ESANN 2026 proceedings, European Symposium on Artificial Neural Networks, Computational Intelligence and Machine Learning. Bruges (Belgium) and online event, 22-24 April 2026, i6doc.com publ., ISBN 9782875870964. Available from http://www.i6doc.com/en/.

Ring-constrained Molecular Graph Generation with Diffusion Models

Davide Rigoni✉, Rana İşlek, and Nicolò Navarin *

University of Padova - Dept of Mathematics "Tullio Levi-Civita"
Via Trieste, 63, 35121 Padova - Italy

Abstract. Designing molecules with specific attributes is vital in drug discovery and materials science. Ring structures are key to a molecule's stability, reactivity, and biological interactions, ensuring that designed compounds are both feasible and synthetically viable, thereby increasing their potential for lab production and therapeutic use. Generative diffusion models have become essential tools for *in silico* molecule generation. However, integrating structural constraints, especially those involving ring structures, remains challenging. This study introduces a method for applying hard ring-related constraints in molecule generation, enhancing synthetic validity and utility, with evaluations on the QM9 dataset.

1 Introduction

Designing molecules with desired properties is a fundamental goal in drug discovery. Generative models have enabled the *de novo* synthesis of candidate molecules *in silico*. Early methodologies are based on variational autoencoders (VAEs) [1, 2, 3], generative adversarial networks (GANs) [4], and normalizing flows [5, 6]. Recently, diffusion [7, 8] models have transformed generative modeling across various domains, including molecule generation. In the *diffusion process*, the model gradually adds random noise to an example in the dataset over several steps, until the data becomes nearly indistinguishable from noise. The *denoising process*, learnt during training, reverses this noise addition step-by-step, reconstructing the original data. Substructures like aromatic rings and functional groups determine important molecular properties such as stability and reactivity. In practical applications, generative models must craft molecules exploring chemical space while meeting structural constraints. Implementing these is challenging due to structural complexity and the need for algorithms that balance exploration and constraint adherence. The conditioning of diffusion [7, 8] models can bias outputs but cannot eliminate violations, highlighting the need

✉Corresponding author: davide.rigoni.1@unipd.it

* Founded by: (i) the National Recovery and Resilience Plan (NRRP), Mission 4 Component 2 Investment 1.3 - Call for tender No. 341 of March 15, 2022 of Italian Ministry of University and Research – NextGenerationEU, project code PE0000013, Concession Decree No. 1555 of October 11, 2022, CUP C63C22000770006, project title "Future AI Research (FAIR) - Spoke 2 Integrative AI - Symbolic conditioning of Graph Generative Models (SymboliG)"; and (ii) Deep Learning in Structured Domains for Functional Neuroimaging Data Analysis - jointly funded by the University of Lausanne and the University of Padua, CUP C93C24004750005. Moreover, authors acknowledge ISCRA for awarding this project access to the LEONARDO supercomputer, owned by the EuroHPC Joint Undertaking, hosted by CINECA.

ESANN 2026 proceedings, European Symposium on Artificial Neural Networks, Computational Intelligence and Machine Learning. Bruges (Belgium) and online event, 22-24 April 2026, i6doc.com publ., ISBN 9782875870964. Available from http://www.i6doc.com/en/.

to transition from soft constraints to hard constraints. Recent studies started to address this gap. GCDM [9] focuses on 3D molecular structure generation by enforcing geometric completeness and equivariance. CoCoGraph [10] targets constrained molecule generation by integrating chemical constraints, such as valence, into the diffusion process. *ConStruct* [11], integrates hard constraints into the graph diffusion process through an edge-absorbing forward process and a projection operator applied at each reverse step. Specifically, the sampling-time projector enforces planarity and acyclicity. EDGE [7] addresses degree distributions and scalability. These studies show the effectiveness of combining diffusion with constraints for flexible, rule-compliant molecular generation.

Despite recent advances, existing models often overlook ring structures, key substructures in drug-like molecules. This work targets this gap by focusing on: *(i)* maximum ring count *(ii)* maximum ring length. Rings strongly influence drug-likeness, stability, and synthetic feasibility. High aromatic ring counts are associated with reduced solubility and poor pharmacokinetic profiles, while very large macrocycles are synthetically challenging [12]. Although *Lipinski's Rule* [13] does not explicitly mention rings, related heuristics penalize excessive ring numbers or complexity. The proposed approach builds on the discrete graph diffusion model ConStruct [11], augmenting it with a projection operator that rejects edge proposals violating constraints. This keeps all intermediate and final graphs feasible, yielding molecules that satisfy specified ring constraints without relying on naive *post-hoc* filtering. The approach targets boolean structural constraints on the QM9 [14] dataset, showing how diffusion models can enforce strict guarantees on molecular rings. Appendices are available online[1].

2 The Approach

Diffusion Model Let molecules be represented as undirected labeled graphs $G = (V, E)$, where V represents atoms and E represents chemical bonds. The ConStruct [11] model, foundational to the proposed approach, operates in a discrete space using categorical variables for atom and bond types, one-hot encoded as $\boldsymbol{X} \in \{0,1\}^{n \times d_X}$ and $\mathbf{E} \in \{0,1\}^{n \times n \times d_E}$, where d_X is the number of atom-type categories and $d_E = c + 1$ includes c bond types plus a "no-bond" category. The forward diffusion process is a T-step categorical Markov chain $q(G_t \mid G_{t-1})$ that progressively corrupts a molecular graph $G_0 \sim \mathcal{D}$ from the dataset, with nodes and edges corrupted independently at each step. Using transition matrices $Q_t^X \in \mathbb{R}^{d_X \times d_X}$ and $Q_t^E \in \mathbb{R}^{d_E \times d_E}$ for atoms and bonds, respectively, transitions are defined as $q_t(x_{i,t} \mid x_{i,t-1}) = Q_t^X[x_{i,t-1}, x_{i,t}]$ and $q_t(e_{ij,t} \mid e_{ij,t-1}) = Q_t^E[e_{ij,t-1}, e_{ij,t}]$. For nodes, the noise process transitions to the marginal of the dataset, while edges transition towards a no-bond state. Specifically, they are defined as: $Q_t^X = \alpha_t I + (1 - \alpha_t)\mathbf{1}_{d_X} m'_X$ and $Q_t^E = \alpha_t^{\text{ABS}} I + (1 - \alpha_t^{\text{ABS}})\mathbf{1}_c e'_E$, where α_t and α_t^{ABS} decreases from 1 to 0, m'_X is the marginal distribution for atom types, e'_E is the one-hot encoding of no-bond state, while $\mathbf{1}_{d_X} \in \{1\}^{d_X}$ and $\mathbf{1}_c \in \{1\}^c$ are vectors filled with ones. In the reverse process, the model recon-

[1] `https://www.math.unipd.it/~drigoni/files/ESANN2026-Constrained-Generation.pdf`

ESANN 2026 proceedings, European Symposium on Artificial Neural Networks, Computational Intelligence and Machine Learning. Bruges (Belgium) and online event, 22-24 April 2026, i6doc.com publ., ISBN 9782875870964. Available from http://www.i6doc.com/en/.

structs G_0 using a parameterized backward Markov chain $p_\theta(G_{t-1} \mid G_t)$. A key feature of ConStruct is the use of sampling-time projectors to enforce constraints. For a given constraint $\mathcal{C}$, the projector $\Pi_\mathcal{C}$ is defined $G_{t-1} = \Pi_\mathcal{C}(\widehat{G}_{t-1}, G_t)$, where $\widehat{G}_{t-1}$ is the provisional graph obtained from the denoising network's predictions. The projector corrects $\widehat{G}_{t-1}$ by rejecting edges that violate $\mathcal{C}$, ensuring $G_{t-1} = G_t \cup \{(i,j) \mid (i,j) \in \widehat{G}_{t-1},\ \mathcal{P}(G_t \cup \{(i,j)\}) = \text{True}\}$, where $\mathcal{P}$ is a boolean predicate associated with the constraint $\mathcal{C}$. Given a predicate $\mathcal{P}$, it is said to be *edge-deletion invariant* if, for any graph $G = (V, E)$ and any subset of edges $\tilde{E} \subseteq E$ such that $G' = (V, \tilde{E})$, $\mathcal{P}(G) = \text{True} \Rightarrow \mathcal{P}(G') = \text{True}$, ensuring that if G_0 satisfies $\mathcal{P}$, all intermediate graphs G_t do as well. ConStruct, by adopting forward transitions to no-bond edges, enforces edge-invariant predicates only at generation time via projectors. The model is trained once, with a single configuration, regardless of the constraints, and the projector guides denoising toward lower-probability yet valid, constraint-consistent regions of the distribution.

Molecular Constraints The proposed approach enforces two structural constraints on the generated molecules: *(i)* the number and *(ii)* the size of rings (simple cycles), to maintain desired chemical properties and synthetic tractability. Because constraints are applied step-by-step during denoising, the results differ from those of naive *post-hoc* filtering. The *first constraint*, "ring count $\leq K$", limits the molecule to at most K distinct simple cycles in its undirected molecular graph G, counting all unique simple cycles rather than those from an arbitrary cycle basis. This *(i)* avoids dependence on a cycle basis and treats isomorphic graphs identically; *(ii)* captures the full cyclic content of fused and polycyclic systems, including larger "outer" rings; *(iii)* is monotone under edge deletion, so removing edges never creates new cycles. The predicate $\mathcal{P}_{\text{ring-count},K}(G)$ is true if and only if the number of simple cycles in G does not exceed K. Indeed, the corresponding projector $\Pi_\mathcal{C}$ ensures that the evolving graph never exceeds K cycles. Any newly proposed bond that would cause the number of simple cycles to surpass K is rejected and stored in a global blocklist, maintaining the constraint throughout sampling. The *second constraint*, "maximum ring length $\leq L$", bounds ring size to enforce chemically plausible, synthetically feasible rings; e.g., $L = 6$ allows benzene but excludes larger macrocycles like cycloheptane. The predicate $\mathcal{P}_{\text{ring-length},L}(G)$ holds if every simple cycle in G has a length of at most L. This property is also invariant under edge deletion, as deleting edges can only remove cycles or split them, never creating longer ones. In the reverse diffusion process, the projector $\Pi_\mathcal{C}$ tests each proposed edge insertion and rejects any bond that would create a cycle longer than L atoms. This maintains the constraint along the entire sampling trajectory, so every intermediate graph G_t and the final molecule satisfy the maximum ring length bound.

3 Experimental Assessment

Experimental Setting The *ConStruct-based* configuration is adopted without changes to the training objective or architecture. At evaluation, $N = 10K$

ESANN 2026 proceedings, European Symposium on Artificial Neural Networks, Computational Intelligence and Machine Learning. Bruges (Belgium) and online event, 22-24 April 2026, i6doc.com publ., ISBN 9782875870964. Available from http://www.i6doc.com/en/.

molecules are sampled per seed across 5 seeds ($50K$ molecules total). The mean and standard deviation across seeds are presented. To isolate the effect of symbolic constraints at generation time, the same trained model and diffusion hyperparameters as the baseline are used, where only sampling-time projectors are enabled. A single projector is activated per experiment. Implementation details are reported in Appendix A, while the code can be publicly accessed online[2].

Dataset and Evaluation Metrics This study consider QM9 [14], a benchmark of $130K$ organic molecules with up to 9 heavy atoms. Molecules are represented as undirected graphs with explicit hydrogen atoms to ensure accurate valence accounting. The dataset is split in training (75%), validation (15%), and test (10%) sets, with no additional preprocessing beyond RDKit[3] sanitization. Ring statistics are computed from all simple cycles in the graph representation. Generated sets are evaluated using standard metrics from de novo molecular generation [15]: *(i)* Validity (%): fraction of samples RDKit can sanitize; *(ii)* Uniqueness (%): fraction of non-duplicate valid molecules; *(iii)* Novelty (%): fraction of unique valid molecules not in the training set; *(iv)* Fréchet ChemNet Distance (FCD): distributional distance via ChemNet embeddings; *(v)* Disconnected (%): percentage of graphs generated with disconnected components.

Results and Discussion Table 1 presents the performance metrics when considering the constraint on the number of rings, i.e., in varying K, while Figure 1 reports some qualitative examples. Results with more values of K are reported in Appendix B. Note that $K = 9$ covers over 96% of the QM9 dataset, with a maximum ring count of 35. Imposing a strict cap on ring count alters the output distribution compared to the QM9 dataset and the unconstrained baseline. In fact, at $K = 0$ — i.e., forcing all molecules to be acyclic — the FCD rises to 10.66, indicating a large distributional divergence. Despite this, the model maintains 99.63% Validity, although Uniqueness drops to 81.86% due to the restricted chemical space. As K increases (constraints less stringent), the metrics gradually align with the unconstrained version. Validity is consistently high, while Novelty is higher under strict constraints, reaching 90.68% at $K = 0$, suggesting that acyclic constraints encourage exploration beyond the training distribution. Novelty stabilizes at $\approx 88.5\%$, indicating no loss in novel outputs with moderate constraints. Restricting molecules to contain only a small number of cycles decreases the number of generated graphs with disconnected components. Table 2 reports the percentage of outputs with the exact ring count for each K. As K increases, the distribution shifts toward higher ring counts, approaching the unconstrained setting, i.e., matching exactly ConStruct [11]. In summary, loose ring-count constraints (large K) barely affect the distribution, whereas strict ones (small K) remove the high-ring tail and shift its mass to lower bins.

Table 3 shows the performances when forcing all rings to be at most of size L, while Figure 2 reports some qualitative examples. Additional results are in

[2]https://github.com/ranaislek/ConStruct-Thesis
[3]`https://www.rdkit.org/`

ESANN 2026 proceedings, European Symposium on Artificial Neural Networks, Computational Intelligence and Machine Learning. Bruges (Belgium) and online event, 22-24 April 2026, i6doc.com publ., ISBN 9782875870964. Available from http://www.i6doc.com/en/.

K	FCD	Uniq. (%)	Novelty (%)	Validy (%)	Disc. (%)
0	10.66 ± 0.02	81.86 ± 0.31	90.68 ± 0.23	99.63 ± 0.03	1.37 ± 0.13
1	1.81 ± 0.01	93.19 ± 0.31	87.29 ± 0.13	99.27 ± 0.14	1.83 ± 0.16
2	1.41 ± 0.01	94.51 ± 0.37	87.22 ± 0.17	99.17 ± 0.14	1.90 ± 0.15
3	1.08 ± 0.01	96.54 ± 0.17	88.55 ± 0.10	98.97 ± 0.11	2.10 ± 0.09
9	1.03 ± 0.01	97.33 ± 0.14	88.74 ± 0.09	98.67 ± 0.11	2.14 ± 0.13
Unc.	1.03 ± 0.01	97.41 ± 0.15	88.72 ± 0.09	98.45 ± 0.09	2.14 ± 0.12

Table 1: Core metrics with ring count $\leq K$

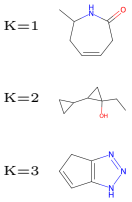

Fig. 1: Examples

K	0-ring	1-ring	2-rings	3-rings	4-rings	5-rings	6-rings	7-rings	8-rings	9-rings	≥ 10
0	100.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000
1	10.996	89.004	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000
2	10.996	72.252	16.752	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000
3	10.996	39.966	10.496	38.542	0.000	0.000	0.000	0.000	0.000	0.000	0.000
9	10.996	39.966	7.234	25.012	2.224	0.012	7.904	6.462	0.180	0.010	0.000
Unc.	10.996	39.966	7.234	25.010	2.072	0.012	6.222	5.130	0.138	3.220	0.000
QM9	10.233	39.323	7.369	24.453	2.706	0.026	5.841	6.606	0.329	0.002	3.114

Table 2: Ring count distributions for constraint ring counts $\leq K$ (%)

Appendix B. Small L values cause noticeable distribution shifts. At $L = 3$, FCD rises to 5.18 from 1.03 of the unconstrained version. High FCDs indicate the model is being pushed into parts of chemical space underrepresented in QM9. Validity still remains high, and Novelty actually increases, indicating the model explores novel chemical structures when constrained to small rings. As L relaxes, the model's outputs align with the unconstrained version. Table 4 shows the conditional distribution of the largest ring lengths in the outputs when enforcing a given L. The ring-length and ring-count projectors perfectly enforce the constraint in all cases.

4 Conclusions and Future Work

This study investigates controlling molecular structure during generation without model retraining. Experiments on the QM9 dataset show that sampling-time projectors achieve perfect compliance with ring constraints while maintaining high chemical validity. Future work will expand the framework to include

L	FCD	Uniq. (%)	Novely (%)	Validy (%)	Disc. (%)
0	10.66 ± 0.02	81.86 ± 0.31	90.68 ± 0.23	99.63 ± 0.03	1.37 ± 0.13
3	5.18 ± 0.03	90.76 ± 0.44	92.35 ± 0.30	99.45 ± 0.06	1.57 ± 0.13
4	3.95 ± 0.01	94.39 ± 0.28	92.95 ± 0.30	98.88 ± 0.06	1.72 ± 0.10
9	1.03 ± 0.01	97.41 ± 0.15	88.72 ± 0.09	98.45 ± 0.09	2.14 ± 0.12
Unc.	1.03 ± 0.01	97.41 ± 0.15	88.72 ± 0.09	98.45 ± 0.09	2.14 ± 0.12

Table 3: Core metrics with max ring length $\leq L$

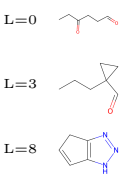

Fig. 2: Examples

ESANN 2026 proceedings, European Symposium on Artificial Neural Networks, Computational Intelligence and Machine Learning. Bruges (Belgium) and online event, 22-24 April 2026, i6doc.com publ., ISBN 9782875870964. Available from http://www.i6doc.com/en/.

L	0-atom	3-atoms	4-atoms	5-atoms	6-atoms	7-atoms	8-atoms	9-atoms	≥ 10
0	100.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000
3	54.342	45.658	0.000	0.000	0.000	0.000	0.000	0.000	0.000
4	38.518	29.278	32.204	0.000	0.000	0.000	0.000	0.000	0.000
9	10.996	10.768	10.724	26.590	19.022	11.722	7.956	2.222	0.000
Unc.	10.996	10.768	10.724	26.590	19.022	11.722	7.956	2.222	0.000
QM9	10.233	10.591	11.073	26.934	20.457	11.526	6.921	2.266	0.000

Table 4: Maximum ring length distributions for constraint ring length $\leq L$ (%)

chemical property constraints, such as Lipinski's Rule of Five, requiring more advanced projectors to evaluate properties like molecular weight and logP.

References

[1] Diederik P Kingma and Max Welling. Auto-encoding variational bayes. *arXiv preprint arXiv:1312.6114*, 2013.

[2] Davide Rigoni, Nicolò Navarin, and Alessandro Sperduti. Conditional constrained graph variational autoencoders for molecule design. In *IEEE SSCI*, pages 729–736. IEEE, 2020.

[3] Davide Rigoni, Navarin Nicolo, and Sperduti Alessandro. A systematic assessment of deep learning models for molecule generation. In *ESANN*, pages 547–552, 2020.

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

[5] Maksim Kuznetsov and Daniil Polykovskiy. Molgrow: A graph normalizing flow for hierarchical molecular generation. In *AAAI*, volume 35, pages 8226–8234, 2021.

[6] Davide Rigoni, Sachithra Yaddehige, Nicoletta Bianchi, Alessandro Sperduti, Stefano Moro, and Cristian Taccioli. Tumflow: An ai model for predicting new anticancer molecules. *International Journal of Molecular Sciences*, 25(11):6186, 2024.

[7] Xiaohui Chen, Jiaxing He, Xu Han, and Li-Ping Liu. Efficient and degree-guided graph generation via discrete diffusion modeling. *Proc. of Machine Learning Research*, 2023.

[8] Han Huang, Leilei Sun, Bowen Du, and Weifeng Lv. Conditional diffusion based on discrete graph structures for molecular graph generation. In *AAAI*, 2023.

[9] Alex Morehead and Jianlin Cheng. Geometry-complete diffusion for 3d molecule generation and optimization. *Communications Chemistry*, 7(1):150, 2024.

[10] Manuel Ruiz-Botella, Marta Sales-Pardo, and Roger Guimerà. A collaborative constrained graph diffusion model for the generation of realistic synthetic molecules. *arXiv preprint arXiv:2505.16365*, 2025.

[11] Manuel Madeira, Clement Vignac, Dorina Thanou, and Pascal Frossard. Generative modelling of structurally constrained graphs. *NeurIPS*, 37:137218–137262, 2024.

[12] Guy W Bemis and Mark A Murcko. The properties of known drugs. 1. molecular frameworks. *Journal of medicinal chemistry*, 39(15):2887–2893, 1996.

[13] Christopher A Lipinski, Franco Lombardo, Beryl W Dominy, and Paul J Feeney. Experimental and computational approaches to estimate solubility and permeability in drug discovery and development settings. *Advanced drug delivery reviews*, 23(1-3):3–25, 1997.

[14] Raghunathan Ramakrishnan, Pavlo O Dral, Matthias Rupp, and O Anatole Von Lilienfeld. Quantum chemistry structures and properties of 134 kilo molecules. *Scientific data*, 1(1):1–7, 2014.

[15] Nathan Brown, Marco Fiscato, Marwin HS Segler, and Alain C Vaucher. Guacamol: benchmarking models for de novo molecular design. *JCIM*, 59(3):1096–1108, 2019.