IMPLICIT NEURAL REPRESENTATIONS OF MOLECULAR VECTOR-VALUED FUNCTIONS

Jirka Lhotka
LTS2
École Polytechnique Fédérale de Lausanne
Lausanne, Switzerland

Daniel Probst
LTS2
École Polytechnique Fédérale de Lausanne
Lausanne, Switzerland

Bioinformatics Group
Wageningen University & Research
Wageningen, The Netherlands
daniel.probst@wur.nl

1 INTRODUCTION

Molecules have various computational representations, including numerical descriptors, strings, graphs, point clouds, and surfaces. Each representation method enables the application of various machine learning methodologies from linear regression to graph neural networks paired with large language models (Serra et al., 2001; He et al., 2024).

To complement existing representations, we introduce the representation of molecules through vector-valued functions, or n-dimensional vector fields, that are parameterized by neural networks, which we denote molecular neural fields. Unlike surface representations, molecular neural fields capture external features and the hydrophobic core of macromolecules such as proteins. Compared to discrete graph or point representations, molecular neural fields are compact, resolution independent and inherently suited for interpolation in spatial and temporal dimensions (Sitzmann et al., 2019; Lombardi et al., 2019; Jiang et al., 2020). These properties inherited by molecular neural fields lend themselves to tasks including the generation of molecules based on their desired shape, structure, and composition, and the resolution-independent interpolation between molecular conformations in space and time. Here, we provide a framework and proofs-of-concept for molecular neural fields, namely, the parametrization and superresolution reconstruction of a protein-ligand complex using an auto-decoder architecture and the embedding of molecular volumes in latent space using an auto-encoder architecture.

2 RELATED WORK

Since their introduction by Sitzmann et al. (2020), implicit neural networks have been widely used in computer vision for processing images, videos and 3D scenes (Essakine et al., 2024). More recently, Mehta et al. (2021) have introduced modulated periodic activations, enabling generalizable high-fidelity implicit neural representations.

Recent approaches in representation learning of 3-dimensional molecular structures have mainly focused on graph-related methods (Zhou et al., 2022; Moon et al., 2023). For the representation of proteins, learning on the molecule's surfaces has proven successful (Gainza et al., 2020). (Gainza et al., 2020) have used implicit neural networks to generate the surfaces of metal-organic frameworks through signed-distance fields. (Kilgour et al., 2024) have introduced multi-type point clouds-representation learning based on an equivariant graph neural network.

3 METHODOLOGY

A Python library was implemented to sample a vector-valued function $f(x, y, z)$, mapping from $\mathbb{R}^3$ (molecular structure) to $\mathbb{R}^d$ (molecular properties). Each dimension in $\mathbb{R}^d$ represents one or more atom types (e.g. hydrogens, carbons, or oxygens) based on an atomic radius (van der Waals radius as

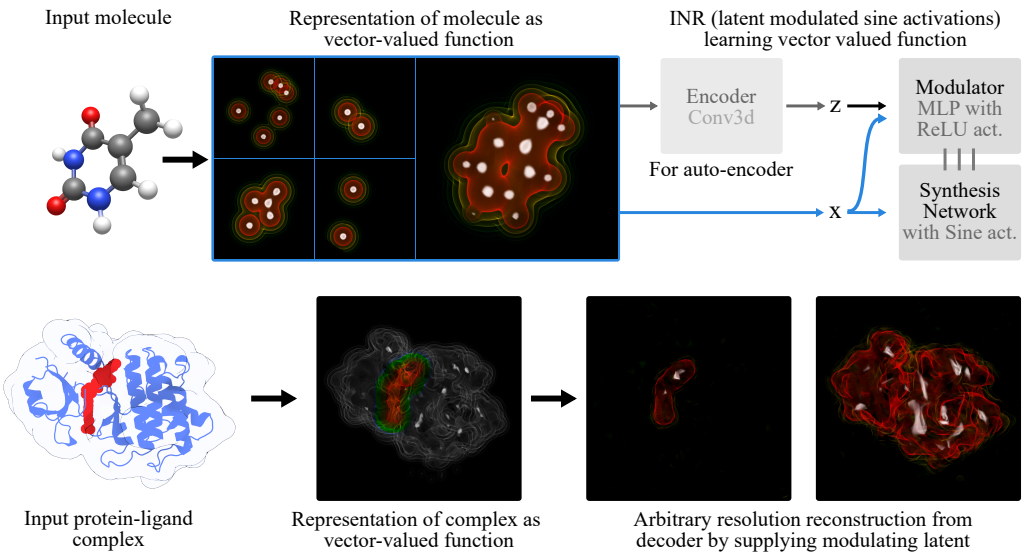

Figure 1: **Upper row**: Workflow and architecture overview of the presented methodology. For a given input molecule, a vector-valued function is computed and sampled on a 3D grid. Based on these samples, a neural network is trained that parametrizes the molecule using a modulated synthesis network with sine activations introduced by Mehta et al. (2021). **Lower row**: Based on the above workflow and architecture, a protein-ligand complex is encoded and parameterized by a modulated neural network in an auto-decoder setting. The two rightmost images show the reconstruction of the ligand and protein using the learned latent.

default) or an atomistic property (e.g. electronegativity or Awigner-Seitz electron density). Values in $\mathbb{R}^d$ are computed by the sum of radial basis functions scaled by a radius or property centered at the origin of each atom with given type(s) in the molecule. Section A.1 describes the radial basis function used. The result of this approach is shown in the second step of the upper row of Figure 1. The neural network architecture is based on an implementation of the modulated periodic activation-based architecture by Mehta et al. (2021), extended with the ability to process d-dimensional data in order to parameterize vector-valued functions and a 3D convolutional encoder. In this study, the input vector-valued functions of the molecules were uniformly sampled using a $32 \times 32 \times 32$ grid. In cases where only SMILES were available, conformers were generated using RDKit.

4 Preliminary Results and Conclusion

For our initial experiment, we trained the proposed architecture as an auto-decoder on a type 2 inhibitor binding to unphosphorylated interleukin-1 receptor-associated kinase 4 (PDB ID `6EGA`). The complex was learned by sampling a $32 \times 32 \times 32$ grid of its vector-valued function mapping $\mathbb{R}^3 \rightarrow \mathbb{R}^6$ (hydrogens, carbons, and Awigner-Seitz electron density of the ligand and protein) and successfully reconstructed using the conditioning latent with a PSNR (peak signal-to-noise ratio) of 38.5. Following the super-resolution approach from computer vision, we generated an up-scaled version with a resolution of $128 \times 128 \times 128$ (PSNR $= 27.2$) compared to a ground truth sampled at the same higher resolution (Wang et al., 2021). A visual comparison of the results is shown in Appendix Figure 2. Next, we parameterized multiple conformers of a molecular complex, the solution structure of a dimeric lactose DNA-binding domain complexed with a DNA sequence (PDB ID `1OSL`), using a single modulated network trained as an auto-encoder (see architecture in Figure 1). An animation of the interpolation through latent space by linearly interpolating between latents can be found in the GitHub repository linked below. Finally, we used the proposed architecture as an auto-encoder trained on the FreeSolv ($n = 642$) data set. We sampled a vector-valued function mapping $\mathbb{R}^3 \rightarrow \mathbb{R}^{10}$ (all occurring atom types and Awigner-Seitz electron density). The reconstruction quality is similar to the up-scaling experiment with an average PSNR of 28.7. Upon visual inspection of the latent space, we note that the latents represent the shape, size, and basic physicochemical proper-

ties of encoded molecules (Appendix Figure 3). In Appendix Figure 4, we show that the model can also smoothly interpolate between randomly chosen molecules when trained on a larger data set.

Molecular neural fields have a vast potential for applications in machine learning for life sciences. Learning directly on electron density maps from X-ray crystallographic experiments or Cryo-EM density maps are near-future applications. Beyond that, the representation allows for the simple integration of organic and non-organic molecular information and the parametrization of physico-chemical valid navigation between conformers. While challenges remain, we believe this work to be valuable to the broader machine learning community in life sciences. All code, models, and data are available at `https://anonymous.4open.science/r/DB1B`.

MEANINGFULNESS STATEMENT

We believe that a meaningful representation of life must achieve a balance between accuracy, generalizability, and explainability. However, computational representations often encounter intrinsic trade-offs in simultaneously optimizing these interdependent properties. The framework we present in this work is a step towards further achieving this balance and reducing the impact of these trade-offs. Furthermore, new representations are meaningful if they enable researchers to explore new methodologies. By representing molecules as vector-valued functions and parametrizing them using neural networks, we open up a new range of neural network architectures and approaches to the life sciences community.

REFERENCES

Amer Essakine, Yanqi Cheng, Chun-Wun Cheng, Lipei Zhang, Zhongying Deng, Lei Zhu, Carola-Bibiane Schönlieb, and Angelica I. Aviles-Rivero. Where Do We Stand with Implicit Neural Representations? A Technical and Performance Survey, November 2024.

P. Gainza, F. Sverrisson, F. Monti, E. Rodolà, D. Boscaini, M. M. Bronstein, and B. E. Correia. Deciphering interaction fingerprints from protein molecular surfaces using geometric deep learning. *Nature Methods*, 17(2):184–192, February 2020. ISSN 1548-7105. doi: 10.1038/s41592-019-0666-6.

Yinhan He, Zaiyi Zheng, Patrick Soga, Yaozhen Zhu, yushun Dong, and Jundong Li. Explaining Graph Neural Networks with Large Language Models: A Counterfactual Perspective for Molecular Property Prediction, October 2024.

Chiyu Max Jiang, Avneesh Sud, Ameesh Makadia, Jingwei Huang, Matthias Nießner, and Thomas Funkhouser. Local Implicit Grid Representations for 3D Scenes, March 2020.

Michael Kilgour, Mark Tuckerman, and Jutta Rogal. Multi-Type Point Cloud Autoencoder: A Complete Equivariant Embedding for Molecule Conformation and Pose, July 2024.

Stephen Lombardi, Tomas Simon, Jason Saragih, Gabriel Schwartz, Andreas Lehrmann, and Yaser Sheikh. Neural Volumes: Learning Dynamic Renderable Volumes from Images. *ACM Transactions on Graphics*, 38(4):1–14, August 2019. ISSN 0730-0301, 1557-7368. doi: 10.1145/3306346.3323020.

Ishit Mehta, Michaël Gharbi, Connelly Barnes, Eli Shechtman, Ravi Ramamoorthi, and Manmohan Chandraker. Modulated Periodic Activations for Generalizable Local Functional Representations, April 2021.

Kisung Moon, Hyeon-Jin Im, and Sunyoung Kwon. 3D graph contrastive learning for molecular property prediction. *Bioinformatics*, 39(6):btad371, June 2023. ISSN 1367-4811. doi: 10.1093/bioinformatics/btad371.

J. R. Serra, P. C. Jurs, and K. L. E. Kaiser. Linear Regression and Computational Neural Network Prediction of Tetrahymena Acute Toxicity for Aromatic Compounds from Molecular Structure. *Chemical Research in Toxicology*, 14(11):1535–1545, November 2001. ISSN 0893-228X. doi: 10.1021/tx010101q.

Vincent Sitzmann, Justus Thies, Felix Heide, Matthias Nießner, Gordon Wetzstein, and Michael Zollhöfer. DeepVoxels: Learning Persistent 3D Feature Embeddings, April 2019.

Vincent Sitzmann, Julien N. P. Martel, Alexander W. Bergman, David B. Lindell, and Gordon Wetzstein. Implicit Neural Representations with Periodic Activation Functions, June 2020.

Zhihao Wang, Jian Chen, and Steven C. H. Hoi. Deep Learning for Image Super-Resolution: A Survey. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 43(10):3365–3387, October 2021. ISSN 1939-3539. doi: 10.1109/TPAMI.2020.2982166.

Gengmo Zhou, Zhifeng Gao, Qiankun Ding, Hang Zheng, Hongteng Xu, Zhewei Wei, Linfeng Zhang, and Guolin Ke. Uni-Mol: A Universal 3D Molecular Representation Learning Framework. In *The Eleventh International Conference on Learning Representations*, September 2022.

A APPENDIX

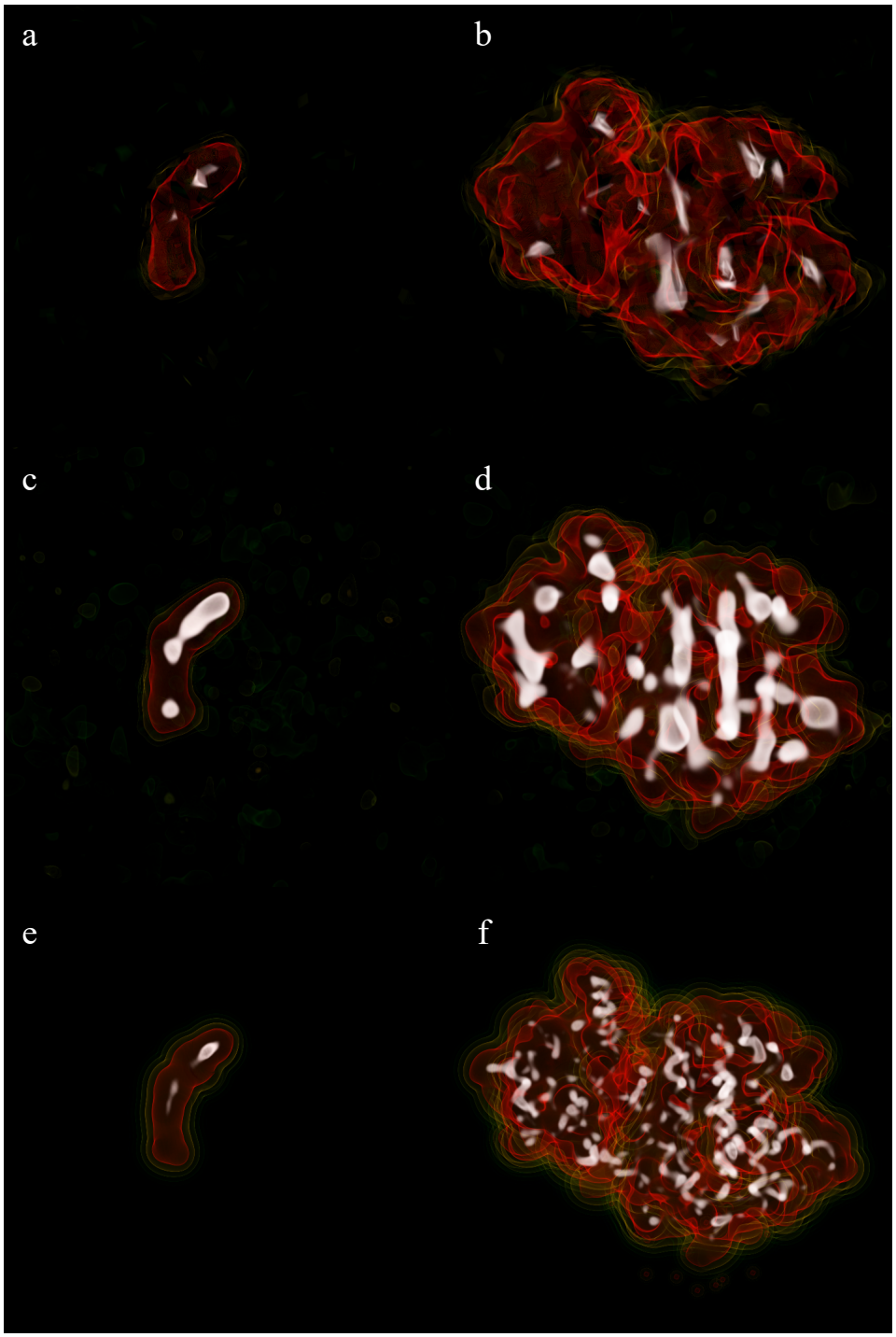

Figure 2: Super-resolution decoding of protein-ligand complex (separated channels). The top row (a, b) shows a density plot with a resolution $32 \times 32 \times 32$. The middle row (c, d) shows the up-scaled version with a resolution of $128 \times 128 \times 128$. The bottom row (e, f) shows the ground truth at a resolution of $128 \times 128 \times 128$. The peak signal-to-noise ratio of the up-scaled version and the ground truth is 75.7.

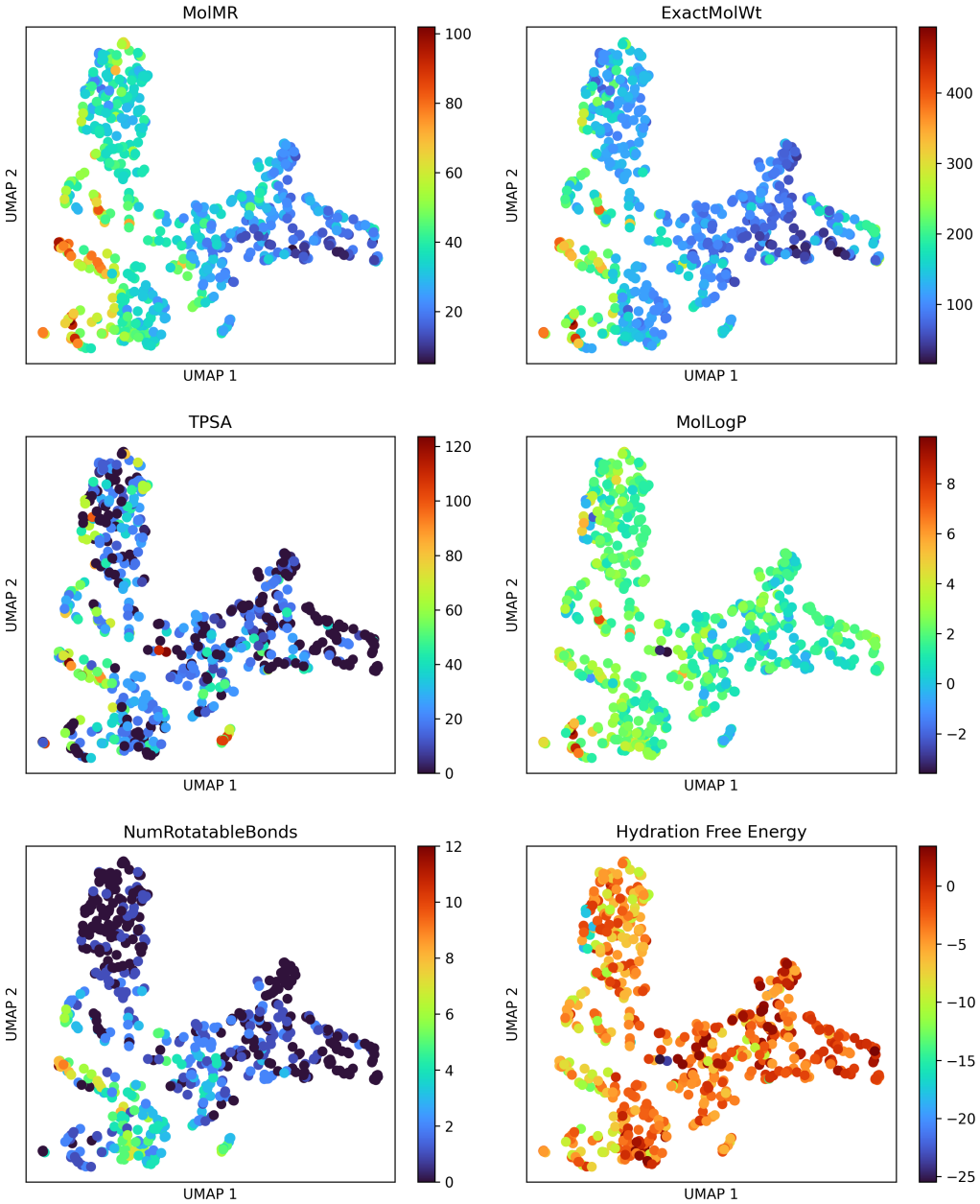

Figure 3: Embedding of molecular neural field latents of the FreeSolv data set. The scatter plots are colored by the Wildman–Crippen molar revlectivity value (MolMR), the exact molecular weight (ExactMolWt), the topological polar surface area (TPSA), the Wildman–Crippen LogP value (MolLogP), the number of rotatable bonds (NumRotatableBonds), and the hydration free energy. The hydration free energy was provided with the FreeSolv data set, all other values were calculated using RDKit. On visual inspection, the plots show that the latent space is meaningful in terms of molecular structure and shape.

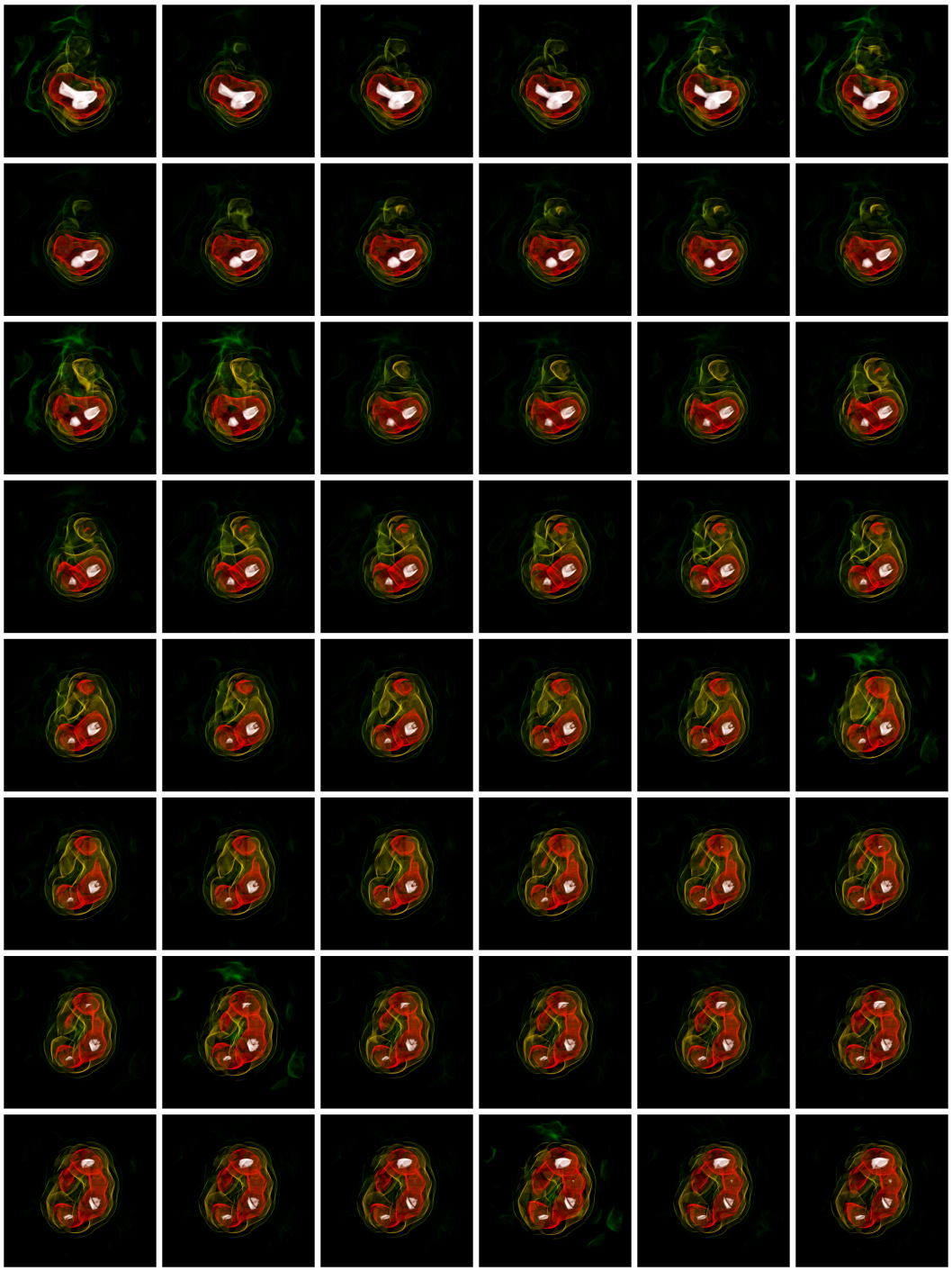

Figure 4: Smoothly interpolating between two molecules parametrized by the auto-encoder trained on the FreeSolv data set.

A.1 EQUATIONS

The vector-valued function f mapping from $\mathbb{R}^3$ to $\mathbb{R}^1$ for a given atom type t is sampled at point (x, y, z) using the following function:

$$f(x, y, z) = \sum_{a_i \in M_t} \exp\left(-\beta\left(\frac{d(a_i, (x, y, z))}{r_a^2} - 1\right)\right) \tag{1}$$

Where a_i are the coordinates for atom $a_i \in M_t$, β a user-chosen scaling factor, $d(a_i, (x, y, z))$ the squared Euclidean distance between a_i and (x, y, z), and r_a the chosen radius property for atom a_i. As default, the radius property for atom field calculations is the Van der Waals radius.