INERTIAL FRAGMENT MATCHING: SCALABLE SHAPE-CONDITIONED MOLECULAR GENERATION VIA DESCRIPTOR DECOMPOSITION

Anonymous authors
Paper under double-blind review

ABSTRACT

Prior work has demonstrated that three-dimensional molecular structures can be generated effectively using an equivariant diffusion model conditioned on the ordered principal moments of inertia as a compact and interpretable shape descriptor. While guidance of the generation process by this coarse-grained representation enables high shape similarity for relatively small molecules, its limitations become increasingly apparent for larger molecular structures. We introduce Inertial Fragment Matching (IFM), an inference framework that exploits the additivity and reference-point translatability of the moment-of-inertia (MOI) tensor while using its ordered principal moments as the model-facing conditioning signal. IFM enables fragment-based molecular generation while preserving global shape consistency, substantially improving generation quality and scalability for generation of larger molecules. IFM achieves an average shape similarity of 0.698, corresponding to a 27.8% improvement over the conventional generation workflow. Under fixed-fragment generation conditions, IFM further attains a fragment integration success rate of 0.638, a 73.8% improvement over conventional inpainting-based approaches. Additionally, the impact of IFM on molecular quality, structural fidelity, and inference efficiency is evaluated. We present a proof of concept for context-window-independent molecular generation by combining inertial fragment matching with chain merging, demonstrating the potential to generate linear molecular chains of arbitrary size. Finally we formulate a general transport-additive fragment matching beyond the molecule generation task.

1 INTRODUCTION

We consider shape-conditioned molecule generation: given a target three-dimensional shape representation $\boldsymbol{s}^\star$, generate a molecule $M = (\mathcal{G}, \boldsymbol{X})$, consisting of a molecular graph $\mathcal{G}$ and atomic coordinates $\boldsymbol{X}$, such that M is chemically valid and its shape is close to $\boldsymbol{s}^\star$. The $\boldsymbol{s}^\star$ may be derived from a reference molecule or specified independently as an arbitrary shape, while the generated molecule need not preserve the reference graph, composition, or atom identities. This formulation reflects a reverse-design setting in which molecular structure is generated to satisfy a prescribed geometric constraint.

Such tasks naturally arise in structure-based molecular design, conformational analysis, reaction and materials design, where molecular geometry can constrain or promote interactions, accessibility, and reactivity (Gebauer et al., 2022). Shape-directed generation can therefore reduce the need for manual trial-and-error molecule design by enabling automated exploration of chemically valid candidates under explicit geometric constraints.

Prior work has shown that the principal components of the moment of inertia (MOI) tensor provide a compact, interpretable, and computationally inexpensive descriptor for shape-guided molecular generation (Sapegin et al., 2025). Despite its low dimensionality, conditioning an equivariant generative model on this descriptor can yield strong shape similarity for small molecules. Moreover, because MOI descriptors are defined for general three-dimensional mass distributions, they provide a flexible conditioning signal that is not restricted to molecular-only reference geometries. However, the limitations of this coarse-grained descriptor become increasingly apparent as molecular

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.

size increases. In particular, a drop in average shape similarity for molecules containing more than approximately 20 heavy atoms is observed in Sapegin et al. (2025). This degradation is expected: the MOI descriptor captures only global inertial properties, making it too coarse to finely resolve the geometric detail of larger molecular structures.

At the same time, the physical definition of MOI results in useful properties that can be exploited beyond direct whole-molecule conditioning. In particular, MOI tensors are additive when fragment contributions are expressed with respect to a common reference point, and they can be translated between reference points using the parallel-axis relationship (Abdulghany, 2017). These properties suggest a fragment-based strategy: decompose a complex target shape into smaller inertial contributions, express each fragment descriptor in its local principal inertial frame, generate fragments with high local shape similarity, transform them back into a global frame and then merge them while preserving consistency with the global MOI context.

This perspective motivates Inertial Fragment Matching (IFM), an inference framework designed to extend MOI-guided molecular generation beyond the regime where whole-molecule conditioning remains sufficiently descriptive. Rather than relying on a single coarse descriptor to guide the generation of an entire large molecule, IFM uses MOI additivity and reference-point translatability to distribute the shape constraint across fragments. This allows the model to operate on smaller, more tractable substructures while maintaining global shape consistency after assembly. As molecular size increases, the mapping from global MOI descriptors to molecular structures becomes increasingly many-to-one. IFM reduces this ambiguity by conditioning generation on locally informative fragment-level descriptors. IFM also provides a natural framework for fragment-constrained generation tasks, including fragment growth and linker design (Hadfield et al., 2022; Imrie et al., 2020; Igashov et al., 2024). We hypothesize that in contrast to inpainting-based methods (Schneuing et al., 2022), it introduces smaller perturbations to the generative distribution, enabling more consistent handling of structural constraints.

A useful way to interpret IFM is as an instance of a more general principle: fragment-based generation becomes descriptor-consistent whenever the conditioning descriptor is not only informative, but also compatible with the decomposition and recomposition operations of the object class. MOI is effective in our setting because it is a transport-additive descriptor: fragment-level inertial descriptors can be re-expressed under rigid transformations and summed in a common reference frame to recover the global descriptor. This property allows the global shape constraint to be factorized into local constraints, making it possible to generate smaller substructures independently while retaining a mathematically controlled relation to the target whole-object descriptor.

This observation clarifies that the key ingredient of IFM is not molecule-specific chemistry alone, nor the use of a generic low-dimensional embedding, but the existence of a descriptor with a known algebra of decomposition, transport, and recomposition. Moment-of-inertia tensors provide a physically interpretable example of such a descriptor. The same abstraction may apply to other object classes whenever an admissible fragmentation, a placement operation, and a transport-additive descriptor can be defined. At the same time, this formulation also makes explicit the limitations of the approach: additivity allows descriptors of parts to be recomposed into a descriptor of the whole, but it does not in general allow unique recovery of the parts from the descriptor of the whole.

Since generation proceeds iteratively, fragment matching can affect energetic properties, and structural diversity. In this work, we formally introduce IFM, describe a practical implementation, and analyze its strengths and limitations in both de novo and fragment-constrained molecular generation scenarios.

2 Results and Discussion

2.1 Application of IFM to molecule generation tasks

IFM is implemented as an inference-time fragment generation and merging procedure. Given a reference structure, we first decompose it into fragments and compute a local MOI descriptor for each fragment in its principal inertial frame. The rigid transformations relating each local fragment frame to the global reference frame are stored. We then generate each fragment independently conditioned on its local MOI descriptor. The generated fragments are subsequently transformed back into the

global frame using the stored rotations and translations, producing an initial assembled structure. To obtain a chemically coherent molecule, we run a final global-MOI-conditioned merging stage using the generative model conditioned on the total reference MOI. This merging stage is initialized from a late denoising step, typically the final 10% from overall steps, and can be performed in either a clamped mode (while accounting for center of mass drift during injection as described in Schneuing et al. (2022)), where fragment coordinates are reinjected at every denoising step, or a relaxed mode, where fragments are allowed to move during joint denoising. The illustration of the IFM inference algorithm is provided in Figure 1. A detailed formal description of the approach is provided in Appendix Methods.

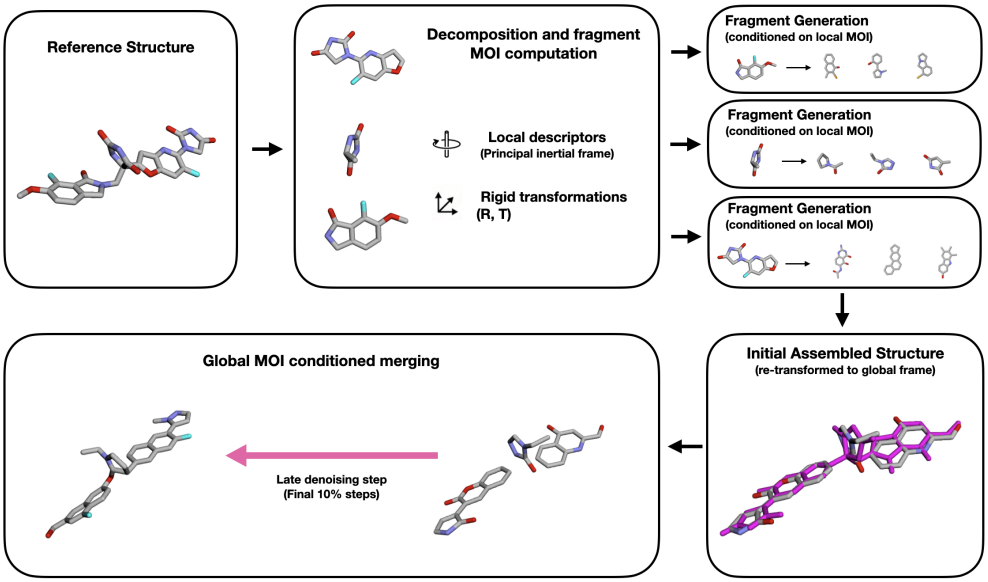

Figure 1: The illustration of the IFM inference algorithm.

IFM was compared with conventional whole-molecule generation. Results averaged over each generated dataset are presented in Table 1. The main evaluation metrics as a function of molecular size, measured by the number of heavy atoms, are shown in Figure 2. As expected, application of IFM resulted in a substantial increase in average shape similarity. Figure 2**a** shows that this improvement is primarily driven by a pronounced increase in shape similarity for larger molecules, where conventional whole-molecule MOI conditioning becomes increasingly limited. However, this improvement in shape similarity comes at the cost of more strained geometries. For conventional generation without deterministic energy optimization, the average strain energy was $866.9\,\mathrm{kJ\,mol^{-1}}$ ($207.2\,\mathrm{kcal\,mol^{-1}}$). In contrast, IFM-generated molecules had a substantially higher average strain energy of $11\,452.9\,\mathrm{kJ\,mol^{-1}}$ ($2737.3\,\mathrm{kcal\,mol^{-1}}$) before optimization. Applying constrained geometry optimization with MMFF94 (Halgren, 1996) to the IFM-generated structures reduced the average strain energy to $43.9\,\mathrm{kJ/mol}$ ($10.5\,\mathrm{kcal\,mol^{-1}}$). This optimization also decreased the average shape similarity from 0.698 to 0.575; however, the optimized IFM structures still retained a 5.3% improvement over the conventional generation baseline. The increase in strain may be attributed to the disturbance introduced by the initial fragment assembly and the relatively low number of merging steps applied.

The chemical similarity to the reference structures as a function of the size of the molecule, expressed in the number of heavy atoms, is shown in Figure 2**b**. IFM-generated molecules exhibited slightly higher chemical similarity to the references compared to conventionally generated molecules. This is expected as variation in smaller fragments in the input dataset (and even in theory) is less — so you get more chances of generating a similar fragment. Nevertheless, the absolute similarity values remained below 0.15, indicating that both methods preserve substantial chemical diversity.

Table 1: Comparison of conventional whole-molecule generation and relaxed IFM.

Metric	Conventional Generation	IFM
Average Shape Similarity	0.546	**0.698**
Average Chemical Similarity	0.110	0.119

Table 2: Comparison of inpainting and clamped IFM in fixed-fragment generation.

Metric	Inpainting	IFM
Average Shape Similarity	0.515	**0.694**
Average Chemical Similarity	0.162	0.288
Fragment Integration Success rate	0.367	**0.638**

Fixed-fragment generation is especially important in fragment growth and linker design workflows (Hadfield et al., 2022; Imrie et al., 2020; Igashov et al., 2024). In this setting, IFM shows a clear improvement over inpainting-based generation. As reported in Table 2, the average fragment integration success rate increased from 0.367 for inpainting to 0.638 for IFM (73.8% increase). This improvement is accompanied by a higher chemical similarity to the reference structures (Figure 3**b**), indicating a more successful preservation of the target fragment. IFM also increased shape similarity in the fixed-fragment setting (Figure 3**a**), consistent with the improvement observed under relaxed generation. As before, the gain in shape similarity was associated with a higher geometric strain, but this can be reduced through the constrained geometry optimization protocol. Because IFM samples each fragment separately and performs fragment merging only during the final 10% of the generation process, the merging operation is introduced relatively late. We hypothesize that this late-stage merging perturbs the learned distribution less than the enforcement of the constraint throughout the full trajectory, allowing the model to adapt more effectively to the fixed-fragment setting and achieve substantially higher fragment integration success rates. Although fragment displacement was not quantified directly, qualitative observations indicate that IFM results in less drift from the target position than analogous inpainting-based strategies, thereby better preserving the fragment's position relative to the reference structure. This behaviour may be explained by the necessity to adjust the centroid of the fixed fragment during generation to maintain equivariance (Schneuing et al., 2022). In conventional inpainting, these adjustments occur throughout denoising and may accumulate over more successive steps, leading to greater positional drift.

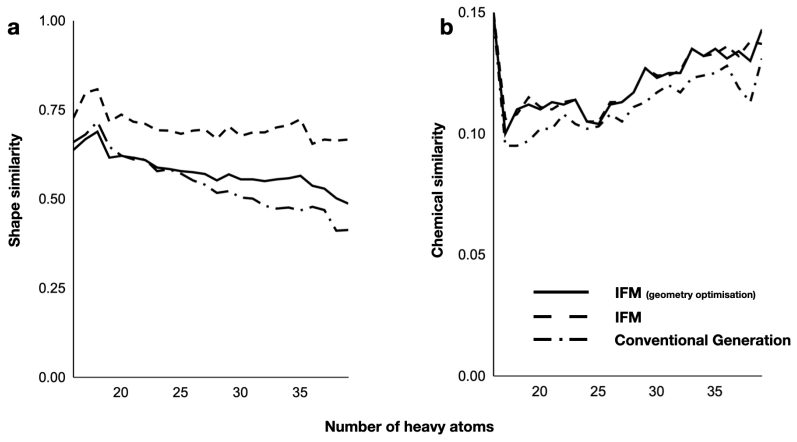

Figure 2: Shape similarity (**a**) and chemical similarity (**b**) as functions of the number of heavy atoms for relaxed IFM and conventional whole-molecule generation.

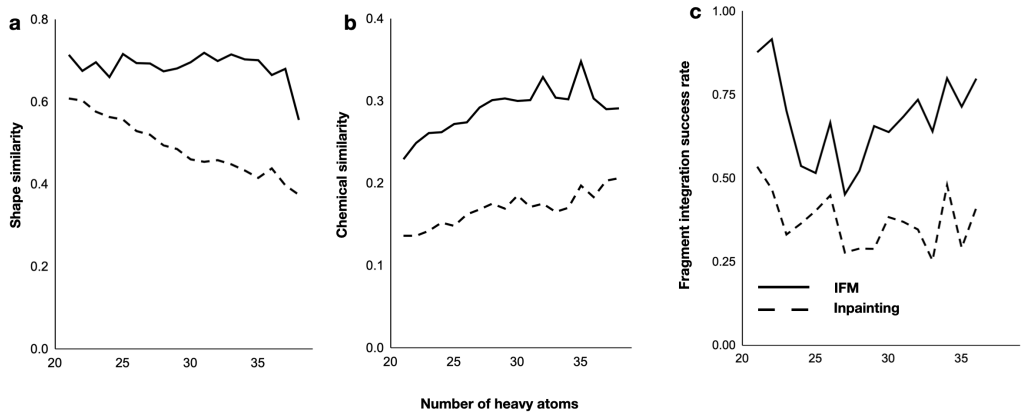

Figure 3: Shape similarity (**a**), chemical similarity (**b**), and fragment integration success rate (**c**) as functions of the number of heavy atoms for clamped IFM and inpainting.

Because IFM showed outstanding performance in the fixed-fragment generation setting, we hypothesized that it could be used to extend generation beyond the model's nominal heavy-atom limit, or context window. To test this, we introduced a simple chain merging procedure in which large molecules are generated iteratively by merging individually generated molecular fragments. At each step, the terminal region of the growing molecule is fixed, while a new fragment being attached is adjusted by the model, and merged onto the existing structure.

As a proof of concept, we applied this procedure using Bavdegalutamide (Snyder et al., 2025) as a reference structure. This example demonstrates that IFM can generate chemically valid molecules containing 58 heavy atoms with a high shape similarity to the reference, despite the model context window being limited to 42 atoms. A representative example as well as the chain merging workflow is shown in Figure 4.

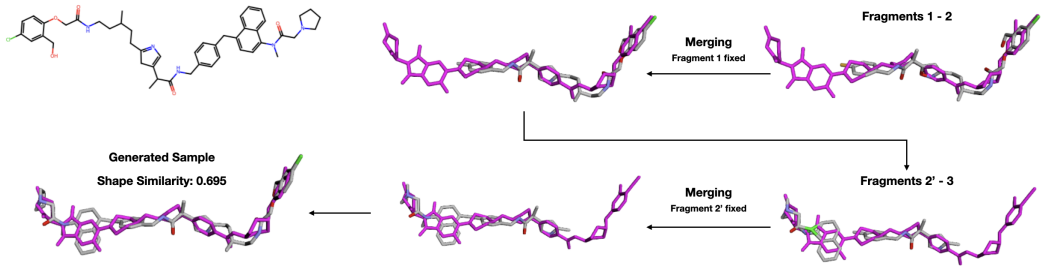

Figure 4: Application of IFM with chain merging to generate a molecular structure exceeding the model's context window.

Although several chemically valid large molecules were successfully generated, these experiments also revealed limitations of the current chain merging procedure. Reliable production use would require further refinement to improve validity, consistency, and robustness across sequential merging steps. Because fixed-fragment generation is not perfectly accurate, small errors can accumulate during iterative merging, reducing final sample quality. The most common failure modes observed were implausible geometries, disconnected fragments, and drift in the direction of chain growth. One practical way to mitigate these issues is to reduce the number of chain merging steps by using larger fragments, thereby limiting the accumulation of errors across the generation trajectory.

It is also important to note that, due to the simplicity of the current implementation, the present version of chain merging is limited to linear molecular architectures. However, this is a limitation of the current implementation rather than of the underlying principle. In principle, IFM-based chain merging could be extended to more general molecular topologies, provided that the merging procedure is designed appropriately and the model is trained or fine-tuned for the corresponding generation or merging task. Nevertheless, these results provide a promising proof of concept, demonstrating that IFM can extend molecular generation beyond the original context window and produce valid large molecules through iterative fragment merging.

2.2 Beyond MOI: learned transport-additive descriptors and transport-additive fragment matching

Although we demonstrate Fragment Matching for molecular generation, the underlying principle applies more generally. We refer to this generalization as Transport-Additive Fragment Matching. If an object x is represented by a measure μ_x over a domain Ω, descriptors of the form

$$\Phi(x) = \left(\int_\Omega \phi_1(u)\, d\mu_x(u), \ldots, \int_\Omega \phi_d(u)\, d\mu_x(u) \right)$$

are additive under decompositions of μ_x into fragment measures. These raw coefficients are additive when all fragments are expressed in a common frame. Exact finite-dimensional transport additionally requires the span of $\{\phi_j\}_{j=1}^d$ to be closed under the relevant placement transformations. Polynomial moments, Fourier coefficients, Gaussian radial-basis features, and spherical-harmonic descriptors can all be viewed as examples of this construction under suitable choices of basis functions ϕ_j. The MOI tensor corresponds to a low-order polynomial moment descriptor with a particularly simple physical interpretation and a closed-form transport rule under rigid motions.

This observation suggests a route toward learned fragment-compatible descriptors. For a training example consisting of a whole object x and an admissible decomposition $\{(x_k, \boldsymbol{p}_k)\}_{k=1}^K$, define

$$\Delta_\theta \left(x, \{(x_k, \boldsymbol{p}_k)\}_{k=1}^K \right) = \Phi_\theta(x) - \sum_{k=1}^K T_{\theta,\boldsymbol{p}_k} \Phi_\theta(x_k).$$

An approximate transport-additivity objective is

$$\mathcal{L}_{\mathrm{add}}(\theta) = \mathbb{E}_{\left(x, \{(x_k, \boldsymbol{p}_k)\}_{k=1}^K\right) \sim \mathcal{D}} \left[\left\| \Delta_\theta \left(x, \{(x_k, \boldsymbol{p}_k)\}_{k=1}^K \right) \right\|_2^2 \right].$$

When the learned transport maps are endomorphisms representing a placement group with identity element e, they may additionally be regularized by

$$T_{\theta,pq} \approx T_{\theta,p} T_{\theta,q}, \qquad T_{\theta,e} \approx \mathrm{Id}.$$

Unlike generic neural embeddings, such descriptors are explicitly constrained to support fragment-level recomposition.

However, increasing descriptor expressivity introduces a trade-off. Low-dimensional descriptors such as MOI are interpretable and stable but many-to-one; richer descriptors can preserve more geometric detail but may become higher-dimensional, less robust, and closer to storing the object itself. Therefore, the relevant design question is not whether an additive descriptor can be defined in principle, but whether it captures the fragment-level constraints needed for generation while remaining compact, stable, and compatible with the generative model.

3 Conclusion

IFM leverages the translation properties and additivity of the moment of inertia, allowing molecular shape information to be decomposed across fragments and recombined during generation. This makes IFM a general strategy that can, in principle, be applied to any generative model capable of producing molecular structures under MOI-based shape constraints.

In our experiments, IFM improved shape-conditioned molecular generation, especially for larger molecules where conventional whole-molecule conditioning becomes less effective. In fixed-fragment generation, IFM outperformed inpainting, increasing fragment integration success rate

from 0.367 to 0.638, while also improving shape and chemical similarity. Finally, chain merging provided a proof of concept for generating molecules beyond the model context window, producing valid 58-heavy-atom molecules despite a 42-atom limit. While further refinement is needed to improve robustness and reduce error accumulation, IFM offers a promising route toward scalable, shape-guided molecular generation. More broadly, the transport-additive fragment-matching framework derived from IFM may enable the underlying principle to be extended beyond molecular generation

AI USE STATEMENT

In this work, we used generative AI tools to assist with copyediting and refining the clarity of the selected sections of this manuscript. Generative AI tools were not used for image preparation, data analysis, or code writing. All generated text was thoroughly reviewed, verified, and edited by the authors. We take responsibility for the final content of this work, including text, claims or artifacts produced with the aid of generative AI.

REFERENCES

A. R. Abdulghany. Generalization of parallel axis theorem for rotational inertia. *American Journal of Physics*, 85(10):791–795, 2017. doi: 10.1119/1.4994835. URL https://doi.org/10.1119/1.4994835.

CCDC. Ccdc virtual screening set, 2026. URL https://www.ccdc.cam.ac.uk/support-andresources/downloads/.

Niklas W. A. Gebauer, Michael Gastegger, Stefaan S. P. Hessmann, Klaus-Robert Müller, and Kristof T. Schütt. Inverse design of 3D molecular structures with conditional generative neural networks. *Nature Communications*, 13(1):973, 2022. doi: 10.1038/s41467-022-28526-y. URL https://doi.org/10.1038/s41467-022-28526-y.

J. A. Grant and B. T. Pickup. Gaussian shape methods. In Wilfred F. van Gunsteren, Paul K. Weiner, and Andrew J. Wilkinson (eds.), *Computer Simulation of Biomolecular Systems*, volume 3, pp. 150–176. Springer Netherlands, 1997. doi: 10.1007/978-94-017-1120-3_5. URL https://doi.org/10.1007/978-94-017-1120-3_5.

Thomas E. Hadfield, Fergus Imrie, Andy Merritt, Kristian Birchall, and Charlotte M. Deane. Incorporating target-specific pharmacophoric information into deep generative models for fragment elaboration. *Journal of Chemical Information and Modeling*, 62(10):2280–2292, 2022. doi: 10.1021/acs.jcim.1c01311. URL https://doi.org/10.1021/acs.jcim.1c01311.

Thomas A. Halgren. Merck molecular force field. I. basis, form, scope, parameterization, and performance of MMFF94. *Journal of Computational Chemistry*, 17(5–6):490–519, 1996. doi: 10.1002/(SICI)1096-987X(199604)17:5/6⟨490::AID-JCC1⟩3.0.CO;2-P. URL https://doi.org/10.1002/(SICI)1096-987X(199604)17:5/6<490::AID-JCC1>3.0.CO;2-P.

Ilia Igashov, Hannes Stärk, Clément Vignac, Arne Schneuing, Victor Garcia Satorras, Pascal Frossard, Max Welling, Michael Bronstein, and Bruno Correia. Equivariant 3D-conditional diffusion model for molecular linker design. *Nature Machine Intelligence*, 6(4):417–427, 2024. doi: 10.1038/s42256-024-00815-9. URL https://doi.org/10.1038/s42256-024-00815-9.

Fergus Imrie, Anthony R. Bradley, Mihaela van der Schaar, and Charlotte M. Deane. Deep generative models for 3D linker design. *Journal of Chemical Information and Modeling*, 60(4):1983–1995, 2020. doi: 10.1021/acs.jcim.9b01120. URL https://doi.org/10.1021/acs.jcim.9b01120.

Greg Landrum, Paolo Tosco, Brian Kelley, et al. rdkit/rdkit: Release_2023.09.5. Zenodo, 2024. URL https://doi.org/10.5281/zenodo.10633624. Version Release_2023_09_5, computer software.

Membrizard. ml_conformer_generator. GitHub repository, 2026. URL https://github.com/Membrizard/ml_conformer_generator. Accessed 2026-08-13.

Denis Sapegin, Fedor Bakharev, Dmitriy Krupenya, Azamat Gafurov, Konstantin Pildish, and Joseph C. Bear. Moment of inertia as a simple shape descriptor for diffusion-based shape-constrained molecular generation. *Digital Discovery*, 4(10):2927–2941, 2025. doi: 10.1039/D5DD00318K. URL https://doi.org/10.1039/D5DD00318K.

Denis Sapegin, Azamat Gafurov, and Fedor Bakharev. Membrizard/ml_conformer_generator: Sphynx [Stable]. Zenodo, 2026. URL https://doi.org/10.5281/zenodo.19862065. Version v0.4.3, computer software.

Arne Schneuing, Charles Harris, Yuanqi Du, Kieran Didi, Arian Jamasb, Ilia Igashov, Weitao Du, Carla Gomes, Tom Blundell, Pietro Lio, Max Welling, Michael Bronstein, and Bruno Correia. Structure-based drug design with equivariant diffusion models. *arXiv preprint arXiv:2210.13695*, 2022. doi: 10.48550/arXiv.2210.13695. URL https://doi.org/10.48550/arXiv.2210.13695. Version 3, revised 2024.

Lawrence B. Snyder, T. K. Neklesa, R. R. Willard, D. A. Gordon, J. Pizzano, N. Vitale, K. Robling, M. A. Dorso, W. Moghrabi, S. Landrette, R. Gedrich, S. H. Lee, I. C. A. Taylor, and J. G. Houston. Preclinical evaluation of Bavdegalutamide (ARV-110), a novel PROteolysis TArgeting Chimera androgen receptor degrader. *Molecular Cancer Therapeutics*, 24(4):511–522, 2025. doi: 10.1158/1535-7163.MCT-23-0655. URL https://doi.org/10.1158/1535-7163.MCT-23-0655.

A Appendix: Methods

A.1 Inertial Fragment Matching

A.1.1 Formulation

We consider a shape-conditioned molecular generation problem. Let a molecular conformation be denoted by $M = (\mathcal{G}, \boldsymbol{X})$, where $\mathcal{G}$ is the molecular graph and $\boldsymbol{X} = (\boldsymbol{x}_1, \ldots, \boldsymbol{x}_N) \in \mathbb{R}^{N\times 3}$ contains the atomic coordinates. Given a target shape descriptor $\boldsymbol{s}^\star$, the goal is to generate a chemically valid conformation $\widehat{M} = (\widehat{\mathcal{G}}, \widehat{\boldsymbol{X}})$ whose three-dimensional shape is close to the target. The target descriptor $\boldsymbol{s}^\star$ may be derived from a reference molecule or specified independently from an arbitrary three-dimensional shape; the generated molecule is not required to preserve the reference graph, scaffold, composition, or atom identities.

For an atom set $\mathbb{A}$ with atomic masses m_i, define its total mass and center of mass as

$$m_{\mathbb{A}} = \sum_{i\in\mathbb{A}} m_i, \qquad \boldsymbol{c}_{\mathbb{A}} = \frac{1}{m_{\mathbb{A}}} \sum_{i\in\mathbb{A}} m_i \boldsymbol{x}_i.$$

The moment-of-inertia tensor about an arbitrary reference point $\boldsymbol{o}$ is

$$\mathbf{I}_{\boldsymbol{o}}(\mathbb{A}) = \sum_{i\in\mathbb{A}} m_i \left(\|\boldsymbol{x}_i - \boldsymbol{o}\|^2 \boldsymbol{I}_3 - (\boldsymbol{x}_i - \boldsymbol{o})(\boldsymbol{x}_i - \boldsymbol{o})^\top \right).$$

The model-facing descriptor is computed from the tensor about the center of mass:

$$\mathbf{I}_{\boldsymbol{c}_{\mathbb{A}}}(\mathbb{A}) = \mathbf{Q}_{\mathbb{A}} \operatorname{diag}(\lambda_{\mathbb{A},1}, \lambda_{\mathbb{A},2}, \lambda_{\mathbb{A},3}) \mathbf{Q}_{\mathbb{A}}^\top, \qquad \lambda_{\mathbb{A},1} \le \lambda_{\mathbb{A},2} \le \lambda_{\mathbb{A},3}.$$

The conditioning descriptor is the ordered vector

$$\boldsymbol{s}(\mathbb{A}) = (\lambda_{\mathbb{A},1}, \lambda_{\mathbb{A},2}, \lambda_{\mathbb{A},3}).$$

Both whole molecules and fragments are generated in a principal-inertial frame. The full tensor, rather than its eigenvalue vector, is used for rotation, change of origin, and recomposition.

The key physical property exploited by IFM is that MOI tensors are additive when expressed around a common reference point. If a molecule or target shape is decomposed into fragments $\mathbb{A}_1, \ldots, \mathbb{A}_K$, then

$$\mathbf{I}_{\boldsymbol{o}}(\mathbb{A}) = \sum_{k=1}^{K} \mathbf{I}_{\boldsymbol{o}}(\mathbb{A}_k).$$

Furthermore, the MOI tensor of each fragment can be re-expressed around a different reference point using the parallel-axis theorem (Abdulghany, 2017). Define

$$\boldsymbol{D}(\boldsymbol{c}, \boldsymbol{o}) = \|\boldsymbol{c} - \boldsymbol{o}\|^2 \boldsymbol{I}_3 - (\boldsymbol{c} - \boldsymbol{o})(\boldsymbol{c} - \boldsymbol{o})^\top.$$

Let $\boldsymbol{c}_k$ be the center of mass of a target fragment $\mathbb{A}_k$, let $\mathbf{I}_k^c$ denote its centered MOI tensor expressed in the local principal-inertial frame, and let $\boldsymbol{R}_k$ rotate that frame into the global frame. Its contribution about a global reference point $\boldsymbol{o}$ is

$$\mathbf{I}_{\boldsymbol{o},k} = \boldsymbol{R}_k \mathbf{I}_k^c \boldsymbol{R}_k^\top + m_k \boldsymbol{D}(\boldsymbol{c}_k, \boldsymbol{o}),$$

where m_k is the fragment mass.

For a generated fragment $\widehat{\mathbb{A}}_k$, let $\widehat{\mathbf{I}}_k^c$ and $\widehat{m}_k$ denote its centered local MOI tensor and mass after recentering the fragment at its own center of mass. After applying the stored placement $(\boldsymbol{R}_k, \boldsymbol{c}_k)$, the center of mass of the assembled generated structure is

$$\widehat{\boldsymbol{c}} = \frac{\sum_{k=1}^K \widehat{m}_k \boldsymbol{c}_k}{\sum_{k=1}^K \widehat{m}_k}.$$

The centered global MOI tensor of the initial assembled structure is therefore

$$\widehat{\mathbf{I}}^{\mathrm{IFM}} = \sum_{k=1}^K \left(\boldsymbol{R}_k \widehat{\mathbf{I}}_k^c \boldsymbol{R}_k^\top + \widehat{m}_k \boldsymbol{D}(\boldsymbol{c}_k, \widehat{\boldsymbol{c}}) \right).$$

IFM can then be viewed as an inference-time fragment-matching problem. In the present implementation, the fragment descriptors and placements are obtained from a decomposable reference geometry; the three-dimensional vector $\boldsymbol{s}^\star$ alone does not determine this decomposition. Each fragment is generated independently in its local principal frame by conditioning the generative model on $\boldsymbol{s}_k^\star$. Candidate fragments are then translated, rotated, and merged with the aim of bringing the ordered principal moments of the recomposed global tensor close to $\boldsymbol{s}^\star$. In this view, conventional MOI-guided generation directly samples an entire molecule from a single global descriptor $\boldsymbol{s}^\star$. IFM instead distributes the global shape constraint across smaller fragment-level descriptors, generates locally shape-consistent fragments, and uses MOI additivity together with reference-point translatability to recombine them under a shared global inertial frame. This allows the method to retain the simplicity and interpretability of MOI conditioning while reducing the loss of geometric detail that occurs when a single coarse descriptor is used for increasingly large molecular structures.

A.1.2 IFM AS TRANSPORT-ADDITIVE DESCRIPTOR FACTORIZATION

We now give a more abstract formulation of IFM that separates the descriptor-level principle from its MOI instantiation. Let $\mathcal{X}$ be a class of objects equipped with an admissible assembly operation

$$x = \mathcal{A}\left((x_1, \boldsymbol{p}_1), \ldots, (x_K, \boldsymbol{p}_K)\right),$$

where $x_1, \ldots, x_K \in \mathcal{X}$ are fragments and $\boldsymbol{p}_k$ are placement variables, such as translations, rotations, attachment points, or more general alignment metadata. Let $\mathbb{V}_{\mathrm{loc}}$ and $\mathbb{V}_{\mathrm{glob}}$ be finite-dimensional local and global descriptor spaces, with descriptor maps

$$\Phi_{\mathrm{loc}} : \mathcal{X} \to \mathbb{V}_{\mathrm{loc}}, \qquad \Phi_{\mathrm{glob}} : \mathcal{X} \to \mathbb{V}_{\mathrm{glob}}.$$

We call the descriptor transport-additive if, for every admissible assembly, there exist placement-dependent maps

$$T_{\boldsymbol{p}_k} : \mathbb{V}_{\mathrm{loc}} \to \mathbb{V}_{\mathrm{glob}}$$

such that

$$\Phi_{\mathrm{glob}}\left(\mathcal{A}\left((x_1, \boldsymbol{p}_1), \ldots, (x_K, \boldsymbol{p}_K)\right)\right) = \sum_{k=1}^K T_{\boldsymbol{p}_k} \Phi_{\mathrm{loc}}(x_k).$$

For the MOI instantiation, take

$$\mathbb{V}_{\mathrm{loc}} = \mathrm{Sym}(3) \times \mathbb{R}, \qquad \mathbb{V}_{\mathrm{glob}} = \mathrm{Sym}(3),$$

and represent a local fragment by

$$\Phi_{\mathrm{loc}}(\mathbb{A}_k) = (\mathbf{I}_k^c, m_k),$$

where $\mathbf{I}_k^c$ is its centered MOI tensor in the local frame. For placement $\boldsymbol{p}_k = (\boldsymbol{R}_k, \boldsymbol{c}_k, \boldsymbol{o})$, define

$$T_{\boldsymbol{p}_k}(\mathbf{I}_k^c, m_k) = \boldsymbol{R}_k \mathbf{I}_k^c \boldsymbol{R}_k^\top + m_k \boldsymbol{D}(\boldsymbol{c}_k, \boldsymbol{o}).$$

The global descriptor is

$$\Phi_{\mathrm{glob}}(\mathbb{A}) = \mathbf{I}_{\boldsymbol{o}}(\mathbb{A}),$$

and therefore

$$\mathbf{I}_{\boldsymbol{o}}(\mathbb{A}) = \sum_{k=1}^{K} T_{\boldsymbol{p}_k}(\mathbf{I}_k^c, m_k).$$

This is the concrete transport-additive structure exploited by IFM.

It is important to distinguish the tensor used by IFM from the low-dimensional conditioning vector used by the generative model. The model is conditioned on the ordered principal moments

$$\boldsymbol{s}(\mathbb{A}) = (\lambda_1, \lambda_2, \lambda_3)$$

obtained by diagonalizing the centered tensor $\mathbf{I}_{\boldsymbol{c}_{\mathbb{A}}}(\mathbb{A})$. Additivity holds for the full MOI tensor, not for its eigenvalue vector. IFM therefore uses principal moments as a compact model-facing conditioning signal, while physically valid recomposition is carried out in full tensor form.

A.1.3 Inference algorithm

IFM is implemented as an inference-time fragment generation and merging procedure. Given a reference structure, we first decompose it into fragments and compute a local MOI descriptor for each fragment in its principal inertial frame. The rigid transformations relating each local fragment frame to the global reference frame are stored. We then generate each fragment independently conditioned on its local MOI descriptor. The generated fragments are subsequently transformed back into the global frame using the stored rotations and translations, producing an initial assembled structure. To obtain a chemically coherent molecule, we run a final global-MOI-conditioned merging stage using the generative model conditioned on the total reference MOI. This merging stage is initialized from a late denoising step, typically the final 10% from denoising trajectory, and can be performed in either a clamped mode (while accounting for center of mass drift during injection as described by Schneuing et al. (2022)), where fragment coordinates are reinjected at every denoising step, or a relaxed mode, where fragments are allowed to move during joint denoising. The illustration of the IFM inference algorithm is provided in Figure 1.

While our formulation is instantiated for molecular references, it is not restricted to them. Any three-dimensional shape with a well-defined mass or density distribution can, in principle, be decomposed into fragments with corresponding MOI tensors, allowing the same inertial matching procedure to be applied beyond molecular references.

A.1.4 Chain Merging

We implement and evaluate a proof-of-concept extension of IFM for sequential molecular growth, which we refer to as chain merging. In contrast to basic IFM, where all fragments are merged simultaneously under a global MOI context, chain merging attaches fragments one at a time. At each step, the existing partial molecule part is kept fixed, while a newly generated fragment is merged into it using the generative model. This preserves the geometry of the previously generated structure and allows generation to proceed beyond the context size of the base model.

In practice, repeated local fragment generation and merging enables the construction of extended, linear molecular structures from local shape descriptors. This provides a route toward context-window-independent molecular generation, since each merge only requires the current partial molecule and the incoming fragment rather than the full target structure at once. However, the procedure also introduces a trade-off: because fragments are added sequentially, local geometric or chemical errors can accumulate over multiple merging steps, and the final structure is not optimized under a single global MOI constraint as in simultaneous IFM. These results demonstrate that IFM is not limited to one-shot fragment assembly, but can also serve as a mechanism for iterative molecular growth. The chain merging algorithm is illustrated in Figure 4.

A.2 INPAINTING

When considering generation with a fixed fragment we compare IFM with inpainting-based strategies. As described in Schneuing et al. (2022). During inpainting the fragment is noised to a necessary stage and then injected with an account to a COM drift of the whole molecule, to prevent discrepancies during generation a polynomial blending and resampling is applied for generation with an injected fragment.

A.3 GENERATIVE MODEL

Inertial Fragment Matching is model-agnostic in the sense that it can be applied to any molecular generative model capable of conditioning on a moment-of-inertia descriptor. In this work, we use MLConformerGenerator (Sapegin et al., 2026; Membrizard, 2026) as the base generative model. The original model was trained on ChEMBL-derived molecular structures containing 15–39 heavy atoms, following the training procedure described in Sapegin et al. (2025).

Since IFM relies on independent fragment generation, the underlying model must be capable of generating structures smaller than those included in the original training distribution. Therefore, an additional version of MLConformerGenerator was trained on an expanded dataset of molecules and fragments containing 6–39 heavy atoms. The training procedure followed Sapegin et al. (2025), using the same base dataset as the original model, supplemented with fragment-level examples generated by the splitting procedure described in Section A.4.

The augmented dataset was constructed by applying the automatic molecular splitting algorithm described in Section A.4 to the full training set. The resulting fragments were filtered to retain structures containing 6–15 heavy atoms. Unique fragments were canonicalized as SMILES strings and appended to the original molecular dataset. This augmentation exposes the model to both complete molecular structures and chemically meaningful substructures, enabling it to generate local fragments required by IFM while preserving compatibility with whole-molecule MOI-conditioned generation.

Bond prediction is carried out by AdjMatSeer (Sapegin et al., 2025) as part of MLConformerGenerator model.

A.4 MOLECULE SPLITTING

We center around generating from molecules as references because their fragmentation is relatively easy. We define an automatic size-aware splitting by breaking single non-ring bonds in the molecule. The breaking proceeds iteratively until the reference structure is not broken into pieces of acceptable size. It is fully understood that it is not general, but it works in most cases. IFM does not care about the fragmentation algorithm, nor is restricted to molecule references. The splitting applied here is one example of many.

A.5 PERFORMANCE EVALUATION

Inertial Fragment Matching was evaluated against conventional whole-molecule generation and inpainting-based fragment conditioning. Separate benchmark datasets were constructed for the relaxed and clamped generation settings, containing 993 and 799 reference molecules, respectively. The datasets were compiled from the unique molecules found in CCDC Virtual Screening SetCCDC (2026) by applying the fragmentation algorithm described in Section A.4 to each molecule.

For the relaxed IFM benchmark, molecules were retained if they could be decomposed into more than two fragments, with each fragment constrained to contain between 6 and 15 heavy atoms. For the clamped generation benchmark, molecules were restricted to decompositions into two fragments, each containing between 6 and 20 heavy atoms. In all clamped generation experiments, the first fragment was treated as the fixed input fragment, while the remaining molecular structure was generated conditionally.

All generation experiments were performed using 100 denoising steps. For relaxed IFM, fragment merging was initialized during the final 10 denoising steps. In the clamped setting, both IFM and the inpainting baseline used polynomial blending of the fixed fragment coordinates with a blending

power of 3. For inpainting, four resampling steps (Schneuing et al., 2022) were additionally applied during generation.

To ensure a consistent comparison, the same base generative model, denoising schedule, and molecular preprocessing pipeline were used across all methods.

A.6 METRICS

A.6.1 SHAPE SIMILARITY

Before computing shape similarity, the generated and reference structures were rigidly aligned on the basis of their Gaussian shape moments. The shape quadrupole was used to define the principal shape axes and obtain a common orientation of the corresponding Gaussian molecular volume distributions.

Shape similarity was then evaluated using a Tanimoto-like coefficient based on the Gaussian volume overlap:

$$S_{\text{shape}}(A, B) = \frac{V_{AB}}{V_A + V_B - V_{AB}},$$

where V_A and V_B denote the Gaussian molecular volumes of structures A and B, and V_{AB} denotes their overlap volume after shape-moment-based alignment. Gaussian shape moments, molecular volumes, and overlap volumes were computed following Grant & Pickup (1997).

A.6.2 CHEMICAL SIMILARITY

The chemical similarity was defined as the Tanimoto coefficient between bit-Morgan fingerprints (2048 bits, 2 hops) of a generated molecule and a corresponding reference conformer.

A.6.3 STRAIN

The strain energy was evaluated relative to a locally energy-minimized conformation. First, the total energy of the model-generated conformation, E_0, was calculated using the MMFF94 force field (Halgren, 1996) as implemented in RDKit (Landrum et al., 2024). Where specified, the generated conformation was subsequently subjected to constrained geometry optimization using the same force field, with a maximum of 1000 optimization steps, an energy tolerance of 10^{-8}, and positional restraints allowing a maximum displacement of $2.0\,\text{Å}$ with a force constant of $20.0\,\text{kcal mol}^{-1}\,\text{Å}^{-2}$. A corresponding local-minimum conformation was obtained by unconstrained MMFF94 energy minimization. The energy of this optimized geometry, E_{opt}, was taken as the reference local-minimum energy. The conformational strain energy was then defined as

$$E_{\text{strain}} = E_0 - E_{\text{opt}}.$$

A.6.4 FRAGMENT INTEGRATION SUCCESS RATE

Fragment integration was defined as a boolean metric: If a target fragment SMILES can be found in the generated molecule (while ignoring bond orders and chirality — so molecule graph only) we count that as a successful fragment incorporation. The substructure check was performed using RDKit (Landrum et al., 2024) .