Supplemental Material

Gotta be *SAFE*: A New Framework for Molecular Design

Emmanuel Noutahi* Cristian Gabellini
Valence Labs *Valence Labs*

Michael Craig Jonathan S.C. Lim Prudencio Tossou
Valence Labs *MILA / Valence Labs* *Valence Labs*

1 Additional figures

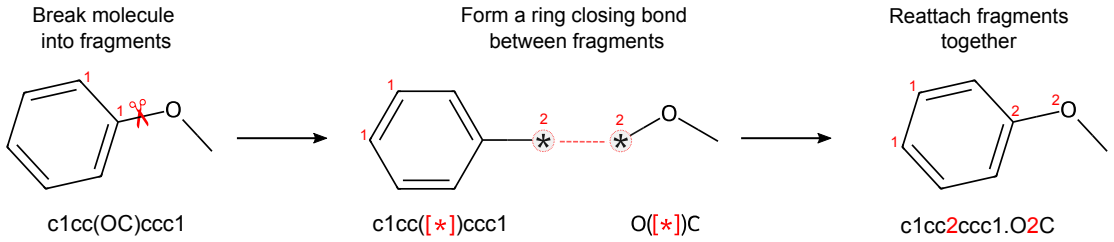


Figure S1: Example encoding of a SMILES string into a SAFE representation. The left panel shows the breaking a bond by the BRICS algorithm. The middle panel shows the addition of attachment points and the ring closing bond connecting the two fragments. The right panel shows the reattached fragments and the final SAFE representation.

2 Comparison between SAFE and Group SELFIES

Both SAFE and Group SELFIES are molecular string representations capable of encoding fragments. In SAFE, fragments are denoted in groups of SMILES tokens separated by dots, while in Group SELFIES, fragments are tokens from a pre-defined grammar of chemical motifs (such as a token representing a toluene fragment). To compare their performance, we trained SAFE-GPT-20M and GSELFIES-GPT-20M on the MOSES dataset and evaluated them in *de novo* molecule generation. We generated 10,000 molecules from each model and analyzed the distribution of molecular properties within these two sets to assess their efficacy.

As seen in Figure S4, molecules generated by SAFE-GPT-20M tend to exhibit higher QED (Quantitative Estimate of Drug-likeness) scores, indicating higher degree of drug-likeness, and lower SA (Synthetic Accessibility) scores, indicating better synthetic feasibility.

We further investigate the differences in the molecules generated by the two models by comparing the distributions of the largest ring size of each molecule. As shown on Figure S5, the model trained using the Group SELFIES notation frequently generate molecules with large and unstable ring structures.

We did not make further experiments and comparisons for the fragment-constrained generation tasks (such as linker design and scaffold decoration) as non-trivial adaptations would have to be made to the Group SELFIES notation, training process and molecular sampling, which could be explored in future works.

*Corresponding author. Email: emmanuel@valencelabs.com

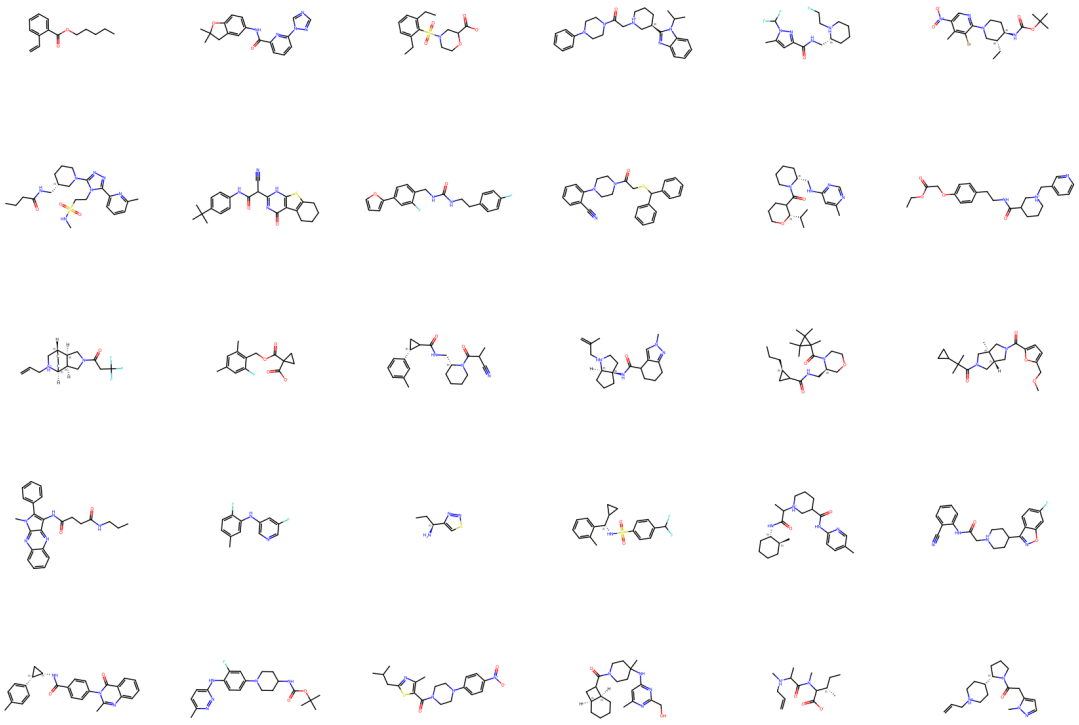


Figure S2: Randomly selected samples of *de novo* generated molecules using SAFE.

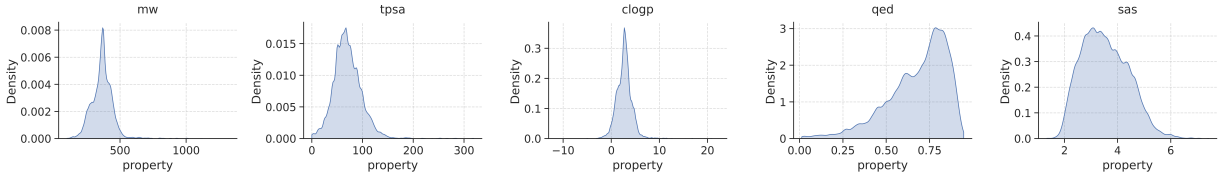


Figure S3: The molecular property distribution for 10,000 molecules generated with SAFE-GPT demonstrates that SAFE-GPT can generate molecules with diverse physicochemical properties, spanning beyond traditional drug-like molecules.

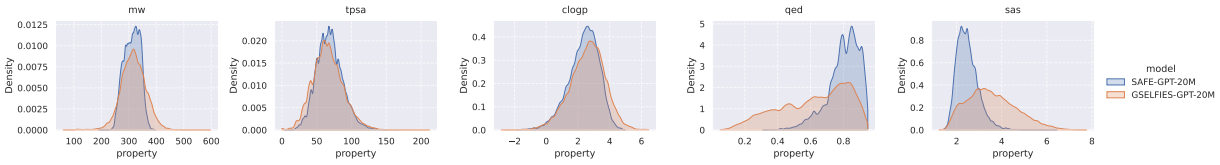
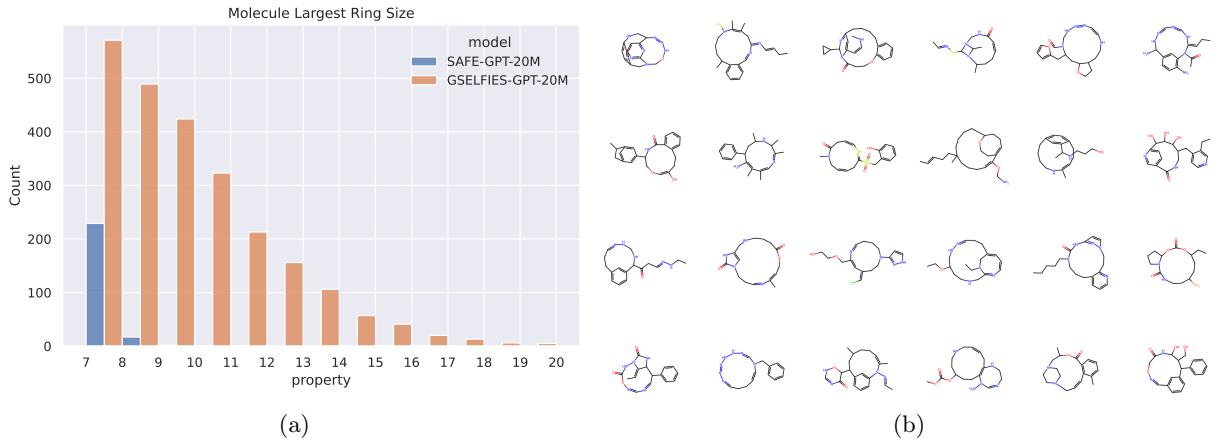


Figure S4: The molecular property distribution of molecules generated with SAFE-GPT-20M compared against molecules generated with GSELFIES-GPT-20M.

3 Optimizing CNS penetration for EGFR inhibitors

Most existing small molecule treatments struggle to effectively penetrate the central nervous system (CNS) due to difficulties in breaching the blood-brain barrier (BBB). Notably, three well-known EGFR inhibitors (afatinib, gefitinib, and erlotinib), all sharing the same scaffold, exhibit generally low CNS penetration rates, with reported values respectively falling below 1%, in the range of 1%–3%, and in the range of 3%–6%. The ability of a small molecule to penetrate the CNS is often associated with specific physicochemical properties such as CLogD, TPSA, and Molecular Weight. Various scoring systems have been developed to assess this ability. Notably, our findings indicate a correlation between the CNS MPO



(a) (b)

Figure S5: Distribution of the largest ring size (> 6 atoms) count in molecules generated with SAFE-GPT-20M compared against molecules generated with GSELFIES-GPT-20M. **(a)** GSELFIES-GPT-20M tends to generate molecules with ring sizes exceeding 8 atoms more frequently. **(b)** Examples of large ring molecules produced by GSELFIES-GPT-20M, illustrating their tendency towards non-druglike and chemically unstable structures.

score [Wager et al., 2016] and the experimental penetration rates for these three EGFR inhibitors.

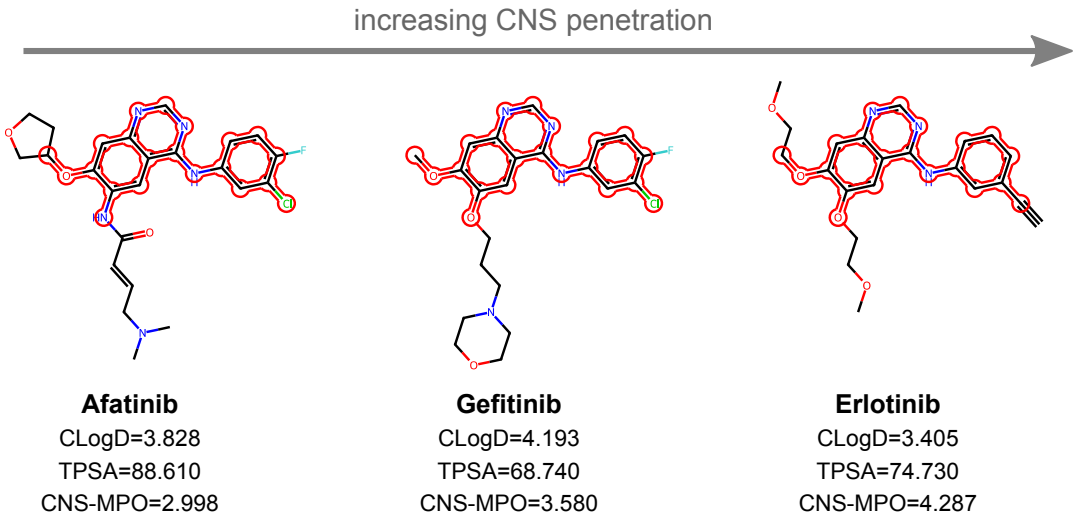


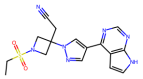
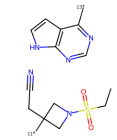
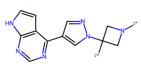
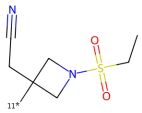
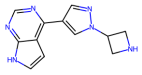
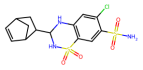
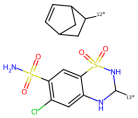
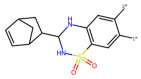
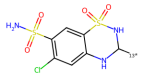
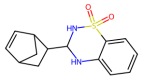
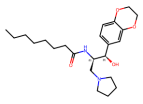
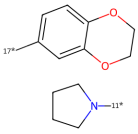
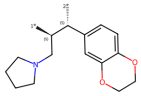
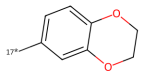
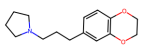
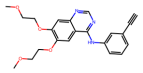
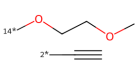
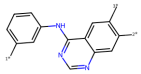
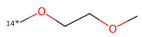
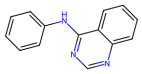
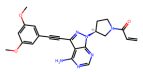
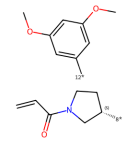
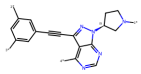
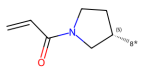
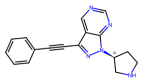
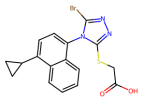
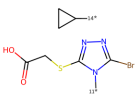
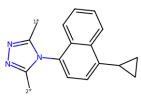
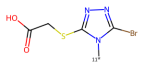
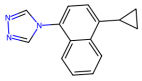
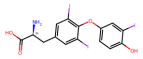
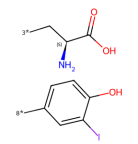
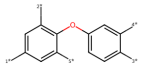
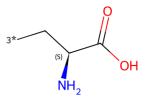
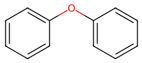
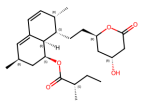
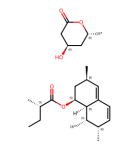
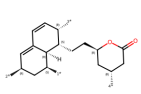
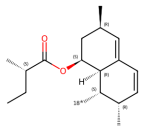
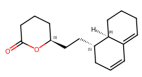
Figure S6: Existing EGFR inhibitors and their CNS profile

4 Fragment-constrained design results

We uses a set of 10 drugs, including **Cyclothiazide**, **Maribavir**, **Spirapril**, **Baricitinib**, **Eliglustat**, **Erlotinib**, **Futibatinib**, **Lesinurad**, **Liothyronine**, and **Lovastatin**. These drugs were chosen as the basis for our fragment-constrained generative design tasks. From each drug, we extracted the main scaffold with attachment points, fragments that serve as side chains, a starting motif, and a core substructure. These components were then respectively used as input for scaffold decoration, linker design / scaffold morphing, motif extension, and superstructure generation, each with its specific objective. The details of the selected drugs and their corresponding inputs for each task can be found in Table S1. It should be noted that linker design and scaffold morphing are two very similar tasks that share the same inputs. In our implementation, the only difference between them lies in the constraints imposed during sampling. For linker design, we employ a constrained beam search to ensure the presence of every fragment in the final molecules. In contrast, for scaffold morphing, new molecules are generated from each fragment with

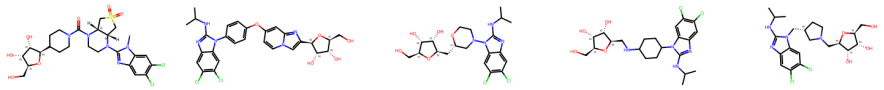
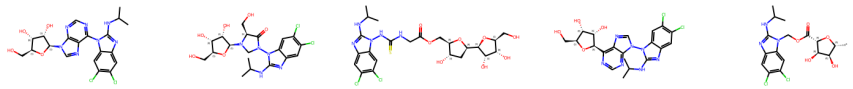
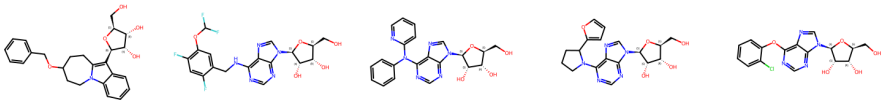
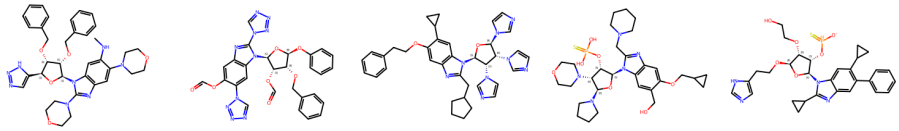
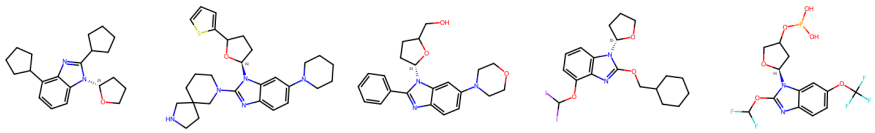
connectivity constraints, after which the scaffold is inferred and linked to the other fragments.

Table S1: List of 10 known drugs and corresponding inputs used by SAFE-GPT for the fragment-constrained benchmark.

Name	Structure	Linker Design*	Scaffold Decoration	Motif Extension	Superstructure
BARICITINIB					
CYCLOTHIAZIDE					
ELIGLUSTAT					
ERLOTINIB					
FUTIBATINIB					
LESINURAD					
LIOTHYRONINE					
LOVASTATIN					
MARIBAVIR					
SPIRAPRIL					

* the linker design and scaffold morphing task share the same input fragments.

Table S2: Examples of generated samples under fragment-constraints for the Maribavir structure

Task	**Generated samples**
Linker design	
Scaffold morphing	
Motif extension	
Scaffold decoration	
Superstructure	

References

Travis T Wager, Xinjun Hou, Patrick R Verhoest, and Anabella Villalobos. Central nervous system multiparameter optimization desirability: application in drug discovery. *ACS chemical neuroscience*, 7 (6):767–775, 2016.