



Physics-Guided Deep Generative Model For New Ligand Discovery

Dikshant Sagar

dsagar2@calstatela.edu

Department of Computer Science
California State University, Los Angeles
Los Angeles, California, USA

Nida Sheikh

nsheikh@calstatela.edu

Department of Computer Science
California State University, Los Angeles
Los Angeles, California, USA

Ali Risheh

arisheh@calstatela.edu

Department of Computer Science
California State University, Los Angeles
Los Angeles, California, USA

Negin Forouzesh

neginf@calstatela.edu

Department of Computer Science
California State University, Los Angeles
Los Angeles, California, USA

ABSTRACT

Structure-based drug discovery aims to identify small molecules that can attach to a specific target protein and change its functionality. Recently, deep learning has shown great promise in generating drug-like molecules with specific biochemical features and conditioned with structural features. However, they usually fail to incorporate an essential factor: the underlying physics which guides molecular formation and binding in real-world scenarios. In this work, we describe a physics-guided deep generative model for new ligand discovery, conditioned not only on the binding site but also on physics-based features that describe the binding mechanism between a receptor and a ligand. The proposed hybrid model has been tested on large protein-ligand complexes and small host-guest systems. Using the top- N methodology, on average more than 75% of the generated structures by our hybrid model were stronger binders than the original reference ligand. All of them had higher ΔG_{bind} (affinity) values than the ones generated by the previous state-of-the-art method by an average margin of 1.88 kcal/mol. The visualization of the top-5 ligands generated by the proposed physics-guided model and the reference deep learning model demonstrate more feasible conformations and orientations by the former. The future directions include training and testing the hybrid model on larger datasets, adding more relevant physics-based features, and interpreting the deep learning outcomes from biophysical perspectives.

CCS CONCEPTS

• **Computing methodologies** → **Neural networks; Artificial intelligence**; • **Applied computing** → **Physics; Chemistry**.

Permission to make digital or hard copies of all or part of this work for personal or classroom use is granted without fee provided that copies are not made or distributed for profit or commercial advantage and that copies bear this notice and the full citation on the first page. Copyrights for components of this work owned by others than the author(s) must be honored. Abstracting with credit is permitted. To copy otherwise, or republish, to post on servers or to redistribute to lists, requires prior specific permission and/or a fee. Request permissions from permissions@acm.org.

BCB '23, September 3–6, 2023, Houston, TX, USA

© 2023 Copyright held by the owner/author(s). Publication rights licensed to ACM.
ACM ISBN 979-8-4007-0126-9/23/09...\$15.00
<https://doi.org/10.1145/3584371.3613067>

KEYWORDS

Drug discovery, Deep learning, Generative neural networks, Implicit solvent models.

ACM Reference Format:

Dikshant Sagar, Ali Risheh, Nida Sheikh, and Negin Forouzesh. 2023. Physics-Guided Deep Generative Model For New Ligand Discovery. In *14th ACM International Conference on Bioinformatics, Computational Biology and Health Informatics (BCB '23)*, September 3–6, 2023, Houston, TX, USA. ACM, New York, NY, USA, 9 pages. <https://doi.org/10.1145/3584371.3613067>

1 INTRODUCTION

Drug discovery plays a pivotal role in combating diseases, improving patient outcomes, and extending human life expectancy. However, the process of discovering effective drugs has traditionally been a costly, time-consuming, and resource-intensive endeavor, often characterized by a high rate of failure [24]. The interaction between a protein and a ligand, the small molecule that binds to the protein, is a fundamental event in drug action. It involves intricate molecular recognition and dynamic interplay between the ligand and the protein's active site or binding pocket [15]. Traditional experimental techniques such as X-ray crystallography and nuclear magnetic resonance (NMR) spectroscopy have provided valuable insights into the three-dimensional structures of protein-ligand complexes. However, these methods are often laborious, expensive, and challenging to apply to a wide range of proteins and ligands [30]. The emergence of computer-aided drug design has revolutionized the landscape of pharmaceutical discovery, offering innovative tools and methodologies that expedite the drug development process while reducing costs and minimizing risks [13]. Through the application of molecular docking, molecular dynamics simulations, and quantitative structure-activity relationship (QSAR) modeling, researchers can probe the energetics, dynamics, and binding affinity of potential drug candidates within the target protein's binding site. These computational tools can enable the exploration of vast chemical space, facilitating the identification of promising compounds and the optimization of their binding properties [13].

The search space for possible molecular structures is enormous and complex. It can be narrowed down by validating candidate molecules based on their chemical constraints, such as bond orders, molecular conformation, valences, etc. The space gets much smaller

when searching for a valid molecule that fits into a specific binding pocket to fulfill particular purposes. e.g., receptor inhibition in disease treatments and drug delivery mechanisms. Discovering a new molecular structure breaks down into two tasks: (1) sampling promising compounds from a constrained chemical space and (2) validating them to see if they bind on the target site as hypothesized. Step (1) expands over a timeline of 3-5 years. This shortfall creates a need for computational systems that could traverse this restricted chemical search space intelligently while also screening those compounds virtually for the possibility of successful binding. This could result in significant cost savings and faster drug development timelines.

The first work [20] to introduce deep learning in structure-based drug discovery was able to score predefined molecular (ligands) poses to dock into a protein (receptor) binding site using a pose scoring function based on a convolutional neural network. The receptor and the ligand were represented as atomic density grids using a molecular grinding tool [27]. Deep learning approaches since then have been applied to various tasks ranging from pose optimizations [20] to binding affinity predictions [8], all of which are being used for optimal molecular docking. However, these works are only helpful to screen or score structures when candidate structures are available. For producing candidate structures, initial approaches [9, 3, 22] utilized SMILES string notations [29] to leverage generative language models, which were improved via reinforcement learning to lead the generative process toward desired cheminformatic characteristics [10, 18]. Despite all the successes, SMILES strings, due to their non-permutation invariance, fail to capture the full concept of chemical similarity. Additionally, their deficiency in conveying conformational information restricts their usefulness in the field of structure-based drug discovery. To overcome this issue, graph-based molecular representations were used, which leveraged Graph Neural Networks to assimilate features from the molecular structures and produce new structures [23, 6]. Nonetheless, the generated bonds were independent, which led to structures with invalid valences. Also, the graph-matching loss functions are computationally expensive unless approximations are made [23], which leads to a sub-optimally trained model.

Molecular data representations are often handled in a 2D space, which is counterintuitive to what exists in a reality where bonds can rotate to a varied degree giving different conformations of the molecule that can affect its intermolecular interactions, such as binding to a receptor. To overcome these challenges, a 3D representation of a molecule was devised using atomic density grids [25]. Each voxel corresponds to a specific location in space, i.e., they are coordinate frame dependent. They are also permutation invariant, making them computationally less expensive for comparisons. In [21], 3D density grids are employed for training a conditional variational autoencoder with conditional protein receptor and input ligand pairs in order to find novel structures. While this was a significant step toward generating novel drug candidates, they failed to incorporate fundamental physics-based characteristics of the binding, particularly the protein-ligand binding free energy, including the enthalpic (polar, non-polar, and Van der Waals energies) and entropic components. While structural information, such as bond connectivity and atom arrangements, forms the basis for molecular representations, they do not capture the intricate and dynamic nature of

chemical systems. Incorporating physics-based information is essential because it provides insights into the energetics, stability, and reactivity of molecules. Physics-based information encompasses molecular forces, intermolecular interactions, and thermodynamic considerations, which play a crucial role in determining the stability and behavior of molecular structures. By integrating physics-based principles, such as molecular mechanics, quantum mechanics, and statistical thermodynamics, into molecular generation approaches, researchers can better explore the vast chemical space and identify novel, energetically favorable configurations. We hypothesize that along with structural data, physics-based features can improve the conditional effect to improve the quality of the learned latent chemical space and, in turn, generate novel structures with higher binding affinity values, as per our knowledge making this the first study to do so. Hence, our contribution is as follows:

- Aggregate experimental data from receptor-ligand binding simulations to form physics-based features for each protein-molecule pair in the PDBBind dataset.
- Create a hybrid conditional variational autoencoder that utilizes both the structural grids and the physics-based features to improve the quality of the learned chemical latent space.
- Compare and evaluate the generated molecules from the hybrid model using ΔG_{bind} (affinity) values.

2 MATERIALS AND METHODS

2.1 Physics-Based Features

Implicit solvent modeling is one of the most popular computational methods that consider the solvent (usually water) as one continuum component. Within this framework, the calculation of ΔG_{bind} could be conducted more efficiently compared to other computational models, e.g., explicit solvents. Poisson-Boltzmann (PB) and generalized Born (GB) models are the two main classes of implicit solvent models that have been used widely in static and dynamic simulations of protein-ligand interactions [19]. In this work, GB-NSR6 [4, 5] and PBSA [12] in AmberTools20 [2] are used for fast yet accurate calculation of binding free energy (see Table 1). By integrating implicit solvents into the deep learning model, it is more likely to generate feasible and strong binders.

2.2 Atom Type Vector

In order to train the deep neural network, molecular data is converted into a vector such that each atom is a vector and each molecule is a vector of atom-type vectors. We follow the same atom typing scheme as described in [21], where atom types are assigned using a set of N_p atomic property functions p and value ranges for those properties v as shown in Table 2. The atomic properties used here were element (different value ranges for ligands and receptors), aromaticity, H bond donor and acceptor status, and formal charge. For every atom a , a one-hot encoded vector p is created for each property, and then N_p vectors are concatenated to create a final atom type vector $t \in \mathbb{R}^{N_t}$. Hence, we get a 1×18 sized type vector for every atom.

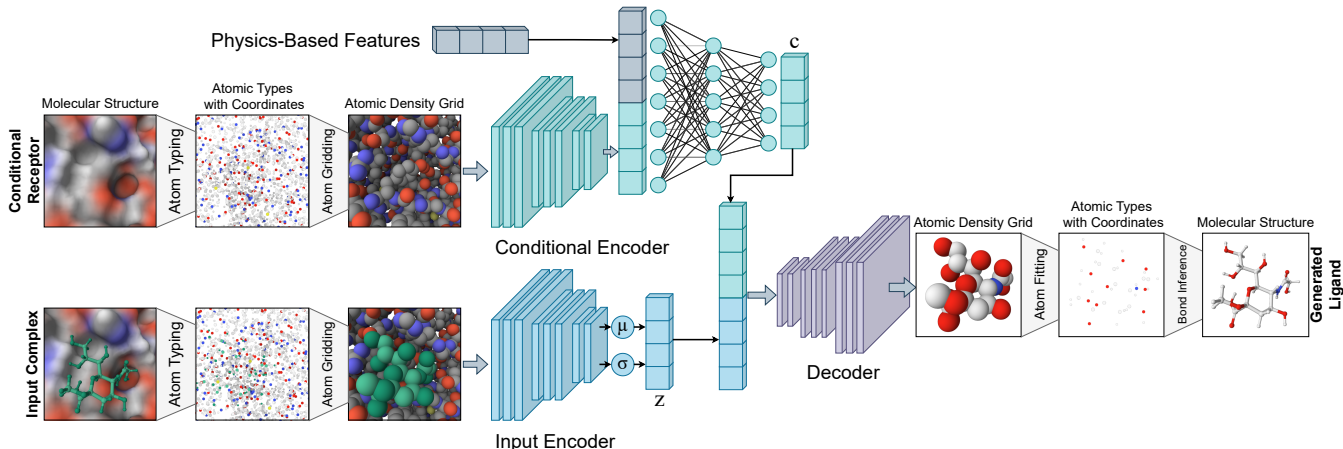


Figure 1: Physics Guided Generative Model Pipeline Overview. First, the input complex of docked protein and ligand are converted to atom-type vectors and subsequently into atomic density grids. Then our physics-guided CVAE model’s encoder branches take the input complex and the protein receptor’s density grids and the physics-based features as inputs. The input encoder produces a probabilistic latent vector sampled from $z \sim N(\mu, \sigma)$, and the conditional encoder gives an encoded vector c , which is then concatenated to z and fed into the decoder to produce an output generated ligand density grid. This density grid is then finally converted to a 3D molecular structure by atom fitting and bond inference algorithms.

Parameter	Description	Method	Count
1-4-EELEC	1-4 Electrostatic energy	GB	3
VDWAALS	Van der Waals energy	PB	3
EELEC	Electrostatic energy	GB&PB	6
ESURF	Non-polar solvation energy	GB	3
EGB	Polar solvation energy	GB	3
ECAVITY	Non-polar solvation free energy	PB	3
EPB	Reaction field energy	PB	3
Etot	Computational calculated $\Delta\Delta G$	GB&PB	6
Enthalpy	Total energy of a system	GB	1
Entropy	Entropy	E^*	1
ΔG_{bind}	Binding free energy	GB	1

Table 1: Physics-based features calculated for complex, protein, and ligand structures using MM/PB(GB)SA tool. * This feature is calculated as the difference between the experimental ΔG_{bind} and computational Enthalpy values. See [1] for details.

2.3 Molecule Density Grids

After atom-typing a molecule, it is essential to select a representation that embodies the molecular 3D spatial characteristics. We utilized a molecular gridding library called libmolgrid [27] that creates a molecular density grid where atoms are represented as continuous densities with truncated Gaussian shapes. Libmolgrid defines the density value of an atom at a grid point by a kernel function $f: \mathbb{R} \times \mathbb{R} \rightarrow \mathbb{R}$ that takes as input the distance d between

Atomic property	Value range	Num. values
Ligand element	B, C, N, O, F, P, S, Cl, Br, I, Fe	11
Receptor element	C, N, O, Na, Mg, P, S, Cl, K, Ca, Zn	11
Aromatic	False, True	2
H-bond acceptor	True	1
H-bond donor	True	1
Formal charge	-1, 0, 1	3

Table 2: Atom typing property functions and their value ranges.

the atom coordinate and the grid point and the atomic radius r :

$$f(d, r) = \begin{cases} e^{-2(\frac{d}{r})^2}, & d \leq 1.5r \\ 0, & d > 1.5r \end{cases} \quad (1)$$

r was fixed to 1.0 Å for all atoms, and the dimension of the cubic grid to 23.5 Å with 0.5 Å resolution to maintain consistency with [21], which results in spatial dimensions of $N_X = N_Y = N_Z = 48$. Also, N is the total number of atoms. To save computational resources, only the atoms that fit within the spatial extent of the grid are represented.

2.4 Atom Fitting and Bond Inference

As our generative model is trained with density grid format data, its predictive output is also a density grid. Now the problem remains of converting a reference density grid G_{ref} back into a discrete 3D molecular structure, which does not have an analytical solution [21] and is solved with the following optimization problem:

$$T^*, C^* = \arg \min_{T, C} ||G_{ref} - g(T, C)||^2 \quad (2)$$

where g is the function to convert a molecule’s atom type vector T and atomic coordinate vector C into density grid G . The initial locations of atoms can be found by selecting the grid points with the largest density values. By using libmolgrid, we can compute the grid representation of an atomic structure and backpropagate a gradient from grid values to atomic coordinates. We used the algorithm defined in [21] that combines iterative atom detection with gradient descent to find the best set of atoms that fit that reference density grid. Once the atoms and their coordinates are known, the only thing left is to assign bonds between the atoms to form valid molecules. This is achieved by a bond inference algorithm which is based on customized bond perception routines implemented in OpenBabel [17]. It uses a sequence of inference rules that add bond information and hydrogens while trying to satisfy the constraints defined by the atom types.

2.5 Deep Generative Model

The main reference [21] proposed a generative deep learning model based on a conditional variational autoencoder (CVAE) [26], which consisted of an input grid encoder, a conditional receptor grid encoder, and a ligand grid decoder. The objective was to learn a sample from a distribution $p(lig|rec)$, where rec is the binding site density grid and lig is the density grid of the ligand that binds to it. Latent sample z was drawn from a standard normal distribution under the assumption that the binding interactions might follow it as a prior. In the generative process, they first drew a sample $z \sim p(z)$ and then generated $lig_{gen} \sim p_{\theta}(lig|z, c)$, where p_{θ} is the decoder neural network and, c is the encoding of the receptor from the conditional encoder.

2.6 Evaluation Metric

To compare the quality of the generated ligands by the original and our hybrid methods, we employed a metric called the ΔG_{bind} , which is the binding affinity value between the receptor and ligand and refers to the change in Gibbs free energy associated with the binding of a ligand to a receptor or target molecule (See Fig.2). A negative ΔG_{bind} value indicates a favorable binding interaction, suggesting a stronger affinity between the ligand and the receptor. Conversely, a positive ΔG_{bind} value indicates a weaker or unfavorable binding interaction [7]. We utilized the GNINA package [14] to calculate the affinity values for the receptor-ligand pairs generated by both methods and then compared the values for the Top-N ligands. Top-N refers to the top N generated ligands that have the highest affinity values.

2.7 Dataset

PDBBind: The PDBBind database [28] is a free and widely used resource in the field of computational biochemistry and drug discovery. It serves as a comprehensive collection of experimentally determined protein-ligand complexes obtained from the Protein Data Bank (PDB). The dataset contains detailed structural information about the interactions between proteins and small molecules,

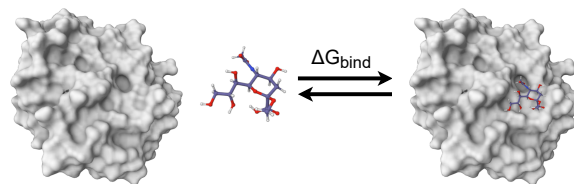


Figure 2: Visualizing the ΔG_{bind} between a receptor (protein) and a ligand (molecule).

including their three-dimensional coordinates, binding affinities, and other relevant properties. The PDBBind dataset is valuable for a range of research tasks, such as developing and validating scoring functions for virtual screening, understanding protein-ligand binding mechanisms, and training machine learning models for structure-based drug design. In this work, we use a subset of the PDBBind-v19 known as the refined set, which has undergone additional processing and filtering to improve its quality and reliability for research purposes. Among the original 3,562 receptor-ligand complexes, 2,728 pairs had all the required features and experimental values available. This dataset was split into training and testing sets randomly with a ratio of 80:20, where the whole testing set was used to negate the possibility of overfitting and evaluate the model’s training performance in terms of Reconstruction Loss, KL Divergence Loss, and Steric Loss (See Fig.4). Randomly selected conditional receptors from the test set were also used to generate candidate ligand structures from the predicted density grids, which were evaluated using GNINA by calculating their ΔG_{bind} values with respect to the conditional receptors (See Fig.5).

Host-guest systems: The small and rigid host-guest systems [16] introduce chemical hosts (size: ~ 100 non-hydrogen atoms) with pockets that enable strong binding to the corresponding compounds, called guests. Hosts bind their guests via the same basic forces that proteins used to bind their ligands, so they can serve as simple test systems for computational models of non-covalent binding. Moreover, their small size and, in many cases, their rigidity can make it feasible to sample all relevant conformations. In this work, structures named α -cyclodextrin and β -cyclodextrin are selected from host-guest systems to test the transferability of the proposed hybrid model by generating guest candidates for these two host molecules.

3 RESULTS AND DISCUSSION

3.1 Hybrid Model

This paper demonstrates how physics-based features could improve a deep generative model’s ability to create novel higher-affinity ligands conditioned to a receptor protein. The rules of physics govern the universe, and they sure govern how molecules are formed and how they interact with each other. With this hypothesis, we created a hybrid model improving upon the work done by [21] to include physics-based features in the conditional input to improve the quality of the learned latent chemical space with the CVAE. To achieve this, we created another branch conditional encoder (See Fig.3) for the physics-based features described in Section 2.7. This encoder maps the raw features concatenated with the output of

the existing receptor conditional encoder to a dimension of 128. This allows the model to learn how to fuse conditional receptor encoding with the physics-based features through backpropagation during training.

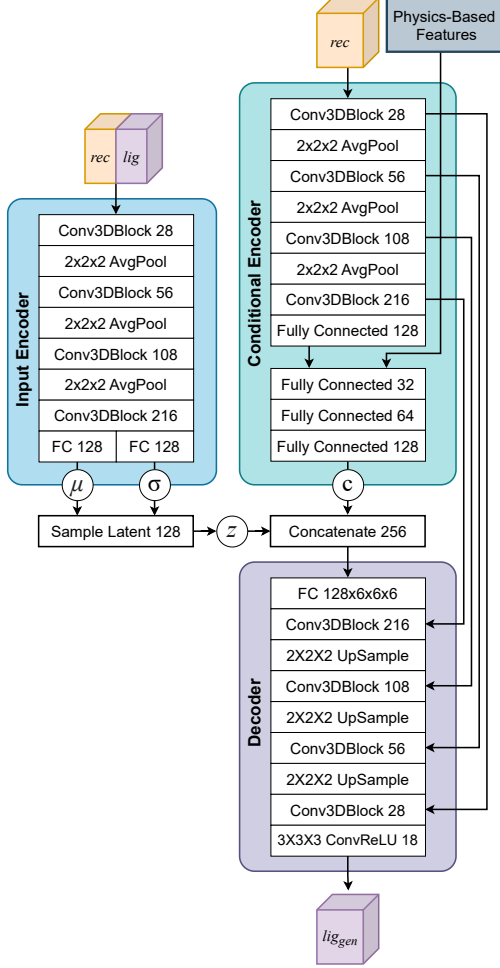


Figure 3: Schematic Architecture of the Physics-Guided Generative Model

The final receptor and physics-based feature conditional encoding concatenate with the input encoding in the same way as the original work. Using this modified architecture design also enables us to utilize previously trained weights for the rest of the segments of the model via transfer learning. Consequently, this gives a great starting point for the weight initialization of our hybrid model, helping us reach loss saturation and end training quickly. Now, the objective becomes learning a sample from a distribution $p(lig|rec, feat)$ where *lig*, *rec*, and *feat* are the ligand density grid, receptor density grid, and physics-based features, respectively. Hence, we sample $z \sim p(z)$ and then generate $lig_{gen} \sim p_{\theta}(lig|z, c)$, where p_{θ} is the same decoder neural network, and c is the new fused conditional encoding of the receptor density grid and the physics-based features given by $c \sim p_{\theta_c}(c|rec, feat)$ where p_{θ_c}

is the new modified conditional encoder neural network parameterized by θ_c . The input encoder maps a protein-ligand (*rec*, *lig*) complex to a set of means and standard deviations (μ, σ) defining latent variables, which are sampled to produce a latent vector z . The conditional encoder maps the same conditional receptor *rec* and the physics-based features *feat* to a conditional encoding vector c . The latent vector and conditional vector are concatenated and provided to the decoder, which maps them to a generated ligand density grid *lig_{gen}*. The input encoder and conditional encoder consist of 3D convolutional blocks with leaky ReLU activation functions and residual connections [11] (See Fig. 3).

Due to the difficult nature of estimating the naive maximum likelihood to compute the latent posterior probability, $p_{\theta}(z|rec, lig)$, we followed the method as described in original work [21] to learn an approximate input encoder model $q_{\phi}(lig|z, c)$ of the posterior distribution which can be trained by the following two objectives :

$$L_{recon} = -\log p_{\theta}(lig|z, c) \propto \frac{1}{2} ||lig - lig_{gen}||^2 \quad (3)$$

$$L_{KL} = D_{KL}(q_{\phi}(z|lig, c)||p(z)) \quad (4)$$

L_{recon} is the reconstruction loss term which maximizes the probability that the latent samples from the approximate posterior distribution $z \sim q_{\phi}(z|rec, lig, feat)$ are decoded as close to the original ligand density *lig* that was provided during the forward pass. L_{KL} is the Kullback–Liebler (KL) divergence loss that forces the learned latent space probability distribution to be as close as possible to a standard normal distribution, i.e., $p(z) = N(0, 1)$. With the joint optimization of both these terms, we are able to learn a latent space that follows a normal distribution, and we end up training a decoder that can decode these latent vectors sampled from a normal distribution into realistic ligand densities.

Following the original work [21], we also included another loss term called the **Steric Loss** that minimized steric clash in terms of the overlap between the generated ligand density and the receptor density. The loss value is calculated by first summing across the grid channels, then multiplying the receptor and ligand density at each point:

$$L_{Steric} = \left\langle \sum_i^{N_T} rec_i, \sum_i^{N_T} lig_{gen,i} \right\rangle \quad (5)$$

Hence, the final loss objective for the complete model becomes :

$$L = \lambda_{recon} L_{recon} + \lambda_{KL} L_{KL} + \lambda_{steric} L_{steric} \quad (6)$$

The loss weights were kept consistent with [21] at $\lambda_{recon} = 4.0$, $\lambda_{KL} = 0.1$, and $\lambda_{steric} = 1.0$, with the KL divergence loss weight increased to 1.6 after 20,000 iterations. The model was fine-tuned using the RMSProp optimizer with a learning rate of 10^{-5} for 50,000 iterations and a batch size of 4.

3.2 Protein-Ligand Complexes

To train and finetune the model to incorporate physics-based features, a separate conditional feature encoder branch was designed

into the model’s architecture (see Fig.3). We observed that finetuning the hybrid model on the PDBBind dataset with the physics-based feature encoder did not destabilize training, probably due to the same domain of the PDBBind dataset (protein-ligand complexes), which did not result in a drastic covariate shift in the model’s initialized weights. Also, the new feature encoder rather evidently improved the quality of the latent chemical space and its closeness to a standard normal distribution as the KL divergence loss was further reduced during finetuning. The test set losses saturated in a similar fashion to the losses on the training set and hence did not lead to overfitting (see Fig.4). The reconstruction loss increases slightly after 20,000 iterations when we make the model focus more on making the latent distribution closer to a standard normal by increasing the KL Divergence loss weight.

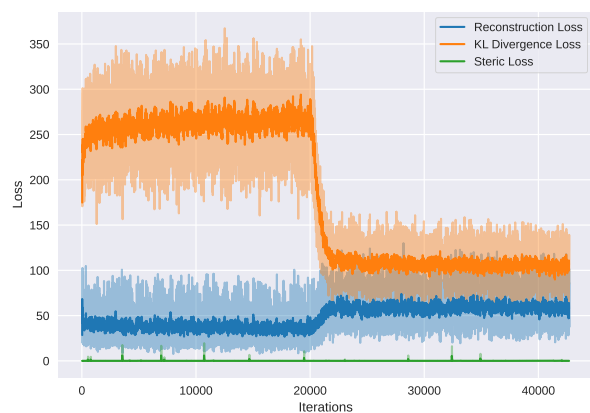


Figure 4: Hybrid-model’s reconstruction, KL divergence, and steric loss values on the test set during finetuning.

To compare the training quality of the two models, from a machine learning engineer’s purview, the loss values paint a clear picture. However, as molecular scientists, we are interested in the generated molecule’s efficiency or its affinity to the binding receptor. Therefore, to compare generated ligands from the two models, we selected the top-5 ligands with the highest affinities. It is observed that our hybrid model produced new ligands in which the top-5 generated molecules had higher binding affinities to the receptor protein than the reference ligand (from PDBBind) as compared to the original model [21] in which only, on average less than 5% of the top-5 had binding affinities higher than the reference ligand (See Fig.5). Also, this figure demonstrates that our physics-guided hybrid model unfailingly generated ligands with more promising ΔG_{bind} values closer to (Fig. 5c) and higher than (Fig. 5a, 5b, and 5d) the reference ligand. Whereas the original model had an inferior performance overall, with no generated ligand having a ΔG_{bind} value higher than the reference ligand.

In Fig 6a, we plot the distribution of the affinity values for generated structures by both models for the same receptor sites, and it clearly indicates that our model consistently generated structures

with higher binding affinity values than the original model. This improves our hypothesis and the effect of physics-based information guiding the learning and generating process.

The top-5 generated ligands by the original and hybrid models are demonstrated in Fig.7. It is observed that docked ligands found by the hybrid model have more feasible conformation and orientation inside the protein binding pocket. The corresponding ΔG_{bind} values confirm stronger binders introduced by the hybrid model.

3.3 Host-Guest Systems

The accuracy of the original and hybrid models to generate new ligands is also tested on host-guest systems. Due to the unavailability of a large dataset for host-guest molecule structures, we could not finetune the model for this use case. Instead, we tested the model trained on the PDBBind dataset. New structures are generated as before for the host molecules, and the top-5 of them are compared. Similar to the case in the previous subsection, Fig.6b visualizes the distribution of the affinity values of the generated structures by the two models for host-guest systems. It is observed that our model for this use case also generated structures with higher affinity values on average than the ones generated by the original model. Fig.8a plots the binding affinity values of the top-5 generated ligands by the two models for the α -cyclodextrin host molecule. We can observe that, in this case, both generated guest molecule groups had binding affinities higher than the reference ligand. However, the structures from our hybrid model had overall higher binding affinities than the ones generated from the original model. In Fig.8b, we observe a similar trend in which we plot the same metrics for the structures generated by both the models for the β -cyclodextrin host molecule and the generated guest molecules from our hybrid model again had higher binding affinities than the reference molecule, and the structures generated by the original model.

4 CONCLUSION

In this paper, we demonstrated that physics-based information could guide a deep generative model to predict higher-quality structures for a conditional receptor protein and that it has immense potential and promise for revolutionizing the field. By combining the power of deep learning algorithms with the fundamental principles of physics, we have been able to improve previous drug discovery techniques. Our hybrid model offers a unique advantage by leveraging the outcomes of implicit solvent models to guide the learning process, enabling the generation of stronger binders and reducing the need for extensive experimental data. The incorporation of physical features, such as electrostatic energy and Van der Waals force interactions, provides a more comprehensive understanding of the underlying mechanisms governing drug-target interactions. Furthermore, the utilization of deep learning models allows for the analysis of large and complex datasets, enabling the extraction of valuable insights from vast amounts of information. By employing these hybrid models, researchers can identify potential drug candidates more efficiently, saving time and resources in the drug development process. It is important to note that this physics-guided model is still in its infancy, and several challenges remain, such as the availability of more extensive, high-quality datasets and the designing of more generalizable models,

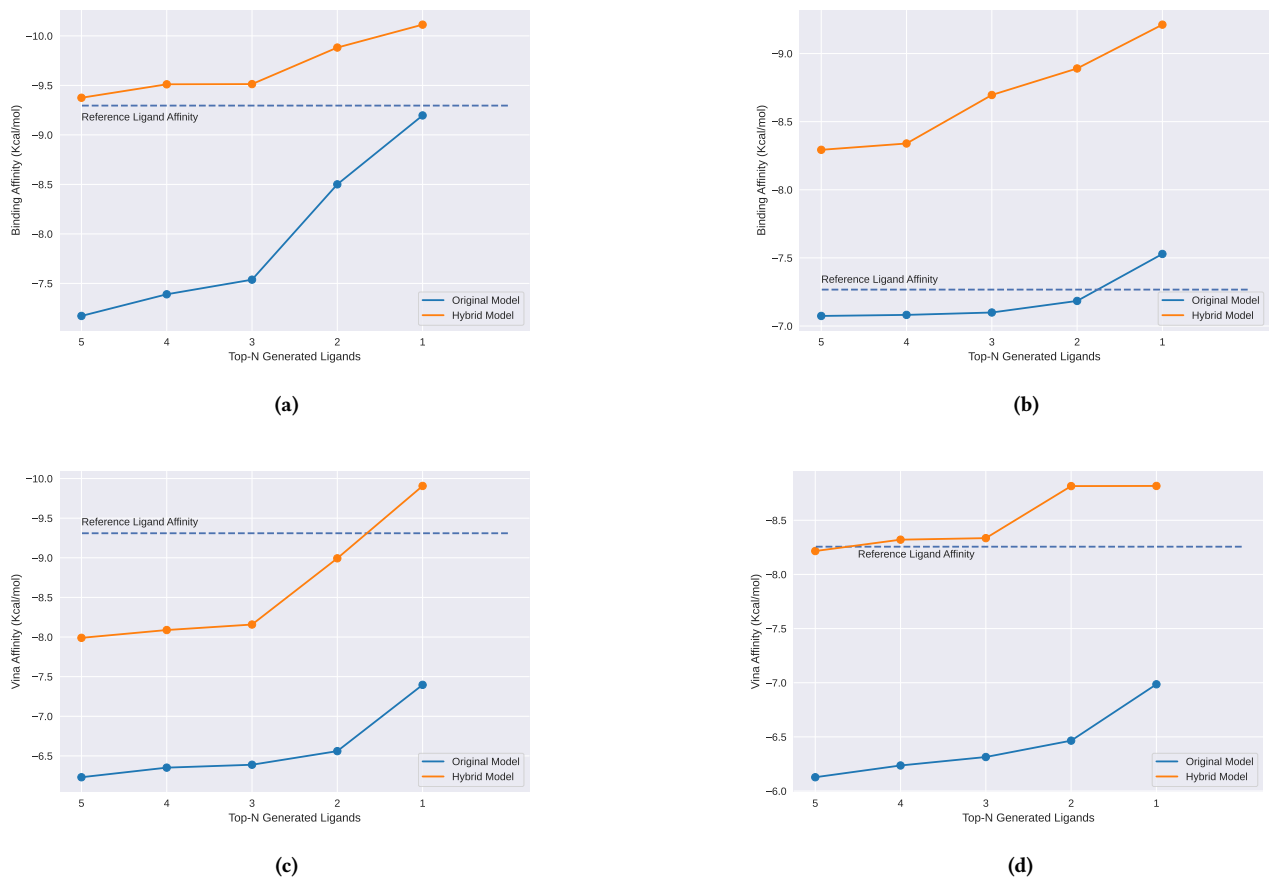


Figure 5: Top-5 generated ligands by the original model and our physics-based hybrid model for PDBBind receptor proteins - (a) 4hy1 [Topoisomerase IV, subunit B], (b) 1igb [beta-d-glucan glucohydrolase isoenzyme exo1], (c) 3lea [catalytic domain of TACE], (d) 1bjv [BETA-TRYPSIN] and their respective binding affinities (Kcal/mol).

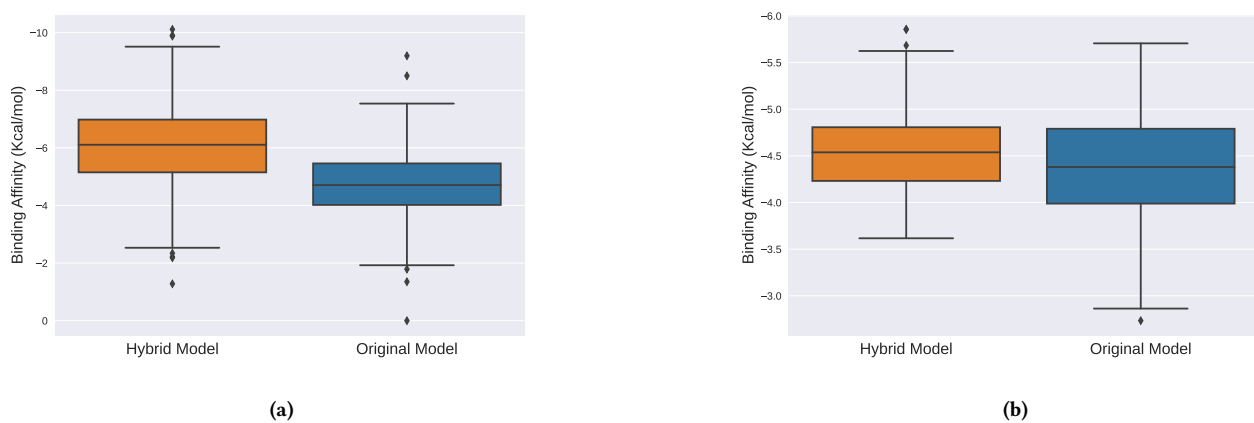
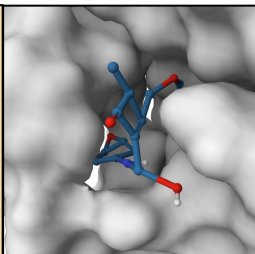
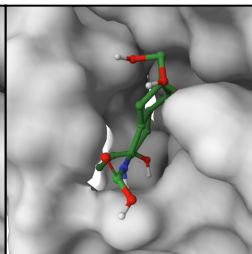
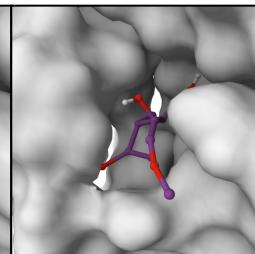
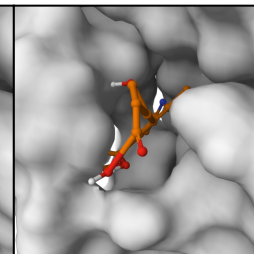
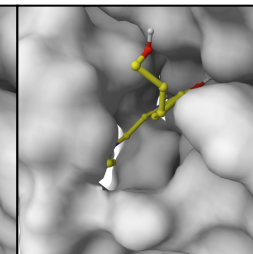


Figure 6: Box plots visualizing the distribution of binding affinities of the generated structures on (a) 4hy1, 1igb, 3lea and 1bjv protein receptors from PDBBind and (b) α -cyclodextrin and β -cyclodextrin host-guest systems by the original and the hybrid model.

Hybrid Model					
ΔG_{bind}	-9.21197 Kcal/mol	-8.89039 Kcal/mol	-8.69573 Kcal/mol	-8.33969 Kcal/mol	-8.29325 Kcal/mol
RMSD	0.52517	0.36858	0.54820	0.29304	0.59767
	C ₉ H ₁₇ N ₃ O ₅	C ₁₂ H ₁₉ NO ₇	C ₁₀ H ₂₂ O ₅	C ₁₁ H ₁₆ O ₇	C ₁₁ H ₁₈ N ₂ O ₃

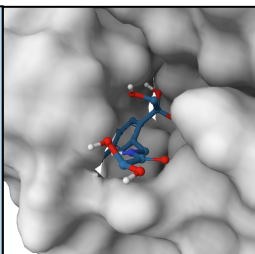
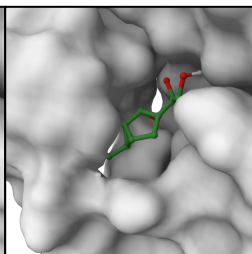
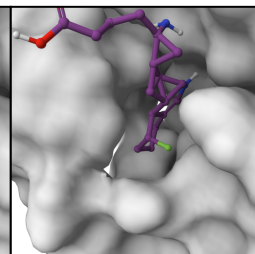
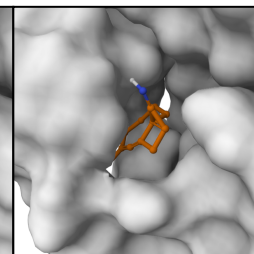
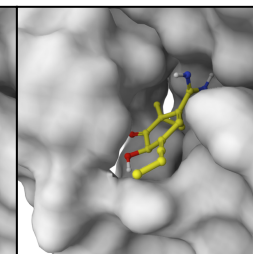
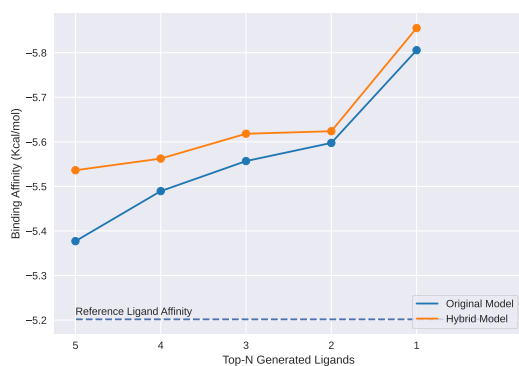
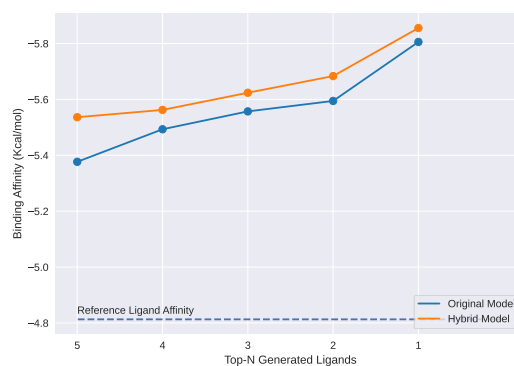
Original Model					
ΔG_{bind}	-7.52882 Kcal/mol	-7.18406 Kcal/mol	-7.09909 Kcal/mol	-7.08157 Kcal/mol	-7.07384 Kcal/mol
RMSD	0.53825	1.21968	0.78007	1.02037	0.62058
	C ₁₂ H ₁₆ FNO ₆	C ₉ H ₁₄ O ₄	C ₁₈ H ₂₃ FN ₂ O ₂	C ₁₁ H ₁₄ N ₂ O	C ₁₁ H ₁₈ O ₃

Figure 7: Visualisation of the Top-5 generated ligands inside the receptor pocket by our physics-guided deep generative model and the original model for the PDBBind protein 1igb [beta-d-glucan glucohydrolase isoenzyme exo1] with their decreasing binding affinities (left to right) and their RMSD values.



(a)



(b)

Figure 8: Top-5 generated guest molecules (ligands) by the original model and our physics-based hybrid model for Host system molecule (receptor) - (a) α -cyclodextrin, (b) β -cyclodextrin, and their respective binding affinities (Kcal/mol).

which we hope to tackle in future works. The code and trained weights for our physics-guided deep generative model are available at <https://github.com/dikshantsagar/PhyMolCVAE>.

ACKNOWLEDGMENTS

This research has been funded by NIH R16GM146633 and NSF 2216858 awarded to N.F.

REFERENCES

- [1] Sahar Cain, Ali Rishesh, and Negin Forouzes. 2022. A physics-guided neural network for predicting protein-ligand binding free energy: from host-guest systems to the pdbname database. *Biomolecules*, 12, 7. doi: 10.3390/biom12070919.
- [2] David A Case et al. 2021. *Amber 2021*. University of California, San Francisco.
- [3] Peter Ertl, Richard Lewis, Eric Martin, and Valery Polyakov. 2017. In silico generation of novel, drug-like chemical matter using the lstm neural network. *arXiv preprint arXiv:1712.07449*.

- [4] Negin Forouzesh, Saeed Izadi, and Alexey V Onufriev. 2017. Grid-based surface generalized born model for calculation of electrostatic binding free energies. *Journal of Chemical Information and Modeling*, 57, 10, 2505–2513.
- [5] Negin Forouzesh, Abhishek Mukhopadhyay, Layne T Watson, and Alexey V Onufriev. 2020. Multidimensional global optimization and robustness analysis in the context of protein-ligand binding. *Journal of Chemical Theory and Computation*.
- [6] Justin Gilmer, Samuel S Schoenholz, Patrick F Riley, Oriol Vinyals, and George E Dahl. 2017. Neural message passing for quantum chemistry. In *International conference on machine learning*. PMLR, 1263–1272.
- [7] Michael K Gilson and Huan-Xiang Zhou. 2007. Calculation of protein-ligand binding affinities. *Annu. Rev. Biophys. Biomol. Struct.*, 36, 21–42.
- [8] Joseph Gomes, Bharath Ramsundar, Evan N Feinberg, and Vijay S Pande. 2017. Atomic convolutional networks for predicting protein-ligand binding affinity. *arXiv preprint arXiv:1703.10603*.
- [9] Rafael Gómez-Bombarelli et al. 2018. Automatic chemical design using a data-driven continuous representation of molecules. *ACS central science*, 4, 2, 268–276.
- [10] Gabriel Lima Guimaraes, Benjamin Sanchez-Lengeling, Carlos Outeiral, Pedro Luis Cunha Farias, and Alán Aspuru-Guzik. 2017. Objective-reinforced generative adversarial networks (organ) for sequence generation models. *arXiv preprint arXiv:1705.10843*.
- [11] Kaiming He, Xiangyu Zhang, Shaoqing Ren, and Jian Sun. 2016. Deep residual learning for image recognition. In *Proceedings of the IEEE conference on computer vision and pattern recognition*, 770–778.
- [12] Ray Luo, Laurent David, and Michael K Gilson. 2002. Accelerated poisson-boltzmann calculations for static and dynamic systems. *J. Comput. Chem.*, 23, 13, 1244–1253.
- [13] Garland R Marshall. 1987. Computer-aided drug design. *Annual review of pharmacology and toxicology*, 27, 1, 193–213.
- [14] Andrew T McNutt, Paul Francoeur, Rishal Aggarwal, Tomohide Masuda, Rocco Meli, Matthew Ragoza, Jocelyn Sunseri, and David Ryan Koes. 2021. Glna 1.0: molecular docking with deep learning. *Journal of cheminformatics*, 13, 1, 1–20.
- [15] Nikita Mishra and Negin Forouzesh. 2012. Protein-ligand binding with applications in molecular docking. In *Algorithms and Methods in Structural Bioinformatics*. Springer, 1–16.
- [16] David L Mobley and Michael K Gilson. 2017. Predicting binding free energies: frontiers and benchmarks. *Annual review of biophysics*, 46, 531–558.
- [17] NM O’Boyle. 2011. Banck m. james ca morley c. vandermeersch t. hutchison gr open babel: an open chemical toolbox. *J. Cheminf.* 3, 1, 33.
- [18] Marcus Olivecrona, Thomas Blaschke, Ola Engkvist, and Hongming Chen. 2017. Molecular de-novo design through deep reinforcement learning. *Journal of cheminformatics*, 9, 1, 1–14.
- [19] Alexey Onufriev. 2008. Chapter 7 - implicit solvent models in molecular dynamics simulations: a brief overview. *Annual Reports in Computational Chemistry* 4, 125–137. Ralph A. Wheeler and David C. Spellmeyer, (Eds.) doi: [https://doi.org/10.1016/S1574-1400\(08\)00007-8](https://doi.org/10.1016/S1574-1400(08)00007-8).
- [20] Matthew Ragoza, Joshua Hochuli, Elisa Idrobo, Jocelyn Sunseri, and David Ryan Koes. 2017. Protein–ligand scoring with convolutional neural networks. *Journal of chemical information and modeling*, 57, 4, 942–957.
- [21] Matthew Ragoza, Tomohide Masuda, and David Ryan Koes. 2022. Generating 3d molecules conditional on receptor binding sites with deep generative models. *Chemical science*, 13, 9, 2701–2713.
- [22] Marwin HS Segler, Thierry Kogej, Christian Tyrchan, and Mark P Waller. 2018. Generating focused molecule libraries for drug discovery with recurrent neural networks. *ACS central science*, 4, 1, 120–131.
- [23] Martin Simonovsky and Nikos Komodakis. 2018. Graphvae: towards generation of small graphs using variational autoencoders. In *Artificial Neural Networks and Machine Learning–ICANN 2018: 27th International Conference on Artificial Neural Networks, Rhodes, Greece, October 4–7, 2018, Proceedings, Part I* 27. Springer, 412–422.
- [24] SITNFlash. 2020. Modern drug discovery: why is the drug development pipeline full of expensive failures? en-US. (Apr. 2020). <https://sitn.hms.harvard.edu/flash/2020/modern-drug-discovery-why-is-the-drug-development-pipeline-full-of-expensive-failures/>.
- [25] Miha Skalic, Davide Sabbadin, Boris Sattarov, Simone Sciabola, and Gianni De Fabritiis. 2019. From target to drug: generative modeling for the multimodal structure-based ligand design. *Molecular pharmaceutics*, 16, 10, 4282–4291.
- [26] Kihyuk Sohn, Honglak Lee, and Xinchen Yan. 2015. Learning structured output representation using deep conditional generative models. *Advances in neural information processing systems*, 28.
- [27] Jocelyn Sunseri and David R Koes. 2020. Libmolgrid: graphics processing unit accelerated molecular gridding for deep learning applications. *Journal of Chemical Information and Modeling*, 60, 3, 1079–1084.
- [28] Renxiao Wang, Xueliang Fang, Yipin Lu, and Shaomeng Wang. 2004. The pdbbind database: collection of binding affinities for protein–ligand complexes with known three-dimensional structures. *Journal of medicinal chemistry*, 47, 12, 2977–2980.
- [29] David Weininger. 1988. Smiles, a chemical language and information system. 1. introduction to methodology and encoding rules. *Journal of chemical information and computer sciences*, 28, 1, 31–36.
- [30] Ian Young. 1984. Nuclear magnetic resonance imaging. *Electronics and Power*, 30, 3, 205–210.