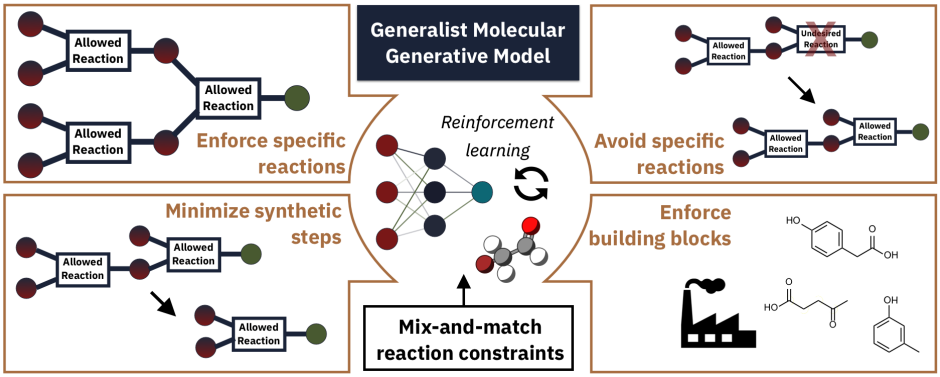
Generative molecular design with steerable and granular synthesizability control

Víctor Sabanza-Gil*[1,2,3], Jeff Guo*[1,2], Zlatko Jončev[1], Jeremy Luterbacher[2,3], Philippe Schwaller[1,2]

[1]Laboratory of Artificial Chemical Intelligence, EPFL (Switzerland), [2]NCCR Catalysis, EPFL (Switzerland), [3]Laboratory of Sustainable and Catalytic Processes, EPFL (Switzerland)

victor.sabanzagil@epfl.ch

Molecular generative design is crucial for identifying and developing novel molecules with targeted physicochemical and biological properties [1]. Despite recent progress facilitated by advancements in generative artificial intelligence, translating computationally designed molecules into experimentally viable candidates remains challenging [2]. A critical bottleneck in the validation process is the requirement to synthesize these computationally proposed molecular structures in laboratory settings. Current generative models often overlook synthetic feasibility, resulting in molecules that, while theoretically promising, present substantial practical difficulties in experimental synthesis [3]. In this work, we introduce a generative molecular design framework that allows selective and granular control over multiple elements in the predicted synthetic routes of the generated molecules [4]. By combining a pretrained language model with a policy gradient optimization algorithm coupled to a retrosynthesis model, the model learns to enforce the presence of pre-defined reactions and/or building blocks, or to avoid the inclusion of selected transformations in the generated molecule's final synthetic route. The steerable control over the synthetic parameters allows high flexibility and is more amenable to experimental validation. Overall, this method opens a new capability for accelerating the discovery of promising molecules with tailored properties



[1] Du, Y., Jamasb, A. R., Guo, J., et al. (2024). Machine learning-aided generative molecular design. *Nature Machine Intelligence, 6*, 589–604.
[2] Bilodeau, C., Jin, W., Jaakkola, T., Barzilay, R., & Jensen, K. F. (2022). Generative models for molecular discovery: Recent advances and challenges. *WIREs Computational Molecular Science, 12*:e1608.
[3] Gao, W., & Coley, C. (2020). The synthesizability of molecules proposed by generative models. *Journal of Chemical Information and Modeling, 60* (12), 5714–5723.
[4] Sabanza-Gil, V., Guo. J, Jončev, Z., Luterbacher, J., Schwaller, P.; Generative molecular design with steerable and granular synthesizability control, https://arxiv.org/abs/2505.08774#