

Table of Contents for Supporting Information

Article Title: Molecular Sonification for Molecule to Music Information Transfer

Contents:

Supplementary Figure S1: Key Occurrence in DrugBank

Supplementary Figure S2: SELFIES encoding/decoding algorithm

Supplementary Figure S3: Large molecules encoded by SELFIES

Supplementary Figure S4: Four-point interpolation of SELFIES encoded drugs

Journal: Submitted to Digital Discovery

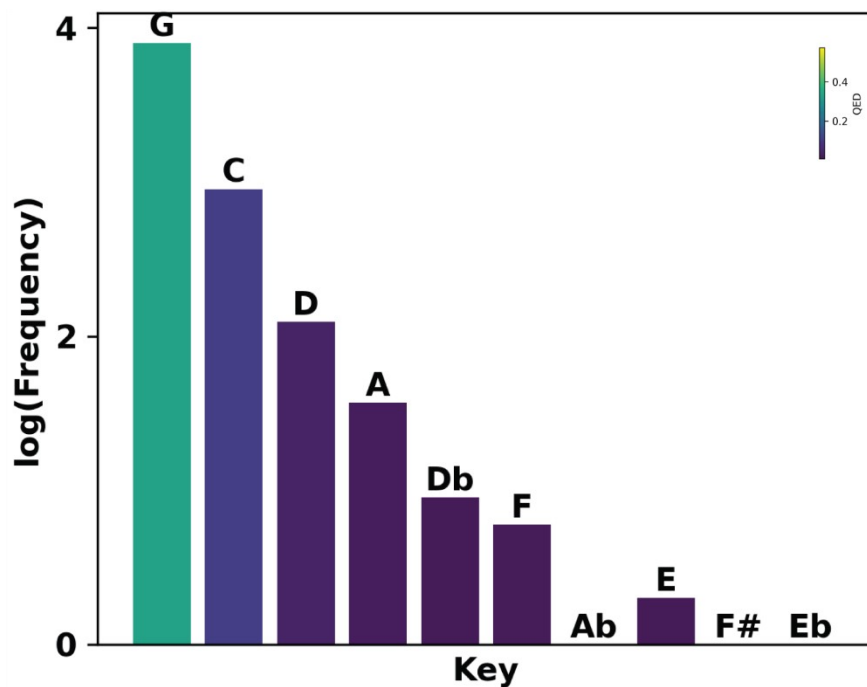


Figure S1 – frequency of occurrence of the key of drugs converted to music in DrugBank, colored by average QED.

SELFIES Token	MIDI Shift (Major Scale)	MIDI Value in key of G	MIDI Value in key of C
[C]	+0	55	48
[=C]	+2	57	50
[Branch1_1]	+4	59	52
[Ring1]	+5	60	53
[O]	+7	62	55
[Branch1_2]	+9	64	57
[N]	+11	66	59
[=O]	+12	67	60

Figure S2 – The MIDI shifts correspond to SELFIES tokens as ranked by their popularity in DrugBank. As such, lower MIDI shifts away from the encoded key of the molecule indicate the more popular the SELFIES token.

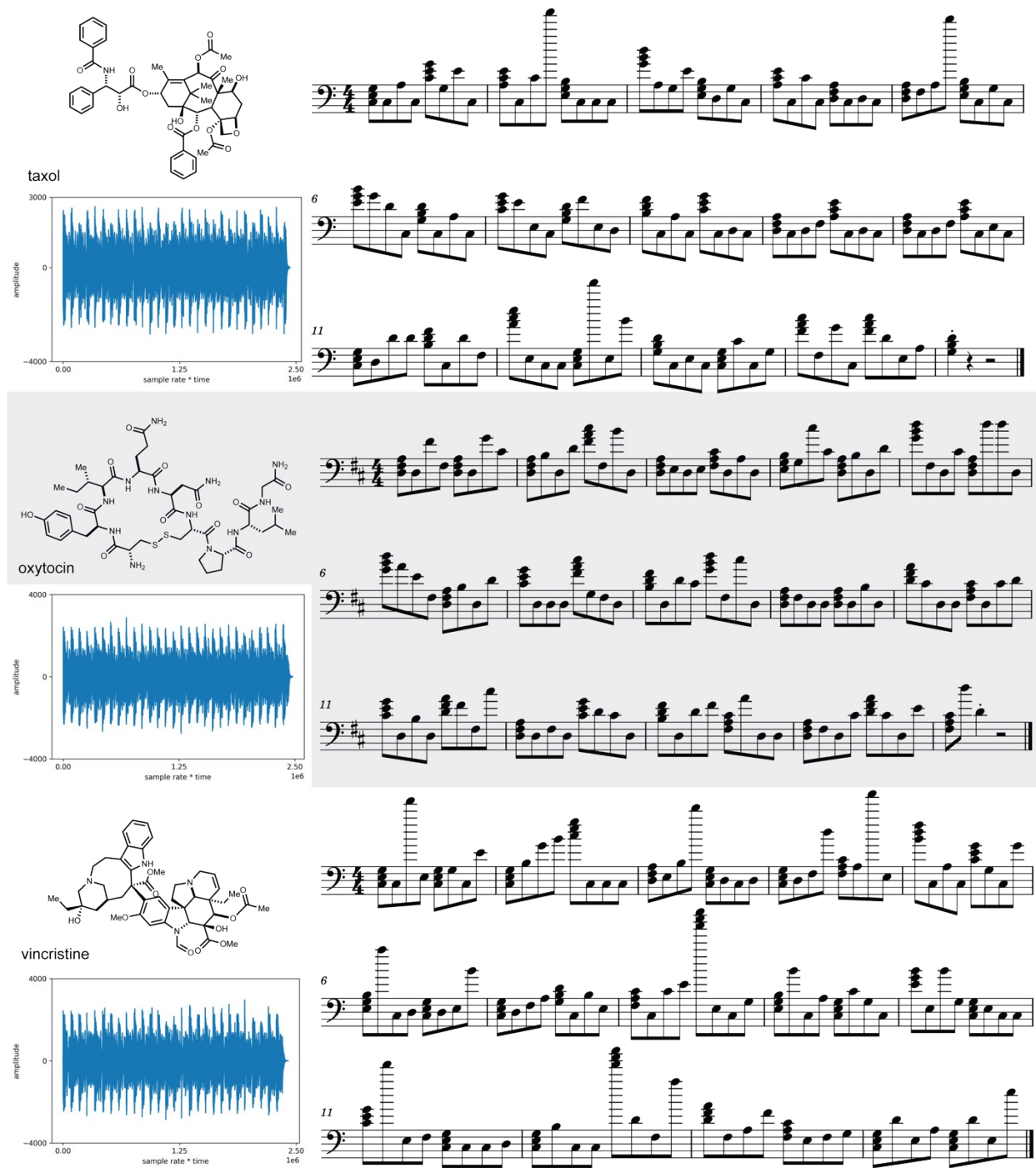


Figure S3 – SAMPLES examples of large molecules taxol, oxytocin, and vincristine.

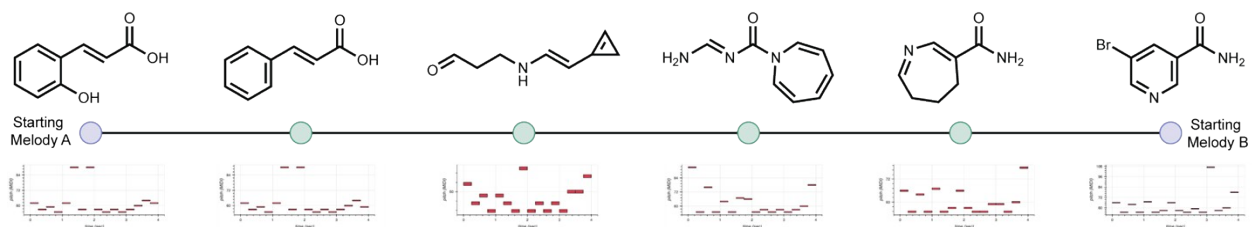


Figure S4 – Four point CROSSFADE interpolation between two drug molecules.