Molecular Sonification for Molecule to Music Information Transfer.

Babak Mahjour¹, Jordan Bench^{1,2}, Rui Zhang², Jared Frazier^{1,3}, Tim Cernak^{1,2*}

- 1. Department of Medicinal Chemistry, University of Michigan, Ann Arbor, MI, USA
- 2. Department of Chemistry, University of Michigan, Ann Arbor, MI, USA
- 3. Department of Computer Science, Middle Tennessee State University, Murfreesboro, TN, USA

*Corresponding Author: Tim Cernak; Email: tcernak@med.umich.edu

Abstract

Organic chemical structures encode information about a molecule's atom and bond arrangement. The most established way to encode a molecular structure is through line drawing, although other representations based on graphs, strings, one-hot encoded labels, or fingerprint arrays are critical to the computational study of molecules. Here we show that music is a highly dimensional information storage medium that can be used to encode molecular structure. The resultant method allows a molecular structure to be heard as a musical composition, where the key of the music is based on the molecular properties and the melody is based on the atom and bond arrangement. This allows for a molecular generation approach that leverages modern artificial intelligence tactics for music generation.

Introduction

The representation of chemical structures is critical to the study and invention of functional molecules. Organic molecules are classically described as line drawings¹, where all atoms and their corresponding bonds are drawn on paper or on a computer. Other simple molecular representations include molecular formulae, IUPAC names or CAS numbers, which require little memory and are machine readable, but carry minimal information. Molecules can also be represented as graphs, with atoms as nodes and bonds as edges. By encoding atomic coordinates and connectivities line by line, the topology of molecules can be embedded as a graph on a computer for rendering, editing and analysis. The transmission of molecular information into machine-readable formats has invited new molecular structure representations, such as SMILES², SMARTS³, InChI keys⁴, DeepSMILES⁵, and SELFIES⁶. These representations are cheap to store in memory and provide valuable structural information for rapid lookup and comparison. While these aforementioned representations have been useful for inputting molecules into computers, and encoding structural and stereochemical information, they are one-dimensional string representations that are more difficult for human users to interpret and interact with than the classic line drawing representation of molecules. To adapt them for machine learning and data science algorithms, SMILES and other strings are typically converted to vector representations via molecular fingerprints such as Extended Connectivity Fingerprints⁷ (ECFP), Morgan Fingerprints⁸, atom-pair fingerprints⁹, and others. This dimensionality expansion is a core tactic in the analysis of virtual chemical libraries or predictions of molecular properties. Other high-dimensionality fingerprint representations, such as physics-based descriptors¹⁰ or physicochemical descriptors^{11,12}, are also common. While computers can easily parse molecular information from these representations, interactivity with human users is difficult with the fingerprint-based information media. In addition, once converted to such a fingerprint, the molecule is typically no longer uniquely revertible to its atom-bond representation.¹³

Music is a high-dimensional information storage medium that maximizes both human and computer interactivity, interpretability, and creativity. We considered that music could be used for storage of molecular information. The encoding of molecules as music is particularly intriguing since the multiple dimensions of music can allow encoding of many molecular properties.¹⁴ Music is also highly interactive both for humans and for computers. Musicians can control many parameters that can embed information about a molecule, such as tempo, rhythm, notes, key, instrument, effects, etc. If molecules could be

encoded as music, opportunities would emerge for visual-to-audio sensory substitution, for instance providing blind chemists new ways to interact with molecules.¹⁵ Contemporary chemistry and drug discovery leverage artificial intelligence¹⁶ (AI) and there has meanwhile been an explosion of AI methods in the study and creation of music¹⁷, so we were excited by the prospects of merging modern chemistry machine learning (ML) techniques with recent ML techniques for music. Our initial impetus was to explore how music could be used as a creative medium to generate new molecules, but in the course of our studies we have learned that molecules likewise can provide an inspiration and creative outlet for the generation of new music.

Sonification is the encoding of non-musical information as music and provides a means to encode information in many musical dimensions, while simultaneously providing a new means of interactivity.¹⁸ A variety of information sources have been sonified, such as visual art¹⁹, the architecture of spider webs²⁰, infrared spectra of chemicals²¹, amino acid sequences²², air pollution²³, fire²⁴, and many more²⁵⁻³². The SELFIES representation provided a viable input for molecular sonification, both for the encoding of molecules into a melody and the construction of new molecules via performance on a musical instrument such as the piano. We developed a workflow for transferring molecules into music, and vice versa, which we call Sonic Architecture for Molecule Production and Live-Input Encoding Software, or SAMPLES. A python script enables direct interactivity with a piano keyboard via the musical instrument digital interface (MIDI) format.

Results

Encoding: To create a melody based on a molecular structure, the key and the sequence of notes are derived from its physicochemical properties and its SELFIES sequence, respectively. To determine the key, the physicochemical properties of a molecule — such as logP, molecular weight, and number of hydrogen bond donors and acceptors — are summed, and the final number is projected into the integer space between 1 and 12, with each bin corresponding to a particular key. The sequence of notes is determined from a one-to-one mapping between the SELFIES token of the molecule and multi-octave steps in the major scale (see Supporting Information for the specific mappings and key distribution used in this study). By adding the MIDI value of the melody's key to the MIDI shifts that correspond to notes of the major scale (derived from the SELFIES tokens of the molecule), the final melody is produced. Every fourth note was converted to a major chord to increase the texture of the music.



Figure 1. Workflow for SAMPLES. Molecules are first assigned a musical key based on aggregate chemical properties, then converted into a sequence of notes based on SELFIES encoding. MusicVAE is trained on a collection of sonified molecules to formulate the chemical/musical latent space. The latent embedding of molecular music can then be sampled, decoded by the MusicVAE decoder, then converted back into a molecular structure by SAMPLES

Decoding: The MIDI shifts are reverse calculated for each key and converted into a molecular structure. As such, multiple structures are generated (one for each key) for the same MIDI sequence. Each structure is then hashed into a key using the original key encoding algorithm. If the hashed key matches the key used in the reverse calculation, the molecular structure is decoded. It is guaranteed that at least one decoding key will match a hashed key for any MIDI generated from SAMPLES.

A demonstration of the SAMPLES encoding function is shown in Figure 2. Ammonia (1) appears as a single note while benzene (2) generates a slightly more complex musical composition. The unity of these two molecules produces aniline (3) whose musical sequence highly resembles the concatenation of the two musical sequences of 1 and 2. Expansion of 3 into indole (4) creates a slightly more complex melody owing to both the increased molecular size and the additional information content required to describe a ring fusion between the 5- and 6-membered rings. In the reverse direction, songs are readily translated to

molecules, such as **5**, which is produced from the song "Twinkle, twinkle little star" when played in the key of D flat.



Figure 2. SAMPLES translates molecules into music. **A**. The generation of increasingly complex molecules from **1** through **4** corresponds to increasing musical complexity. Each line shows the molecular structure, the corresponding musical score, and a waveform of the MIDI output. Audio recordings are available in the Supporting Information and can be quickly retrieved by scanning the QR code with a mobile device. **B**. In the reverse direction, the song "Twinkle, Twinkle Little Star" produces molecule **5**.

SAMPLES is readily scaled to more complex and drug-like molecules. Tolmetin (6) and ketolorac (7) create a rich and textured musical composition. Meanwhile, tabersonine (8) and vindoline (9) provide complex melodies. Scaling to large complex molecules, such as taxol, oxytocin or vincristine (see Supporting Information) required no modifications and generated nuanced euphonic melodies.



Figure 3. SAMPLES is amenable to encoding complex molecules. The pair of similar molecules 6 and 7 have SAMPLES compositions that are distinct from another similar pair of molecules 8 and 9.

Case Studies

To showcase the utility of this novel algorithm, four experimental case studies are presented. Using our approach, molecular properties can be heard. For instance, the songs generated from molecules that pass the Lipinski rules³³ can be auditorily distinguished from those that fail the Lipinski rules based on the musical key. This is largely because the molecule's aggregate physicochemical properties were hashed to the musical key, with the most popular physicochemical property fingerprints from the pharmaceutical database DrugBank hashing to the most popular song keys from the music database Spotify.³⁴ The concept of molecular similarity is of high importance to molecular invention, such as in selecting molecules with comparable functional properties for drug discovery. We were curious to explore if

SAMPLES generated from molecules with high Tanimoto similarity³⁵ (fingerprint based) would sound similar, appreciating that both molecular similarity and musical similarity are difficult to define.³⁶ Indeed, we deemed the SAMPLES of codeine (**10**) and morphine (**11**) to sound similar to each other while the SAMPLES of sulfamethoxazole (**12**) and sulfadoxin (**13**) likewise sound similar, while the pair of **10** and **11** sounded distinct from the pair of **12** and **13** (Figure 4).



Figure 4. tSNE embedding of 11,159 drugs from DrugBank (2,048-bit Morgan Fingerprints of radius 2), colored by their SAMPLES musical key. Similar molecules 10 and 11 have SAMPLES outputs that are distinct from other similar molecules 12 and 13.

Our second experiment investigates the generation of molecules via modification of the music domain. A key motivator for our research was the ability to generate new molecules through the interactivity of a piano keyboard, or other musical hardware or software. This was made possible in SAMPLES through the application of SELFIES, which enable editing of string bits while consistently producing valid molecular structures. Thus, starting from morphine (11), the musical score could be modified one note at a time (Figure 5) to generate new chemical structures 14–16 bearing a clear relationship to 11 but with noticeably modified bond and atom architecture. Note that SAMPLES may generate undefined stereocenters.



Figure 5. Molecular editing in SAMPLES generates distinct but related molecules. The manual editing of single notes in the SAMPLES of 11 leads to 14, 15, or 16.

Having demonstrated the feasibility of molecular generation using SAMPLES, we explored the ability of modern machine learning methods developed for music generation as tools for molecule generation. In this third case study, we applied the melody mixing function of MusicVAE³⁷ using MIDI melodies derived from SAMPLES as inputs. Using MusicVAE, two melodies could be blended to generate an interpolated melody, and that new melody could be translated back to a molecular structure using SAMPLES, thus creating a new molecule that was a "blend" of the two input molecules (Figure 6). We call this function CROSSFADE. The blending of musical compositions is an established practice, with considerable

hardware and software to support the musical blending process. While algorithms that generate new molecules by blending the structures or properties of input molecules are known³⁸, we are intrigued by the interactivity offered by CROSSFADE. As an example, glutamic acid (**17**) and acetylcholine (**18**) were CROSSFADEd to produce **19**, **20** and **21** CROSSFADE to **22** and similar results are obtained for **23–28**. A four-step interpolation is shown in the Supporting Information.



Figure 6. CROSSFADE merges SAMPLES with the melody mixing function of MusicVAE to create interpolated molecules based on two input molecules.

As a final experiment, to take the editing of the molecules on the keyboard a step further, and to demonstrate the human-interactivity enabled by the SAMPLES algorithm, a human created a monophonic composition inspired by SAMPLES-generated music (Figure 7), which was decoded to molecule **29**. It was necessary to exert some human bias into the musical composition, based on the composer's knowledge of chemistry and SAMPLES, since generating a molecule that is as carboniferous as most drugs and natural products requires bias towards the key's tonic note, in this case C, since that is mapped to the carbon atom.



Figure 7. A human created music composition leading to 29.

Visual-to-audio sensory substitution allows users to interpret information sonically using sources not traditionally stored as sound.¹⁵ To test SAMPLES in this context, a survey was given to 32 undergraduate students wherein they were presented with the SAMPLES song of a single survey molecule and then asked to listen to additional SAMPLES songs of four distinct test molecules and select the most similar sounding song in a multiple-choice quiz format. Among the four test molecules, one molecule had a higher Tanimoto similarity to the survey molecule. For three of the four surveyed pairs, the students selected the molecule with highest Tanimoto similarity based on their interpretation of the similarity between SAMPLES songs. No visual information was given. Thus, students listening to the song of 12 selected 13 as having highest sonic similarity, instead of 33-35. Students listening to the song of 6 selected 7 as having highest sonic similarity, instead of 36-38. In both cases, the molecular Tanimoto similarity, based on Morgan fingerprints, was highest for the most popular musical survey response. A similar result was observed comparing 8 to 9, versus 39-40. However, the songs for morphine (10) and codeine (11), were not determined to have the highest musical similarity despite the apparent molecular similarity between these molecules. Instead, the song for compound 30 was chosen most frequently (17 out of 32 students) as being most similar to 10. Further studies are required to understand why 10's SAMPLES sounds more like **30** than **11**. Nonetheless, our collective results highlight that visual-to-audio sensory substitution may enable molecular interpretation when visual information cannot be used.



Figure 8. **A)** Survey results from 32 participants. Each participant was given the SAMPLES encoded melody of four survey molecules. For each survey molecule, without knowledge of the name or structure of any molecule, each participant was asked to choose the most similar melody from a selection of four other SAMPLES encoded drugs. Survey responses are cross examined against the Tanimoto similarity between each test molecule and survey molecule for each question. B) Structures of survey and test molecules for each question. One structure for each set of test molecules was chosen to have high similarity to the respective question's survey molecule to serve as the 'correct answer'.

Discussion

We report an alternative means of encoding organic molecules through music. The resultant melodies allow a human to interact with molecular structures through musical hardware and software via note editing, insertion and deletion, as well as produce molecular structures through original compositions. One transfer learning application for which the current study may be used is music generation. The motivation for machine learning for content generation is its generality, that is no formal grammar or rules must be specified for such a model to generate content.¹⁷ Transforming molecules into music provides a rich collection of musical data that can be used to train music generation models, as seen with MusicVAE. Particularly, sequence to sequence (seq2seq) models, such as recurrent neural networks³⁹, allow for the interconnection of domains containing data signals with variable lengths such as text, music, and machine-readable molecular representations based on structure. Seg2seg models can learn a fixed length embedding of variable length signals that can be used for classification tasks and direct mathematical comparison. For instance, word2vec⁴⁰ and GloVe⁴¹ provide pretrained word embeddings that have been learned from massive text corpuses such as Wikipedia or Twitter. In a molecular context, variational autoencoders have been used to learn the distribution of molecular features, such as SELFIES tokens, to provide a continuous embedding of molecular space.⁴² SAMPLES provides an avenue to directly connect molecules to content-generating machine learning models in the music domain. Computational exploration and interpolation within the melodies described herein is possible, generating new molecules that sound and look similar to existing molecules. This highlights the possibility of leveraging music-based artificial intelligence for molecular design.

Materials and Methods

Sonification and visualization code was written in Python (version 3.7.12). All Python dependencies were installed using pip, version 21.1.3. SELFIES (version 1.0.0) was utilized to encode molecules into string format. RDKIT (version 2021.9.2.1) was utilized to calculate physicochemical properties of molecules for key hashing. Magenta (version 2.1.3) provided tools to manipulate MIDI files and train MusicVAE. Fluidsynth (version 2.2.3, installed via apt-get) was used to convert MIDI into wav format. Music21 (version 5.5.0) was used to create and read MIDI files. Matplotlib (version 3.2.2) was used to create plots and graphs. Sklearn was used to calculate the tSNE dimensionality reduction. Drug structures were collected from DrugBank Release Version 5.1.8 (2021-01-03).

Acknowledgments

The authors thank the University of Michigan College of Pharmacy for startup funds, as well as the NSF-Interdisciplinary REU Program (NSF-1851985) for funding J.F.

Conflicts of interest

The Cernak Lab has received research funding or in-kind donations from MilliporeSigma, Relay Therapeutics, Janssen Therapeutics, Entos, Inc., SPT Labtech and Merck & Co., Inc. T.C. holds equity in Scorpion Therapeutics, and is a co-Founder and equity holder of Entos, Inc.

References

- 1. "Organic molecules are classically described as line drawings,": Kekulé, A. Sur la constitution des substances aromatiques. Bulletin de la Société Chimique de Paris, **3**, 98-110 (1865).
- 2. D. Weininger, SMILES, a chemical language and information system. 1. Introduction to methodology and encoding rules. J. Chem. Inf. Model. **28**, 31–36 (1988).

- 3. Daylight Chemical Information Systems, Daylight Theory: SMARTS A Language for Describing Molecular Patterns. Available at https://www.daylight.com/dayhtml/doc/theory/theory.smarts.html. Accessed October 23, 2021.
- 4. S. Heller, A. McNaught, S. Stein, D. Tchekhovskoi, I. Pletnev, InChI the worldwide chemical structure identifier standard. J. Cheminform. **5**, 7 (2013).
- N. O'Boyle, A. Dalke, DeepSMILES: An Adaptation of SMILES for Use in Machine-Learning of Chemical Structures. ChemRxiv (2018). https://doi.org/10.26434/chemrxiv.7097960 (accessed 4 November 2021).
- M. Krenn, F. Häse, A. Nigam, P. Friederich, A. Aspuru-Guzik, Self-referencing embedded strings (SELFIES): A 100% robust molecular string representation. Machine Learning: Science and Technology 1, 045024 (2020).
- 7. D. Rogers, M. Hahn, Extended-Connectivity Fingerprints. J. Chem. Inf. Model. 50, 742–754 (2010).
- 8. H. L. Morgan, The Generation of a Unique Machine Description for Chemical Structures-A Technique Developed at Chemical Abstracts Service. Journal of Chemical Documentation **5**, 107–113 (1965).
- R. E. Carhart, D. H. Smith, R. Venkataraghavan, Atom pairs as molecular features in structure-activity studies: definition and applications. Journal of Chemical Information and Computer Sciences 25, 64– 73 (1985).
- 10. S. Boobier, D. R. J. Hose, A. John Blacker, B. N. Nguyen, Machine learning with physicochemical relationships: solubility prediction in organic solvents and water. Nat. Commun. **11**, (2020).
- 11. P. S. Kutchukian, et al., Chemistry informer libraries: a chemoinformatics enabled approach to evaluate and advance synthetic methods. Chem. Sci. **7** (2016).
- 12. F. Pereira, J. Aires-de-Sousa, Machine learning for the prediction of molecular dipole moments obtained by density functional theory. J. Cheminform. **10**, 43 (2018).
- 13. T. Le, R. Winter, F. Noé, D.-A. Clevert, Neuraldecipher reverse-engineering extended-connectivity fingerprints (ECFPs) to their molecular structures. Chemical Science **11** (2020).
- 14. M. Kumbar, Musical Chemistry: Integrating Chemistry and Music. Journal of Chemical Education **84**, 1933 (2007).
- 15. V. Tóth, L. Parkkonen, "Autoencoding sensory substitution," Aalto University, Espoo, Finland. (2019).
- 16. Z. J. Baum, et al., Artificial Intelligence in Chemistry: Current Trends and Future Directions. Journal of Chemical Information and Modeling **61** (2021).
- 17. J.-P. Briot, G. Hadjeres, F.-D. Pachet, Deep Learning Techniques for Music Generation (Springer International Publishing, 2020) https://doi.org/10.1007/978-3-319-70163-9.
- 18. S. Barrass, et al., The Sonification Handbook, T. Hermann, A. Hunt, J. G. Neuhoff, Eds., 1st Ed. (Logos Publishing House, 2011).
- M. Muller-Eberstein, N. van Noord, Translating Visual Art Into Music in 2019 IEEE/CVF International Conference on Computer Vision Workshop (ICCVW), (IEEE, 2019) https://doi.org/10.1109/ICCVW.2019.00378.
- 20. Amercian Chemical Society, Making music from spider webs (2021). Available at https://phys.org/news/2021-04-music-spider-webs.html. Accessed October 23, 2021.
- 21. N. Garrido, et al., The Sound of Chemistry: Translating Infrared Wavenumbers into Musical Notes. Journal of Chemical Education **97**, 703–709 (2020).
- C.-H. Yu, Z. Qin, F. J. Martin-Martinez, M. J. Buehler, A Self-Consistent Sonification Method to Translate Amino Acid Sequences into Musical Compositions and Application in Protein Design Using Artificial Intelligence. ACS Nano 13 (2019).
- 23. P. Sarmento, Data from "Final Assignment for ECS7012P: Music and Audio Programming." Github. Available at https://github.com/otnemrasordep/sonification-bela. Deposited 5 November 2020.
- 24. M. Milazzo, M. J. Buehler, Designing and fabricating materials from fire using sonification and deep learning. iScience **24** (2021).
- 25. N. W. Tay, et al., Protein music of enhanced musicality by music style guided exploration of diverse amino acid properties. Heliyon **7** (2021).
- 26. Su, et al., Interactive exploration of a hierarchical spider web structure with sound. Journal on Multimodal User Interfaces (2021) https://doi.org/10.1007/s12193-021-00375-x.
- D. K. F. Meijer, I. Jerman, A. v. Melkikh, V. I. Sbitnev, "Biophysics of Consciousness: A Scale-Invariant Acoustic Information Code of a Superfluid Quantum Space Guides the Mental Attribute of the Universe" in (2021) https://doi.org/10.1007/978-981-15-7253-1_8.

- 28. T. J. Mitchell, et al., Towards molecular musical instruments in Proceedings of the 15th International Conference on Audio Mostly, (ACM, 2020) https://doi.org/10.1145/3411109.3411143.
- 29. M. Groß, Die Musik der Proteine. Nachrichten aus der Chemie 67 (2019).
- 30. S. L. Franjou, M. Milazzo, C.-H. Yu, M. J. Buehler, Sounds interesting: can sonification help us design new proteins? Expert Review of Proteomics **16** (2019).
- 31. Borgonovo, G. Haus, 1986. Sound Synthesis by Means of Two-Variable Functions: Experimental Criteria and Results. Computer Music Journal **10**, 57–71 (1986).
- 32. M. A. Garcia-Ruiz, J. R. Gutierrez-Pulido, An overview of auditory display to assist comprehension of molecular information. Interacting with Computers **18** (2006).
- 33. C. A. Lipinski, Lead- and drug-like compounds: the rule-of-five revolution. Drug Discov. Today Technol. **1**, 337–341 (2004).
- 34. K. Ning, Most used keys on Spotify. Available at https://forum.bassbuzz.com/t/most-used-keys-onspotify/5886 (Accessed 4 November 2021)
- 35. N. C. Chung, B. Miasojedow, M. Startek, A. Gambin, Jaccard/Tanimoto similarity test and estimation methods for biological presence-absence data. BMC Bioinformatics **20**, 644 (2019).
- 36. G. Maggiora, M. Vogt, D. Stumpfe, J. Bajorath, Molecular similarity in medicinal chemistry, Journal of Medicinal Chemistry **57**, 3186–3204 (2014).
- 37. Roberts, J. Engel, C. Raffel, C. Hawthorne, D. Eck, A Hierarchical Latent Vector Model for Learning Long-Term Structure in Music (2018).
- 38. J. Besnard, et al., Automated design of ligands to polypharmacological profiles. Nature **492**, 215–220 (2012).
- 39. M. H. S. Segler, T. Kogej, C. Tyrchan, M. P. Waller, Generating Focused Molecule Libraries for Drug Discovery with Recurrent Neural Networks. ACS Cent. Sci. 4, 120–131 (2018).
- 40. T. Mikolov, I. Sutskever, K. Chen, G. Corrado, J. Dean, Distributed Representations of Words and Phrases and their Compositionality (2013).
- J. Pennington, R. Socher, C. Manning, Glove: Global Vectors for Word Representation in Proceedings of the 2014 Conference on Empirical Methods in Natural Language Processing (EMNLP), (Association for Computational Linguistics, 2014) https://doi.org/10.3115/v1/D14-1162.
- 42. R. Gómez-Bombarelli, et al., Automatic Chemical Design Using a Data-Driven Continuous Representation of Molecules. ACS. Cent. Sci. **4**, 268–276 (2018).
- 43. Google LLC, S. Sidor, Magenta: Music and Art Generation with Machine Intelligence. Available at https://github.com/magenta/magenta. Deposited 8 September 2020.