ALIGNING MOLECULES AND FRAGMENTS IN A SHARED EMBEDDING SPACE FOR RL-BASED MOLECULE GENERATION

Youngkuk Kim[1] **Yinhua Piao**[2] **Sangseon Lee**[3] **Sun Kim**[1,4,5,6]
[1]Department of Computer Science and Engineering, Seoul National University
[2]Bio-MAX Institute, Seoul National University
[3]Department of Artificial Intelligence, Inha University
[4]Interdisciplinary Program in Bioinformatics, Seoul National University
[5]Interdisciplinary Program in Artificial Intelligence, Seoul National University
[6]AIGENDRUG Co., Ltd.
{kyk9174844, 2018-27910, sunkim.bioinfo}@snu.ac.kr
{ss.lee}@inha.ac.kr

ABSTRACT

Drug discovery is a complex and resource-intensive process requiring the design of molecules that possess specific chemical and biological properties, such as high binding affinity and drug-likeness. Fragment-based drug discovery (FBDD) has gained prominence as a strategy for efficiently identifying lead compounds by deconstructing molecules into smaller fragments. However, existing approaches face challenges in fully leveraging the relationships between molecules and their constituent fragments, especially in optimizing molecular properties. In this paper, we introduce Molecule-Fragment Representation Alignment space for RL-based Generation (M-FRAG), a novel framework that harmonizes molecule and fragment embeddings in a shared, property-driven space. By aligning fragments with their molecular context, M-FRAG ensures that fragment selection is optimized both for chemical feasibility and the desired molecular properties. Using reinforcement learning, M-FRAG generates chemically realistic molecules optimized for target properties while also providing interpretability for individual fragments during the molecule generation process. Experimental results demonstrate that M-FRAG outperforms existing methods in terms of optimization, diversity, and chemical validity, positioning it as a powerful tool for the efficient and transparent generation of drug-like molecules.

1 INTRODUCTION

Drug discovery is a complex and resource-intensive process that involves designing molecules with desirable chemical and biological properties, such as high binding affinity, drug-likeness, and favorable pharmacokinetics. Given the vastness of chemical space, exhaustively exploring all potential candidates is infeasible, leading to increased interest in computational approaches for efficient molecular design (Jin et al., 2018; Zhavoronkov et al., 2019; De Cao & Kipf, 2018).

Among computational methods, reinforcement learning (RL) (Kaelbling et al., 1996; Wiering & Van Otterlo, 2012; Haarnoja et al., 2018) has emerged as a promising approach for automated molecular generation guided by property optimization (Olivecrona et al., 2017; Zhou et al., 2019; Jeon & Kim, 2020; Goel et al., 2021). While these approaches enhance molecular optimization, they often struggle to maintain chemical realism, leading to the generation of impractical or synthetically infeasible molecules.

Fragments, as chemically meaningful substructures, offer a structured way to generate valid molecules while preserving key chemical and biological characteristics. Thus, Fragment-based Drug Discovery (FBDD) has emerged as a powerful strategy to ensure chemical realism by utilizing fragments that inherently contain chemical structures, thereby simplifying the molecular generation

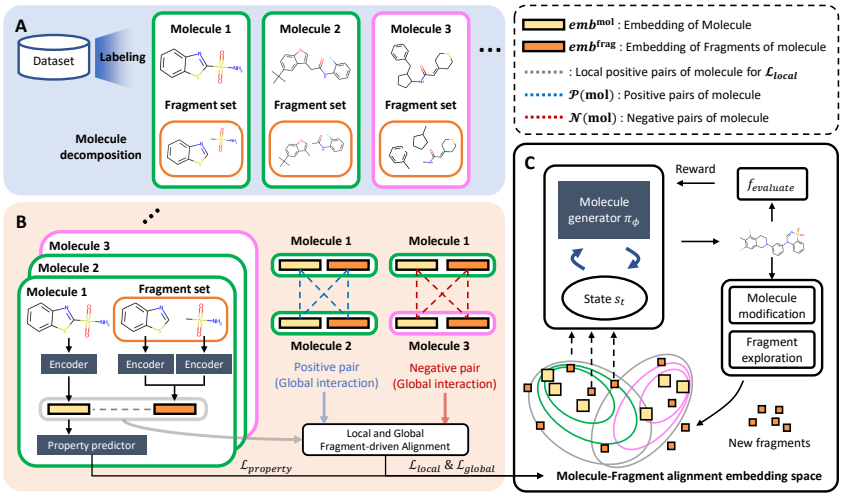

Figure 1: **Overview of M-FRAG. A.** Labeling and molecule decomposition from ZINC250k dataset (Irwin et al., 2012). **B.** Harmonized aligned embedding space construction. **C.** Molecule generation via fragment-based reinforcement learning and evolutionary search.

process (Li, 2020; Jin et al., 2018; 2020a;b; Yang et al., 2021; Xie et al., 2021; Maziarz et al., 2021; Kong et al., 2022; Geng et al., 2023; Lee et al., 2023b). However, these models treat fragments independently, without a property-driven molecular space or understanding of their global context.

Understanding the relationship between fragments and molecular properties is crucial for optimizing molecular design (Jin et al., 2020a; Maziarz et al., 2021; Coley, 2021), yet challenging due to the complex interplay between them. Details of the relationship between fragments and molecular properties are described in Appendix A.

To address this challenge, we propose M-FRAG, which harmonizes molecule and fragment embeddings within a unified, property-driven space. Using metric learning, M-FRAG aligns fragments with molecule embeddings, ensuring that fragment selection optimizes both chemical feasibility and specific property constraints. Experimental results confirm that M-FRAG significantly outperforms state-of-the-art methods in property optimization, demonstrating its effectiveness in drug discovery tasks. By improving both molecular generation quality and interpretability, M-FRAG provides a novel framework for designing high-quality molecules in a property-driven space.

2 METHOD

In this section, we introduce the details of M-FRAG for fragment-based molecule generation. As shown in Figure 1, M-FRAG follows two steps: (1) constructing a harmonized embedding space for molecules and fragments, where (2) we generate molecules via a fragment-based RL framework.

2.1 HARMONIZED EMBEDDING SPACE CONSTRUCTION

To effectively integrate molecule and fragment representations, our method constructs a harmonized embedding space that captures their intrinsic relationships. By aligning these representations, we aim to generate a structured space that accurately reflects molecular properties. First, we describe the molecule property predictor for property-driven alignment in Sec. 2.1.1. Next, we outline the local and global fragment-driven alignment in Sec. 2.1.2. Finally, we introduce the joint optimization strategy for harmonizing the embedding space in Sec. 2.1.3.

2.1.1 MOLECULE PROPERTY PREDICTOR FOR PROPERTY-DRIVEN ALIGNMENT

To construct a harmonized embedding space where molecules and fragments are meaningfully integrated, it is crucial to ensure that molecular properties are well-represented. However, identifying

how individual fragments contribute to the overall molecular property remains challenging. Since fragments can appear in various molecular contexts, their isolated impact on molecular properties is not always straightforward. To guide the alignment process, we incorporate a molecular property predictor, which provides a reference by estimating the whole molecule's property.

Given a dataset $\mathcal{D} = \{(G_i, Y_i)\}_{i=1}^N$, where each molecular graph G_i is associated with a property value $Y_i \in \mathbb{R}$, we aim to align molecules and fragments within a shared space. Each molecular graph $G_i = (X_i, V_i, E_i)$ consists of a set of atoms (nodes) V_i, a set of chemical bonds (edges) E_i, and a node feature matrix $X_i \in \mathbb{R}^{|V_i| \times d}$ that encodes atomic properties. Here, $|V_i|$ represents the number of atoms in G_i, and d denotes the feature dimension.

To effectively capture molecular properties while maintaining computational efficiency, we employ a lightweight predictor g_θ that predicts the molecular property Y from the molecular graph G. The predictor consists of a Message Passing Neural Network (MPNN) (Gilmer et al., 2017) as an encoder for extracting graph representations and a Multi-Layer Perceptron (MLP) for property prediction. First, the MPNN processes the molecular graph $G_i = (X_i, V_i, E_i)$ to compute node embeddings: $[z_1, \ldots, z_l, \ldots, z_{|V_i|}]^\top = \text{MPNN}_\theta(X_i, V_i, E_i)$, where $z_l \in \mathbb{R}^{d_h}$ represents the hidden embedding of node v_l, and d_h is the dimension of the node embedding v_l.

Next, a readout function g_{readout} aggregates the node embeddings to obtain a molecule embedding emb_i^{mol}: $emb_i^{\text{mol}} = g_{\text{readout}}([z_1, \ldots, z_l, \ldots, z_{|V_i|}]^\top) \in \mathbb{R}^{d_h}$, where g_{readout} is a mean pooling operator, and $emb_i^{\text{mol}} \in \mathbb{R}^{d_h}$ denotes embedding of the molecule G_i.

The molecular embedding emb_i^{mol} is then passed through MLP_θ to predict the molecular property $\hat{Y}_i \in \mathbb{R}$: $\hat{Y}_i = g_\theta(G_i) = \text{MLP}_\theta(emb_i^{\text{mol}})$. The predictor g_θ is trained by minimizing the mean squared error (MSE) loss between the predicted and actual molecular property Y_i: $\mathcal{L}_{\text{property}} = \frac{1}{N}\sum_{i=1}^N \left(Y_i - \hat{Y}_i\right)^2$.

2.1.2 Local and Global Fragment-driven Alignment

To effectively align molecules and fragments, it is essential to consider both local and global perspectives when structuring the embedding space (Kaya & Bilge, 2019; Chen et al., 2020; Wang et al., 2022; Khosla et al., 2020). The local perspective aims to position fragments close to their parent molecule, preserving structural and chemical coherence. Meanwhile, the global perspective arranges fragment representations according to molecular property similarities, ensuring that fragments from molecules with similar properties are positioned near each other, while those from molecules with distinct properties remain farther apart. Given a molecule $G_i = (X_i, V_i, E_i)$, we decompose it into a set of m fragments, denoted as $\mathcal{F}_i = \{F_j\}_{j=1}^m$, using BRICS (Degen et al., 2008). Each fragment $F_j = (X_j, V_j, E_j)$ consists of a subset of the molecular nodes and edges, i.e., $V_j \subset V_i$ and $E_j \subset E_i$.

To obtain fragment representations $emb_j \in \mathbb{R}^{d_h}$, we apply the same MPNN encoder used in g_θ: $emb_j = g_{\text{readout}}(\text{MPNN}_\theta(F_j)), \quad j = 1, \ldots, m$. The pooled fragment representation emb_i^{frag} for molecule G_i is then computed as: $emb_i^{\text{frag}} = \frac{1}{m}\sum_{j=1}^m emb_j, where \quad emb_i^{\text{frag}} \in \mathbb{R}^{d_h}$.

Local Alignment Since both emb_i^{mol} and emb_i^{frag} should encode similar molecular properties, we enforce their consistency using a local alignment loss $\mathcal{L}_{\text{local}}$: $\mathcal{L}_{\text{local}} = \sum_{i=1}^N \|emb_i^{\text{mol}} - emb_i^{\text{frag}}\|_2^2$. This objective ensures that the pooled fragment representation remains aligned with the molecular representation, preserving the underlying molecular property distribution.

Global Alignment Fragments should be positioned in the embedding space based on molecular property similarities. Each molecule G_i is associated with a ground-truth property value Y_i, and its corresponding pooled fragment representation emb_i^{frag} inherits a similar property from its parent molecule.

To enforce global contrastive alignment, we define positive and negative fragment pairs based on a molecular property threshold δ. Let G_a and G_b be two molecules. A pair is assigned as: Positive pair: $\mathcal{P}(a) = \{emb_b \mid |Y_a - Y_b| < \delta\}$, Negative pair: $\mathcal{N}(a) = \{emb_b \mid |Y_a - Y_b| \geq \delta\}$. where emb_b refers to either emb_b^{mol} or emb_b^{mol}.

Table 1: **Novel hit ratio (%) results.** The results are the means and the standard deviations of 3 runs. The results for the baselines except for RationaleRL and PS-VAE are taken from Lee et al. (2023b). The best results are highlighted in bold.

	Target protein				
	parp1	**fa7**	**5ht1b**	**braf**	**jak2**
REINVENT (Olivecrona et al., 2017)	0.480 (± 0.344)	0.213 (± 0.081)	2.453 (± 0.561)	0.127 (± 0.088)	0.613 (± 0.167)
Graph GA (Jensen, 2019)	4.811 (± 1.661)	0.422 (± 0.193)	7.011 (± 2.732)	3.767 (± 1.498)	5.311 (± 1.667)
MORLD (Jeon & Kim, 2020)	0.047 (± 0.050)	0.007 (± 0.013)	0.880 (± 0.735)	0.047 (± 0.040)	0.227 (± 0.118)
HierVAE (Jin et al., 2020a)	0.553 (± 0.214)	0.007 (± 0.013)	0.507 (± 0.278)	0.207 (± 0.220)	0.227 (± 0.127)
RationaleRL (Jin et al., 2020b)	4.267 (± 0.450)	0.900 (± 0.098)	2.967 (± 0.307)	0.000 (± 0.000)	2.967 (± 0.196)
FREED (Yang et al., 2021)	4.627 (± 0.727)	1.332 (± 0.113)	16.767 (± 0.897)	2.940 (± 0.359)	5.800 (± 0.295)
PS-VAE (Kong et al., 2022)	1.644 (± 0.389)	0.478 (± 0.140)	12.622 (± 1.437)	0.367 (± 0.047)	4.178 (± 0.933)
MOOD (Lee et al., 2023a)	7.017 (± 0.428)	0.733 (± 0.141)	18.673 (± 0.423)	5.240 (± 0.285)	9.200 (± 0.524)
GEAM (Lee et al., 2023b)	40.567 (± 0.827)	20.711 (± 1.873)	38.489 (± 0.350)	27.900 (± 1.822)	42.950 (± 1.117)
M-FRAG (ours)	**40.881** (± 0.115)	**26.511** (± 1.433)	**43.243** (± 0.294)	**32.104** (± 3.255)	**46.542** (± 1.320)

Using these positive and negative pairs, we apply supervised contrastive learning (Khosla et al., 2020) to both representations. The contrastive loss function $\mathcal{L}_{\text{global}}$ is defined as: $\mathcal{L}_{\text{global}} = \sum_{a=1}^{N} -\log \frac{\sum_{emb_p \in \mathcal{P}(a)} \exp(\text{sim}(emb_a, emb_p)/\tau)}{\sum_{emb_n \in \mathcal{N}(a)} \exp(\text{sim}(emb_a, emb_n)/\tau)}$. where emb_a refers to either emb_a^{mol} or emb_a^{mol}. τ is a temperature parameter. emb_p and emb_n represent positive and negative embeddings from $\mathcal{P}(a)$ and $\mathcal{N}(a)$, respectively.

Final Alignment Loss The final alignment loss integrates both local and global alignment constraints, formulated as $\mathcal{L}_{\text{alignment}} = \lambda_{\text{local}}\mathcal{L}_{\text{local}} + \lambda_{\text{global}}\mathcal{L}_{\text{global}}$, where λ_{local} and λ_{global} are hyperparameters controlling the relative importance of local and global alignment.

Through this dual-level optimization, fragments are embedded in a structured manner within the molecular space, ensuring that they reflect both their intrinsic molecular relationships and broader chemical property distributions.

2.1.3 Joint Optimization for Harmonizing the Space

Through joint optimization, we align fragments within a molecule's property-driven embedding space by accounting for both local and global interactions with the molecule. The final loss function integrates alignment constraints with property predictor, formulated as $\mathcal{L} = \lambda_{\text{property}}\mathcal{L}_{\text{property}} + \lambda_{\text{alignment}}\mathcal{L}_{\text{alignment}}$ where $\lambda_{\text{property}}$ and $\lambda_{\text{alignment}}$ are hyperparameters that control the relative importance of property predictor and alignment.

2.2 Molecule Generation via Fragment-based Reinforcement Learning and Evolutionary Search

To enable optimized molecule generation, we integrate a reinforcement learning (RL) framework for fragment-based molecule assembly with an evolutionary strategy for dynamic vocabulary expansion. This approach ensures that molecules are constructed in a structured and property-aligned manner while also allowing for continuous adaptation to improve diversity and molecular properties. A key component of this framework is the scoring function, which utilizes the harmonized embedding space trained via g_θ to evaluate fragment relevance with respect to molecular properties. Details of molecule generation process via fragment-based RL and evolutionary search are described in Appendix C.

3 Experiments

To evaluate the effectiveness of our proposed method, M-FRAG, we conduct a series of experiments on multi-objective molecule optimization tasks that closely resemble real-world drug discovery applications. The primary objective is to generate novel molecules that optimize multiple molecule properties while maintaining diversity and synthesizability.

3.1 Optimization of Binding Affinity under QED, SA, and Novelty Constraints

Datasets & Evaluation Metrics We evaluate M-FRAG on five docking score (DS) (Trott & Olson, 2010; Alhossary et al., 2015) optimization tasks (parp1, fa7, 5ht1b, braf, jak2) while ensuring that generated molecules satisfy drug-likeness (QED) (Bickerton et al., 2012), synthetic accessibility (SA) (Ertl & Schuffenhauer, 2009), and novelty constraints. The evaluation considers multiple criteria, including the Novel Hit Ratio, Novel Top-5% DS, and Novelty. Detailed metric definitions, computation methods and datasets are provided in the Appendix D.1.

Baselines. To validate the effectiveness of M-FRAG, we compare its performance against several state-of-the-art molecule generation models: REINVENT (Olivecrona et al., 2017), GraphGA (Jensen, 2019), MORLD (Jeon & Kim, 2020), HierVAE (Jin et al., 2020a), RationaleRL (Jin et al., 2020b), FREED (Yang et al., 2021), PS-VAE (Kong et al., 2022), MOOD (Lee et al., 2023a), and GEAM (Lee et al., 2023b). Details of baselines are described in Appendix D.2.

Results. Experimental results show that M-FRAG consistently outperforms all baselines across multiple evaluation criteria. As shown in Table 1, M-FRAG achieves a significantly higher novel hit ratio, demonstrating its ability to balance molecular optimization with novelty constraints. Additionally, Table 3 confirms that M-FRAG maintains strong docking score performance, suggesting that the molecule-fragment alignment space effectively guides molecular design while enabling broader chemical space exploration.

Unlike conventional methods that often face a trade-off between optimization and diversity (Gao et al., 2022), M-FRAG effectively mitigates this issue through its structured fragment alignment mechanism. Also, our method efficiently generates acceptable molecules that meet the chemical realism and pharmacochemical suitability required in drug design, while also demonstrating excellent performance in terms of Validity and Uniqueness. Details of baselines are described in Appendix D.3.

3.2 Ablation Study on Alignment and Modification

To analyze the contribution of each key component in M-FRAG, we conduct an ablation study focusing on the role of molecule-fragment alignment in the embedding space and the impact of molecule modification through genetic algorithms in Figure 3. Details of the experiment are described in Appendix D.4

3.3 Case Study: Property-driven Molecule Generation in the Aligned Embedding Space

To further illustrate the effectiveness of M-FRAG, we analyze a specific molecule generation process within the aligned embedding space about jak2 target protein. As shown in Figure 4, the model selects fragments in a structured manner to progressively refine molecule properties, leading to an optimized final molecule. Details of the experiment are described in Appendix D.5

4 Conclusion

We have introduced M-FRAG, a novel framework for fragment-based molecule generation, guided by reinforcement learning that utilizes a hamonized embedding space for molecules and fragments. By aligning molecule and fragment embeddings within a property-driven space, M-FRAG improves the exploration of fragment importance and enables the generation of molecules that are both chemically realistic and optimized for desired properties. Our approach outperforms existing methods in terms of optimization, diversity, and chemical validity, and it provides a more interpretable path for understanding how fragment selection contributes to molecular property optimization.

In future work, we aim to extend M-FRAG to handle larger, more complex molecular datasets and incorporate additional property constraints. This will further enhance the applicability of our method to real-world drug discovery tasks, offering a more efficient and interpretable approach to generating novel drug candidates.

REFERENCES

Amr Alhossary, Stephanus Daniel Handoko, Yuguang Mu, and Chee-Keong Kwoh. Fast, accurate, and reliable molecular docking with quickvina 2. *Bioinformatics*, 31(13):2214–2216, 2015.

Jonathan B Baell and Georgina A Holloway. New substructure filters for removal of pan assay interference compounds (pains) from screening libraries and for their exclusion in bioassays. *Journal of medicinal chemistry*, 53(7):2719–2740, 2010.

G Richard Bickerton, Gaia V Paolini, Jérémy Besnard, Sorel Muresan, and Andrew L Hopkins. Quantifying the chemical beauty of drugs. *Nature chemistry*, 4(2):90–98, 2012.

Ting Chen, Simon Kornblith, Mohammad Norouzi, and Geoffrey Hinton. A simple framework for contrastive learning of visual representations. In *International conference on machine learning*, pp. 1597–1607. PMLR, 2020.

Connor W Coley. Defining and exploring chemical spaces. *Trends in Chemistry*, 3(2):133–145, 2021.

Nicola De Cao and Thomas Kipf. Molgan: An implicit generative model for small molecular graphs. *arXiv preprint arXiv:1805.11973*, 2018.

Jorg Degen, Christof Wegscheid-Gerlach, Andrea Zaliani, and Matthias Rarey. On the art of compiling and using'drug-like'chemical fragment spaces. *ChemMedChem*, 3(10):1503, 2008.

Peter Ertl and Ansgar Schuffenhauer. Estimation of synthetic accessibility score of drug-like molecules based on molecular complexity and fragment contributions. *Journal of cheminformatics*, 1:1–11, 2009.

Wenhao Gao, Tianfan Fu, Jimeng Sun, and Connor Coley. Sample efficiency matters: a benchmark for practical molecular optimization. *Advances in neural information processing systems*, 35: 21342–21357, 2022.

Zijie Geng, Shufang Xie, Yingce Xia, Lijun Wu, Tao Qin, Jie Wang, Yongdong Zhang, Feng Wu, and Tie-Yan Liu. De novo molecular generation via connection-aware motif mining. *arXiv preprint arXiv:2302.01129*, 2023.

Justin Gilmer, Samuel S Schoenholz, Patrick F Riley, Oriol Vinyals, and George E Dahl. Neural message passing for quantum chemistry. In *International conference on machine learning*, pp. 1263–1272. PMLR, 2017.

Manan Goel, Shampa Raghunathan, Siddhartha Laghuvarapu, and U Deva Priyakumar. Molegular: molecule generation using reinforcement learning with alternating rewards. *Journal of Chemical Information and Modeling*, 61(12):5815–5826, 2021.

Tuomas Haarnoja, Aurick Zhou, Pieter Abbeel, and Sergey Levine. Soft actor-critic: Off-policy maximum entropy deep reinforcement learning with a stochastic actor. In *International conference on machine learning*, pp. 1861–1870. PMLR, 2018.

John J Irwin, Teague Sterling, Michael M Mysinger, Erin S Bolstad, and Ryan G Coleman. Zinc: a free tool to discover chemistry for biology. *Journal of chemical information and modeling*, 52 (7):1757–1768, 2012.

Siddhant M Jayakumar, Wojciech M Czarnecki, Jacob Menick, Jonathan Schwarz, Jack Rae, Simon Osindero, Yee Whye Teh, Tim Harley, and Razvan Pascanu. Multiplicative interactions and where to find them. In *International conference on learning representations*, 2020.

Jan H Jensen. A graph-based genetic algorithm and generative model/monte carlo tree search for the exploration of chemical space. *Chemical science*, 10(12):3567–3572, 2019.

Woosung Jeon and Dongsup Kim. Autonomous molecule generation using reinforcement learning and docking to develop potential novel inhibitors. *Scientific reports*, 10(1):22104, 2020.

Wengong Jin, Regina Barzilay, and Tommi Jaakkola. Junction tree variational autoencoder for molecular graph generation. In *International conference on machine learning*, pp. 2323–2332. PMLR, 2018.

Wengong Jin, Regina Barzilay, and Tommi Jaakkola. Hierarchical generation of molecular graphs using structural motifs. In *International conference on machine learning*, pp. 4839–4848. PMLR, 2020a.

Wengong Jin, Regina Barzilay, and Tommi Jaakkola. Multi-objective molecule generation using interpretable substructures. In *International conference on machine learning*, pp. 4849–4859. PMLR, 2020b.

Leslie Pack Kaelbling, Michael L Littman, and Andrew W Moore. Reinforcement learning: A survey. *Journal of artificial intelligence research*, 4:237–285, 1996.

Mahmut Kaya and Hasan Şakir Bilge. Deep metric learning: A survey. *Symmetry*, 11(9):1066, 2019.

Prannay Khosla, Piotr Teterwak, Chen Wang, Aaron Sarna, Yonglong Tian, Phillip Isola, Aaron Maschinot, Ce Liu, and Dilip Krishnan. Supervised contrastive learning. *Advances in neural information processing systems*, 33:18661–18673, 2020.

Thomas N Kipf and Max Welling. Semi-supervised classification with graph convolutional networks. *arXiv preprint arXiv:1609.02907*, 2016.

Xiangzhe Kong, Wenbing Huang, Zhixing Tan, and Yang Liu. Molecule generation by principal subgraph mining and assembling. *Advances in Neural Information Processing Systems*, 35:2550–2563, 2022.

Stephen J Lane, Drake S Eggleston, Keith A Brinded, John C Hollerton, Nicholas L Taylor, and Simon A Readshaw. Defining and maintaining a high quality screening collection: the gsk experience. *Drug discovery today*, 11(5-6):267–272, 2006.

Seul Lee, Jaehyeong Jo, and Sung Ju Hwang. Exploring chemical space with score-based out-of-distribution generation. In *International Conference on Machine Learning*, pp. 18872–18892. PMLR, 2023a.

Seul Lee, Seanie Lee, Kenji Kawaguchi, and Sung Ju Hwang. Drug discovery with dynamic goal-aware fragments. *arXiv preprint arXiv:2310.00841*, 2023b.

Qingxin Li. Application of fragment-based drug discovery to versatile targets. *Frontiers in molecular biosciences*, 7:180, 2020.

Krzysztof Maziarz, Henry Jackson-Flux, Pashmina Cameron, Finton Sirockin, Nadine Schneider, Nikolaus Stiefl, Marwin Segler, and Marc Brockschmidt. Learning to extend molecular scaffolds with structural motifs. *arXiv preprint arXiv:2103.03864*, 2021.

Marcus Olivecrona, Thomas Blaschke, Ola Engkvist, and Hongming Chen. Molecular de-novo design through deep reinforcement learning. *Journal of cheminformatics*, 9:1–14, 2017.

George Papadatos, Mark Davies, Nathan Dedman, Jon Chambers, Anna Gaulton, James Siddle, Richard Koks, Sean A Irvine, Joe Pettersson, Nicko Goncharoff, et al. Surechembl: a large-scale, chemically annotated patent document database. *Nucleic acids research*, 44(D1):D1220–D1228, 2016.

Oleg Trott and Arthur J Olson. Autodock vina: improving the speed and accuracy of docking with a new scoring function, efficient optimization, and multithreading. *Journal of computational chemistry*, 31(2):455–461, 2010.

Yuyang Wang, Jianren Wang, Zhonglin Cao, and Amir Barati Farimani. Molecular contrastive learning of representations via graph neural networks. *Nature Machine Intelligence*, 4(3):279–287, 2022.

Marco A Wiering and Martijn Van Otterlo. Reinforcement learning. *Adaptation, learning, and optimization*, 12(3):729, 2012.

Yutong Xie, Chence Shi, Hao Zhou, Yuwei Yang, Weinan Zhang, Yong Yu, and Lei Li. Mars: Markov molecular sampling for multi-objective drug discovery. *arXiv preprint arXiv:2103.10432*, 2021.

Soojung Yang, Doyeong Hwang, Seul Lee, Seongok Ryu, and Sung Ju Hwang. Hit and lead discovery with explorative rl and fragment-based molecule generation. *Advances in Neural Information Processing Systems*, 34:7924–7936, 2021.

Alex Zhavoronkov, Yan A Ivanenkov, Alex Aliper, Mark S Veselov, Vladimir A Aladinskiy, Anastasiya V Aladinskaya, Victor A Terentiev, Daniil A Polykovskiy, Maksim D Kuznetsov, Arip Asadulaev, et al. Deep learning enables rapid identification of potent ddr1 kinase inhibitors. *Nature biotechnology*, 37(9):1038–1040, 2019.

Zhenpeng Zhou, Steven Kearnes, Li Li, Richard N Zare, and Patrick Riley. Optimization of molecules via deep reinforcement learning. *Scientific reports*, 9(1):10752, 2019.

A Fragment-Molecule Relationship

A molecule consists of multiple fragments, making it unclear which one influences a specific property in Figure 2A. At the same time, a single fragment can even appear in multiple molecules in Figure 2B. These many-to-many relationships complicate the attribution of molecular properties to individual fragments, as their collective influence remains difficult to model effectively.

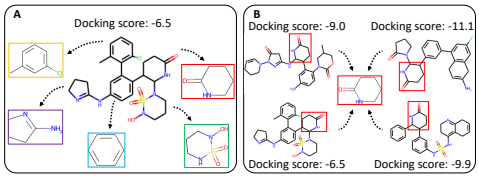

Figure 2: **Fragment-Molecule Relationship A.** A single molecule has a specific property, such as a docking score, but consists of multiple fragments, making it unclear which fragment contributes to it. **B.** A fragment appears in multiple molecules with different scores, complicating the attribution of molecular properties to individual fragments.

B Related Work

B.1 Reinforcement Learning in Molecular Generation

Reinforcement learning (RL) has been increasingly applied to molecular generation, as it offers an effective way to explore the vast chemical space. In RL-based models, molecules are generated by optimizing a reward function that evaluates the chemical properties of the molecule. For instance, REINVENT (Olivecrona et al., 2017) and MORLD (Jeon & Kim, 2020) use RL to generate molecules optimized for certain properties. However, these methods often suffer from the challenge of reward sparsity, particularly when multiple property constraints are considered.

B.2 Fragment-based Drug Discovery

Fragment-based drug discovery (FBDD) has been widely adopted in molecular generative models to explore chemical space efficiently. RationaleRL (Jin et al., 2020b) proposes a method where molecules are generated by assembling substructures, that are identified as key contributors to desired properties. This approach allows for better interpretability and ensures that the generated molecules are more likely to satisfy the target properties. Our method builds on the strengths of RL and rationale-based generation by incorporating a unified alignment of molecule and fragment embeddings, improving both the diversity and quality of the generated molecules while maintaining property optimization. Many FBDD methods, such as PS-VAE (Kong et al., 2022) and FREED (Yang et al., 2021), focus on the assembly of fragments to generate valid molecules. These models typically extract fragments from predefined libraries or via heuristic rules, without considering

the target chemical properties during the extraction process. GEAM (Lee et al., 2023b) introduced a more dynamic approach by updating fragment vocabularies during generation, but the lack of alignment between molecules and fragments makes it difficult to precisely capture and interpret the contribution of each fragment to the overall molecule properties. Unlike these methods, our approach introduces a novel method for harmonizing molecule and fragment embeddings within a shared space, allowing for more efficient and interpretable molecule generation.

C MOLECULE GENERATION PROCESS

C.1 FRAGMENT ASSEMBLY VIA SOFT ACTOR-CRITIC (SAC)

We model fragment assembly as a reinforcement learning problem, where an agent iteratively constructs a molecule graph by selecting and attaching fragments. The Soft Actor-Critic (SAC) algorithm (Haarnoja et al., 2018) is employed to optimize the policy, balancing exploration and exploitation through entropy regularization. The RL agent models molecule generation as a Markov decision process (MDP), At each time step t, the agent observes the molecule state s_t and selects an action a_t, which includes: (1) Choosing an attachment site on the existing molecule structure. (2) Selecting a fragment F from the vocabulary $\mathcal{VOCAB}$. (3) Determining the attachment site on the chosen fragment.

The molecule state s_t is encoded using a graph convolutional network (GCN) (Kipf & Welling, 2016). $H_t = \text{GCN}(G_t)$ represents node embeddings of the partially constructed molecule G_t. The overall molecule embedding is obtained via sum pooling $h_{G_t} = \sum H_t$. he policy network π_ϕ consists of three sub-networks responsible for different fragment selection steps:

$$p_{\pi_1}(\cdot|s_t) = \pi_{\phi,1}(Z_1), \quad Z_1 = f_1(h_{G_t}, H_{\text{att}}), \tag{1}$$

$$p_{\pi_2}(\cdot|a_1, s_t) = \pi_{\phi,2}(Z_2), \quad Z_2 = f_2(Z_{1,a_1}, \text{ECFP}(\mathcal{VOCAB})), \tag{2}$$

$$p_{\pi_3}(\cdot|a_{1:2}, s_t) = \pi_{\phi,3}(Z_3), \quad Z_3 = f_3(\sum \text{GCN}(F_{a_2}), H_{\text{att},F_{a_2}}), \tag{3}$$

where H_{att} represents node embeddings of attachment sites, and f_1, f_2, f_3 employ multiplicative interactions for efficient feature fusion (Jayakumar et al., 2020).

The SAC objective function is defined as:

$$J(\pi) = \sum_t \mathbb{E}_{(s_t,a_t)\sim\rho_\pi} \left[r(s_t, a_t) + \alpha \mathcal{H}(\pi(\cdot|s_t)) \right], \tag{4}$$

where $r(s_t, a_t)$ is the reward function f_{evaluate} evaluating molecule structural and property alignment, and $\mathcal{H}(\pi(\cdot|s_t))$ is an entropy regularization term with a temperature parameter α. The reward function is designed to ensure optimization of target properties (Trott & Olson, 2010; Alhossary et al., 2015), drug-likeness (QED) (Bickerton et al., 2012), synthetic accessibility (SA) (Ertl & Schuffenhauer, 2009).

C.2 MOLECULE MODIFICATION AND FRAGMENT EXPLORATION

While fragment assembly facilitates molecule construction from a predefined vocabulary, relying solely on a fixed set of fragments may limit structural diversity. To address this, we integrate an evolutionary strategy that dynamically modifies and expands the vocabulary.

A genetic algorithm (GA) (Jensen, 2019) is utilized to introduce novel fragments through mutation and crossover operations. The GA iteratively selects parent molecules from the generated set and produces offspring molecules containing new fragments. The process follows: (1) Selection: Choose parent molecules based on molecule property scores. (2) Crossover: Combine molecule fragments from two parents. (3) Mutation: Introduce structural variations in selected fragments.

Each newly generated fragment is assigned a score based on its alignment with molecule properties. g_θ represents the fragment score.

To ensure the fragment vocabulary remains optimized, we iteratively update $\mathcal{VOCAB}$ by adding high-value fragments while filtering out underperforming ones:

$$\mathcal{VOCAB}_{t+1} = \mathcal{VOCAB}_t \cup \{F \mid g_\theta(F) > \delta\}, \tag{5}$$

where δ define the thresholds for inclusion and removal.

By integrating reinforcement learning for structured fragment assembly with an evolutionary approach for vocabulary expansion, our framework enables property-driven molecule generation that continuously adapts to optimize desired properties.

D Experimental Details

D.1 Datasets & Evaluation Metrics

ZINC250k, a widely used molecular dataset derived from the ZINC database, contains commercially available drug-like molecules. ZINC250k consists of 249,455 molecular graphs, where each molecule is represented as a graph with atoms as nodes and bonds as edges. The dataset statistics are summarized in Table 2.

Table 2: Statistics of the ZINC250k dataset.

Statistic	**Value**
Number of molecules	249455
Average number of nodes (atoms)	23.15
Average number of edges (bonds)	24.90
Average number of fragments per molecule	5.35

The dataset provides a diverse set of molecules, making it well-suited for evaluating fragment-based molecular generation and optimization methods.

To comprehensively assess the performance of M-FRAG, we employ the following evaluation metrics:

Novel Hit Ratio (%) quantifies the percentage of generated molecules that satisfy all optimization constraints, including docking score (DS), drug-likeness (QED), and synthetic accessibility (SA), while maintaining novelty. Formally, it is defined as:

$$\text{Novel Hit Ratio} = \frac{N_{\text{valid}}}{N_{\text{total}}} \times 100, \tag{6}$$

where N_{valid} represents the number of molecules that meet all constraints, and N_{total} is the total number of generated molecules.

Novel Top-5% DS (kcal/mol) measures the mean docking score of the top 5% of unique, novel molecules. Given a set of generated molecules $\mathcal{G}$, we first select those classified as novel and then compute:

$$\text{Top-5\% DS} = \frac{1}{|\mathcal{G}_{5\%}|} \sum_{g \in \mathcal{G}_{5\%}} DS(g), \tag{7}$$

where $\mathcal{G}_{5\%}$ denotes the top 5% molecules with the lowest (most negative) docking scores.

Novelty (%) represents the fraction of generated molecules that are structurally distinct from the training set. Novelty is determined based on the maximum Tanimoto similarity $S_{\text{max}}(g)$ of each generated molecule g to any molecule in the training dataset:

$$\text{Novelty} = \frac{N_{\text{novel}}}{N_{\text{total}}} \times 100, \tag{8}$$

where N_{novel} is the number of generated molecules satisfying $S_{\text{max}}(g) < \tau$, and $\tau = 0.4$ is used as the novelty threshold following prior studies (Irwin et al., 2012).

These metrics ensure a comprehensive evaluation of M-FRAG's ability to generate molecules that not only optimize binding affinity but also maintain drug-like properties and structural diversity.

D.2 Baseline Models

To provide a fair comparison, we evaluate M-FRAG against several state-of-the-art molecule generation models. Below, we briefly describe each baseline model:

REINVENT (Olivecrona et al., 2017) is a reinforcement learning (RL) model that generates molecules in SMILES format using a recurrent neural network (RNN) trained via policy gradient methods.

Graph GA (Jensen, 2019) is a genetic algorithm-based approach that applies predefined crossover and mutation rules to optimize molecular structures directly in graph form.

MORLD (Jeon & Kim, 2020) utilizes deep Q-learning to optimize molecular properties through reinforcement learning, ensuring that generated molecules satisfy predefined objectives.

HierVAE (Jin et al., 2020a) is a hierarchical variational autoencoder (VAE) that learns a structured latent space for molecule generation, capturing both molecular topology and fine-grained atomic details.

RationaleRL (Jin et al., 2020b) applies reinforcement learning to extract key molecular substructures (rationales) and extend them to generate novel drug-like molecules.

FREED (Yang et al., 2021) is a fragment-based RL model that constructs molecules by iteratively assembling fragments selected from a predefined vocabulary.

PS-VAE (Kong et al., 2022) integrates principal subgraph mining with variational autoencoding to improve the structural diversity and chemical validity of generated molecules.

MOOD (Lee et al., 2023a) employs a diffusion-based generative model to generate molecules with enhanced novelty while maintaining property constraints.

GEAM (Lee et al., 2023b) combines reinforcement learning and genetic algorithms to iteratively refine molecular structures and update the fragment vocabulary.

These models serve as comparative baselines to evaluate the efficacy of M-FRAG in generating high-quality molecules while optimizing molecular properties.

D.3 Additional Experiments about a Trade-off between Optimization and Diversity

The alignment mechanism enables M-FRAG to conduct extensive exploration without deviating from molecule property constraints, ultimately leading to the discovery of a greater number of hit molecules. The impact of exploration is particularly evident in Table 4, where M-FRAG demonstrates a substantial improvement in molecule diversity.

By ensuring that fragments and molecules are meaningfully positioned within the embedding space, M-FRAG not only enhances molecule diversity but also maintains optimization performance, allowing for the generation of structurally diverse yet highly optimized molecules. These results validate M-FRAG's ability to achieve a more comprehensive exploration of chemical space while maintaining property-driven molecule design.

Compared to GEAM (Lee et al., 2023b), which integrates reinforcement learning and genetic algorithms, M-FRAG benefits from its embedding space alignment, where molecule fragments are positioned according to their chemical relevance and contribution to molecule properties. This approach allows the model to maintain high docking scores while simultaneously enhancing QED and synthetic accessibility. Additionally, unlike methods such as MORLD, which achieves high novelty scores at the expense of optimization performance, M-FRAG demonstrates strong performance across all key metrics including novelty. By effectively aligning molecule and fragment representations, M-FRAG offers a robust framework for generating novel, drug-like molecules with high optimization potential, making it a promising approach for real-world drug discovery applications.

We also report three widely used pharmacochemical filter scores – Glaxo (Lane et al., 2006), SureChEMBL (Papadatos et al., 2016), PAINS (Baell & Holloway, 2010) on Table 5 with the target fa7 protein. These scores are defined as the ratio of accepted, valid molecules to total generated molecules, with the filters rejecting compounds containing functional groups deemed inappropriate for drugs (i.e., toxic or reactive groups). The higher the quality scores, the higher the likelihood the molecule will be an acceptable drug. We also report the ratio of valid molecules to total generated molecules (validity) and the ratio of unique molecules among valid generated molecules (uniqueness).

Table 3: **Novel top 5% docking score (kcal/mol) results.** The results are the means and the standard deviations of 3 runs. The results for the baselines except for RationaleRL and PS-VAE are taken from Lee et al. (2023b). The best results are highlighted in bold.

	Target protein				
	parp1	**fa7**	**5ht1b**	**braf**	**jak2**
REINVENT (Olivecrona et al., 2017)	-8.702 (± 0.523)	-7.205 (± 0.264)	-8.770 (± 0.316)	-8.392 (± 0.400)	-8.165 (± 0.277)
Graph GA (Jensen, 2019)	-10.949 (± 0.532)	-7.365 (± 0.326)	-10.422 (± 0.670)	-10.789 (± 0.341)	-10.167 (± 0.576)
MORLD (Jeon & Kim, 2020)	-7.532 (± 0.260)	-6.263 (± 0.165)	-7.869 (± 0.650)	-8.040 (± 0.337)	-7.816 (± 0.133)
HierVAE (Jin et al., 2020a)	-9.487 (± 0.278)	-6.812 (± 0.274)	-8.081 (± 0.252)	-8.978 (± 0.525)	-8.285 (± 0.370)
RationaleRL (Jin et al., 2020b)	-10.663 (± 0.086)	-8.129 (± 0.048)	-9.005 (± 0.155)	No hit found	-9.398 (± 0.076)
FREED (Yang et al., 2021)	-10.579 (± 0.104)	-8.378 (± 0.044)	-10.714 (± 0.183)	-10.561 (± 0.080)	-9.735 (± 0.022)
PS-VAE (Kong et al., 2022)	-9.978 (± 0.091)	-8.028 (± 0.050)	-9.887 (± 0.115)	-9.637 (± 0.049)	-9.464 (± 0.129)
MOOD (Lee et al., 2023a)	-10.865 (± 0.113)	-8.160 (± 0.071)	-11.145 (± 0.042)	-11.063 (± 0.034)	-10.147 (± 0.060)
GEAM (Lee et al., 2023b)	**-12.891** (± 0.158)	-9.890 (± 0.116)	-12.374 (± 0.036)	-12.342 (± 0.095)	-11.816 (± 0.067)
M-FRAG (ours)	-12.688 (± 0.115)	**-10.093** (± 1.433)	**-12.985** (± 0.294)	**-12.412** (± 3.255)	**-12.016** (± 1.320)

Table 4: **Novelty (%) results.** The results are the means and the standard deviations of 3 runs. The results for the baselines except for RationaleRL and PS-VAE are taken from Lee et al. (2023b). The best results are highlighted in bold.

	Target protein				
	parp1	**fa7**	**5ht1b**	**braf**	**jak2**
REINVENT (Olivecrona et al., 2017)	9.894 (± 2.178)	10.731 (± 1.516)	11.605 (± 3.688)	8.715 (± 2.712)	11.456 (± 1.793)
MORLD (Jeon & Kim, 2020)	**98.433** (± 1.189)	**97.967** (± 1.764)	**98.787** (± 0.743)	**96.993** (± 2.787)	**97.720** (± 0.995)
HierVAE (Jin et al., 2020a)	60.453 (± 17.165)	24.853 (± 15.416)	48.107 (± 1.988)	59.747 (± 16.403)	85.200 (± 14.262)
RationaleRL (Jin et al., 2020b)	9.300 (± 0.354)	9.802 (± 0.166)	7.133 (± 0.141)	0.000 (± 0.000)	7.389 (± 0.220)
FREED (Yang et al., 2021)	74.640 (± 2.953)	78.787 (± 2.132)	75.027 (± 5.194)	73.653 (± 4.312)	75.907 (± 5.916)
PS-VAE (Kong et al., 2022)	60.822 (± 2.251)	56.611 (± 1.892)	57.956 (± 2.181)	57.744 (± 2.710)	58.689 (± 2.307)
MOOD (Lee et al., 2023a)	84.180 (± 2.123)	83.180 (± 1.519)	84.613 (± 0.822)	87.413 (± 0.830)	83.273 (± 1.455)
GEAM (Lee et al., 2023b)	88.611 (± 3.107)	89.378 (± 2.619)	84.222 (± 2.968)	90.322 (± 3.467)	89.222 (± 1.824)
M-FRAG (ours)	92.990 (± 1.424)	92.610 (± 1.775)	88.067 (± 1.964)	92.763 (± 1.074)	90.127 (± 2.103)

Our method efficiently generates acceptable molecules that meet the chemical realism and pharmacochemical suitability required in drug design, while also demonstrating excellent performance in terms of Validity and Uniqueness. This proves that our model not only satisfies the qualitative requirements for drug candidates but is also highly effective in discovering a diverse range of valid molecules.

D.4 Ablation Study on Alignment and Modification

To analyze the contribution of each key component in M-FRAG, we conduct an ablation study focusing on the role of molecule-fragment alignment in the embedding space and the impact of molecule modification through genetic algorithms. As shown in Figure 3, we evaluate four different configurations of our model across five target proteins: (1) full M-FRAG with both alignment and modification, (2) without modification, (3) without alignment, and (4) without both alignment and modification. The target proteins used in this analysis include parp1, fa7, 5ht1b, braf, and jak2, allowing us to assess the performance of each configuration across a diverse set of molecular properties.

Effect of Molecule-Fragment Alignment. When this alignment mechanism is removed, the performance drops significantly, as observed in Figure 3. Without alignment, the reinforcement learning (RL) agent operates in an unstructured space, making it more difficult to identify high-quality molecule fragments that contribute effectively to the desired molecule properties.

Interplay Between Alignment and Modification. In M-FRAG, the genetic algorithm is used for fragment exploration and molecule modification. To isolate the effect of modification, we disable it in an experiment while retaining fragment generation. With alignment, removing modification has little impact and can even improve performance. However, without alignment, the model becomes reliant on modification, and performance drops significantly when it is absent. These results highlight that alignment enables M-FRAG to generate high-quality molecules without heavy reliance on modification. As genetic algorithms inherently introduce randomness, alignment provides a more property-driven approach, reinforcing its importance in interpretable and effective molecule design.

Table 5: **Quality Scores of the Models with the target fa7.** Quality scores are defined as the ratio of valid molecules to total generated molecules, with filters rejecting compounds containing toxic or reactive groups. Higher quality scores indicate a higher likelihood of drug acceptability. We also report the ratio of valid molecules (validity) and unique molecules among valid ones (uniqueness).

	Glaxo	SureChEMBL	PAINS	Validity	Uniqueness
REINVENT (Olivecrona et al., 2017)	0.832 (± 0.034)	0.747 (± 0.040)	0.842 (± 0.034)	0.872 (± 0.028)	0.990 (± 0.007)
MORLD (Jeon & Kim, 2020)	0.578 (± 0.010)	0.145 (± 0.018)	0.816 (± 0.008)	**1.000** (± 0.000)	**1.000** (± 0.001)
HierVAE (Jin et al., 2020a)	0.975 (± 0.004)	0.795 (± 0.007)	0.893 (± 0.011)	**1.000** (± 0.000)	0.131 (± 0.003)
FREED (Yang et al., 2021)	**0.996** (± 0.001)	0.808 (± 0.049)	0.991 (± 0.002)	**1.000** (± 0.000)	0.723 (± 0.135)
M-FRAG (ours)	0.931 (± 0.012)	**0.832** (± 0.005)	**0.992** (± 0.001)	**1.000** (± 0.000)	0.962 (± 0.011)

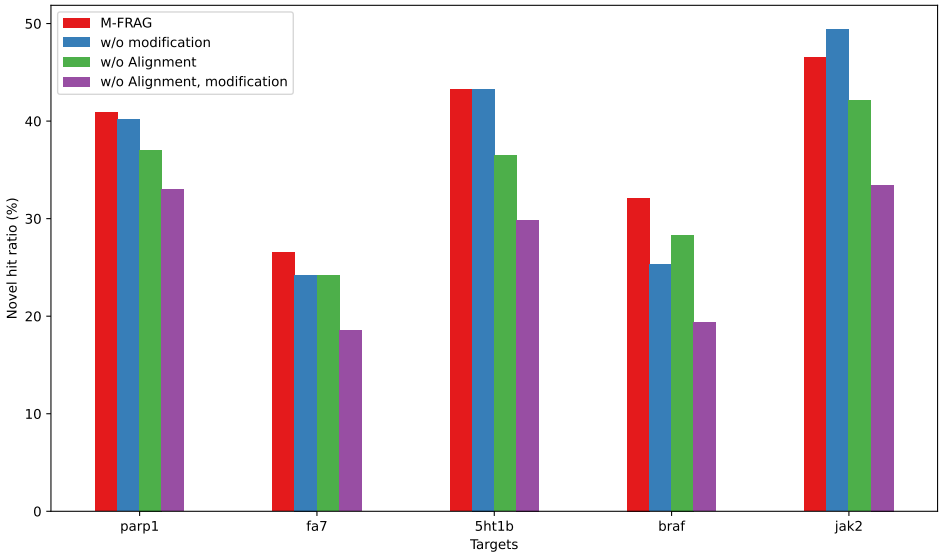

Figure 3: **Ablation study on M-FRAG's alignment and modification.** The Y-axis represents hit ratio (%) across five target proteins. The figure compares four settings: full M-FRAG, without modification, without alignment, and without both. Removing alignment significantly reduces performance, highlighting its key role in optimization. While modification aids in an unstructured space, alignment reduces its necessity, enabling efficient molecule generation.

D.5 Case Study: Property-driven Molecule Generation in the Aligned Embedding Space

To further illustrate the effectiveness of M-FRAG, we analyze a specific molecule generation process within the aligned embedding space about jak2 target protein. As shown in Figure 4, the model selects fragments in a structured manner to progressively refine molecule properties, leading to an optimized final molecule.

Progressive Optimization of Molecule Properties. Initially, the molecule is positioned in a region of the embedding space with low QED. As fragment selection progresses, it moves towards an optimized region where Docking Score, QED, and SA are all high. This transformation is guided by the molecule-fragment alignment, ensuring a systematic improvement in molecule properties.

Fragment-Based Interpretability in Embedding Space. By utilizing fragment-molecule interactions, M-FRAG provides a framework where molecular generation is interpreted by the relative positioning and movement of fragments and molecules in the embedding space. Each selected fragment shifts the molecule toward regions with improved Docking Score, QED, and SA, allowing a stepwise understanding of property refinement.

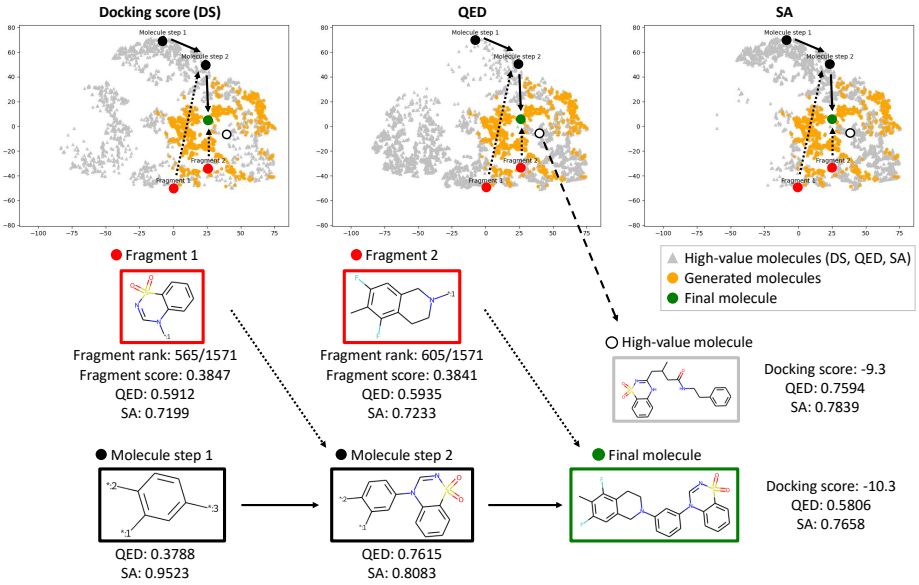

Figure 4: **Visualization of M-FRAG's molecule generation in embedding space.** The scatter plot shows molecule properties (docking score, QED, SA), with dataset molecules in gray and generated ones in orange. Molecule steps (black) and selected fragments (red) guide construction toward the final molecule (green). Below, actual structures illustrate the stepwise generation process.