Density-guided AlphaFold reveals unmodeled alternative turn conformations in protein structures

Anonymous Authors[1]

Abstract

Protein structures derived from X-ray crystallography often underrepresent local conformational heterogeneity despite clear support from experimental electron density. Here, we introduce an electron density–guided AlphaFold approach for generating local multiconformer ensembles of β-turn regions directly from crystallographic density maps. The method combines density-guided conformational sampling with dihedral-angle-based classification to filter ensembles and identify distinct clusters of backbone conformations consistent with the experimental density. Across multiple case studies, ED-guided modeling recovers known alternative states, resolves cases of ambiguously deposited altlocs, and uncovers additional conformations absent from deposited PDB models. These results suggest that experimental density-guided structure generation can enable large-scale studies of local conformational variability across the PDB, providing a route to systematically characterize heterogeneous structural states that are only partially captured in the PDB.

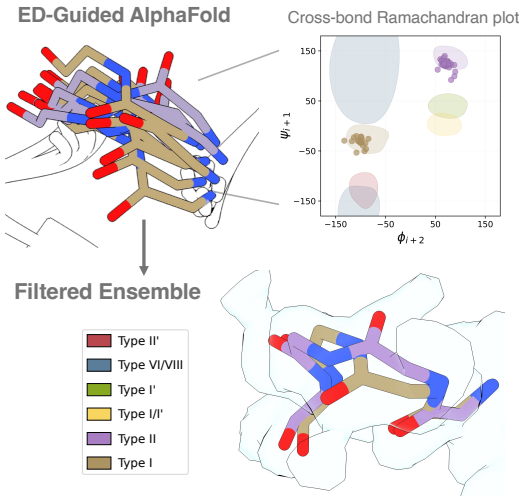

Figure 1. **Local ensemble modeling of β-turn heterogeneity using electron density–guided AlphaFold.** ED-guided AlphaFold generates multiple locally distinct backbone conformations, capturing alternative β-turn geometries. These conformations populate discrete regions in cross-bond Ramachandran space, corresponding to discrete β-turn states. Candidate conformations are filtered to obtain a minimal ensemble that best fits F_o while preserving the conformational heterogeneity. The resulting ensemble captures discrete alternative turn states within a shared density envelope, enabling local multiconformer modeling of β-turn dynamics.

1. Introduction

Macromolecular structures deposited in the Protein Data Bank (PDB) are typically represented as single static models, despite the fact that crystallographic data arise from ensemble-averaged measurements over many conformational states. Electron density maps therefore encode not only a dominant structure but also contributions from alternative conformations and dynamic fluctuations (Austin et al., 1975; Frauenfelder et al., 1979). Traditional representations of heterogeneity rely on atomic displacement parameters (B-factors), which provide a compact description of positional uncertainty but cannot capture anharmonic or multimodal conformational distributions (Trueblood et al., 1996; Burnley et al., 2012).

To more explicitly model heterogeneity, crystallographic refinement introduced discrete multiconformer representations, in which atoms or residues are modeled by a small number of alternative conformations ("altlocs") with associated occupancies. While such models reveal that conformational heterogeneity is widespread (DePristo et al., 2004; Richardson & Richardson, 2013), their identification is often manual and strongly dependent on the initial structural model, leading to systematic under-representation of alternative states. Automated approaches such as qFit address this by sampling local conformational space around an initial model (Van Den Bedem et al., 2009; Keedy et al., 2015; Riley et al., 2021), showing that many regions are better described by multiple, often functionally relevant, states (Fraser et al., 2009). However, local sampling remains biased toward the input model and may fail to recover

[1]Anonymous Institution, Anonymous City, Anonymous Region, Anonymous Country. Correspondence to: Anonymous Author <anon.email@domain.com>.

Submitted to the 2026 Workshop on Generative and Agentic AI for Biology (ICML 2026). Do not distribute.

conformations separated by large backbone rearrangements or discontinuous density features.

Recent work has instead formulated structure determination as an inverse problem, using AlphaFold3 (Abramson et al., 2024)[1] as a generative structural prior conditioned directly on X-ray crystallographic electron density (Maddipatla et al., 2025). Conformational ensembles are inferred as samples from density-consistent posterior distributions rather than by sampling around an initial model (Maddipatla et al., 2024; 2026). Interpreting the heterogeneity captured by these ensembles, however, requires mapping the generated conformations onto discrete structural states that can be compared across proteins.

β-turns provide a natural setting for this problem: their conformational states are defined by backbone dihedral angles, and a recent large-scale molecular dynamics study identified six distinct β-turn types accessible to these motifs (Zhang et al., 2026). Here, we apply the framework of Maddipatla et al. (2025) to β-turns and ask whether density-guided ensembles concentrate around a single dominant turn state, separate into multiple discrete clusters, or exhibit broader variability within the β-turn landscape. By mapping generated structures onto turn states and retaining density-supported representatives, we summarize each ensemble as a compact set of experimentally supported β-turn conformations, and characterize how this heterogeneity is distributed across crystallographic structures.

Contributions. We use β-turns as a discrete, dihedral-based representation of local backbone heterogeneity, enabling generated conformations to be compared across structures through shared turn-state labels (Fig. 1). We show that electron density–guided AlphaFold3 recovers alternative β-turn conformations directly from crystallographic density, including states absent from deposited PDB structures and qFit-derived ensembles (Figs. 2 and 5). Across independent random seedings, the turn-type distributions of the relaxed ensembles remain consistent and the filtered ensembles converge to similar compositions, indicating reproducible sampling and stable selection (Fig. 4). Together, these results support large-scale density-guided analysis of local conformational variability across experimentally determined protein structures.

2. Methods

2.1. Identification of β-turns potentially harboring alternate conformations

Identifying β-turns that harbor true conformational heterogeneity is non-trivial because alternative states are systematically underrepresented in deposited PDB models. Simply querying existing altloc annotations would restrict our evaluation to already-solved cases, preventing us from testing the method's ability to discover unmodeled states. To establish a robust benchmark, we first identified known altlocs modeled in the PDB at turn locations. Then, to find regions potentially harboring *unmodeled* alternative conformations, we leveraged structural variations observed across homologous crystal structures.

To achieve this, a non-redundant set of X-ray crystallographic proteins was assembled from PDB entries with refined models available in PDB-REDO (Joosten et al., 2014) using the following criteria: (i) resolution $\leq$ 2.5 Å; (ii) $R_{\text{free}} \leq 0.3$; and (iii) a defined expression system (*E. coli*). Structures were grouped by UniProt identifier (Consortium, 2015), and representative reference chains were selected based on resolution and sequence completeness.

For each reference chain, homologous chains were identified using BLAST (Altschul et al., 1990) and structurally aligned (Kabsch, 1976). Closely matching structures were retained based on sequence similarity and backbone RMSD. Candidate alternate local conformations were identified by comparing short aligned backbone windows between structures. This procedure operates on the premise that if two homologous structures exhibit different backbone conformations in the exact same local context, their respective electron densities may actually contain a mixture of both states. To ensure these variations arise from intrinsic dynamics rather than divergent sequence or crystal packing artifacts, matched windows were strictly required to have: (i) identical local sequence; (ii) comparable residue interaction environments, computed using Arpeggio (Jubb et al., 2017); and (iii) a candidate β-turn, defined as a four-residue backbone segment (i–$i+3$) with C_α atoms of residues i and $i+3$ separated by less than 7.5 Å.

2.2. Ensemble generation using electron density-guided AlphaFold3

For each PDB entry, we generated an observed electron density map F_{o} in physical units of $\text{e}^-/\text{Å}^3$ using Lang et al. (2014), since PDB deposited maps are typically mean-centered and lack an absolute scale. Using `phenix.map_box` (Adams et al., 2010), we carved a map box around the corresponding PDB-REDO model with 10 Å padding. To focus the analysis on the local β-turn motif, density outside a 5 Å neighborhood of the turn residues was masked. We then fit an ensemble $\boldsymbol{\mathcal{X}} = \{\mathbf{X}^1, \ldots, \mathbf{X}^n\}$ of $n = 64$ structural models to this local density using the density-guided procedure of Maddipatla et al. (2025).

From the initial ensemble, we discarded structures with stereochemical artifacts, defined as structures in which any bonded-atom distance exceeded 2.1 Å or any nonbonded interatomic distance was less than 1.1 Å. The remaining

[1] AlphaFold, AF3, and AlphaFold3 are used interchangeably.

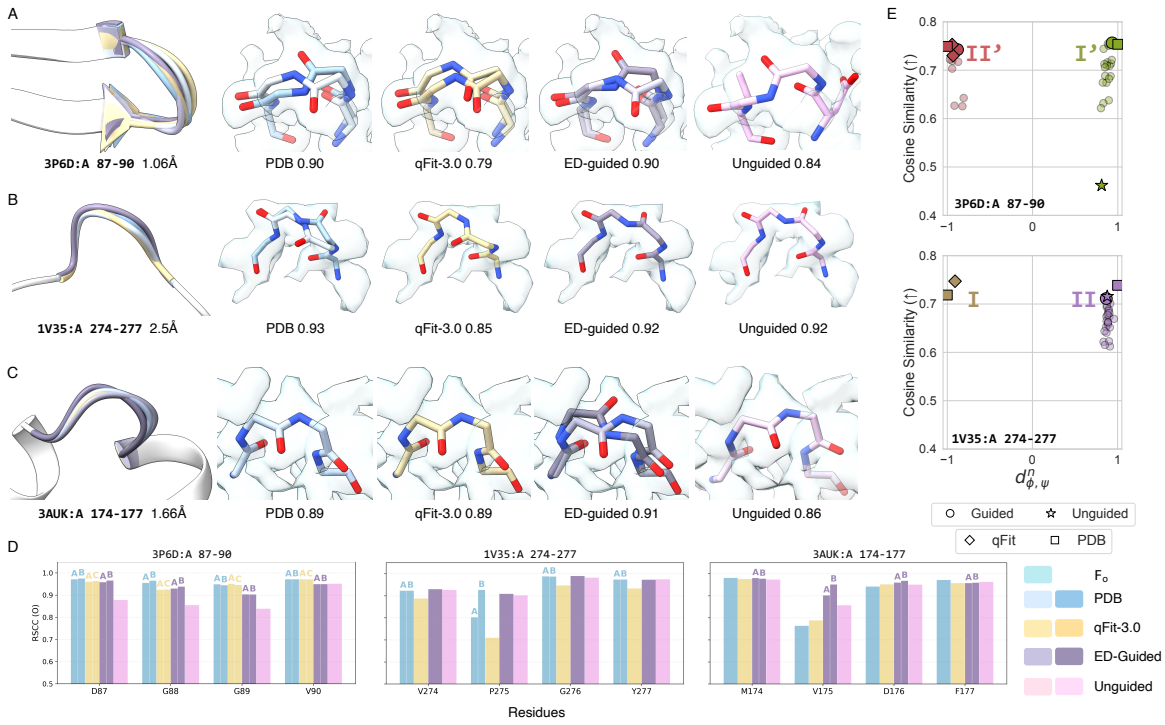

Figure 2. **Electron density–guided AlphaFold recovers and reveals unmodeled alternative local backbone conformations consistent with experimental electron density.** (A–C) Representative local β-turn regions comparing deposited PDB conformations, qFit3.0, electron density–guided AlphaFold (ED-guided), and unguided AlphaFold3 models. Numbers indicate local cosine similarity to the density. (A) `3P6D:A`, residues 87–90, where ED-guided AlphaFold recovers deposited altloc states. (B) `1V35:A`, residues 274–277, where deposited altlocs collapse into a single dominant density-supported state. (C) `3AUK:A`, residues 174–177, where ED-guided AlphaFold reveals alternative conformations not modeled in the deposited structure. (D) Per-residue backbone oxygen RSCC values for (A–C), showing that ED-guided conformations achieve comparable or improved agreement with experimental density at heterogeneous sites. Cyan surfaces indicate observed electron density (F_o); blue, yellow, purple, and pink correspond to PDB, qFit3.0, ED-guided, and unguided models. (E) Relative toroidal dihedral-angle distance between ED-guided conformations and deposited PDB altloc states, with -1 and $+1$ corresponding to altloc states B and A. The y-axis reports cosine similarity to the experimental electron density. Points are colored by β-turn type; symbols denote modeling methods (circles, ED-guided; stars, unguided; diamonds, qFit; squares, PDB). Faded circles represent all ED-guided conformations, whereas opaque circles indicate the filtered ensemble.

structures were relaxed in vacuum using the AMBER force field (Wang et al., 2004) to correct minor bond-length deviations and improve local structural consistency.

After relaxation, we retained structures that both adopted a β-turn geometry and fit the local density. For each relaxed structure, we computed a backbone-only calculated density map F_c using the N, C_α, C, and O atoms of the four β-turn residues. The calculated map was evaluated on the same local mask as F_o; details of the F_c calculation are provided in Appendix B.1. Each member was scored by the masked cosine similarity between F_c and F_o. Structures that failed the β-turn geometric criterion or whose cosine similarity fell in the bottom 40% of the relaxed ensemble were discarded.

Each surviving ensemble member was then assigned to one of the six β-turn types defined by Zhang et al. (2026). For a turn spanning residues i through $i+3$, we represented its backbone geometry by the four central dihedral angles $\boldsymbol{\theta} = (\phi_{i+1}, \psi_{i+1}, \phi_{i+2}, \psi_{i+2}) \in \mathbb{T}^4$ in structure $\mathbf{X}$. Each angle (in degrees) has a 180° periodicity. Using the six reference β-turn centroids $\{\boldsymbol{\mu}_k\}_{k=1}^6$ from Zhang et al. (2026), each ensemble member $\mathbf{X}^m$ (with dihedral vector $\boldsymbol{\theta}_m$) was assigned to its nearest centroid:

$$c_m = \arg\min_{k\in\{1,\dots,6\}} d_{\text{torus}}(\boldsymbol{\theta}_m, \boldsymbol{\mu}_k),$$

where d_{torus} is the squared shortest-path distance on the 4D torus $\mathbb{T}^4$, defined for vectors $\boldsymbol{a}, \boldsymbol{b}$ as:

$$d_{\text{torus}}(\boldsymbol{a}, \boldsymbol{b}) = \sum_{j=1}^{4} \left[\min\left(|a_j - b_j|, 2\pi - |a_j - b_j|\right)\right]^2.$$

To filter weakly supported assignments, we applied silhouette-based pruning to the surviving members $\boldsymbol{\mathcal{M}} \subset \boldsymbol{\mathcal{X}}$. Excluding singletons (which were treated as outliers), we computed the silhouette coefficient $s(m)$ for each member:

$$s(m) = \frac{b(m) - a(m)}{\max\{a(m), b(m)\}},$$

where $a(m)$ is the mean distance from $\boldsymbol{\theta}_m$ to other members in its assigned cluster $\boldsymbol{\mathcal{M}}_{c_m}$, and $b(m)$ is the mean distance to the nearest competing cluster.

We iteratively removed the smallest active cluster until two remained, selecting the partition that maximized the mean silhouette $\bar{s} = |\boldsymbol{\mathcal{M}}|^{-1} \sum_{\mathbf{X}^m \in \boldsymbol{\mathcal{M}}} s(m)$. The clusters in this optimal partition represent the robustly recovered β-turn states. Finally, for each recovered state, the member with the highest local density cosine similarity was selected as the representative structure.

3. Results

3.1. Density-guided modeling captures and extends β-turn conformational states

Across representative β-turn regions (Fig. 2), ED-guided models accurately recapitulate known alternative conformations (altlocs), resolve ambiguous cases where the density supports a single dominant state, and reveal alternative backbone states not modeled in deposited structures. In the case of `1V35:A` (Fig. 2B), the deposited structure contains two altloc conformations, whereas ED-guided modeling converges to a single dominant state. Despite resolving the bimodal deposited model into a unimodal ensemble, the resulting ED-guided conformation maintains comparable cosine similarity to the experimental density, with only a slight decrease relative to the deposited altloc model. Per-residue backbone oxygen RSCC analysis further supports this interpretation (Fig. 2D), as one deposited altloc state exhibits substantially reduced RSCC values, suggesting weaker experimental support for that conformation. Together, these results indicate that ED-guided modeling preferentially selects the conformation most strongly supported by the density, acting as a robust structural filter rather than blindly reproducing potentially weakly supported alternative states.

Crucially, ED-guided modeling also discovers previously unmodeled conformations that better explain the density in cases where deposited models appear incomplete (Fig. 2C and 5). Across multiple β-turn regions, ED-guided ensembles reveal alternative backbone states that are not modeled in the deposited PDB structures (Table 1) yet remain strongly supported by the experimental density. In several cases, these conformations extend beyond both deposited altloc annotations and qFit-derived ensembles, demonstrating the method's unique capacity to uncover hidden conformational heterogeneity systematically underrepresented in current structural models.

To systematically compare across methods, we next evaluated conformational agreement with experimental density and dihedral-angle similarity to deposited PDB structures (Fig. 2E and 3). ED-guided conformations span a broader region of dihedral space while maintaining high cosine similarity to the density, indicating that the method captures both structural diversity and density consistency. Quantitative comparisons further support this trend, with ED-guided models achieving higher or comparable cosine similarity relative to PDB and qFit models, while maintaining consistent crystallographic agreement as measured by R_{work} and R_{free} (Tables 2–4), confirming that our localized four-residue perturbations do not degrade the overall model viability.

Together, these results demonstrate that ED-guided modeling not only refines and validates historically modeled conformational heterogeneity but significantly extends it by uncovering experimentally supported, yet previously hidden, alternative states.

3.2. ED-guided ensembles are robust across independent runs

Finally, we assessed the robustness of ED-guided ensembles across independent random seedings (Fig. 4). Across multiple runs, the distributions of β-turn types in the relaxed ensembles remain consistent, and the filtered ensembles converge to similar compositions, indicating stable ensemble selection. This reproducibility suggests that the observed conformational preferences are not driven by stochastic initialization but reflect underlying density-supported structural states.

4. Conclusion

We presented an electron density–guided AlphaFold framework for local multiconformer modeling of β-turn regions from X-ray crystallographic density maps. By combining density-guided conformational sampling with ensemble filtering, the method recovers experimentally supported alternative backbone conformations while maintaining agreement with experimental density and crystallographic metrics. Across representative systems, ED-guided modeling reproduced known altloc states, identified cases where deposited multiconformer models were weakly supported by the density, and revealed previously unmodeled alternative conformations consistent with the experimental map. Quantitative comparisons against deposited PDB structures and qFit-derived ensembles further showed that ED-guided conformations achieve comparable or improved local density agreement while capturing broader conformational diversity. In addition, ensemble compositions remained consistent across independent random seedings, indicating robust and reproducible conformational sampling. Together, these results demonstrate that integrating experimental density directly into structure prediction models allows a more accurate and complete characterization of local conformational heterogeneity and provide a scalable framework for density-guided ensemble modeling in structural biology.

Impact Statement

This work introduces a machine learning framework for density-guided modeling of local protein conformational heterogeneity from experimental structural data. The primary goal of this research is to advance methods for integrating generative protein structure prediction with experimental observations in structural biology. Potential downstream applications include improved characterization of protein dynamics, conformational variability, and experimentally supported multiconformer ensembles, which may contribute to future advances in molecular biology, protein engineering, and drug discovery. We do not anticipate direct negative societal consequences arising from this work beyond the general considerations associated with computational modeling and machine learning research.

References

Abramson, J., Adler, J., Dunger, J., Evans, R., Green, T., Pritzel, A., Ronneberger, O., Willmore, L., Ballard, A. J., Bambrick, J., et al. Accurate structure prediction of biomolecular interactions with alphafold 3. *Nature*, 630 (8016):493–500, 2024.

Adams, P. D., Afonine, P. V., Bunkóczi, G., Chen, V. B., Davis, I. W., Echols, N., Headd, J. J., Hung, L.-W., Kapral, G. J., Grosse-Kunstleve, R. W., et al. Phenix: a comprehensive python-based system for macromolecular structure solution. *Biological crystallography*, 66(2): 213–221, 2010.

Agirre, J., Atanasova, M., Bagdonas, H., Ballard, C. B., Baslé, A., Beilsten-Edmands, J., Borges, R. J., Brown, D. G., Burgos-Mármol, J. J., Berrisford, J. M., et al. The ccp4 suite: integrative software for macromolecular crystallography. *Biological Crystallography*, 79(6):449–461, 2023.

Altschul, S. F., Gish, W., Miller, W., Myers, E. W., and Lipman, D. J. Basic local alignment search tool. *Journal of molecular biology*, 215(3):403–410, 1990.

Austin, R. H., Beeson, K., Eisenstein, L., Frauenfelder, H., and Gunsalus, I. Dynamics of ligand binding to myoglobin. *Biochemistry*, 14(24):5355–5373, 1975.

Brock, C. P., Hahn, T., Wondratschek, H., Müller, U., Shmueli, U., Prince, E., Authier, A., Kopskỳ, V., Litvin, D., Arnold, E., et al. International tables for crystallography volume a: Space-group symmetry, 2016.

Burnley, B. T., Afonine, P. V., Adams, P. D., and Gros, P. Modelling dynamics in protein crystal structures by ensemble refinement. *elife*, 1:e00311, 2012.

Consortium, U. Uniprot: a hub for protein information. *Nucleic acids research*, 43(D1):D204–D212, 2015.

DePristo, M. A., De Bakker, P. I., and Blundell, T. L. Heterogeneity and inaccuracy in protein structures solved by x-ray crystallography. *Structure*, 12(5):831–838, 2004.

Fraser, J. S., Clarkson, M. W., Degnan, S. C., Erion, R., Kern, D., and Alber, T. Hidden alternative structures of proline isomerase essential for catalysis. *Nature*, 462 (7273):669–673, 2009.

Frauenfelder, H., Petsko, G. A., and Tsernoglou, D. Temperature-dependent x-ray diffraction as a probe of protein structural dynamics. *Nature*, 280(5723):558–563, 1979.

Joosten, R. P., Long, F., Murshudov, G. N., and Perrakis, A. The pdb_redo server for macromolecular structure model optimization. *IUCrJ*, 1(4):213–220, 2014.

Jubb, H. C., Higueruelo, A. P., Ochoa-Montaño, B., Pitt, W. R., Ascher, D. B., and Blundell, T. L. Arpeggio: a web server for calculating and visualising interatomic interactions in protein structures. *Journal of molecular biology*, 429(3):365–371, 2017.

Kabsch, W. A solution for the best rotation to relate two sets of vectors. *Foundations of Crystallography*, 32(5): 922–923, 1976.

Keedy, D. A., Fraser, J. S., and van den Bedem, H. Exposing hidden alternative backbone conformations in x-ray crystallography using qfit. *PLoS computational biology*, 11(10):e1004507, 2015.

Lang, P. T., Holton, J. M., Fraser, J. S., and Alber, T. Protein structural ensembles are revealed by redefining x-ray electron density noise. *Proceedings of the National Academy of Sciences*, 111(1):237–242, 2014.

Maddipatla, A., Bojan, N. S., Bojan, M., Masalitin, V., Vedula, S., Schanda, P., Marx, A., and Bronstein, A. M. Experiment-guided alphafold3 resolves accurate protein ensembles. *bioRxiv*, pp. 2025–10, 2025.

Maddipatla, A., Vedula, S., Bronstein, A. M., and Marx, A. Density-guided alphafold3 uncovers unmodelled conformations in β_2-microglobulin. *bioRxiv*, pp. 2026–02, 2026.

Maddipatla, S. A., Sellam, N. B., Vedula, S., Marx, A., and Bronstein, A. Generative modeling of protein ensembles guided by crystallographic electron densities. *arXiv preprint arXiv:2412.13223*, 2024.

Prince, E. *International Tables for Crystallography, Volume C: Mathematical, physical and chemical tables*. Springer Science & Business Media, 2004.

Richardson, J. S. and Richardson, D. C. Doing molecular biophysics: Finding, naming, and picturing signal within complexity. *Annual review of biophysics*, 42:1–28, 2013.

Riley, B. T., Wankowicz, S. A., De Oliveira, S. H. P., Van Zundert, G. C. P., Hogan, D. W., Fraser, J. S., Keedy, D. A., and Van Den Bedem, H. qFit: Protein and ligand multiconformer modeling for x-ray crystallographic and single-particle cryo-EM density maps. *Protein Science*, 30(1):270–285, 2021. ISSN 0961-8368, 1469-896X. doi: 10.1002/pro.4001. URL https://onlinelibrary.wiley.com/doi/10.1002/pro.4001.

Trueblood, K., Bürgi, H.-B., Burzlaff, H., Dunitz, J., Gramaccioli, C., Schulz, H., Shmueli, U., and Abrahams, S. Atomic dispacement parameter nomenclature. report of a subcommittee on atomic displacement parameter nomenclature. *Foundations of Crystallography*, 52(5):770–781, 1996.

Van Den Bedem, H., Dhanik, A., Latombe, J.-C., and Deacon, A. M. Modeling discrete heterogeneity in x-ray diffraction data by fitting multi-conformers. *Biological crystallography*, 65(10):1107–1117, 2009.

Wang, J., Wolf, R. M., Caldwell, J. W., Kollman, P. A., and Case, D. A. Development and testing of a general amber force field. *Journal of computational chemistry*, 25(9): 1157–1174, 2004.

Zhang, S., Maddipatla, S. A., Vedula, S., Marx, A., and Bronstein, A. M. The turn less taken: Investigating patterns in β-turn dynamics using large-scale molecular dynamics data. *bioRxiv*, 2026. doi: 10.64898/2026.05.07.721674. URL https://www.biorxiv.org/content/early/2026/05/08/2026.05.07.721674.

A. Appendix

A.1. Tables

Table 1. **Detection of distinct β-turn altloc states across structures and methods.** A turn altloc is annotated as present when two distinct β-turn states, assigned according to the turn-type classification of Zhang et al. (2026), are detected across the corresponding ensemble. The lower section of the table indicates cases in which distinct β-turn altloc states are detected by a method despite not being annotated in the deposited PDB model.

		Turn Altloc Detected			
Structure	**Residue Range**	**PDB**	**qFit**	**Guided AF3**	**AF3**
1GOU:A	65–68	✓	✗	✓	✓
1IX9:A	134–137	✓	✗	✓	✗
1M7S:A	121–124	✓	✗	✓	✗
1N6P:A	90–93	✓	✗	✓	✗
1QOP:A	155–158	✓	✗	✓	✗
1V35:A	274–277	✓	✗	✗	✗
1X9I:B	27–30	✓	✗	✓	✗
1YWA:A	100–103	✓	✗	✓	✗
2B49:A	833–836	✓	✗	✓	✗
2DEI:A	177–180	✓	✗	✓	✗
2HC0:B	2149–2152	✓	✗	✓	✗
2HNX:A	119–122	✓	✗	✓	✗
2HS1:A	15–18	✓	✗	✓	✗
2HZI:A	388–391	✓	✗	✗	✗
2IMH:A	110–113	✓	✗	✓	✓
2JIN:A	73–76	✓	✗	✓	✗
2NS0:A	20–23	✓	✗	✓	✗
2OYZ:A	20–23	✓	✗	✓	✗
2OYZ:A	56–59	✓	✗	✗	✗
2R9X:B	262–265	✓	✗	✓	✗
3P6D:A	87–90	✓	✗	✓	✗
1AB0:A	119–122	✗	✗	✓	✓
1ERZ:A	161–164	✗	✗	✓	✗
2XUR:A	38–41	✗	✗	✓	✗
3AUK:A	174–177	✗	✗	✓	✗
3E0U:A	156–159	✗	✗	✓	✗
3WAL:A	42–45	✗	✗	✓	✗
4M8I:A	233–236	✗	✗	✓	✓
4N0D:A	164–167	✗	✗	✓	✗
4X7B:A	174–177	✗	✗	✓	✗

Table 2. **Quantitative evaluation of cosine similarity (↑).** Entries highlighted in **green** correspond to ensembles that outperform all baseline methods, excluding the PDB, while entries highlighted in **blue** outperform all baselines, including the PDB. The lower section of the table indicates cases in which distinct β-turn altloc states are detected by a method despite not being annotated in the deposited PDB model.

PDB ID	Residue Range	PDB	qFit	Guided AF3	Unguided AF3
1GOU:A	65–68	0.761	0.743	0.795	**0.813**
1IX9:A	134–137	0.903	0.825	**0.889**	0.761
1M7S:A	121–124	0.872	0.834	0.861	**0.883**
1N6P:A	90–93	0.912	0.898	**0.913**	0.887
1QOP:A	155–158	0.864	0.826	**0.873**	0.822
1V35:A	274–277	0.927	0.914	**0.920**	0.905
1X9I:B	27–30	0.875	0.820	**0.874**	0.847
1YWA:A	100–103	0.921	0.905	**0.918**	0.803
2B49:A	833–836	0.862	0.827	**0.856**	0.758
2DEI:A	177–180	0.894	0.857	**0.876**	0.803
2HC0:B	2149–2152	0.816	**0.831**	0.820	0.814
2HNX:A	119–122	0.884	0.864	**0.887**	0.881
2HS1:A	15–18	0.892	0.882	**0.895**	0.682
2HZI:A	388–391	0.833	**0.794**	0.792	0.588
2IMH:A	110–113	0.890	0.866	0.887	**0.912**
2JIN:A	73–76	0.514	0.502	**0.538**	0.529
2NS0:A	20–23	0.917	0.897	**0.922**	0.849
2OYZ:A	20–23	0.865	0.843	**0.845**	0.821
2OYZ:A	56–59	0.864	**0.852**	0.844	0.840
2R9X:B	262–265	0.910	0.818	**0.835**	0.777
3P6D:A	87–90	0.899	0.891	**0.903**	0.847
1AB0:A	119–122	0.936	**0.936**	0.930	0.897
1ERZ:A	161–164	0.934	0.935	**0.948**	0.944
2XUR:A	38–41	0.914	0.878	**0.927**	0.909
3AUK:A	174–177	0.895	0.883	**0.913**	0.886
3E0U:A	156–159	0.933	**0.937**	0.919	0.661
3WAL:A	42–45	0.853	0.852	**0.868**	0.820
4M8I:A	233–236	0.848	0.812	**0.855**	0.654
4N0D:A	164–167	0.922	0.908	**0.925**	0.816
4X7B:A	174–177	0.915	**0.912**	0.910	0.798

Table 3. **Quantitative evaluation of R_{work} (↓).** Entries highlighted in **green** correspond to ensembles that outperform all baseline methods, excluding the PDB, while entries highlighted in **blue** outperform all baselines, including the PDB. The lower section of the table indicates cases in which distinct β-turn altloc states are detected by a method despite not being annotated in the deposited PDB model.

PDB ID	Residue Range	PDB	qFit	Guided AF3	Unguided AF3
1GOU:A	65–68	0.164	0.165	**0.164**	0.165
1IX9:A	134–137	0.154	0.154	**0.154**	0.155
1M7S:A	121–124	0.169	0.169	**0.169**	0.169
1N6P:A	90–93	0.159	**0.159**	0.160	0.160
1QOP:A	155–158	0.156	0.156	**0.156**	0.157
1V35:A	274–277	0.134	0.135	**0.135**	0.136
1X9I:B	27–30	0.164	0.165	**0.164**	0.164
1YWA:A	100–103	0.168	0.168	**0.168**	0.171
2B49:A	833–836	0.219	**0.219**	0.220	0.221
2DEI:A	177–180	0.162	**0.162**	0.162	0.163
2HC0:B	2149–2152	0.195	**0.195**	0.196	0.197
2HNX:A	119–122	0.168	0.168	**0.168**	0.170
2HS1:A	15–18	0.179	0.178	**0.179**	0.184
2HZI:A	388–391	0.164	0.164	**0.164**	0.167
2IMH:A	110–113	0.155	0.155	**0.155**	0.156
2JIN:A	73–76	0.173	**0.175**	0.184	0.187
2NS0:A	20–23	0.183	0.184	**0.183**	0.192
2OYZ:A	20–23	0.203	**0.203**	0.205	0.206
2OYZ:A	56–59	0.203	**0.203**	0.203	0.204
2R9X:B	262–265	0.150	0.151	**0.150**	0.151
3P6D:A	87–90	0.184	0.185	**0.184**	0.188
1AB0:A	119–122	0.189	0.189	**0.189**	0.192
1ERZ:A	161–164	0.140	0.140	**0.140**	0.141
2XUR:A	38–41	0.180	0.180	**0.180**	0.180
3AUK:A	174–177	0.151	0.151	**0.151**	0.153
3E0U:A	156–159	0.225	**0.225**	0.227	0.238
3WAL:A	42–45	0.189	**0.189**	0.190	0.190
4M8I:A	233–236	0.180	0.182	**0.180**	0.185
4N0D:A	164–167	0.156	**0.156**	0.157	0.159
4X7B:A	174–177	0.186	**0.186**	0.187	0.189

Table 4. **Quantitative evaluation of R_{free} (↓).** Entries highlighted in **green** correspond to ensembles that outperform all baseline methods, excluding the PDB, while entries highlighted in **blue** outperform all baselines, including the PDB. The lower section of the table indicates cases in which distinct β-turn altloc states are detected by a method despite not being annotated in the deposited PDB model.

PDB ID	Residue Range	PDB	qFit	Guided AF3	Unguided AF3
1GOU:A	65–68	0.187	0.188	**0.187**	0.187
1IX9:A	134–137	0.164	0.164	**0.162**	0.162
1M7S:A	121–124	0.206	0.206	**0.206**	0.206
1N6P:A	90–93	0.190	0.191	**0.191**	0.193
1QOP:A	155–158	0.176	0.176	**0.176**	0.176
1V35:A	274–277	0.179	**0.179**	0.180	0.180
1X9I:B	27–30	0.181	0.182	**0.181**	0.182
1YWA:A	100–103	0.177	0.178	**0.177**	0.180
2B49:A	833–836	0.251	0.252	**0.252**	0.253
2DEI:A	177–180	0.203	**0.202**	0.205	0.205
2HC0:B	2149–2152	0.223	0.223	**0.223**	0.227
2HNX:A	119–122	0.206	0.206	**0.205**	0.208
2HS1:A	15–18	0.191	0.192	**0.190**	0.197
2HZI:A	388–391	0.197	0.197	**0.197**	0.199
2IMH:A	110–113	0.193	**0.193**	0.194	0.193
2JIN:A	73–76	0.229	**0.230**	0.244	0.241
2NS0:A	20–23	0.234	0.234	**0.233**	0.241
2OYZ:A	20–23	0.248	**0.251**	0.260	0.258
2OYZ:A	56–59	0.248	**0.247**	0.248	0.250
2R9X:B	262–265	0.187	0.188	**0.188**	0.189
3P6D:A	87–90	0.204	0.204	**0.204**	0.206
1AB0:A	119–122	0.240	**0.239**	0.240	0.242
1ERZ:A	161–164	0.170	0.170	**0.169**	0.170
2XUR:A	38–41	0.229	0.229	**0.228**	0.228
3AUK:A	174–177	0.175	0.175	**0.175**	0.177
3E0U:A	156–159	0.276	**0.277**	0.281	0.290
3WAL:A	42–45	0.236	0.237	**0.236**	0.241
4M8I:A	233–236	0.214	**0.209**	0.214	0.219
4N0D:A	164–167	0.183	**0.182**	0.183	0.185
4X7B:A	174–177	0.242	**0.242**	0.242	0.244

A.2. Figures

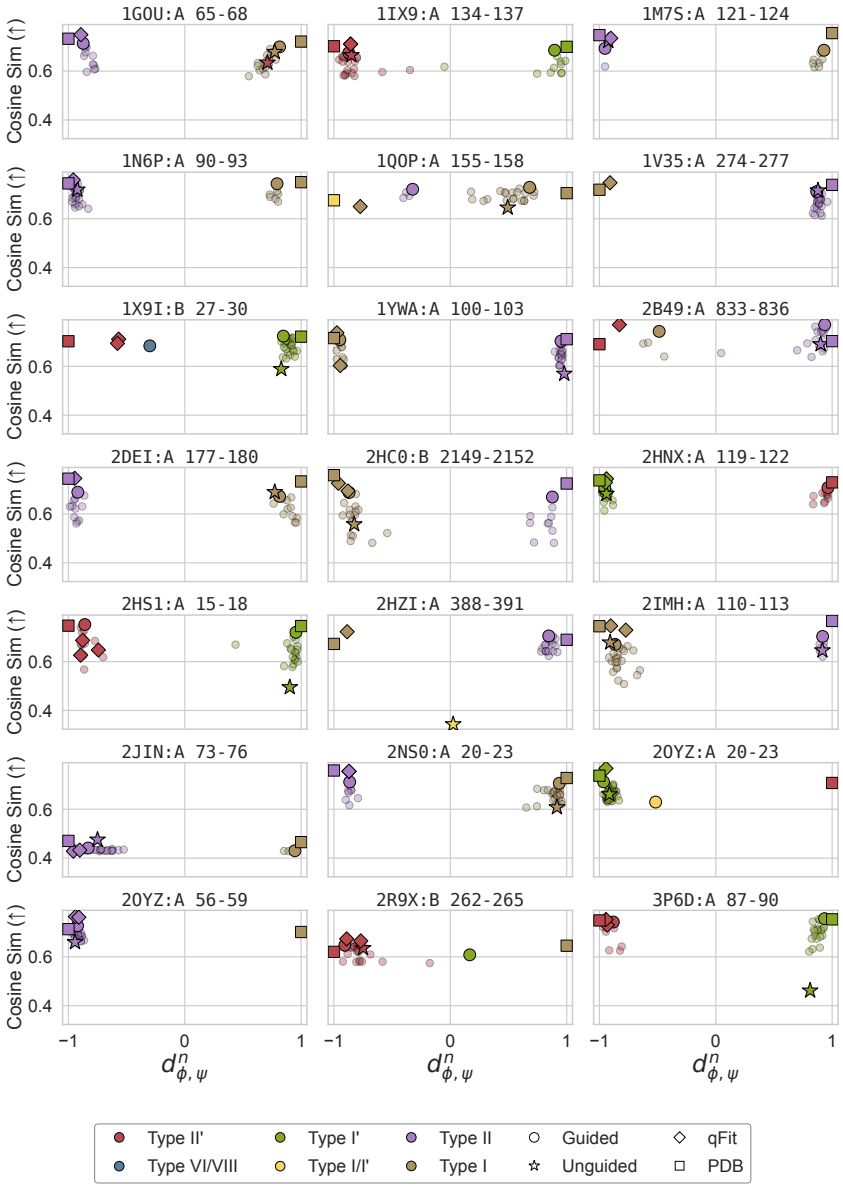

Figure 3. **Comparison of predicted conformational states and agreement with experimental electron density across representative β-turn regions.** Each panel corresponds to a selected PDB structure and residue range. Points are plotted in the space of cross-bond dihedral similarity ($d^n_{\phi,\psi}$, X-axis) versus cosine similarity to the experimental electron density (Y-axis). Colors indicate β-turn types assigned to each conformation. Symbols denote different modeling approaches: circles, electron density–guided (ED-guided) AlphaFold; stars, unguided AlphaFold; diamonds, qFit; and squares, deposited PDB conformations. Semi-transparent circles represent the full set of conformations generated by ED-guided AlphaFold, whereas opaque circles indicate the subset retained after ensemble filtering.

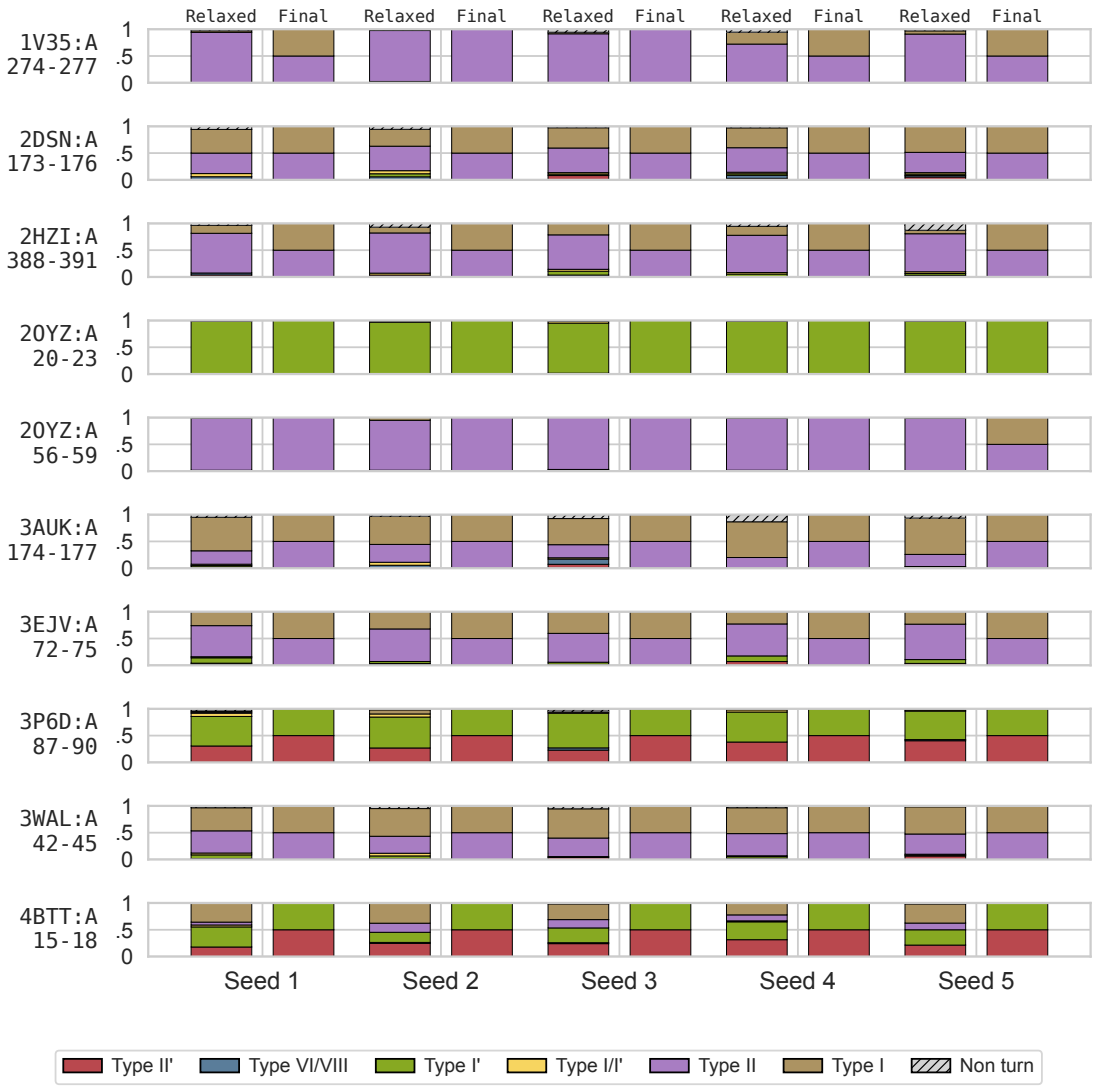

Figure 4. **Consistency of ED-guided ensembles across independent random seedings.** For each protein and residue range, five independent ED-guided runs (seeds) are shown. Within each seed, the left bar represents the distribution of β-turn types across the relaxed ensemble, whereas the right bar shows the composition after ensemble filtering. Colors denote β-turn types: $\mathtt{II}'$ (red), $\mathtt{VI/VIII}$ (blue), $\mathtt{I}'$ (green), $\mathtt{I/I}'$ (yellow), $\mathtt{II}$ (purple), and $\mathtt{I}$ (beige); hatched regions indicate non-turn conformations. Across seeds, both relaxed and filtered ensembles exhibit consistent turn-type distributions, indicating that ED-guided modeling yields reproducible conformational preferences and robust ensemble selection.

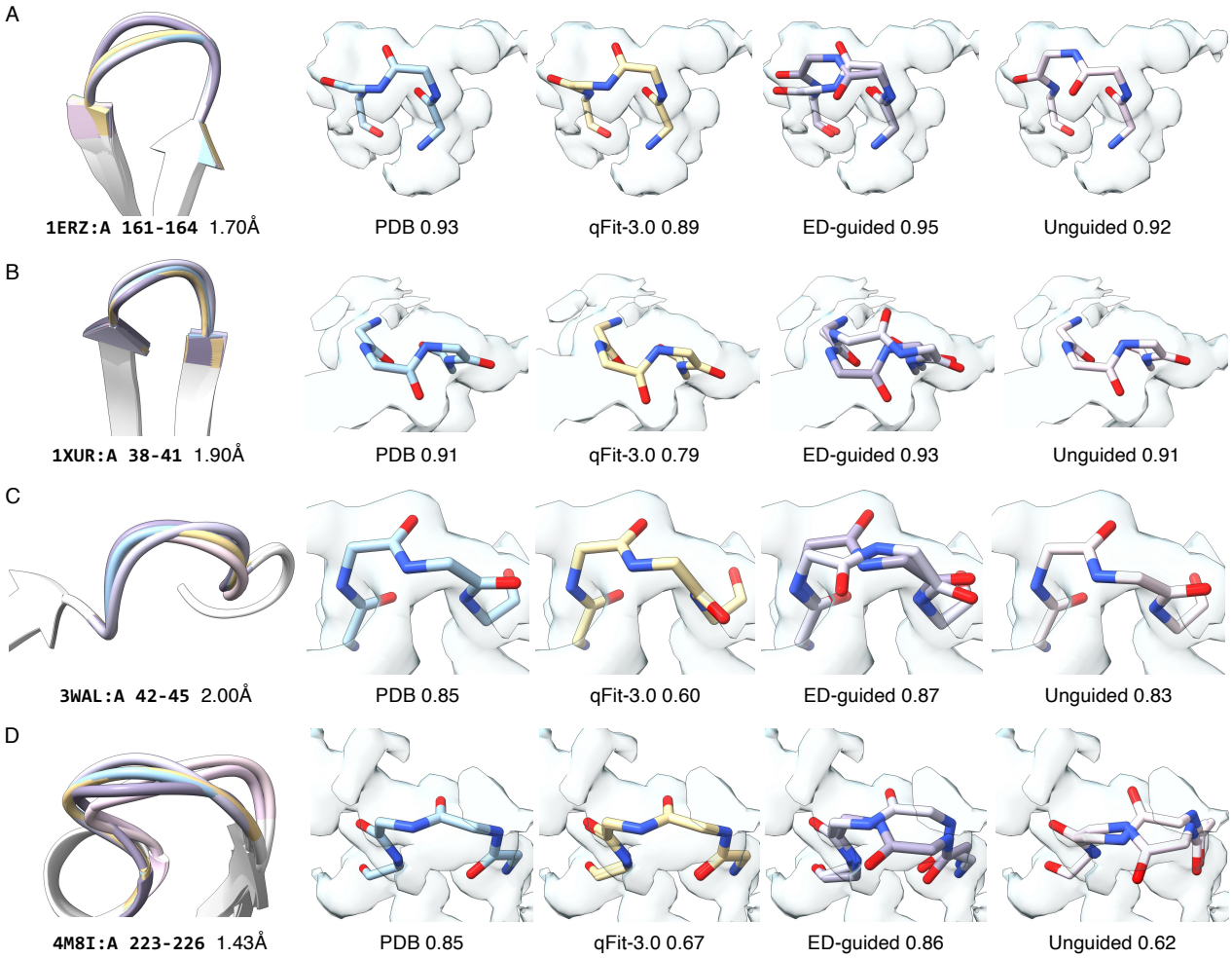

Figure 5. **Electron density–guided AlphaFold reveals previously unmodeled alternative backbone conformations consistent with experimental density.** (A–D) Representative β-turn regions comparing deposited PDB conformations (blue), qFit multiconformer models (yellow), electron density–guided (ED-guided) AlphaFold (purple), and unguided AlphaFold3 (pink) predictions. For each example, the left panel shows local backbone superpositions, and the right panels show the corresponding conformations fitted into the experimental electron density. Numbers below each model indicate cosine similarity to the experimental density. In multiple cases, ED-guided AlphaFold identifies alternative conformations not modeled in the deposited PDB structures, providing improved agreement with the experimental density

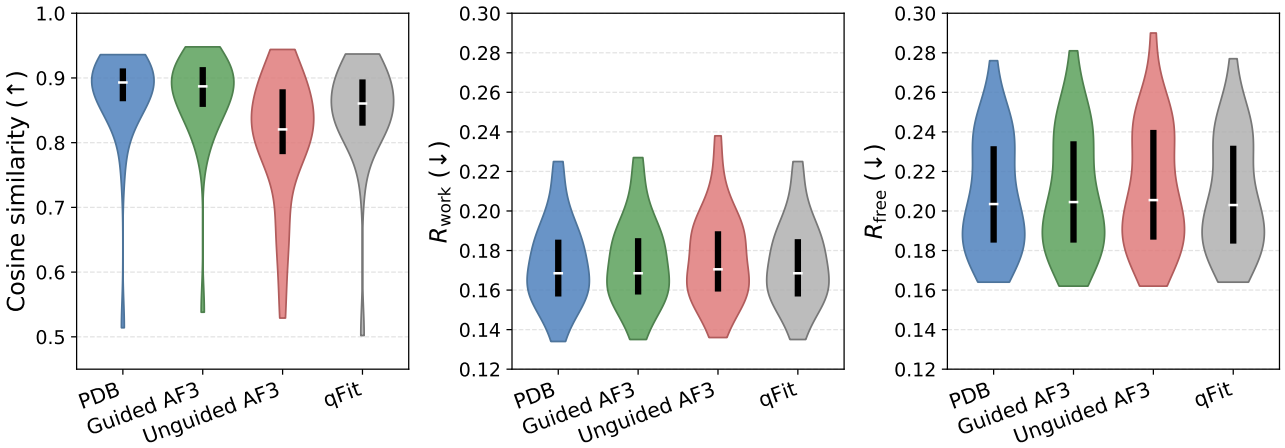

Figure 6. **Summary of X-ray crystallographic benchmark performance.** Distribution of cosine similarity between F_o and F_c (left), R_{work} (middle), and R_{free} (right) across the evaluated benchmarks. Full per-structure results are provided in Tables 2–4.

B. Additional Details

B.1. X-ray crystallography model

Let $\boldsymbol{\mathcal{X}} = \{\mathbf{X}^k\}_{k=1}^n$ be an ensemble of n conformers and conformer $\mathbf{X}^k$ have atomic coordinates $\{\mathbf{x}_i^k\}$ and isotropic B-factors B_i^k. The observed electron density is denoted as $F_\mathrm{o} : \mathbb{R}^3 \to \mathbb{R}$. In order to restrict map comparisons around a region of interest, we use a spatial mask.

Forward model. To compute the crystallographic likelihood term, we first calculate the theoretical electron density map $F_\mathrm{c} : \mathbb{R}^3 \to \mathbb{R}$ for all structures in ensemble $\boldsymbol{\mathcal{X}}$. F_c can be calculated at any spatial coordinate location $\boldsymbol{\xi} \in \mathbb{R}^3$. For a single conformer $\mathbf{X}^k = (\mathbf{x}_1^k \dots \mathbf{x}_m^k)$, F_c is a sum over kernel-density estimates built from six spherical Gaussians centered at the symmetry operation (Brock et al., 2016) of every atom.

$$F_\mathrm{c}(\boldsymbol{\xi}; \mathbf{X}^k, \mathbf{a}) = \sum_{q=1}^{N_\mathrm{s}} \sum_{i=1}^{m} \sum_{j=1}^{6} a_{i,j} \cdot \left(\frac{4\pi}{b_{i,j} + B_i^k}\right)^{3/2} \exp\left(\frac{-4\pi^2}{b_{i,j} + B_i^k} \cdot \|(\mathbf{R}_q \mathbf{x}_i^k + \mathbf{t}_q) - \boldsymbol{\xi}\|_2^2\right) \tag{1}$$

Here, N_s is the number of symmetry operations (Brock et al., 2016), m is the number of atoms in the asymmetric unit of conformer $\mathbf{X}^k$, $\mathbf{R}_q \in SO(3)$ is the rotation matrix of symmetry operation q, $\mathbf{t}_q \in \mathbb{R}^3$ is the translation vector of the symmetry operation q, $a_{i,j}$ and $b_{i,j}$ are tabulated atomic form-factor coefficients for each heavy atom $\mathbf{x}_i^k$ (Prince, 2004), and B_i^k is the B-factor which represents the isotropic displacement parameter (Trueblood et al., 1996). Here, we consider the B-factor to be a bandwidth parameter.

B.2. Metrics

B.2.1. R-FACTORS

The R-factor (or crystallographic R-value) is a metric for assessing how well a crystal structure's model explains experimental X-ray diffraction data. The metric measures if observed structure factor amplitudes $|F_\mathrm{obs}|$ match the calculated structure factor amplitudes $|F_\mathrm{calc}|$. Formally,

$$R = \frac{\sum_\mathbf{h} \big||F_\mathrm{obs}(\mathbf{h})| - |F_\mathrm{calc}(\mathbf{h})|\big|}{\sum_\mathbf{h} |F_\mathrm{obs}(\mathbf{h})|}$$

Here, $\mathbf{h}$ denotes a Miller index triple (h, k, l) which denotes a discrete reflection in the reciprocal (Fourier) space. To avoid overfitting, R_free value is calculated in a similar manner, but using a small validation subset $\mathbf{T}$. Hence,

$$R_\mathrm{free} = \frac{\sum_{\mathbf{h}\in\mathbf{T}} \big||F_\mathrm{obs}(\mathbf{h})| - |F_\mathrm{calc}(\mathbf{h})|\big|}{\sum_{\mathbf{h}\in\mathbf{T}} |F_\mathrm{obs}(\mathbf{h})|}$$

It is important to note that $|F_\mathrm{obs}|$ and $|F_\mathrm{calc}|$ are structure factor amplitudes in Fourier domain. For each experiment, we report both the R and R_free values as calculated by `REFMAC5` after refinement from CCP4 (Agirre et al., 2023).

B.2.2. REAL-SPACE CORRELATION COEFFICIENT

The real-space correlation coefficient (RSCC) measures the alignment between a structural model and the observed experimental electron density map. RSCC is calculated directly in real space by comparing observed electron density $F_\mathrm{o}(\boldsymbol{\xi})$ with calculated electron density $F_\mathrm{c}(\boldsymbol{\xi})$ at coordinates $\boldsymbol{\xi} \in \mathbb{R}^3$. Formally, RSCC for a grid of voxels $\boldsymbol{\mathcal{G}}$ around an atom is given by,

$$\mathrm{RSCC} = \frac{\sum_{\boldsymbol{\xi}\in\boldsymbol{\mathcal{G}}} (F_\mathrm{o}(\boldsymbol{\xi}) - \overline{F_\mathrm{o}})(F_\mathrm{c}(\boldsymbol{\xi}) - \overline{F_\mathrm{c}})}{\sqrt{\sum_{\boldsymbol{\xi}\in\boldsymbol{\mathcal{G}}} (F_\mathrm{o}(\boldsymbol{\xi}) - \overline{F_\mathrm{o}})^2} \sqrt{\sum_{\boldsymbol{\xi}\in\boldsymbol{\mathcal{G}}} (F_\mathrm{c}(\boldsymbol{\xi}) - \overline{F_\mathrm{c}})^2}}$$

Where, $\overline{F_\mathrm{o}} = \mathbb{E}_{\boldsymbol{\mathcal{G}}}[F_\mathrm{o}]$, and $\overline{F_\mathrm{c}} = \mathbb{E}_{\boldsymbol{\mathcal{G}}}[F_\mathrm{c}]$. For our use case, we used `phenix.real_space_correlation` to measure the RSCC for each atom (and altloc) in our protein.

B.2.3. COSINE SIMILARITY

Cosine similarity helps us report a single global alignment score between the observed density F_o and the calculated density F_c over the β-turn region. First, we identify the set of spatial coordinates $\boldsymbol{\xi} \in \mathbb{R}^3$ by selecting all voxels from F_o that lie

within 5Å of atoms in the relevant residue range of the reference PDB structures. Formally,

$$\text{Cosine Similarity} = \frac{\sum_{\boldsymbol{\xi}} F_{\text{o}}(\boldsymbol{\xi}) F_{\text{c}}(\boldsymbol{\xi})}{\sqrt{\sum_{\boldsymbol{\xi}} F_{\text{o}}^2(\boldsymbol{\xi})} \cdot \sqrt{\sum_{\boldsymbol{\xi}} F_{\text{c}}^2(\boldsymbol{\xi})}}$$

where $F_{\text{o}}(\boldsymbol{\xi})$ and $F_{\text{c}}(\boldsymbol{\xi})$ denote vectors of density values at the selected voxel coordinates.

B.3. Dihedral Bimodal distribution of β-turns

For visualization, we define how close a relaxed structure $\mathbf{X}^m \in \boldsymbol{\mathcal{X}}$ is to each deposited alternate conformation by a single signed scalar $d^n_{\phi,\psi} \in [-1, +1]$. Let $\boldsymbol{\theta}_m$ be the member's dihedral vector at the β-turn and $\boldsymbol{\theta}_A, \boldsymbol{\theta}_B$ the corresponding dihedral vectors of the PDB deposited altloc A and altloc B conformations. We define

$$d_A = d_{\text{torus}}(\boldsymbol{\theta}_m, \boldsymbol{\theta}_A), \qquad d_B = d_{\text{torus}}(\boldsymbol{\theta}_m, \boldsymbol{\theta}_B).$$

The signed normalized distance is

$$d^n_{\phi,\psi} = \text{sign}\big(d_A - d_B\big) \left(1 - \frac{\min\big(d_A, d_B\big)}{\max\big(d_A, d_B\big)}\right) \tag{2}$$

By construction, $d^n_{\phi,\psi} = -1$ when the member's dihedral angles coincide with altloc A's backbone dihedral angles, $+1$ when it coincides with altloc B's backbone dihedral angles. Intermediate magnitudes interpolate monotonically between the altloc-anchored extremes. The resulting normalized dihedral distances are shown in Figure 3 and panels E of Figure 2. Across these examples, electron density–guided AlphaFold3 recovers ensemble members near both deposited β-turn conformations, whereas unguided AlphaFold3 and qFit typically populate only one of the two altloc-associated turn states.