Scalable Low-Energy Molecular Conformer Generation with Quantum Mechanical Accuracy

Anonymous Author(s)
Affiliation
Address
email

Abstract

Molecular geometry is crucial for biological activity and chemical reactivity; however, computational methods for generating 3D structures are limited by the vast scale of conformational space and the complexities of stereochemistry. Here we present an approach that combines an expansive dataset of molecular conformers with generative diffusion models to address this problem. We introduce ChEMBL3D, which contains over 250 million molecular geometries for 1.8 million drug-like compounds, optimized using AIMNet2 neural network potentials to a near-quantum mechanical accuracy with implicit solvent effects included. This dataset captures complex organic molecules in various protonation states and stereochemical configurations. We then developed LoQI, a stereochemistry-aware diffusion model that learns molecular geometry distributions directly from this data. Through graph augmentation, LoQI accurately generates molecular structures with targeted stereochemistry, representing a significant advance in modeling capabilities over previous generative methods. The model outperforms traditional approaches, achieving up to tenfold improvements in energy accuracy and effective recovery of optimal conformations. Benchmark tests on complex systems, including macrocycles and flexible molecules, as well as validation with crystal structures, show LoQI can perform low energy conformer search efficiently. The model code and dataset will be available before the workshop.

1 Introduction

Molecular conformation plays a fundamental role in determining the behavior and function of chemical compounds across diverse domains, from materials science to pharmacology. A prototypical example of the critical role that molecular conformation plays is in specificity of protein active site, where the three-dimensional (3D) geometry of a ligand must complement the shape and electrostatics of the binding pocket to enable favorable interactions, such as hydrogen bonds and hydrophobic contacts [1, 2]. Poorly fitting conformers are known to cause steric clashes or suboptimal binding, leading to a reduction in affinity and efficacy. Stereochemistry adds another layer of complexity, as many biological targets are chiral and can distinguish between stereoisomers, often resulting in significant differences in potency or toxicity. One exemplary case is thalidomide, where only one enantiomer is therapeutically safe. Despite advances in computational methods, generating accurate low-energy conformers at thousands-to-millions scale remains prohibitively expensive, limiting their use in high-throughput settings and creating a persistent bottleneck in structure-based molecular design [3].

Modern conformational analysis must account for both thermodynamic stability and kinetic accessibility. For example, studies have shown that many bioactive conformations lie close to the global energy minimum, emphasizing the importance of understanding the conformational landscape [4]. Reliable generation of low-energy diverse conformers anchored to realistic energy landscapes is

therefore critical for downstream applications such as docking, virtual screening, and binding-free energy estimation [5–8].

Classical approaches for conformer generation rely primarily on heuristic algorithms and force fields. Examples include single-conformer generators such as CORINA, as well as ensemble-based methods like ETKTG, OMEGA [5], ConfGen [9], Frog2 [10], and Multiconf-DOCK [11]. Despite their efficiency, these methods frequently struggle with molecules characterized by significant flexibility, including macrocycles and "beyond rule-of-five" chemotypes [12]. As a consequence, classical force-field-based methods often yield significant errors in energy rankings and incomplete coverage of conformational ensembles, particularly as molecular complexity increases [13, 14].

Recent advances in machine-learned interatomic potentials (MLIPs), such as AIMNet2 [15], ANI [16], MACE [17], and UMA [18] have shown great promise in addressing these limitations. MLIPs can offer near-density functional theory (DFT) accuracy at substantially reduced computational cost, thereby enabling rapid and reliable geometry optimization.

In parallel, deep generative modeling techniques have gained widespread adoption in computational chemistry [19]. These models have been mostly trained on the GEOM dataset [20], which was created with the CREST software [21] based on semi-empirical quantum methods. One of the models, Torsional Diffusion [22] introduced a diffusion-based framework that operates within the torsion space. Other group of methods directly model atomic coordinates, the group including GeoDiff [23], Molecular Conformer Fields (MCF)[24], Diffusion Molecular Transformer (DMT)[25], and AvgFlow [26]. However, all of these models completely disregard stereochemistry, making them poorly suited for chiral or stereochemically rich molecules.

In this work, we combine the strengths of MLIP energy evaluation with AIMNet2 and advanced generative models to develop a robust and widely applicable workflow for conformer generation. Specifically, we introduce **ChEMBL3D**, an extensive dataset containing over 250 million AIMNet2-optimized conformers for approximately 1.8 million unique molecules covering both charged and neutral states within an implicit solvent environment. We present **LoQI**(**Lo**w-energy **Q**M **I**nformed conformer generative model), a stereochemistry-aware, equivariant generative model trained end-to-end on ChEMBL3D. In extensive tests on molecules, LoQI outperformed conventional cheminformatics tools and generative models while explicitly encoding stereochemistry.

2 Methods

2.1 ChEMBL3D dataset

The ChEMBL3D dataset was curated from the ChEMBL v34 database [27]. Initial processing involved generating protomers in neutral and non-ionized states using OpenEye FixpKa, and subsequently merging these with the original structures. OpenEye's Flipper tool has been used to enumerate stereocenters with missing annotation. Reference topology structures, including hydrogens but without explicit 3D coordinates, were created with Open Babel [28]. The molecules were sorted by the number of atoms and the training conformers were generated using the OpenEye Omega classic [5], followed by structural optimization using the FIRE algorithm [29] with a variant of the AIMNet2 [15] neural network potential, trained to reproduce the DFT energies within a CPCM solvation model. Unlike previous datasets that rely on classical force fields or semi-empirical quantum computations, all conformations were optimized under implicit solvation conditions to near DFT accuracy using AIMNet2.

The conformer dataset initially comprised approximately 505 million conformers prior to post-processing. Downstream steps included topology validation, energy-based filtering, and deduplication, resulting in a final dataset of 252 million processed conformers for 2,417,498 unique structures (representing 1,786,517 distinct ChEMBL molecules). Our deduplication procedure is based on the relative changes in the distance matrix, following the assumption that only one enantiomer of each pair is retained, in accordance with the protocol described in the work of Pracht [30]. Aftere this filtering, the dataset contains 187 million conformers within a 6 kcal/mol energy window, 94 million within 2.5 kcal/mol, and 38 million within 1 kcal/mol. ChEMBL3D is orders of magnitude larger than previous 3D conformer datasets (e.g., GEOM), and to our knowledge it is the first to incor-

porate both an implicit solvent model and exhaustive stereochemistry/protomer enumeration at this scale. This makes it a uniquely comprehensive benchmark for generative models.

ChEMBL3D should be regarded as a collection of unique, relaxed conformers for ChEMBL molecules rather than a thermodynamical ensemble (e.g., GEOM-Drugs). Since a lot of recent generative molecular models employs only a limited number of conformations per molecule during training [25, 26, 31–34], we believe that providing a large-scale dataset of conformers with near QM accuracy will benefit the research community and facilitate further advances in generative modeling.

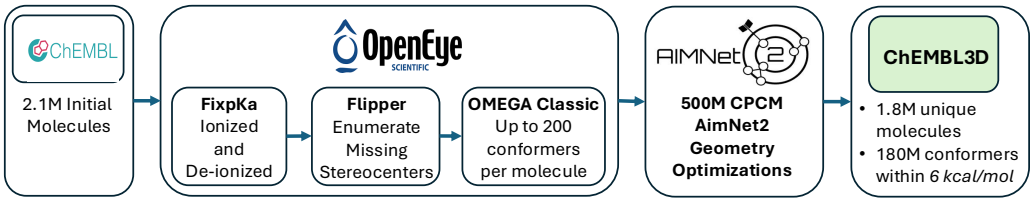

Figure 1: Overview of the ChEMBL3Ddataset construction workflow. Molecules from ChEMBL v34 were expanded via protomer generation (OpenEye FixpKa), conformer generation (OpenEye Omega Classic), and stereoisomer enumeration (OpenEye Flipper). Conformers were geometry-optimized with AIMNet2 and filtered for broken topologies and duplicates. The final ChEMBL3Ddataset contains approximately 1.8 million unique molecules and 180 millions conformers within a 6 kcal/mol energy window.

2.2 Model Architecture and Training Objective

LoQI utilizes the recently introduced Megalodon architecture from [33], which shows great results for 3D molecule generation. This architecture combines a Diffusion Transformer (DiT) [35] layer, which processes invariant features, with an equivariant graph neural network (EGNN) [36] that updates atomic coordinates. Specifically, each transformer module integrates discrete atomic features, such as atom types and charges, and bond features, such as bond types and geometric features (pairwise distances and dot products). These combined features are processed through DiT layer which represents multi-head attention with adaptive layer normalization. After DiT layer, the resulting atom and bond features serve as additional inputs for coordinate updates performed by a structure layer based on EGNN. This structure-update step employs distance-based EGNN updates enhanced by a cross-product term refining atomic coordinates.

In adapting the Megalodon framework, our approach simplifies the prediction task by exclusively focusing on atomic coordinates, using atom types, charges, and bonds solely as inputs. This design choice capitalizes on Megalodon's demonstrated strengths in accurately predicting molecular coordinates, thus streamlining computational efficiency without explicit modeling of discrete bond or atom-type outputs.

Our training procedure employs a diffusion-based approach consistent with few recent developments in generative modeling [31, 32]. Specifically, we utilize weighted cosine noise schedules within a standard denoising diffusion probabilistic model (DDPM). Comprehensive technical details and exact model configurations can be found in the original Megalodon paper [33] and implementation. Remarkably, we observe that for conformer generative task, only 25 diffusion steps are sufficient to reach peak performance.

3 Results

The LoQI model has been trained to generate lowest energy conformer from ChEMBL3D utilizing a diffusion training objective. Rather than navigating the space with predefined rules, LoQI learns directly to produce chemically valid, low-energy 3D geometries. This data-driven modeling strategy is not restricted by predefined rotatable bonds or fragment libraries, which crucially allows our model to efficiently handle the generation of complex molecular geometries. In the following sec-

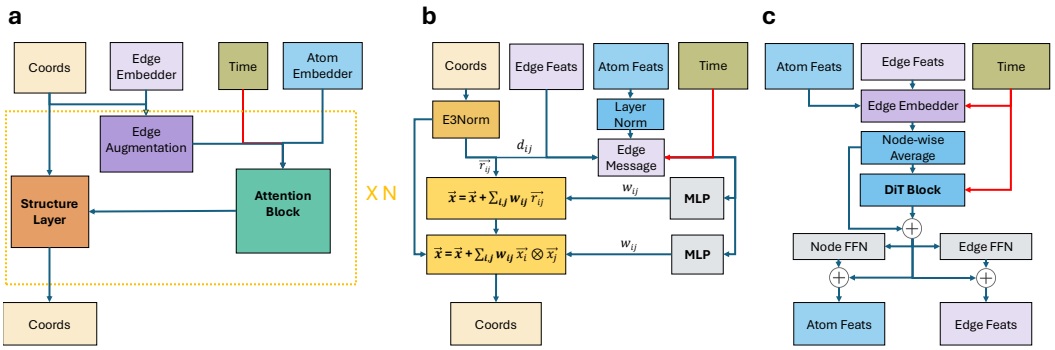

Figure 2: LoQI is a Megalodon-based adaptation [33] tailored for denoising diffusion over 25 steps, where the model iteratively predicts denoised atomic coordinates from progressively noised inputs. (a) The overall neural network architecture processes discrete atom features, bond features, diffusion timestep embeddings, and atomic coordinates; (b) A structure layer based on EGNN performs equivariant coordinate updates using equivariant and invariant features; (c) A DiT-style transformer block updates invariant atom and edge features conditioned on diffusion timestep.

tions, we benchmark LoQI against classical and generative baselines on multiple datasets to assess its performance and generalizability.

3.1 Hold-Out Set Performance

The generalization of LoQI has been rigorously evaluated using a newly constructed hold-out set of 37,075 molecules (including diverse isomeric and protonation states) sampled from ChEMBL3D. None of these molecules were used during training, allowing us to assess performance in a realistic setting, where the model must generate low-energy conformers without relying on memorized patterns or specific heuristics. The results are presented in Table 1.

LoQI consistently outperforms all baselines across key metrics. It achieves a median relative energy of 0.33 kcal/mol, significantly lower than both retrained generative baselines (e.g., 0.45 kcal/mol for DMT-B) and classical methods like RDKit and Omega (3.39 and 1.42 kcal/mol, respectively). Notably, LoQI finds conformers within 0.1 kcal/mol of the minimum of the dataset in 38.5% of molecules, nearly doubling all other methods, without pretraining on our dataset. In other words, generating and optimizing a single conformer using LoQI achieves, in 38.5% cases, results equivalent to optimizing up to 200 conformers produced by OpenEye Omega Classic. Furthermore, in 14.0% of the cases, the generated conformer has a lower energy compared to the reference conformer in ChEMBL3D, highlighting the ability of the model to perform low-energy conformer search.

The generative model has been trained on the AIMNet2 optimized conformers; therefore, it is worth confirming that the conformers with lower energy than those in the dataset are not an artifact or errors of the MLIP model. We selected 88 examples from the test set in which the model, using single-shot generation, identified conformers at least $1 kcal/mol$ lower in energy compared to the dataset reference. To validate these results, we performed density functional theory (DFT) calculations with a CPCM implicit solvent model. Of the 77 molecules that successfully completed DFT geometry optimization, only one had a higher energy (0.8 kcal/mol) compared to the reference. For 56 molecules, DFT optimization confirmed an energy improvement of at least 0.5 kcal/mol, validating the accuracy of our generative model (see Figure 3 for examples and plot). These results demonstrate that LoQI can efficiently search conformers with DFT-level accuracy at computational costs significantly lower than direct DFT-level geometry optimization.

LoQI excels not only on conformational energy metrics but also structurally. Geometry optimization converged of the generated conformers converged in nearly every case, perserving the molecular topology defined by the input structure. Additionally, the median relaxation energy (ΔE_{relax})—the energy adjustment required after post-generation geometry optimization—is minimal ($3.48 kcal/mol$), indicating that LoQI produces conformations remarkably close to true minima, requiring little subsequent refinement. This reduces the need for costly refinement steps, making the approach highly suitable for applications requiring high-throughput, structure-based analysis.

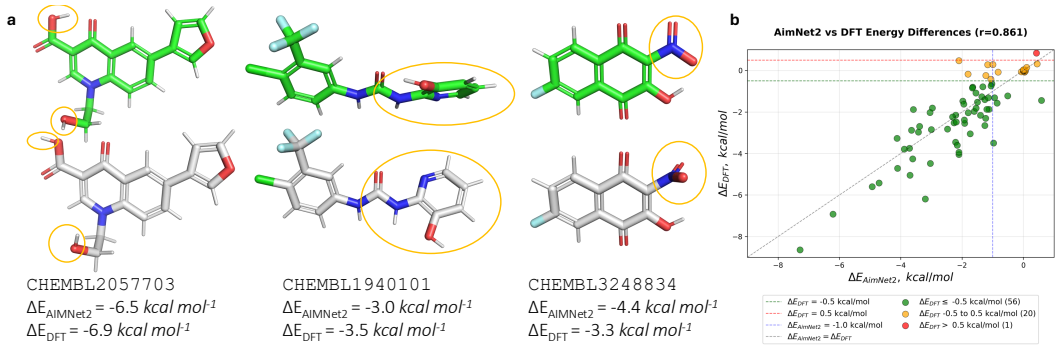

Figure 3: Comparison of AIMNet2 and DFT energies for molecules where LoQI identified a lower-energy conformer via single-shot generation. (a) Three representative examples illustrating structural differences between the ChEMBL3D reference conformers (gray) and the generated conformers (green), along with corresponding energy differences computed using AIMNet2 and DFT. (b) Correlation between energy differences computed using AIMNet2 and those obtained from DFT, both evaluated on DFT-optimized structures across all 77 successfully optimized molecules. The strong agreement highlights the reliability of AIMNet2 for accurately guiding low-energy conformer generation.

Classical methods, while highly reliable in preserving topology (e.g., RDKit at 99.5%), struggle to generate accurate low-energy conformations. RDKit finds near-optimal conformers in only 10.4% of molecules and improves upon the dataset minimum in just 3.1%. Omega performs similarly, reflecting the limitations of heuristic-driven sampling in covering the true conformational landscape.

Among the generative model baselines, retrained DMT-B is the most competitive, achieving 33.4% within 0.1 kcal/mol of the minimum. That model exhibits a substantially higher relaxation energy (43.42 kcal/mol), indicating its generated conformers substantially deviate from true local minima and revealing a fundamental limitation in learning the underlying energy landscape. Models without retraining perform substantially worse, underscoring the value of both the method and the training data.

Together, these results support a broader conclusion: LoQI represents a practical solution to the long-standing problem of efficient low-energy conformer search without relying on heuristics, predefined rotatable bonds, or fragment templates.

Table 1: Performance comparison of low-energy conformer generation methods on the ChEMBL3D hold-out set. Metrics include optimization success, topological fidelity, and energy-based accuracy.

Method	Median $\Delta E^{\text{opt}}_{\text{relative}}$ [1] (kcal/mol)	% $\Delta E^{\text{opt}}_{\text{relative}}$ [1] < 0.1 kcal/mol	% $\Delta E^{\text{opt}}_{\text{relative}}$ [1] < -0.1 kcal/mol	Minimization Converged, %	Topology Preserved, %	Median ΔE_{relax} [2] (kcal/mol)
RDKit	3.39	10.4	3.1	99.5	99.5	64.50
OEOmega	1.42	15.7	13.0	99.8	99.8	13.19
TorsDiff	2.62	15.3	7.7	99.0	97.4	63.09
DMT-B	1.81	19.9	11.5	99.8	99.8	82.70
DMT-L	1.72	20.8	12.4	99.9	99.8	86.43
TorsDiff[3]	4.66	7.7	2.7	94.6	87.3	92.02
DMT-B[3]	0.45	33.4	12.5	99.5	95.7	43.42
LoQI	**0.33**	**38.5**	**14.0**	**100**	**99.9**	**3.48**

3.2 LoQI is a Stereochemistry-Aware Model

A significant limitation of current all-atom generative models operating directly on molecular graphs is the lack of explicit stereochemical encoding. Unlike torsion-based methods such as Torsional Diffusion, which theoretically preserve stereochemistry by modeling conformers in torsional space,

Cartesian coordinate-based models rely only on relative atomic arrangements and element types. As a result, they are unable to recover stereochemical configurations such as E/Z isomerism and R/S chirality, which makes these tools less practical.

We address this challenge by introducing a simple yet effective graph augmentation that encodes stereochemical information directly into the input graph (see Figure 4). For E/Z stereochemistry, we add four auxiliary edges between neighbors of each double bond with possible E/Z isomerism: two "joint" Z-bonds connecting adjacent substituents and two "diagonal" E-bonds connecting across the double bond. This distinguishes stereoisomers based on relative substituent positioning. For R/S stereocenters, the lowest-priority (according to Cahn-Ingold-Prelog rules) atom connects to the remaining three via undirected edges, while the other three form a directed cycle consistent with the stereochemical configuration. These augmentations, combined with directional message passing, provide sufficient context for the neural networks to drive stereochemistry-aware 3D structure generation.

The stereochemical augmentation is both architecture-agnostic and easily generalizable to other stereochemical types, such as axial or planar chirality. It effectively bridges a gap between the classical cheminformatics tools that inherently handle stereochemistry and modern generative approaches.

Table 2 demonstrates the effectiveness of this strategy. LoQI is capable to preserve defined stereochemistry, with 0.956 R/S and 0.992 E/Z accuracy, on par with traditional methods like RDKit and OpenEye Omega, and substantially outperforming all previous generative models. Notably, even retrained diffusion models such as Torsional Diffusion struggle with E/Z stereochemistry, despite our best efforts to enforce correct initialization and labeling. This suggests that torsional generative models, while inherently more structured, still lack reliable mechanisms for the enforcement of stereochemistry during sampling.

The proposed stereochemical graph encoding approach not only enhances LoQI's practical applicability but also represents, to our knowledge, the first successful integration of explicit stereochemical encoding into an all-atom diffusion-based generative model. This advancement significantly expands the scope of generative models for reliable and accurate application across diverse chemical and pharmaceutical use cases.

3.3 Impact of Rotatable Bonds

A critical challenge in molecular conformer generation is the exponential complexity introduced by increasing number of rotatable bonds. To specifically assess performance under this challenge, we created an additional test subset, which we refer to as ChEMBL3D-XL. This subset consists of approximately 300 molecules per number of rotable bonds selected to span up to 25. For each molecule, we generated up to 100 conformers using LoQI, RDKit ETKDG, and OE Omega Classic, followed by optimization with the same neural network potential used during ChEMBL3D creation. We then evaluated the probability of recovering the reference minimum conformer as a function of the number of trials. We define success for this performance assessment as at least one attempt producing a conformer within within 0.1 kcal/mol of the value in the dataset. We also look at the fraction of molecules for which we were able to find better conformation with 0.5 kcal/mol margin.

The results highlight the significant gain in performance of the LoQI model, especially as molecular complexity increases with the number of rotatable bonds. As shown on Figure 5, the model's accuracy improves with increasing numbers of rotatable bonds leading to an order of magnitude higher success rate for molecules with large number of rotable bonds compared to OE Omega and RDKit.

3.4 Fast and Accurate Estimation of Energy Differences Between Bound and Unbound Molecular States

To demonstrate the utility of LoQI in estimating energy differences between ligand-bound and unbound states, we evaluated it on the Platinum Diverse dataset [4], a widely used benchmark containing 2,879 ligand crystal structures. Although this dataset was originally proposed for conformer generation benchmarking, it consists of crystal-bound conformers that often deviate significantly from the lowest-energy gas-phase or solution-phase geometries due to intermolecular interactions and crystal packing effects [4, 37]. Since LoQI is trained to predict low-energy conformers of iso-

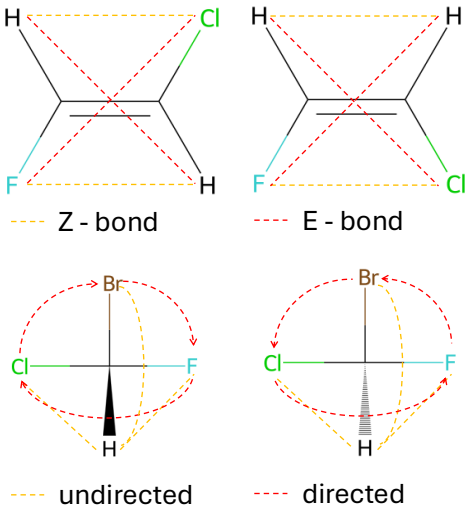

Figure 4: Stereochemistry-aware graph augmentations. **Top:** For E/Z, four auxiliary bonds are added—two *Z-bonds* (joint) and two *E-bonds* (diagonal)—between neighbors of a double bond. **Bottom:** For R/S, the lowest-priority atom connects to the others via undirected edges; the rest form a directed cycle reflecting stereochemistry.

Table 2: Stereochemistry accuracy of conformer generation methods on the ChEMBL3D hold-out set. Reported values indicate the fraction of molecules with correctly generated configurations for R/S and E/Z stereocenters.

Method	**R/S** Accuracy[1]	**E/Z** Accuracy[2]
RDKit	0.993	0.993
OEOmega	0.993	0.985
TorsDiff	0.993	0.734
DMT-B	0.377	0.706
DMT-L	0.385	0.740
TorsDiff[3]	0.993	0.502
DMT-B[3]	0.793	0.990
LoQI	0.956	0.992

[1] R/S accuracy: fraction of molecules with correctly reproduced absolute stereochemistry at tetrahedral centers.
[2] E/Z accuracy: fraction of molecules with correct stereochemistry at double bonds.
[3] Indicates models retrained using the same protocol as LoQI on ChEMBL3D.

lated molecules in CPCM implicit solvent, we do not report standard conformer generation metrics on this set (see Appendix A for further discussion). Instead, we use the energy differences between experimentally observed bound conformers and computationally predicted low-energy structures as a rigorous test of the model performance.

For each ligand, we generated conformational ensembles comprising 500 structures using LoQI and RDKit's ETKDG algorithm. We identified the lowest-energy conformer from each ensemble as the reference minimum. The observed energy differences showed a median of $1.84, \mathrm{kcal}, \mathrm{mol}^{-1}$, consistent with the typical magnitude of the change in energy from the crystal packing (Figure 6a). Several structures exhibited large deviations ($> 10, \mathrm{kcal}, \mathrm{mol}^{-1}$), reflecting substantial structural differences between the crystal-bound and isolated conformational preferences.

Figure 6b illustrates the accuracy tradeoff between the number of sampled conformers and the estimated energy differences between crystal-bound and isolated minima. Notably, chemical accuracy (mean absolute error $< 1, \mathrm{kcal}, \mathrm{mol}^{-1}$) is achieved by LoQI with as few as five sampled conformers, whereas RDKit's ETKDG fails to reach this threshold even after generating 100 conformers. Quantitatively, expanding the conformer ensemble from one to 99 conformers reduces mean absolute prediction error from 2.29 to $0.14, \mathrm{kcal}, \mathrm{mol}^{-1}$ for LoQI, compared to a reduction from 5.24 to $1.31, \mathrm{kcal}, \mathrm{mol}^{-1}$ for RDKit ETKDG. This contrast highlights the capability of generative models to rapidly and accurately capture nuanced conformational energy landscapes that would be difficult to achieve with the rule-based heuristics of traditional methods.

4 Conclusions

In this study, we introduced ChEMBL3D, the largest publicly available dataset of high-quality, DFT-level molecular conformations, comprising over 250 million optimized conformers for 1.8 million molecules with diverse protonation states and stereochemical configurations. Leveraging this dataset, we trained LoQI, a stereochemistry-aware diffusion model tailored explicitly for accurate and computationally efficient low-energy conformer generation. LoQI represents the first generative

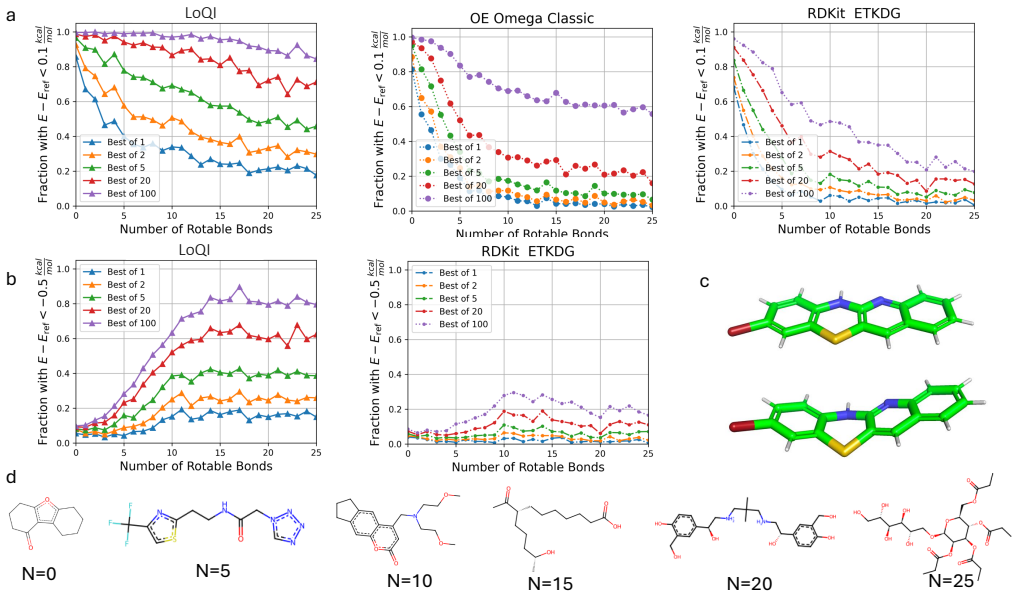

Figure 5: (a) Fraction of molecules for which the LoQI finds the minimum-energy conformer within a specified number of trials (1, 2, 5, 20, 100), as a function of the number of rotatable bonds. (b) Fraction of molecules where LoQI identifies a conformer lower in energy than the ChEMBL3D dataset reference conformer by at least 0.5 kcal/mol. (c) Example of a molecule with zero rotatable bonds for which LoQI finds a lower-energy conformer than those in ChEMBL3D. (d) Representative molecules with different numbers of rotatable bonds, as defined in RDKit.

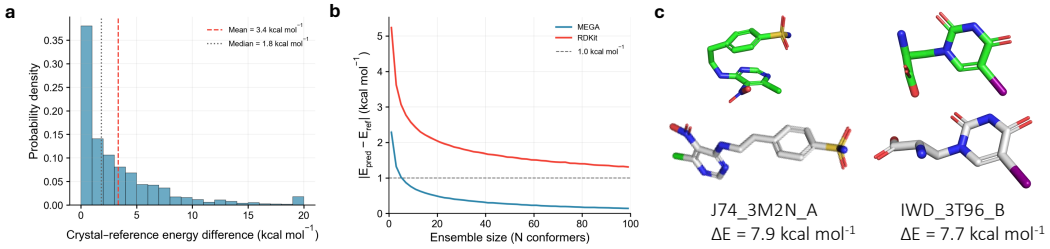

Figure 6: **Neural-network-driven estimation of conformational energy differences on the Platinum Diverse dataset.** (a) Distribution of energy differences between crystal-bound and isolated minimal-energy conformers, highlighting significant energy deviations due to crystal packing constraints. (b) Mean absolute error (MAE) in energy difference estimation versus conformer ensemble size. LoQI rapidly achieves chemical accuracy compared to RDKit's ETKDG. (c) Representative crystal structures demonstrating geometric deviations between crystal-bound (grey) and isolated minimal-energy conformations (green). Crystal-bound structures are consistently more open and strained compared to minima.

model trained on complex molecules sourced from ChEMBL. Critically, ground truth conformers are optimized under realistic conditions with implicit solvent.

The LoQI model overcomes the challenge of conformational sampling and the need for exhaustive multi-conformer searches by directly generating low-energy 3D geometries in a single shot, bypassing the heuristic-driven exploration of vast conformational space. This enables it to consistently produce conformers within 0.1 kcal/mol of the minimum-energy structure for a large fraction of molecules while requiring minimal post-generation optimization. As a result, LoQI effectively integrates the conformational search and low-energy selection steps into a single generative process, eliminating the computational cost of enumerating and optimizing large conformer ensembles.

To facilitate widespread adoption, reproducibility, and community-driven advancements, we openly release both the ChEMBL3D dataset and the LoQI model. The model's code, references to datasets, and other relevant resources will be available at the time of workshop. We believe this resource will significantly benefit researchers and practitioners across various fields, inspiring new methodological developments and fostering deeper collaboration within the computational chemistry and artificial intelligence communities.

References

[1] Nicolas Foloppe and I-Jen Chen. Conformational sampling and energetics of drug-like molecules. *Current medicinal chemistry*, 16(26):3381–3413, 2009.

[2] Richard E Taylor, Yue Chen, Alicia Beatty, David C Myles, and Yiqing Zhou. Conformation-activity relationships in polyketide natural products: A new perspective on the rational design of epothilone analogues. *Journal of the American Chemical Society*, 125(1):26–27, 2003.

[3] Ron Unger and John Moult. Finding the lowest free energy conformation of a protein is an np-hard problem: proof and implications. *Bulletin of mathematical biology*, 55:1183–1198, 1993.

[4] Nils-Ole Friedrich, Agnes Meyder, Christina de Bruyn Kops, Kai Sommer, Florian Flachsenberg, Matthias Rarey, and Johannes Kirchmair. High-quality dataset of protein-bound ligand conformations and its application to benchmarking conformer ensemble generators. *Journal of chemical information and modeling*, 57(3):529–539, 2017.

[5] Paul CD Hawkins, A Geoffrey Skillman, Gregory L Warren, Benjamin A Ellingson, and Matthew T Stahl. Conformer generation with omega: algorithm and validation using high quality structures from the protein databank and cambridge structural database. *Journal of chemical information and modeling*, 50(4):572–584, 2010.

[6] Sereina Riniker and Gregory A Landrum. Better informed distance geometry: using what we know to improve conformation generation. *Journal of chemical information and modeling*, 55 (12):2562–2574, 2015.

[7] Zhen Liu, Tetiana Zubatiuk, Adrian Roitberg, and Olexandr Isayev. Auto3d: Automatic generation of the low-energy 3d structures with ani neural network potentials. *Journal of Chemical Information and Modeling*, 62(22):5373–5382, 2022.

[8] Omer Tayfuroglu, Irem N Zengin, M Serdar Koca, and Abdulkadir Kocak. Deepconf: Leveraging ani-ml potentials for exploring local minima with application to bioactive conformations. *Journal of Chemical Information and Modeling*, 65(6):2818–2833, 2025.

[9] K Shawn Watts, Pranav Dalal, Robert B Murphy, Woody Sherman, Rich A Friesner, and John C Shelley. Confgen: a conformational search method for efficient generation of bioactive conformers. *Journal of chemical information and modeling*, 50(4):534–546, 2010.

[10] Maria A Miteva, Frederic Guyon, and Pierre Tuffi¿ 1/2ry. Frog2: Efficient 3d conformation ensemble generator for small compounds. *Nucleic acids research*, 38(suppl_2):W622–W627, 2010.

[11] Nicolas Sauton, David Lagorce, Bruno O Villoutreix, and Maria A Miteva. Ms-dock: accurate multiple conformation generator and rigid docking protocol for multi-step virtual ligand screening. *BMC bioinformatics*, 9:1–12, 2008.

[12] I-Jen Chen and Nicolas Foloppe. Tackling the conformational sampling of larger flexible compounds and macrocycles in pharmacology and drug discovery. *Bioorganic & medicinal chemistry*, 21(24):7898–7920, 2013.

[13] Dakota Folmsbee and Geoffrey Hutchison. Assessing conformer energies using electronic structure and machine learning methods. *International Journal of Quantum Chemistry*, 121 (1):e26381, 2021.

[14] Ilana Y Kanal, John A Keith, and Geoffrey R Hutchison. A sobering assessment of small-molecule force field methods for low energy conformer predictions. *International Journal of Quantum Chemistry*, 118(5):e25512, 2018.

[15] Dylan Anstine, Roman Zubatyuk, and Olexandr Isayev. Aimnet2: a neural network potential to meet your neutral, charged, organic, and elemental-organic needs. 2024.

[16] Justin S Smith, Olexandr Isayev, and Adrian E Roitberg. Ani-1: an extensible neural network potential with dft accuracy at force field computational cost. *Chemical science*, 8(4):3192–3203, 2017.

[17] Ilyes Batatia, David P Kovacs, Gregor Simm, Christoph Ortner, and Gábor Csányi. Mace: Higher order equivariant message passing neural networks for fast and accurate force fields. *Advances in neural information processing systems*, 35:11423–11436, 2022.

[18] Brandon M Wood, Misko Dzamba, Xiang Fu, Meng Gao, Muhammed Shuaibi, Luis Barroso-Luque, Kareem Abdelmaqsoud, Vahe Gharakhanyan, John R Kitchin, Daniel S Levine, et al. Uma: A family of universal models for atoms. *arXiv preprint arXiv:2506.23971*, 2025.

[19] Dylan M Anstine and Olexandr Isayev. Generative models as an emerging paradigm in the chemical sciences. *Journal of the American Chemical Society*, 145(16):8736–8750, 2023.

[20] Simon Axelrod and Rafael Gomez-Bombarelli. Geom, energy-annotated molecular conformations for property prediction and molecular generation. *Scientific Data*, 9(1):185, 2022.

[21] Philipp Pracht, Stefan Grimme, Christoph Bannwarth, Fabian Bohle, Sebastian Ehlert, Gereon Feldmann, Johannes Gorges, Marcel Müller, Tim Neudecker, Christoph Plett, et al. Crest—a program for the exploration of low-energy molecular chemical space. *The Journal of Chemical Physics*, 160(11), 2024.

[22] Bowen Jing, Gabriele Corso, Jeffrey Chang, Regina Barzilay, and Tommi Jaakkola. Torsional diffusion for molecular conformer generation. *Advances in neural information processing systems*, 35:24240–24253, 2022.

[23] Minkai Xu, Lantao Yu, Yang Song, Chence Shi, Stefano Ermon, and Jian Tang. Geodiff: A geometric diffusion model for molecular conformation generation. *arXiv preprint arXiv:2203.02923*, 2022.

[24] Yuyang Wang, Ahmed A Elhag, Navdeep Jaitly, Joshua M Susskind, and Miguel Angel Bautista. Swallowing the bitter pill: Simplified scalable conformer generation. *arXiv preprint arXiv:2311.17932*, 2023.

[25] Zhiyuan Liu, Yanchen Luo, Han Huang, Enzhi Zhang, Sihang Li, Junfeng Fang, Yaorui Shi, Xiang Wang, Kenji Kawaguchi, and Tat-Seng Chua. Next-mol: 3d diffusion meets 1d language modeling for 3d molecule generation. *arXiv preprint arXiv:2502.12638*, 2025.

[26] Zhonglin Cao, Mario Geiger, Allan Dos Santos Costa, Danny Reidenbach, Karsten Kreis, Tomas Geffner, Franco Pellegrini, Guoqing Zhou, and Emine Kucukbenli. Efficient molecular conformer generation with so (3)-averaged flow matching and reflow. *arXiv preprint arXiv:2507.09785*, 2025.

[27] Barbara Zdrazil, Eloy Felix, Fiona Hunter, Emma J Manners, James Blackshaw, Sybilla Corbett, Marleen de Veij, Harris Ioannidis, David Mendez Lopez, Juan F Mosquera, et al. The chembl database in 2023: a drug discovery platform spanning multiple bioactivity data types and time periods. *Nucleic acids research*, 52(D1):D1180–D1192, 2024.

[28] Noel M O'Boyle, Michael Banck, Craig A James, Chris Morley, Tim Vandermeersch, and Geoffrey R Hutchison. Open babel: An open chemical toolbox. *Journal of cheminformatics*, 3:1–14, 2011.

[29] Erik Bitzek, Pekka Koskinen, Franz Gähler, Michael Moseler, and Peter Gumbsch. Structural relaxation made simple. *Physical review letters*, 97(17):170201, 2006.

[30] Philipp Pracht. Conformational pruning via the permutation invariant root-mean-square deviation of atomic positions. *Journal of Chemical Information and Modeling*, 65(9):4501–4511, 2025.

[31] Tuan Le, Julian Cremer, Frank Noe, Djork-Arné Clevert, and Kristof Schütt. Navigating the design space of equivariant diffusion-based generative models for de novo 3d molecule generation. *arXiv preprint arXiv:2309.17296*, 2023.

[32] Clement Vignac, Nagham Osman, Laura Toni, and Pascal Frossard. Midi: Mixed graph and 3d denoising diffusion for molecule generation. In *Joint European Conference on Machine Learning and Knowledge Discovery in Databases*, pages 560–576. Springer, 2023.

[33] Danny Reidenbach, Filipp Nikitin, Olexandr Isayev, and Saee Paliwal. Applications of modular co-design for de novo 3d molecule generation. *arXiv preprint arXiv:2505.18392*, 2025.

[34] Julian Cremer, Tuan Le, Frank Noé, Djork-Arné Clevert, and Kristof T Schütt. Pilot: equivariant diffusion for pocket-conditioned de novo ligand generation with multi-objective guidance via importance sampling. *Chemical Science*, 15(36):14954–14967, 2024.

[35] Junlong Li, Yiheng Xu, Tengchao Lv, Lei Cui, Cha Zhang, and Furu Wei. Dit: Self-supervised pre-training for document image transformer. In *Proceedings of the 30th ACM international conference on multimedia*, pages 3530–3539, 2022.

[36] Vıctor Garcia Satorras, Emiel Hoogeboom, and Max Welling. E (n) equivariant graph neural networks. In *International conference on machine learning*, pages 9323–9332. PMLR, 2021.

[37] Yu Hamakawa and Tomoyuki Miyao. Understanding conformation importance in data-driven property prediction models. *Journal of Chemical Information and Modeling*, 65(7):3388–3404, 2025.

[38] Jan Rezac, Daniel Bím, Ondrej Gutten, and Lubomir Rulisek. Toward accurate conformational energies of smaller peptides and medium-sized macrocycles: Mpconf196 benchmark energy data set. *Journal of chemical theory and computation*, 14(3):1254–1266, 2018.

A Platinum-Diverse Conformer Search Benchmark

The Platinum Diverse dataset, comprising 2,879 experimentally determined ligand crystal structures [4], is a widely accepted benchmark for evaluating small-molecule conformer generation methods. However, crystal-packing effects and protein-ligand interactions frequently cause bound ligand conformers to deviate substantially from their gas-phase or implicit-solvent energy minima. Our model, trained exclusively on *in silico* conformers—specifically, ChEMBL molecules optimized using the CPCM–AIMNet2 neural network potential to approximate implicit-solvent DFT conditions—is inherently biased toward lower-energy, solution-phase geometries. As such, we anticipated partial but not complete transferability to the more energetically diverse and structurally constrained conformations found in the Platinum Diverse dataset.

Despite these inherent limitations, LoQI consistently yields the lowest mean and median RMSD values across all ensemble sizes evaluated (Table 3). Notably, the model achieves a median RMSD of $0.45\mathring{A}$ at an ensemble size of $N = 250$, surpassing RDKit ($0.52\mathring{A}$) and ETKDG ($0.54\mathring{A}$), and matching performance of computationally intensive genetic-algorithm methods.

These findings demonstrate that a generative model trained exclusively on low-energy solution-phase conformations effectively generalizes to the distinct conformational landscape of protein-bound ligands. More broadly, this benchmark underscores the potential of data-driven 3D generative models to overcome the intrinsic complexity of conformer searches, even under conditions significantly divergent from the training distribution.

Table 3: Arithmetic mean and median RMSD (Å) for the Platinum Diverse dataset. Values in **bold** indicate the best (lowest) performance in each column; values statistically indistinguishable from the best (within ± 0.02 Å) are underlined. Bold and underlined entries represent the LoQI performance when at least one other method is within this tolerance range.

Method	**10**		**50**		**250**	
	Mean	Median	Mean	Median	Mean	Median
Balloon DG	1.10	0.97	1.00	0.86	0.92	0.77
Balloon GA	1.22	1.10	0.90	0.80	0.72	0.63
RDKit	1.00	0.89	0.77	0.64	0.63	0.52
ETKDG	0.98	0.87	0.77	0.66	0.63	0.54
Frog2	1.18	1.19	0.93	0.85	0.75	0.65
Multiconf-DOCK	0.99	0.89	0.84	0.72	0.80	0.69
LoQI	**0.95** $\pm$ 0.01	**0.88**	**0.74** $\pm$ 0.01	**0.65**	**0.58** $\pm$ 0.01	**0.45**

B Performance on Macrocycles and Complex Stereoisomers

Macrocyclic molecules are a particularly demanding class of chemical compounds for conformer generation due to their large ring structures, inherent flexibility, and complex conformational landscapes. We selected six representative macrocycles from the MPCONF196 benchmark dataset [38] to serve as a targeted assessment of LoQI's accuracy and transferability. For each macrocyclic compound, ensembles of 100 conformers were generated, and the lowest-energy conformer most closely resembling the high-level quantum-chemical reference structures (CCSD or MP2 optimized) was identified.

The results, illustrated in Figure 7, demonstrate that LoQI consistently generates accurate macrocyclic conformers closely aligned with reference structures, achieving root-mean-square deviation (RMSD) ranging from approximately 1.0 Å (for smaller macrocycles such as POXTR) to about 1.9 Å (for larger, more flexible cases like SANGLI).

Additionally, we evaluated the performance of our method on taxol, a molecule recognized for its challenging stereochemistry and structural complexity. Figure 7b compares the generated taxol conformer (in green) against its experimental crystal structure obtained from the Cambridge Structural Database (CSD) (in grey). The close alignment, 1.2 Å, underscores our model's robustness in accurately reproducing intricate stereochemical details. These results collectively highlight LoQI's generalizability to complex molecules with numerous stereocenters.

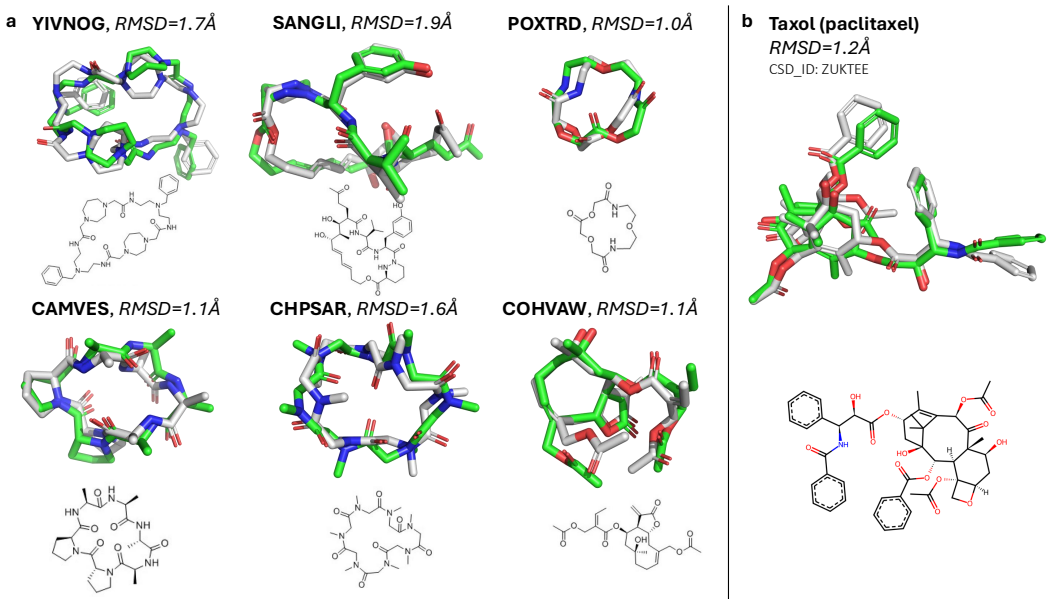

Figure 7: **Performance on challenging macrocycles and complex stereoisomers.** (a) Examples of macrocyclic conformers generated for selected molecules from the MPCONF196 dataset. Generated conformers (green) are overlaid with their reference structures (grey). RMSD values highlight strong structural agreement. (b) Example of taxol, a molecule known for challenging stereochemistry, comparing generated conformer (green) with its experimental crystal structure from CSD (grey).