 2022. doi: 10.1038/s41467-022-31748-9.

Fuyao Huang, Xiaozhu Yu, Kui Xu, and Qiangfeng Cliff Zhang. CryoNet.Refine: A one-step diffusion model for rapid refinement of structural models with cryo-EM density map restraints. In *International Conference on Learning Representations*, 2026. URL https://openreview.net/forum?id=NwzY2yhlme.

Kiarash Jamali, Lukas Käll, Rui Zhang, et al. Automated model building and protein identification in cryo-EM maps. *Nature*, 628:450–457, 2024. doi: 10.1038/s41586-024-07215-4.

John Jumper, Richard Evans, Alexander Pritzel, et al. Highly accurate protein structure prediction with AlphaFold. *Nature*, 596:583–589, 2021. doi: 10.1038/s41586-021-03819-2.

Rohith Krishna, Jue Wang, Woody Ahern, et al. Generalized biomolecular modeling and design with RoseTTAFold All-Atom. *Science*, 384:eadl2528, 2024. doi: 10.1126/science.adl2528.

Alp Kucukelbir, Fred J. Sigworth, and Hemant D. Tagare. Quantifying the local resolution of cryo-EM density maps. *Nature Methods*, 11(1):63–65, 2014. doi: 10.1038/nmeth.2727.

G. Labesse, N. Colloc'h, J. Pothier, and J.-P. Mornon. P-SEA: a new efficient assignment of secondary structure from Cα trace of proteins. *Bioinformatics*, 13(3):291–295, 1997. doi: 10.1093/bioinformatics/13.3.291.

Catherine L. Lawson, Andriy Kryshtafovych, Paul D. Adams, et al. Cryo-EM model validation recommendations based on outcomes of the 2019 EMDataResource challenge. *Nature Methods*, 18:156–164, 2021. doi: 10.1038/s41592-020-01051-w.

Minhuan Li, Jiequn Han, Pilar Cossio, and Luhuan Wu. Robust inference-time steering of protein diffusion models via embedding optimization, 2026a. URL https://arxiv.org/abs/2602.05285.

Minzhang Li, Mingrui Li, Weichen Qin, Qihe Chen, Sixian Shen, Yuan Pei, Jiakai Zhang, and Jingyi Yu. CryoACE: An atom-centric framework for accurate and automated model building in cryo-EM, 2026b. URL https://arxiv.org/abs/2606.31332.

Junwen Liao, Dihan Zheng, Hui Zhang, Linfeng Zhang, Mingxu Hu, and Chenglong Bao. Fine-tuning AlphaFold with limited cryo-EM observations. *Communications Chemistry*, 9:95, 2026. doi: 10.1038/s42004-026-01899-7.

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

Shengchao Liu, Hanchen Wang, Weiyang Liu, Joan Lasenby, Hongyu Guo, and Jian Tang. Pre-training molecular graph representation with 3d geometry. In *International Conference on Learning Representations*, 2022. URL https://openreview.net/forum?id=xQUe1pOKPam.

Shengchao Liu, Weitao Du, Yanjing Li, Zhuoxinran Li, Zhiling Zheng, Chenru Duan, Zhi-Ming Ma, Omar Yaghi, Animashree Anandkumar, Christian Borgs, Jennifer Chayes, Hongyu Guo, and Jian Tang. Symmetry-informed geometric representation for molecules, proteins, and crystalline materials. In *Advances in Neural Information Processing Systems 36: Datasets and Benchmarks Track*, pp. 66084–66101, 2023. doi: 10.52202/075280-2886. URL https://proceedings.neurips.cc

/paper_files/paper/2023/hash/d07379f3acf3af51dfc8598862cadfa0-Abstract-Datasets_and_Benchmarks.html.

Wenzhi Ma, Zhenyu Liu, Jincai Yang, Chan Lu, Hanyu Zhang, and Wenzhi Xiao. From dataset curation to unified evaluation: Revisiting structure prediction benchmarks with PXMeter. *bioRxiv*, 2025. doi: 10.1101/2025.07.17.664878. Preprint.

Advaith Maddipatla, Nadav Sellam Bojan, Meital Bojan, Volodymyr Masalitin, Sanketh Vedula, Paul Schanda, Ailie Marx, and Alex M. Bronstein. Experiment-guided AlphaFold3 resolves measurement-consistent protein ensembles. *Nature Biotechnology*, 2026. doi: 10.1038/s41587-026-03166-5. Published online 29 June 2026.

Valerio Mariani, Marco Biasini, Alessandro Barbato, and Torsten Schwede. lDDT: a local superposition-free score for comparing protein structures and models using distance difference tests. *Bioinformatics*, 29(21):2722–2728, 2013. doi: 10.1093/bioinformatics/btt473.

Claudio Mirabello and Björn Wallner. DockQ v2: improved automatic quality measure for protein multimers, nucleic acids, and small molecules. *Bioinformatics*, 40(10):btae586, 2024. doi: 10.1093/bioinformatics/btae586.

Faruck Morcos, Biman Jana, Terence Hwa, and José N. Onuchic. Coevolutionary signals across protein lineages help capture multiple protein conformations. *Proceedings of the National Academy of Sciences*, 110:20533–20538, 2013. doi: 10.1073/pnas.1315625110.

Takanori Nakane, Abhay Kotecha, Andrija Sente, et al. Single-particle cryo-EM at atomic resolution. *Nature*, 587:152–156, 2020. doi: 10.1038/s41586-020-2829-0.

Saro Passaro, Gabriele Corso, Jeremy Wohlwend, et al. Boltz-2: Towards accurate and efficient binding affinity prediction. *bioRxiv*, 2025. doi: 10.1101/2025.06.14.659707. Preprint.

Eric F. Pettersen, Thomas D. Goddard, Conrad C. Huang, Elaine C. Meng, Gregory S. Couch, Tristan I. Croll, John H. Morris, and Thomas E. Ferrin. UCSF ChimeraX: Structure visualization for researchers, educators, and developers. *Protein Science*, 30(1):70–82, 2021. doi: 10.1002/pro.3943.

Jonas Pfab, Nhut Minh Phan, and Dong Si. DeepTracer for fast de novo cryo-EM protein structure modeling and special studies on CoV-related complexes. *Proceedings of the National Academy of Sciences*, 118(2):e2017525118, 2021. doi: 10.1073/pnas.2017525118.

Grigore Pintilie, Kaiming Zhang, Zhaoming Su, Shanshan Li, Michael F. Schmid, and Wah Chiu. Measurement of atom resolvability in cryo-EM maps with Q-scores. *Nature Methods*, 17:328–334, 2020. doi: 10.1038/s41592-020-0731-1.

Valerio Piomponi, Alberto Cazzaniga, and Francesca Cuturello. Evolutionary constraints guide AlphaFold2 in predicting alternative conformations and inform rational mutation design. *Journal of Chemical Information and Modeling*, 65(18):9459–9468, 2025. doi: 10.1021/acs.jcim.5c01090.

Ali Punjani and David J. Fleet. 3d variability analysis: Resolving continuous flexibility and discrete heterogeneity from single particle cryo-EM. *Journal of Structural Biology*, 213(2):107702, 2021. doi: 10.1016/j.jsb.2021.107702.

Rishwanth Raghu, Axel Levy, Gordon Wetzstein, and Ellen D. Zhong. Multiscale guidance of protein structure prediction with heterogeneous cryo-EM data. In *Advances in Neural Information Processing Systems*, volume 38, 2025. doi: 10.52202/085713-0344. URL https://proceedings.neurips.cc/paper_files/paper/2025/hash/0ee4999d856dc51b6e97b2853e69e9ea-Abstract-Conference.html.

Jay Shenoy, Miro Astore, Axel Levy, Frédéric Poitevin, Sonya M. Hanson, and Gordon Wetzstein. Modeling atomic conformational ensembles of proteins via test-time supervision of Boltz-2 on cryo-EM density maps, 2026. URL https://arxiv.org/abs/2605.09832. Preprint.

Kolja Stahl, Andrea Graziadei, Therese Dau, Oliver Brock, and Juri Rappsilber. Protein structure prediction with in-cell photo-crosslinking mass spectrometry and deep learning. *Nature Biotechnology*, 41:1810–1819, 2023. doi: 10.1038/s41587-023-01704-z.

Kolja Stahl, Robert Warneke, Lorenz Demann, et al. Modelling protein complexes with crosslinking mass spectrometry and deep learning. *Nature Communications*, 15:7866, 2024. doi: 10.1038/s41467-024-51771-2.

B. Su, K. Huang, Z. Peng, A. Amunts, and J. Yang. CryoAtom improves model building for cryo-EM. *Nature Structural & Molecular Biology*, 33:351–361, 2026. doi: 10.1038/s41594-025-01713-3.

Genki Terashi and Daisuke Kihara. De novo main-chain modeling for EM maps using MAINMAST. *Nature Communications*, 9:1618, 2018. doi: 10.1038/s41467-018-04053-7.

Genki Terashi, Xiao Wang, Devashish Prasad, Tsukasa Nakamura, and Daisuke Kihara. DeepMainmast: integrated protocol of protein structure modeling for cryo-EM with deep learning and structure prediction. *Nature Methods*, 21(1):122–131, 2024. doi: 10.1038/s41592-023-02099-0.

Thomas C. Terwilliger, Paul D. Adams, Pavel V. Afonine, and Oleg V. Sobolev. A fully automatic method yielding initial models from high-resolution cryo-electron microscopy maps. *Nature Methods*, 15(11):905–908, 2018. doi: 10.1038/s41592-018-0173-1.

Leonardo G. Trabuco, Elizabeth Villa, Eduard Schreiner, Christopher B. Harrison, and Klaus Schulten. Molecular dynamics flexible fitting: A practical guide to combine cryo-electron microscopy and x-ray crystallography. *Methods*, 49(2):174–180, 2009. doi: 10.1016/j.ymeth.2009.04.005.

Minh H. Tran, Cristina E. Martina, Rocco Moretti, Marcus Nagel, Kevin L. Schey, and Jens Meiler. RosettaHDX: Predicting antibody-antigen interaction from hydrogen-deuterium exchange mass spectrometry data. *Journal of Structural Biology*, 217(1):108166, 2025. doi: 10.1016/j.jsb.2025.108166.

J. Wang, C. Tan, Z. Gao, et al. End-to-end cryo-EM complex structure determination with high accuracy and ultra-fast speed. *Nature Machine Intelligence*, 7:1091–1103, 2025. doi: 10.1038/s42256-025-01056-0.

Xiao Wang, Genki Terashi, and Daisuke Kihara. CryoREAD: de novo structure modeling for nucleic acids in cryo-EM maps using deep learning. *Nature Methods*, 20:1739–1747, 2023. doi: 10.1038/s41592-023-02032-5.

Xiao Wang, Han Zhu, Genki Terashi, Manav Taluja, and Daisuke Kihara. DiffModeler: large macromolecular structure modeling for cryo-EM maps using a diffusion model. *Nature Methods*, 21(12):2307–2317, 2024. doi: 10.1038/s41592-024-02479-0.

Hannah K. Wayment-Steele, Adedolapo Ojoawo, Renee Otten, et al. Predicting multiple conformations via sequence clustering and AlphaFold2. *Nature*, 625:832–839, 2024. doi: 10.1038/s41586-023-06832-9.

Jeremy Wohlwend, Gabriele Corso, Saro Passaro, et al. Boltz-1: Democratizing biomolecular interaction modeling. *bioRxiv*, 2024. doi: 10.1101/2024.11.19.624167. Preprint.

Tianqi Wu, Richard A. Stein, Te-Yu Kao, Benjamin Brown, and Hassane S. Mchaourab. Modeling protein conformational ensembles by guiding AlphaFold2 with double electron electron resonance (DEER) distance distributions. *Nature Communications*, 16:7107, 2025. doi: 10.1038/s41467-025-62582-4.

wwPDB Consortium. EMDB—the electron microscopy data bank. *Nucleic Acids Research*, 52(D1):D456–D465, 2024. doi: 10.1093/nar/gkad1019.

Yuhao Xie, Chengwei Zhang, Shimian Li, et al. Integrating diverse experimental information to assist protein complex structure prediction by GRASP. *Nature Methods*, 22:2362–2374, 2025. doi: 10.1038/s41592-025-02820-1.

Sheng Xu, Qiantai Feng, Lifeng Qiao, et al. Benchmarking all-atom biomolecular structure prediction with FoldBench. *Nature Communications*, 17:442, 2026. doi: 10.1038/s41467-025-67127-3.

Shaoheng Yan, Zian Li, Cai Zhou, Qiaojing Huang, Kai Liu, and Muhan Zhang. Toward better geometric representations for molecule generative models, 2026. URL https://arxiv.org/abs/2605.07693. Preprint.

Sheheryar Zaidi, Michael Schaarschmidt, James Martens, Hyunjik Kim, Yee Whye Teh, Alvaro Sanchez-Gonzalez, Peter Battaglia, Razvan Pascanu, and Jonathan Godwin. Pre-training via denoising for molecular property prediction. In *International Conference on Learning Representations*, 2023. URL https://openreview.net/forum?id=tYIMtogyee.

Chenwei Zhang, Anne Condon, and Khanh Dao Duc. A comprehensive survey and benchmark of deep learning-based methods for atomic model building from cryo-electron microscopy density maps. *Briefings in Bioinformatics*, 26:bbaf322, 2025. doi: 10.1093/bib/bbaf322.

Xi Zhang, Biao Zhang, Lydia Freddolino, and Yang Zhang. CR-I-TASSER: assemble protein structures from cryo-EM density maps using deep convolutional neural networks. *Nature Methods*, 19: 195–204, 2022. doi: 10.1038/s41592-021-01389-9.

Yang Zhang and Jeffrey Skolnick. Scoring function for automated assessment of protein structure template quality. *Proteins: Structure, Function, and Bioinformatics*, 57(4):702–710, 2004. doi: 10.1002/prot.20264.

Ellen D. Zhong, Tristan Bepler, Bonnie Berger, and Joseph H. Davis. CryoDRGN: reconstruction of heterogeneous cryo-EM structures using neural networks. *Nature Methods*, 18:176–185, 2021. doi: 10.1038/s41592-020-01049-4.

Gengmo Zhou, Zhifeng Gao, Qiankun Ding, Hang Zheng, Hongteng Xu, Zhewei Wei, Linfeng Zhang, and Guolin Ke. Uni-Mol: A universal 3d molecular representation learning framework. In *International Conference on Learning Representations*, 2023. URL https://openreview.net/forum?id=6K2RM6wVqKu.