

Adversarial Threshold Neural Computer for Molecular *de Novo* Design

Evgeny Putin,^{*,†,‡} Arip Asadulaev,[‡] Quentin Vanhaelen,[†] Yan Ivanenkov,^{†,§} Anastasia V. Aladinskaya,^{†,||} Alex Aliper,[†] and Alex Zhavoronkov^{*,†,||}

[†]Pharma.AI Department, Insilico Medicine, Inc., Baltimore, Maryland 21218, United States

[‡]Computer Technologies Lab, ITMO University, St. Petersburg 197101, Russia

^{||}Moscow Institute of Physics and Technology (State University), 9 Institutskiy Lane, Dolgoprudny City, Moscow Region 141700, Russian Federation

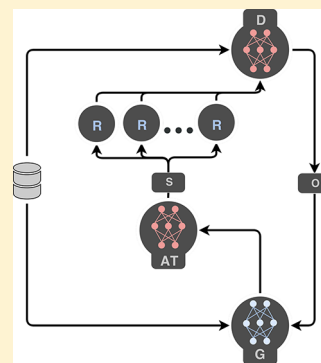
[§]Institute of Biochemistry and Genetics Russian Academy of Science (IBG RAS) Ufa Scientific Centre, Oktyabrya Prospekt 71, 450054 Ufa, Russian Federation

^{||}The Biogerontology Research Foundation, OX1 1RU Oxford, U.K.

Supporting Information

ABSTRACT: In this article, we propose the deep neural network Adversarial Threshold Neural Computer (ATNC). The ATNC model is intended for the *de novo* design of novel small-molecule organic structures. The model is based on generative adversarial network architecture and reinforcement learning. ATNC uses a Differentiable Neural Computer as a generator and has a new specific block, called adversarial threshold (AT). AT acts as a filter between the agent (generator) and the environment (discriminator + objective reward functions). Furthermore, to generate more diverse molecules we introduce a new objective reward function named Internal Diversity Clustering (IDC). In this work, ATNC is tested and compared with the ORGANIC model. Both models were trained on the SMILES string representation of the molecules, using four objective functions (internal similarity, Muegge druglikeness filter, presence or absence of sp³-rich fragments, and IDC). The SMILES representations of 15K druglike molecules from the ChemDiv collection were used as a training data set. For the different functions, ATNC outperforms ORGANIC. Combined with the IDC, ATNC generates 72% of valid and 77% of unique SMILES strings, while ORGANIC generates only 7% of valid and 86% of unique SMILES strings. For each set of molecules generated by ATNC and ORGANIC, we analyzed distributions of four molecular descriptors (number of atoms, molecular weight, logP, and tpsa) and calculated five chemical statistical features (internal diversity, number of unique heterocycles, number of clusters, number of singletons, and number of compounds that have not been passed through medicinal chemistry filters). Analysis of key molecular descriptors and chemical statistical features demonstrated that the molecules generated by ATNC elicited better druglikeness properties. We also performed *in vitro* validation of the molecules generated by ATNC; results indicated that ATNC is an effective method for producing hit compounds.

KEYWORDS: molecular *de novo* design, deep neural network, reinforcement learning, generative adversarial network



INTRODUCTION

The drug discovery industry is facing unprecedented financial pressures principally driven by the cost of bringing a drug to market, which has increased to over US \$1.5 billion.^{1–3} This situation has encouraged the pharmaceutical companies to optimize key steps of the drug discovery pipeline. A major improvement was the development of high-throughput screening (HTS), which led to an unprecedented increase in targets and leads.^{4,5} Using compound libraries, HTS allows to quickly conduct millions of tests in order to rapidly identify active compounds of interest.^{6,7}

While there has been an increase in the number of targets and leads, the number of new molecular entities generated remained stable because of a high attrition rate during clinical phases caused by a selection of leads with inappropriate

physicochemical properties.^{8–10} In order to improve the target-to-lead efforts² to better identify candidates with the highest probability of success,^{11,12} one developed rational drug design approaches that aim at finding new medications based on the knowledge of a biological “druggable” target.^{13–16} Several methods and molecular predictors, such as the Lipinski’s “Rule of 5” (RO5),¹⁷ were established for estimating the “druglike-

Special Issue: Deep Learning for Drug Discovery and Biomarker Development

Received: December 15, 2017

Revised: March 8, 2018

Accepted: March 15, 2018

Published: March 23, 2018

ness". Generally, early selection of candidates requires the design of molecules complementary in shape and charge to the biomolecular target of interest. Furthermore, the determination of the nature and the rates of any active processes that are involved in the absorption, distribution, metabolism, elimination and toxicity (ADMET) of drug candidates are also of primary importance. ADMET profiling and prediction being mostly dependent on molecular descriptors such as ROS.¹¹

Computational techniques have been developed to optimize the prediction of ADMET properties.^{18–21} These methods are used to design *in silico* filters to eliminate compounds with undesirable properties. These filters are widely applied for the assembly of the compound libraries using combinatorial chemistry.^{22–24} The integration of early ADMET profiling of "lead" molecules have contributed to speed-up the "lead" selection further for phase-I trial without large amounts of revenue loss.^{9,25} Today, compounds are added in the libraries on the basis of target-focused design or diversity considerations,^{26,27} and HTS can screen compound libraries to select a subset of compounds whose properties are in agreement with various criteria.^{28–31}

Despite these efforts, the current largest databases of chemical compounds remain small compared to the space spanned by all energetically stable stoichiometric combinations of electrons, atomic nuclei, and topologies in molecules. Considering these factors, it is estimated that there are 10^{60} distinct molecules, 10^{30} that are druglike.⁶ As a result, computational techniques are also being developed for the design of *de novo* druglike compounds³² and for generating large virtual chemical libraries, which can be more efficiently screened for *in silico* drug discovery purposes.

Among the computational techniques available, deep neural networks (DNN) have gained a lot of interest for their ability to do feature extraction and learn rules from training data. Indeed, an outstanding success and high efficacy in pattern recognition and natural language processing have been achieved using DNNs.³³ More interestingly, DNN-based architectures have also been successfully developed for application in various fields of biological and biomedical sciences.^{34–40}

Several DNN-based methods have been proposed for molecular *de novo* design and molecular feature extraction. From a computational perspective, druglike molecular structure can be represented in five ways. As a string in formats like SMILES⁴¹ or InChi,⁴² as a molecular fingerprint,⁴³ as a set of molecular descriptors, such as molecular weight, logP, number of heavy atoms, number of rotatable bonds, etc., as a graph in which atoms are nodes and links are bonds between atoms, or as a 3D electron density map. A method for molecular feature extraction based on an architecture of adversarial autoencoder⁴⁴ (AAE) using the fingerprint representation of the molecules was introduced by Kadurin et al.⁴⁵ and was shown to perform better than variational autoencoder (VAE) models.⁴⁶ Nevertheless, molecular fingerprint has two disadvantages compared to the string representation of the molecule. First, one fingerprint can match several molecules, so there is no one-to-one mapping from a molecule to the fingerprint, and second, the fingerprint representation contains less information about the molecule topology than the string representation. This suggests employing the string representation for molecule generation. For example, a VAE model for learning continuous representation of molecules represented in SMILES format was introduced by Gomez et al.⁴⁷ In this model, the VAE was used to ensure that the latent spaces of the molecules are valid

SMILES strings, and criteria such as QED⁴⁸ (Quantitative Estimation of Druglikeness), logP, and synthetic accessibility score⁴⁹ were added to the VAE loss function to generate more druglike molecules. Nevertheless, the authors pointed out that the usage of VAE does not guarantee the validity of the SMILES strings in the latent space.

Recently, models based on Recurrent Neural Networks (RNNs) have been developed. RNNs are more adapted for data of sequential nature such as a molecule represented as a sequence of characters SMILES. Such architecture was used by Bjerrum et al.⁵⁰ where it was shown that performance was improved when using long short-term memory (LSTM) cells and gated recurrent units (GRU). Nevertheless, as no convenient "generative" method to sample new molecules from the model could be defined, the generation process remained limited. Also, the model did not include any mechanism to assess whether the generated molecules are a valid SMILES string, and the model also lacks criteria to generate targeted molecular structures with desired properties. These simple models illustrate that designing biochemical rules that can be effectively used by the algorithms to identify or to prioritize molecular structures of specific interest is currently an important challenge. In another work,⁵¹ several architectures of RNN-based models were studied, and it was found that standard stochastic optimizers could be improved by using generative models trained on unlabeled data, to perform knowledge-driven optimization. This could allow replacing handcrafted rules currently commonly used in the context of *de novo* drug design by rules learned from data.

The development of generative adversarial networks⁵² (GANs) allowed to investigate this direction in more detail. An approach for sequence generation via deep reinforcement learning (DRL) was proposed.⁵³ The architecture, called SeqGAN (Sequence Generative Adversarial Network), was based on a GAN extended with a RL-based generator. The task of sequence generation is formulated in a reinforcement learning setting where the agent (generator) is given k -th length of already generated sequence and must choose the next $k + 1$ symbol to be generated. An environment (discriminator) returns a reward for the generator, which is estimated as the likelihood of fooling the discriminator. Nevertheless, since the generator outputs discrete values, standard backpropagation cannot be used, and authors proposed to update the generator via a set of policy gradients and a Monte Carlo search based on the expected end reward (reward for the whole sequence) returned from the discriminator. To estimate the action-value function, the REINFORCE⁵⁴ algorithm was applied. An extension of SeqGAN called ORGAN (Objective-Reinforced Generative Adversarial Network) was proposed.^{55,56} This model adds an "objective-reinforced" reward function for particular sequences into the SeqGAN reward loss. The ORGAN's reward function is similar to a filter. When the generated molecules pass the filter, the generator gets more rewards for generating such molecules from the discriminator. It is worth mentioning that an architecture similar to ORGAN but without GAN component and based wholly on RL was also introduced.⁵⁷ The model used an atom-based LSTM with RL term needed to generate molecules with desired properties and was trained on the ChEMBL druglike data set. As for the SeqGAN model, the REINFORCE algorithm was used to find the optimal policy. The results showed that the agent network achieved 98% accuracy of generation of the valid SMILES strings. However, without the GAN component included inside

the model, this approach had the same disadvantages as the model presented by Bjerrum et al.⁵⁰ Further works based on objective functions for molecular design within ORGAN paradigm was done by Lengeling et al.⁵⁸ The proposed architecture, ORGANIC (Objective-Reinforced Generative Adversarial Network for Inverse-design Chemistry), used ROS and other criteria as objective filters to train the ORGAN model on the druglike data set from the ZINC⁵⁹ database. The results showed that the use of different objective reward functions makes it possible to bias the generation process and generates molecules with desired user-specified properties.⁶⁰

In this article, we propose a novel DNN-based architecture called ATNC. This architecture is based on the ORGANIC paradigm, but it includes a specific AT block and uses Differentiable Neural Computer (DNC)⁶¹ as a generator. AT block is a copy of the discriminator in the environment, which lags behind the original discriminator for a given number of training epochs. AT blocks implement the idea of the RL model; it adds the model an ability to simulate environment reactions and filter samples that are most close to receive a positive rewards in a real environment. In addition to the AT block, we introduced a new objective reward function called IDC, which has been designed to allow the generation of more diverse molecular structures. In what follows, we provide a detailed description of the ATNC and compare the new features of this implementation with the ORGANIC model. Then, the performances of the ATNC model are assessed and compared to the ORGANIC model to evaluate the effect of the different objective reward functions on the training as well as on the properties of the generated molecules. The comparison between the IDC-based model and the IS-based model is of special interest. This first comparison is followed by an analysis of the chemical statistical features of the generated molecules. This allows for a better assessment of the applicability of the ATNC and ORGANIC models within the context of modern *in silico* molecular design. Finally, we summarize the results of the *in vitro* validation performed to obtain a preliminary biological evaluation of the molecules generated by ATNC. We conclude with a discussion about the advantages of this new approach and summarize directions for future investigations.

METHOD

Description of the ATNC Model. There are two types of RL methods: model-free and model-based methods.⁶² In model-free methods, the learning of a policy takes place by directly interacting with an environment. While such models have been successfully applied for a lot of RL problems, this method requires carrying out millions of actions in the environment before an optimal policy is found. In order to overcome this issue, model-based approaches have their own internal understanding of the environment. Using this ability, model-based approaches can simulate the environment within themselves, and select the most favorable actions for the trained agent. This ability accelerates the process of finding the optimal policy.

In the RL context, ORGANIC can be seen as a model-free method, where the generator G_θ , which produces the sequence $Y_{1:T} = (y_1, \dots, y_T)$, is the agent, and the discriminator D_ϕ , together with the objective functions $O(Y_{1:T})$, creates the environment. However, ORGANIC sample model actions are very stochastic, especially in the initial stages. Furthermore, all actions are executed in the environment immediately, the problem with

this model is that most of the actions receive negative rewards from the environment. This negatively affects the training of the model and drastically reduces the quality of the generated molecules. This occurs because the model is trained to only generate simply valid molecules and forget the properties of the training data.

To overcome this issue of the ORGANIC architecture and create an end-to-end system similar to the RL-based model, we propose a novel architecture named Adversarial Threshold Neural Computer (ATNC) with a new specific block called adversarial threshold (AT). AT is a copy of the discriminator in the environment, which lags behind the original discriminator for a given number of τ epochs. The main feature of this block is that it can select actions by the agent before they are performed and evaluated in the environment. The overall scheme of the proposed architecture is presented in Figure 1.

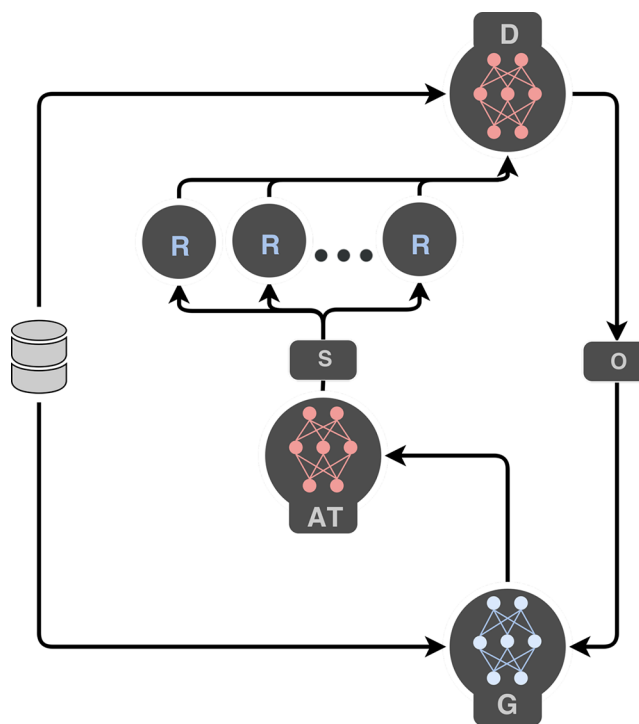


Figure 1. Schematic view of the ATNC model. Red blocks are AT block and the discriminator. Blue blocks are the generator and Monte Carlo roll out models. S is the set of chosen samples (molecules) by AT block. O is some objective reward function. R stands for the read head in the DNC generator.

The first stage of the ATNC training is the pretraining of the θ -parametrized generator G_θ and the ϕ -parametrized discriminator D_ϕ . Using the maximum likelihood estimation (MLE), the generators are pretrained using the same approach as the one described for ORGANIC. The objective of the ATNC generator G_θ is to simultaneously fool the discriminator D_ϕ and to maximize the objective reward function O . This is done using the policy gradient method.⁶³

$$J(\theta) = \sum_{y_1 \in Y} G_\theta(y_1 | s_0) \cdot Q(s_0, y_1) \quad (1)$$

The DNC⁶¹ was used as a generator G_θ within the ATNC architecture. DNC is a type of RNN with external memory. Using the DNC-based model as a SMILES-based molecule

generator G_θ is advantageous because it is able to generate much more longer and complex sequences of $Y_{1:T}$ than LSTM⁶⁴ (see Graves et al.⁶¹ for detailed performance and comparison between DNC and LSTM). After the MLE pretraining, the generator G_θ produces negative samples for the pretraining of the discriminator D_ϕ . The objective of the discriminator is to minimize the cross entropy between the ground truth and the generated sequences. For the pretraining of the discriminator, we used a special parameter ζ that controls how many epochs the discriminator D_ϕ is pretrained on only the valid SMILES sequences Y' . Starting from the RL training, the weights of the pretrained discriminator D_ϕ are copied to AT.

$$\nabla_\phi J(\phi) = \begin{cases} \min_{\phi} E_{Y \sim p_{\text{data}}(Y)} [\log D_\phi(Y)] + E_{Y \sim p_{G_\theta}(Y)} [\log(1 - D_\phi(Y))], & \text{if epoch} < \zeta \\ \min_{\phi} E_{Y' \sim p_{\text{data}}(Y')} [\log D_\phi(Y')] + E_{Y' \sim p_{G_\theta}(Y')} [\log(1 - D_\phi(Y'))], & \text{if epoch} > \zeta \end{cases} \quad (2)$$

On each epoch, the generator G_θ produces K (the batch size of the generator) molecules. Then, AT selects the molecules that most closely match to the training samples. If the number of selected molecules is less than J (the batch size of the discriminator), then the generator G_θ again generates new K molecules. This process will be repeated until the number of selected molecules is equal to J . After the selection of the molecules, a N -times Monte Carlo search is applied to calculate the action value function $Q_{s,a}$ for the partial sequences at the intermediate states. This is done to avoid the long-term reward problem.

$$Q(s = Y_{1:t-1}, a = y_t) = \begin{cases} \frac{1}{N} \sum_{n=1}^N R(Y_{1:T}^n), & Y_{1:T}^n \in MC^{G_\theta}(Y_{1:T}; N), \text{ if } t < T \\ R(Y_{1:T}), & \text{ if } t = T \end{cases} \quad (3)$$

Following the original ORGANIC approach, the total rewards R are computed as the sum of the outputs of the discriminator and the objective reward functions, where the contribution of each component in the sum is regulated by the parameter λ . $R(Y_{1:T}) = \lambda \cdot D_\phi(Y_{1:T}) + (1 - \lambda) \cdot O(Y_{1:T})$.

Finally, the parameters θ of the ATNC generator G_θ can be obtained by using the policy gradient method:⁶³

$$\nabla_\theta J(\theta) \simeq \frac{1}{T} \sum_{t=1}^T E_{Y_t \sim G_\theta(Y_t | Y_{1:t-1})} [\nabla_\theta \log G_\theta(y_t | Y_{1:t-1}) \cdot Q(Y_{1:t-1}, y_t)] \quad (4)$$

The update of the AT weights is controlled by the parameter τ . This parameter allows us to regulate how many epochs the environment model (AT) is lagging behind the original discriminator. In addition to the generator training, AT also influences the training of the discriminator. From the beginning of the RL training, the training samples of the discriminator become selected samples Y'' (from total number of samples generated by the generator) by AT.

$$\begin{aligned} \min_{\phi} E_{Y'' \sim p_{\text{data}}(Y'')} [\log D_\phi(Y'')] \\ + E_{Y'' \sim p_{G_\theta}(Y'')} [\log(1 - D_\phi(Y''))] \end{aligned} \quad (5)$$

AT will try to choose samples, which are very similar to training data, that allow us to filter molecules even during the generation stage. This simplifies the postprocess selection of molecules.

Algorithm 1 Training procedure of the ATNC model

Require: DNC generator policy G_θ ; roll-out policy G_β ; Discriminator D_ϕ ; objective O ; dataset $S = \{Y_{1:T}\}; \lambda, \zeta, \tau$;

```

Initialize  $G_\theta, D_\phi$  with random weights  $\theta, \phi$ ;
Pre-train  $G_\theta$  using MLE on  $S$ ; Pre-train  $D_\phi$  by minimizing cross entropy via generated negative samples by  $G_\theta$  or verified negative samples, and positive  $S$ ;
Initialize AT with  $\phi$  parameters;

1: repeat
2:   for  $g$  in  $g$ -steps do
3:     while Number of samples in the set  $S$  less than  $D_\phi$  batch size do
4:       Generate a sequence  $Y_{1:T} = (y_1, \dots, y_T) \sim G_\theta$ ; additionally writing and reading from the external memory.
5:       Choose by AT and add samples to the set  $S$ 
6:     end while
7:     for  $t$  in  $1:T$  do
8:       Compute  $Q(s = Y_{1:t-1}, a = y_t)$  by Eq.(3)
9:     end for
10:    Update generator parameters  $\theta$  via Eq.(2);
11:    Update roll-out policy parameters  $\beta$ 
12:    for d-steps do
13:      Train  $D_\phi$  for  $k$  epochs by Eq.(5) via only generated by  $G_\theta$  and filtered by AT negative samples and positive  $S$ ;
14:      Update AT parameters by  $\phi_{g-\tau}$ .
15:    end for
16:  end for
17: until convergence

```

Internal Diversity Clustering. It appears that the best way toward the automatic generation of diverse molecules, based on the ORGANIC paradigm, is to elaborate a user-based specification of the objective reward functions. Thus, in this work, we have introduced a new objective reward function named IDC, in addition to the AT module described above. The pseudo code of the IDC is described in Algorithm 2 and can be summarized as follows:

Algorithm 2 Internal diversity clustering procedure

Require: generated pool of molecules P ; fingerprint function F_{func} ; similarity measure S_{measure} ; similarity threshold $S_{\text{threshold}}$ for adding molecule in a cluster; top K molecules selected per cluster

```

1:  $P$  is clustered by its  $F_{\text{func}}$  representation using  $S_{\text{measure}}$  with  $S_{\text{threshold}}$ ;
2: From each cluster only top  $K$  most diverse molecules are selected;
3: Construct a binary reward vector  $\text{rewards}$ , with ones only for the selected top  $K$  generated molecules per cluster
4: return  $\text{rewards}$ 

```

First, a pairwise similarity matrix for the fingerprints of the generated molecules is computed. Then, using this similarity matrix, the molecules are sorted according to their similarity and placed in different clusters. The procedure is as follows: The first cluster is initialized using the first molecule inside the similarity matrix. Then, starting from the second molecule, each one is compared in terms of similarity to the molecules already included in the clusters. If the similarity is larger than a specified threshold, then the molecules is included in the same cluster; otherwise, the molecule is added in a newly initialized cluster. Finally, the top K most diverse molecules per cluster are selected. As a result of the IDC procedure, these molecules form a binary reward vector, where ones are placed only for the selected molecules. The objective reward function IDC is applied on each generation epoch of the ATNC and ORGANIC models. If the value of the parameter K is smaller than the average size of the clusters, the resulting reward vector should be sparse. Such a sparse reward vector can help the model to focus only on the diverse molecules. As a result, one can expect that the model will gain more rewards for generating diverse molecular structures.

RESULTS

Data. A collected subset (CD data set) of 15K druglike molecules, available within the ChemDiv (<http://www.chemdiv.com/>) stock collection, was used as a training data set. The chemical statistical features for the CD data set can be summarized as follows: 0.87 diversity, 644 unique heterocycles, 793 clusters, 3035 singletons, and 18.9 average cluster size. Within the data set, 11690 molecules satisfy RO5. Furthermore, only 21 compounds have not successfully passed through medicinal chemistry filters (MCFs). Internal diversity of the compounds included in the reference data set was calculated based on the extended-connectivity fingerprints⁶⁵ using the Tanimoto distance.⁶⁶ The cluster analysis was carried out with the minimum similarity in a cluster equal to 0.6. Empirically, we have found that the clusters containing less than five compounds were very cross-overlapped, i.e., similar structures were incorrectly included in different clusters. Thus, the minimum number of compounds per cluster was set to 5. All the calculations of chemical statistical features for the CD data set as well as for the generated molecules (see subsection *Chemical Analysis of the Generated Molecules*) were performed using the proprietary ChemoSoft Software by ChemDiv Company.

Settings. To make an objective comparison between the ATNC and ORGANIC models, all the experiments were performed with almost the same hyper-parameters and settings. This was done instead of using the number of training epochs because ATNC converges much faster than ORGANIC. We have used the sequences of SMILES symbols in a one-hot representation as inputs to the ATNC/ORGANIC model. A unique dictionary, encoding symbols in one-hot vector has been created for each data set. The size of the dictionary for the CD data sets was 29; see an example of a one-hot representation of a molecule in the *Supporting Information* Figure S1.

The DNC generator G_θ of the ATNC model consisted of a LSTM controller with 256 neurons, one read head R, and external memory size of 30·30. Eight times Monte Carlo search was used for rolling out. RMSprop was applied to optimize the controller with 1×10^{-3} learning rate and 0.9 momentum. The setting of the discriminator D_ϕ remained the same as for the ORGANIC model. The batch size was set to 64 for both the generator and the discriminator.⁵⁵ The coefficients were as follows: λ (for the reward function) was set to the optimal value of 0.5, ζ (for regulate discriminator pretraining) was set to 0.8, and AT update rate τ was set to 1.

To perform an honest comparison between the two models, we restricted the length of the training SMILES strings to 90. For the ORGANIC model we used 280 and 30 epochs to pretrain the generator and the discriminator, respectively. Two hundred RL epochs were used to train the model. Meanwhile, the ATNC model was trained with only 30 pretraining epochs for the generator and discriminator and also 200 RL epochs. At the end of each epoch, the ORGANIC model sampled 6400 molecules, and 32000 molecules every 10th epoch. The ATNC model sampled 3200 molecules at each epoch.

The ATNC model was implemented in Tensorflow.⁶⁷ We used the original implementation of ORGANIC (<https://github.com/aspuru-guzik-group/ORGANIC>). RDKit software was used to canonize training SMILES strings.⁶⁹ We also used RDKit to calculate all four objective reward functions during the training of the ATNC/ORGANIC model. The training of

the models was conducted using NVIDIA Titan Pascal GPUs with 128 Gb of RAM.

Evaluation of ATNC and ORGANIC Models. *Comparison of the Performances of ATNC and ORGANIC.* The ATNC and ORGANIC models were trained on the CD data sets using four different objective reward functions:⁶⁸

1. Internal Similarity (IS). The IS objective reward function was computed as follows: IS was given a similarity matrix for the pool of generated molecules. It considered each generated molecule and computed the average similarity between this molecule and the other molecules. These averaged similarity values were then used to form the real-valued reward vector of the pool.
2. Proposed IDC (see *Method* section for details).
3. Muegge druglikeness filter⁷⁰ (MU). When a molecule satisfied the MU objective reward, the corresponding bit in the reward vector was updated to one. If the MU criterion was not met, then the bit in the reward vector was set to zero.
4. Presence or absence of sp^3 -rich fragments (SP3). In this work, we used 150 different SP3. When a molecule had at least one sp^3 -rich fragment in its structure, the corresponding bit in the SP3 reward vector was set to one. If this criterion was not met, the bit in the reward vector was set to zero.

Our primary goal was to evaluate the effect of the four different objective reward functions on the training of the models as well as on the properties of the generated molecules. However, it is generally established that the generation of diverse molecular structures remain a central issue for computational-based structure generation. In this work, to address this particular issue, we have introduced the new objective reward function IDC, which rewards the model for generating molecules that are diverse within specific clusters. Thus, in order to assess the improvements provided by the implementation of the IDC, we were especially interested in comparing the performance of the IDC-based model with the IS-based model, as the objective function IS is essentially the average similarity between one molecule and the other generated molecules.

To compare the impact of all four objective reward functions on the training of the models, we pretrained the ORGANIC model with 280 and 30 epochs for the generator and the discriminator, respectively, and we used 200 RL epochs to train the model. All other settings of the model were as reported by Guimaraes et al.⁵⁵ Meanwhile, the ATNC model was trained with only 30 generator and 30 discriminator pretraining epochs and trained with 200 RL epochs. In order to perform a correct comparison of the models, we restricted the length of the training SMILES strings up to 90.

The comparison was done in three steps. First, we investigated the percentage of valid and unique SMILES strings generated by the two models. The percentages of valid SMILES strings were calculated with respect to the total number of generated molecules. Meanwhile, the percentages of unique SMILES strings were calculated with respect to the number of valid generated SMILES strings. Each of the two models was trained with the four objective reward functions presented above. The results of this experiment are shown in the *Table 1*. It should be noted that ATNC generated a different total number of SMILES strings. This is because the AT block acts as

Table 1. Percentages of Valid and Unique SMILES Strings Generated by the ATNC and the ORGANIC Models Using the CD Dataset^a

	ATNC-IDC	ORGANIC-IDC	ATNC-IS	ORGANIC-IS
total	157986	1792000	101652	1792000
% of valid molecules	72	8	71	7
% of unique molecules	77	86	86	91

	ATNC-MU	ORGANIC-MU	ATNC-SP3	ORGANIC-SP3
total	176342	1792000	156605	1792000
% of valid molecules	74	83	74	83
% of unique molecules	76	30	73	22

^aThe percentages of valid molecules were computed with respect to the total number of generated molecules, while the percentages of unique molecules were calculated with respect to the number valid generated SMILES strings.

a filter by selecting a subset of generated molecules. This subset is then handled and evaluated by the environment.

Generally, ATNC generated a larger percentage of unique molecules than ORGANIC; except when the ORGANIC model was used in combination with the IDC and IS objective reward functions. It is also worth mentioning that ATNC elicited a similar percentage of valid SMILES strings on all four objective reward functions while ORGANIC had a dramatically lower percentage of valid SMILES strings, when used with the IDC and IS objective reward functions. More detailed information about the numbers (total, valid, and unique) and

related percentages can be found in Table S1 in the [Supporting Information](#).

Second, we analyzed the average length of the valid SMILES strings generated by the two models for each epoch. The results of this analysis are presented in the [Figure 2](#). For the four objective reward functions, the ATNC model demonstrates a better stability. It was able to generate SMILES strings whose lengths were close to the ones of the training molecules. The behavior of the ORGANIC model was different, the SMILES strings generated after several RL epochs were much smaller.

Third, we studied the distributions of the lengths of the SMILES strings and the distribution of several key molecular descriptors (number of atoms, molecular weight, logP, and TPSA) for the generated molecules. The results are depicted in [Figure 3](#). One can see that the molecules generated by ATNC matched with the distributions of the molecules used for training for the four objective reward functions. However, the molecules generated by ORGANIC failed to preserve the prior distributions of the reference molecules.

Altogether, these results clearly demonstrate the capabilities of the ATNC model to generate unique and valid SMILES strings corresponding to structures that, while being dissimilar in structure, are very similar to molecules included in the training set in terms of key molecular descriptors. Both of these features are of particular importance within the modern drug discovery and development framework.

Chemical Analysis of the Generated Molecules. In order to assess the applicability of the ATNC and ORGANIC models within the context of modern *in silico* molecular design, we studied the chemical statistical features of the generated molecules. For each generated pool (two models and four

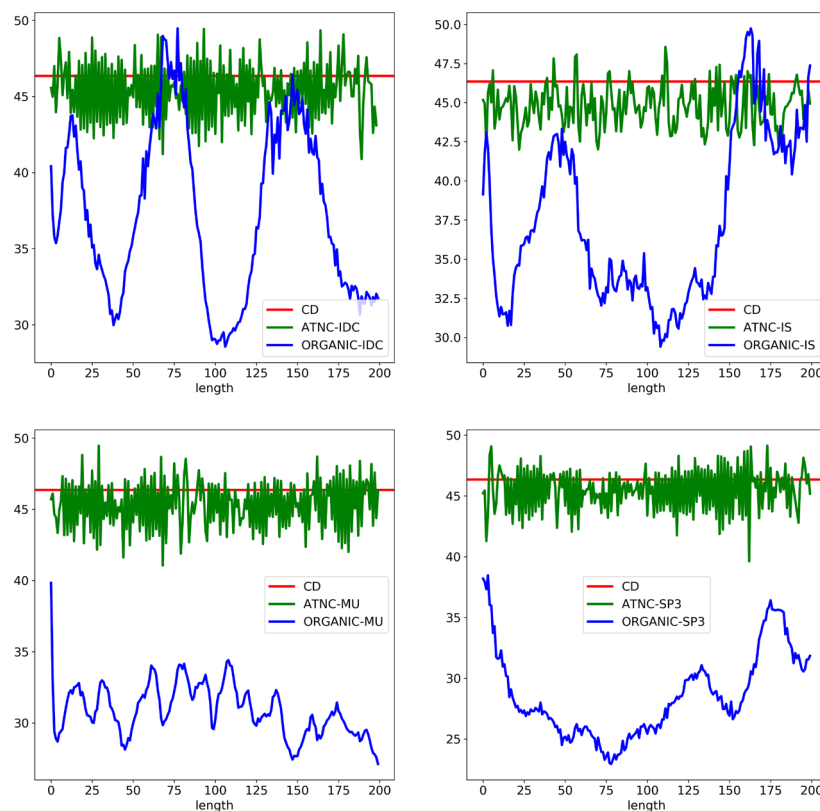


Figure 2. Average lengths of the generated SMILES strings on each training epoch by ORGANIC and ATNC models for the CD data set. Red lines are means in the training data set.

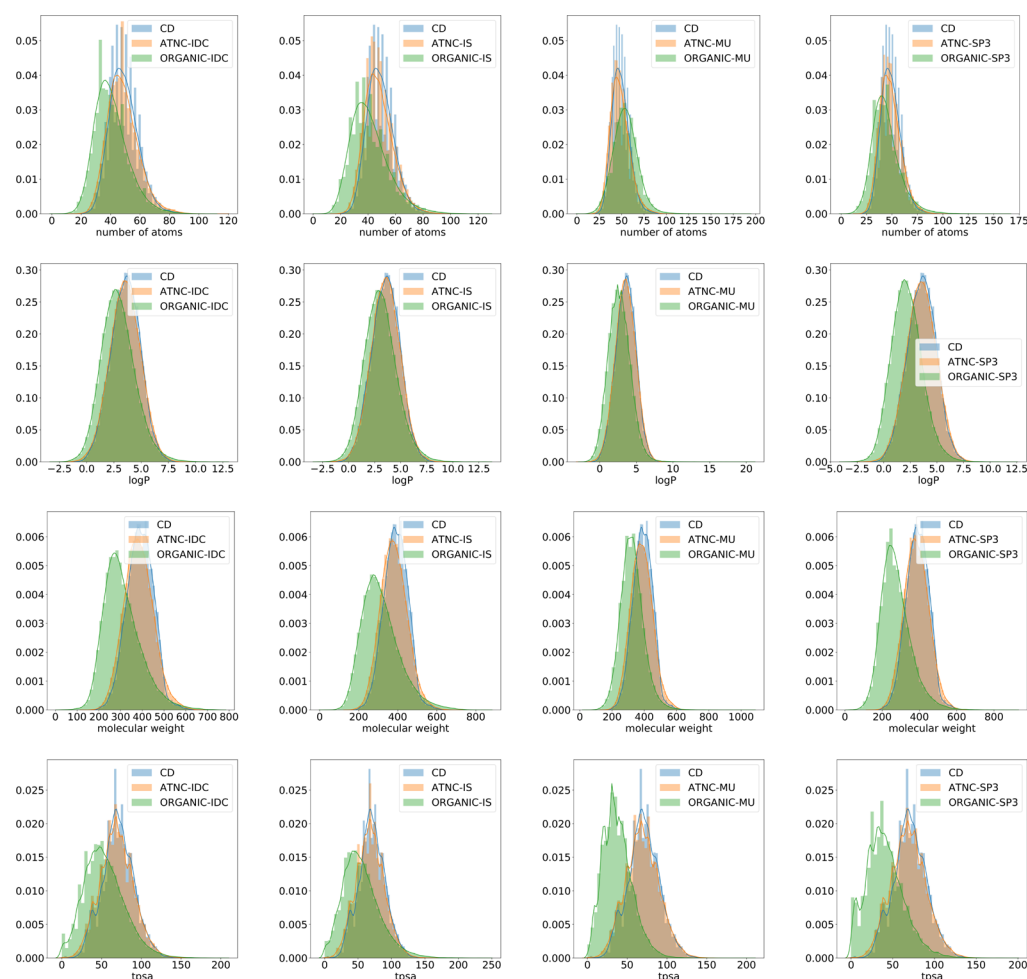


Figure 3. Distributions of the representative molecular descriptors calculated for the generated molecules. The rows correspond to the following molecular descriptors: number of atoms, molecular weight, logP, and tpsa.

objective reward functions for a total of eight pools of generated molecules) we have calculated the five following features: internal diversity, number of unique heterocycles, number of clusters, number of singletons, and number of compounds that have not been passed through MCFs. This analysis was performed with the settings used for the original CD data set (see subsection [Data](#)). In order to exclude the molecules that do not elicit convincing druglikeness from the different pools, we randomly selected 30K structures in each pool that satisfied the MU filter. Indeed, MU is a very severe criterion. This is due to the fact that it has a lower boundary as well as an upper boundary for each chemical descriptor it calculates. As a result, MU is able to cut nondruglike molecules. The results of the analysis are summarized in [Table 2](#).

We have also analyzed the druglikeness of the valid generated molecules. This analysis was performed using five druglikeness filters: RO5, Muegge, Egan, Veber, and Ghooose.^{70–73} The results (cf. Table S2 in the [Supporting Information](#)) showed that the valid molecules generated by ATNC pass all five druglikeness filters, almost two times more often than the valid molecules generated by ORGANIC.

Overall, the results showed that ATNC outperformed ORGANIC. ATNC produced more molecules per cluster, and it had a smaller portion of generated molecules that failed the MCFs (cf. [Discussion](#) for further details).

Table 2. ATNC and ORGANIC Compared Using Four Objective Reward Functions (IDC, IS, MU, and SP3) and Their Respective Performances Assessed with Five Different Features (Diversity, Number of Unique Heterocycles, Number of Clusters, Number of Singletons, and Number of Molecules That Did Not Pass MCFs)^a

	ATNC-IDC	ORGANIC-IDC	ATNC-IS	ORGANIC-IS
diversity	0.87	0.89	0.87	0.89
unique heterocycles	2790	10677	3287	9514
clusters	1285	659	1219	613
singletons	12110	24022	13493	24509
MCFs	947	7269	1213	7081
	ATNC-MU	ORGANIC-MU	ATNC-SP3	ORGANIC-SP3
diversity	0.87	0.8	0.87	0.85
unique heterocycles	2849	1827	2782	2432
clusters	1270	1261	1271	1283
singletons	12167	7600	11608	12254
MCFs	916	14245	961	8822

^aFor each pool of the valid molecules, the analysis was performed on 30K randomly selected molecules that satisfy MU criterion.

In Vitro Validation of Molecules Generated by ATNC. We have performed a preliminary biological evaluation in order

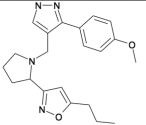
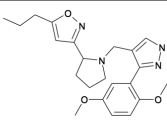
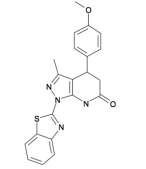
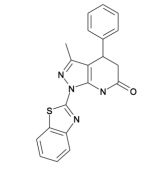
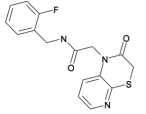
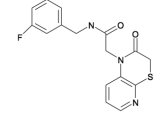
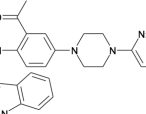
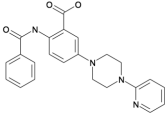
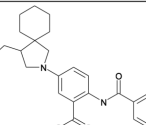
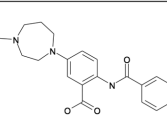
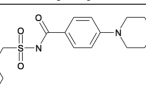
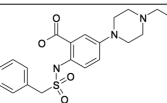
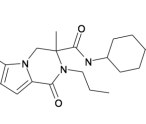
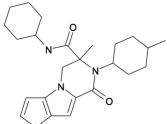
Nº	Generation	Compound	Kinase (example)	Inhibition (%)	Supplier
1			SGK1	101	AKos Consulting & Solutions AKOS005014423
2			Aurora A/B	95	AKos Consulting & Solutions AKOS022051824
3			MAPKAPK2	50	AKos Consulting & Solutions AKOS001851817
4			SGK1	52	AKos Consulting & Solutions AKOS004974309
5			SGK1	58	AKos Consulting & Solutions AKOS004974565
6			SGK1	59	AKos Consulting & Solutions AKOS002001036
7			SGK1	99	AKos Consulting & Solutions AKOS004939857

Figure 4. Representative examples of active compounds with high similarity in structure to the generated analogues.

to obtain an estimation of the efficacy of our approach. The procedure can be summarized as follows: First, we used the training data set of more than 30K small-molecule kinase inhibitors collected from the Thomson Integrity Database (<https://integrity.thomson-pharma.com/integrity/xmlxsl/>). Then, we generated 30K structures based on the ATNC model with IDC as an objective reward function. These generated structures were then filtered using 150 MCFs to exclude compounds containing undesirable moieties. Finally, the remaining 25K structures went through a 2D Tanimoto similarity analysis toward the reference molecules, and the structures with a Tanimoto coefficient greater than 0.7 were excluded from the set.

Subsequently, we performed a clustering analysis to group similar compounds together and also to keep the compounds with the highest diversity in each cluster. Thereby, we reduced the initial virtual chemical space to a more compact and normally distributed space. Then, we performed a thorough search among the patents and applications available in the SciFinder database (<https://origin-scifinder.cas.org>) to eliminate the structures eliciting a low IP factor, including those with kinase activity. The same procedure was performed using the ChEMBL database (<https://www.ebi.ac.uk/chembl/>). As a result, 5K unique structures were included in the final data set.

Using a set of available in-house collections of small-molecule compounds, we carried out a similarity search to select the molecules with the following properties: best Tanimoto score

(>0.7), bearing the same or isosteric scaffold, the same spatial geometry, and a similar substituent profile. Then, based on the search procedure described above, we selected 50 compounds for which no activity against kinase family enzymes had been reported.

These 50 compounds were purchased at a concentration of 10 μ M and screened in triplicate in a panel of different kinases using a routine ADP Hunter assay.⁷⁴ The compounds with similarity score of more than 70% were investigated on different target activities. As shown in Figure 4, seven of the selected molecules showed a good activity against different kinases (only one example is shown), while two compounds demonstrated a relatively high selectivity and a good inhibition potency. Based on these results, we can deduce that our *in silico* approach is capable of producing hit compounds. In the near future, we plan to synthesize several novel compounds from the generated pool and evaluate them against the kinase panel.

DISCUSSION

The comparison of the abilities of ATNC and ORGANIC models to generate unique SMILES strings shows that our engine is able to generate a higher percentage of unique molecules than ORGANIC. This result holds for the MU and SP3 objective reward functions (see Table 1). Although ORGANIC reached higher percentages of unique SMILES strings for the IDC and IS objective reward functions, it obtained percentages of valid SMILES strings almost 10 times

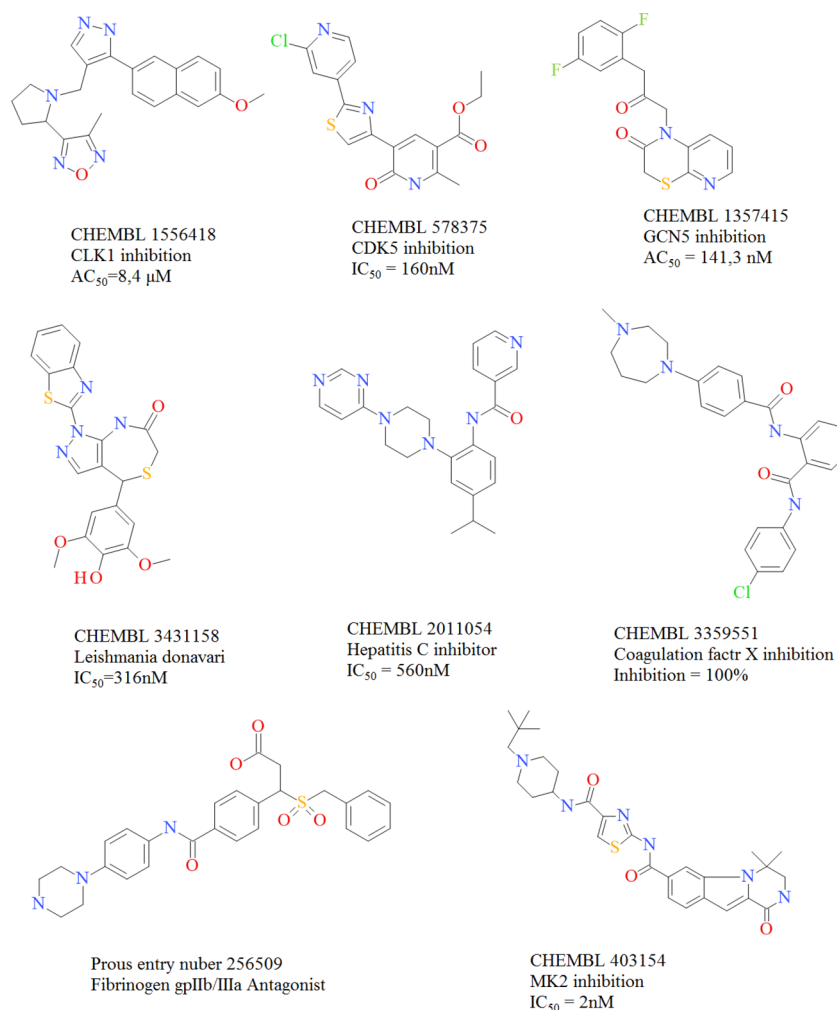


Figure 5. Molecular structures of the top eight compounds identified as potential novel cores of kinase inhibitors.

smaller than ATNC. This is interesting because IDC and IS are the most complex reward functions. Indeed, they integrate information about the interactions of the generated molecules, whereas the MU and SP3 reward functions simply consider generated molecules separately from each other.

The better performance of the ATNC model can be explained by the fact that ATNC uses the AT block and DNC as a generator. DNC helps the model to only select the “interesting” molecules, while the AT block helps the model to reach a good percentage of valid SMILES strings among the generated molecules. As a result, ATNC is able to produce a very stable percentage of valid and unique SMILES strings.

In this analysis, both methods generated pools of structures with a similar diversity. Nevertheless, the structures generated by ATNC are more attractive in terms of drug discovery and development because the number of compounds that successfully passed the MCFs is on average 9.5 times higher than ORGANIC (see Table 2). This is an important factor to take into account when considering the results showing that ORGANIC produced almost three times more unique heterocycles than ATNC (especially for the IDC and IS reward functions). Indeed, this “high” diversity is mainly due to entries that failed MCFs. These “false” heterocycles are represented by strained, unstable, and synthetically inaccessible cycles; such as epoxides, aziridines and azo- or thio-aziridines, thiiranes, diazo- and peroxide-containing cycles, not trivial π –

rich macrocycles, crown-like structures, etc. The structures obtained with the ATNC model show that ATNC is more sensitive to this important issue. For instance, as clearly shown in Figure 4, the structure 5 contains an attractive sp^3 -reach heterocyclic fragment that is positively charged at physiologically relevant pH: a methylamino-substituted spiropyrrolidine. Structures 3 and 7 have benzothiazinone and dihydrothienopyrrolo[1,2-*a*]pyrazinone scaffolds, respectively. These heterocycles are well distributed among biologically active compounds. Moreover, the structures generated by ATNC are quite homogeneous and are better clustered than the ones obtained by ORGANIC (with an acceptable number of singletons and outliers).

A deeper analysis of these compounds was performed using the following databases: Thomson Reuters Integrity, ChEMBL, and PubChem. Potency data have not been found for the exact structures of mentioned above compounds. Therefore, the searching of the nearest structural analogues was performed (see Figure 5). Among seven compounds presented on Figure 4, the highest similarity for protein kinase inhibitor was found for molecule 1. Molecule 1 has a structural core similar to ChEMBL compound (CHEMBL1556418). It contains the same pyrazole-pyrrolidine moiety and 1,2,5-oxadiazole instead of oxazole. This analogue has moderate activity ($IC_{50}=8.4\ \mu\text{M}$) against CLK1 kinase. Molecule 2 is mostly similar to the series of GSK inhibitors of parasitic *Leishmania donovani*; for

example, ChEMBL3431158 (IC_{50} = 316 nM). Among the protein kinase inhibitors, molecule 2 has a structure analogous to the highly active CDK5 inhibitor (ChEMBL578375), with the key binding moiety pyrazolo-pyridin-6-one (instead of 6-oxo-1,6-dihydropyridine).⁷⁵ Molecule 3 is very similar to histone acetyltransferase GCN5 inhibitor (ChEMBL1357415) (IC_{50} = 141.3 nM), with the same 1*h*-Pyrido[2,3-*b*][1,4]-thiazine-1-acetamide core. Molecule 4 is structurally close to the hepatitis C inhibitor (ChEMBL2011054) (IC_{50} = 560 nM).⁷⁶ The similarity search for derivatives of molecule 5 mainly revealed coagulation factor X inhibitors; for example, ChEMBL3359551 with an inhibition value of 100%.⁷⁷ Molecule 6 is similar to the Merck Integrin $\alpha\beta$ 3 (Fibrinogen gpIIb/IIIa) antagonist (integrity entry number: 256509).⁷⁸ Among the protein kinase inhibitors, no analogues of molecules 3–6 have been found. A similarity search for analogues of molecule 7 revealed ChEMBL403154 (IC_{50} = 2 nM), an MAPKAPK2 inhibitor. Thus, we can consider the presented scaffolds as novel cores of kinase inhibitors.

Although ATNC outperforms ORGANIC, it also inherits several limitations from the original approach. First, in both models there is no mechanism controlling the percentage of correctly reconstructed molecules (valid SMILES strings). A possible approach to address this issue is to replace the GAN part of the model with an AAE component. Second, ATNC and ORGANIC are SMILES-based models, and as a result, they cannot correctly employ the fragment-based objective reward functions. This is due to the fact that a SMILES string of a fragment cannot be found in a SMILES string of a molecule because of the SMILES format notation. Thus, training the models using fragment-based objective reward functions could confuse the models, leading to poor results. To get over this limitation one should change the SMILES notation to some special molecule representation, where each molecule will be represented in a unique way. For example, graph representation of molecules does not suffer from this problem and allows us to find each fragment correctly. Third, it must be pointed out that the created environment does not describe which molecule is really good in the sense that it is, for example, difficult to use several objective rewards simultaneously because the different reward functions may conflict with each other. A possible solution to this problem is to use modern multiobjective RL techniques.⁷⁹

■ ASSOCIATED CONTENT

■ Supporting Information

The Supporting Information is available free of charge on the ACS Publications website at DOI: 10.1021/acs.molpharmaceut.7b01137.

Model representation, SMILES, and druglikeness metrics (PDF)

■ AUTHOR INFORMATION

Corresponding Authors

*E-mail: putin@insilicomedicine.com.

*E-mail: alex@biogerontology.org.

ORCID

Quentin Vanhaelen: 0000-0002-4611-2046

Alex Zhavoronkov: 0000-0001-7067-8966

Notes

The authors declare the following competing financial interest(s): All authors are working for Insilico Medicine,

Inc., a private company involved in the development of DNN-based algorithms for molecular structures generation.

■ ACKNOWLEDGMENTS

The authors would like to acknowledge Government of the Russian Federation (grant 074-U01), Ministry of Education and Science of the Russian Federation, government grant 20.9907.2017/VU (for valuable expert opinion, discussion, and manuscript preparation), and Russian Science Foundation 17-74-30012, IBG RAS Ufa (for training sets preparation and analysis).

■ REFERENCES

- (1) Paul, S. M.; Mytelka, D. S.; Dunwiddie, C. T.; Persinger, C. C.; Munos, B. H.; Lindborg, S. R.; Schacht, A. L. How to improve R&D productivity: the pharmaceutical industry's grand challenge. *Nat. Rev. Drug Discovery* **2010**, *9*, 203–214.
- (2) DiMasi, J. A.; Grabowski, H. G.; Hansen, R. W. Innovation in the pharmaceutical industry: new estimates of R&D costs. *Journal of health economics* **2016**, *47*, 20–33.
- (3) Hughes, J. P.; Rees, S.; Kalindjian, S. B.; Philpott, K. L. Principles of early drug discovery. *British journal of pharmacology* **2011**, *162*, 1239–1249.
- (4) Hann, M. M.; Oprea, T. I. Pursuing the leadlikeness concept in pharmaceutical research. *Curr. Opin. Chem. Biol.* **2004**, *8*, 255–263.
- (5) Caraus, I.; Alsuwailam, A. A.; Nadon, R.; Makarenkov, V. Detecting and overcoming systematic bias in high-throughput screening technologies: a comprehensive review of practical issues and methodological solutions. *Briefings Bioinf.* **2015**, *16*, 974–986.
- (6) Macarron, R.; Banks, M.; Bojanic, D.; Burns, D. J.; et al. Impact of high-throughput screening in biomedical research. *Nat. Rev. Drug Discovery* **2011**, *10*, 188–195.
- (7) Chan, J. N.; Nislow, C.; Emili, A. Recent advances and method development for drug target identification. *Trends Pharmacol. Sci.* **2010**, *31*, 82–88.
- (8) Graul, A. I.; Revel, L.; Rosa, E.; Cruces, E. Overcoming the obstacles in the pharma/biotech industry: 2008 Update. *Drug News Perspect.* **2009**, *22*, 39.
- (9) Tareq Hassan Khan, M. Predictions of the ADMET properties of candidate drug molecules utilizing different QSAR/QSPR modelling approaches. *Curr. Drug Metab.* **2010**, *11*, 285–295.
- (10) Yamashita, F.; Hashida, M. In silico approaches for predicting ADME properties of drugs. *Drug Metab. Pharmacokinet.* **2004**, *19*, 327–338.
- (11) Tsaion, K.; Bottlaender, M.; Mabondzo, A. ADDME—Avoiding Drug Development Mistakes Early: central nervous system drug discovery perspective. *BMC Neurol.* **2009**, *9*, S1.
- (12) Waring, M. J.; Arrowsmith, J.; Leach, A. R.; Leeson, P. D.; Mandrell, S.; Owen, R. M.; Pairaud, G.; Pennie, W. D.; Pickett, S. D.; Wang, J.; et al. An analysis of the attrition of drug candidates from four major pharmaceutical companies. *Nat. Rev. Drug Discovery* **2015**, *14*, 475–486.
- (13) Hopkins, A. L.; Groom, C. R. The druggable genome. *Nat. Rev. Drug Discovery* **2002**, *1*, 727–730.
- (14) Lipinski, C.; Hopkins, A. Navigating chemical space for biology and medicine. *Nature* **2004**, *432*, 855–861.
- (15) Kubinyi, H. Drug research: myths, hype and reality. *Nat. Rev. Drug Discovery* **2003**, *2*, 665–668.
- (16) Stromgaard, K.; Krogsgaard-Larsen, P.; Madsen, U. *Textbook of Drug Design and Discovery*; CRC Press, 2009.
- (17) Rishon, G. M. Nonleadlikeness and leadlikeness in biochemical screening. *Drug Discovery Today* **2003**, *8*, 86–96.
- (18) Alqahtani, S. In silico ADME-Tox modeling: progress and prospects. *Expert Opin. Drug Metab. Toxicol.* **2017**, *13*, 1147–1158.
- (19) Ekins, S.; Rose, J. In silico ADME/Tox: the state of the art. *J. Mol. Graphics Modell.* **2002**, *20*, 305–309.

- (20) Kumar, R.; Sharma, A.; Varadwaj, P. K. A prediction model for oral bioavailability of drugs using physicochemical properties by support vector machine. *Journal of natural science, biology, and medicine* **2011**, *2*, 168.
- (21) Gilad, Y.; Nadassy, K.; Senderowitz, H. A reliable computational workflow for the selection of optimal screening libraries. *J. Cheminf.* **2015**, *7*, 61.
- (22) Furka, A.; SEBESTYÉN, F.; ASGEDOM, M.; DIBÓ, G. General method for rapid synthesis of multicomponent peptide mixtures. *Int. J. Pept. Protein Res.* **1991**, *37*, 487–493.
- (23) Lehn, J.-M.; Ramstrom, O. Generation and screening of a dynamic combinatorial library. US Patent App. US 10/220,470, 2001.
- (24) Balakin, K.; Ivanenkov, Y.; Savchuk, N. Compound library design for target families. *Methods Mol. Biol.* **2009**, *575*, 21–46.
- (25) Kapetanovic, I. Computer-aided drug discovery and development (CADD): in silico-chemico-biological approach. *Chem.-Biol. Interact.* **2008**, *171*, 165–176.
- (26) Ganesan, A. Recent developments in combinatorial organic synthesis. *Drug Discovery Today* **2002**, *7*, 47–55.
- (27) Huang, R.; Leung, I. K. Protein-Directed Dynamic Combinatorial Chemistry: A Guide to Protein Ligand and Inhibitor Discovery. *Molecules* **2016**, *21*, 910.
- (28) Yu, H.; Adedoyin, A. ADME-Tox in drug discovery: integration of experimental and computational technologies. *Drug Discovery Today* **2003**, *8*, 852–861.
- (29) Szymański, P.; Markowicz, M.; Mikiciuk-Olasik, E. Adaptation of high-throughput screening in drug discovery to toxicological screening tests. *Int. J. Mol. Sci.* **2012**, *13*, 427–452.
- (30) Zhang, Z.; Guan, N.; Li, T.; Mais, D. E.; Wang, M. Quality control of cell-based high-throughput drug screening. *Acta Pharm. Sin. B* **2012**, *2*, 429–438.
- (31) Patel, D. A.; Patel, A. C.; Nolan, W. C.; Zhang, Y.; Holtzman, M. J. High throughput screening for small molecule enhancers of the interferon signaling pathway to drive next-generation antiviral drug discovery. *PLoS One* **2012**, *7*, e36594.
- (32) Schneider, G.; Fechner, U. Computer-based de novo design of drug-like molecules. *Nat. Rev. Drug Discovery* **2005**, *4*, 649–663.
- (33) LeCun, Y.; Bengio, Y.; Hinton, G. Deep learning. *Nature* **2015**, *521*, 436–444.
- (34) Min, S.; Lee, B.; Yoon, S. Deep learning in bioinformatics. *Briefings Bioinf.* **2016**, bbw068.
- (35) Mamoshina, P.; Vieira, A.; Putin, E.; Zhavoronkov, A. Applications of deep learning in biomedicine. *Mol. Pharmaceutics* **2016**, *13*, 1445–1454.
- (36) Klefogiannis, D.; Kalnis, P.; Bajic, V. B. DEEP: a general computational framework for predicting enhancers. *Nucleic Acids Res.* **2015**, *43*, e6–e6.
- (37) Kim, S. G.; Theera-Ampornpunt, N.; Fang, C.-H.; Harwani, M.; Grama, A.; Chatterji, S. Opening up the blackbox: an interpretable deep neural network-based classifier for cell-type specific enhancer predictions. *BMC Syst. Biol.* **2016**, *10*, 54.
- (38) Aliper, A.; Plis, S.; Artemov, A.; Ulloa, A.; Mamoshina, P.; Zhavoronkov, A. Deep learning applications for predicting pharmacological properties of drugs and drug repurposing using transcriptomic data. *Mol. Pharmaceutics* **2016**, *13*, 2524–2530.
- (39) Xu, Y.; Pei, J.; Lai, L. Deep Learning Based Regression and Multi-class Models for Acute Oral Toxicity Prediction. *J. Chem. Inf. Model.* **2017**, *57*, 2672.
- (40) Lenselink, E. B.; Dijke, N.; Bongers, B.; Papadatos, G.; Vlijmen, H. W.; Kowalczyk, W.; IJzerman, A. P.; Westen, G. J. Beyond the hype: deep neural networks outperform established methods using a ChEMBL bioactivity benchmark set. *J. Cheminf.* **2017**, *9*, 45.
- (41) Weininger, D. SMILES, a chemical language and information system. 1. Introduction to methodology and encoding rules. *J. Chem. Inf. Model.* **1988**, *28*, 31–36.
- (42) Heller, S.; McNaught, A.; Stein, S.; Tchekhovskoi, D.; Pletnev, I. InChI-the worldwide chemical structure identifier standard. *J. Cheminf.* **2013**, *5*, 7.
- (43) Willett, P. *Chemoinformatics and Computational Chemical Biology*; Springer, 2010; pp 133–158.
- (44) Makhzani, A.; Shlens, J.; Jaitly, N.; Goodfellow, I.; Frey, B. Adversarial autoencoders. *arXiv preprint arXiv:1511.05644*, **2015**.
- (45) Kadurin, A.; Nikolenko, S.; Khrabrov, K.; Aliper, A.; Zhavoronkov, A. druGAN: An Advanced Generative Adversarial Autoencoder Model for de Novo Generation of New Molecules with Desired Molecular Properties in Silico. *Mol. Pharmaceutics* **2017**, *14*, 3098.
- (46) Kingma, D. P.; Welling, M. Auto-encoding variational bayes. *arXiv preprint arXiv:1312.6114*, **2013**.
- (47) Gómez-Bombarelli, R.; Duvenaud, D.; Hernández-Lobato, J. M.; Aguilera-Iparraguirre, J.; Hirzel, T. D.; Adams, R. P.; Aspuru-Guzik, A. Automatic chemical design using a data-driven continuous representation of molecules. *ACS Cent. Sci.* **2016**, *4*, 268.
- (48) Bickerton, G. R.; Paolini, G. V.; Besnard, J.; Muresan, S.; Hopkins, A. L. Quantifying the chemical beauty of drugs. *Nat. Chem.* **2012**, *4*, 90–98.
- (49) Ertl, P.; Schuffenhauer, A. Estimation of synthetic accessibility score of drug-like molecules based on molecular complexity and fragment contributions. *J. Cheminf.* **2009**, *1*, 8.
- (50) Bjerrum, E. J.; Threlfall, R. Molecular Generation with Recurrent Neural Networks (RNNs). *arXiv preprint arXiv:1705.04612*, **2017**.
- (51) Cherti, M.; Kegl, B.; Kazakci, A. De Novo Drug Design with Deep Generative Models: an Empirical Study, 2017.
- (52) Goodfellow, I.; Pouget-Abadie, J.; Mirza, M.; Xu, B.; Warde-Farley, D.; Ozair, S.; Courville, A.; Bengio, Y. Generative adversarial nets. *Advances in Neural Information Processing Systems* **2014**, 2672–2680.
- (53) Yu, L.; Zhang, W.; Wang, J.; Yu, Y. SeqGAN: Sequence Generative Adversarial Nets with Policy Gradient. *AAAI arXiv preprint arXiv:1609.05473*, **2017**.
- (54) Williams, R. J. Simple statistical gradient-following algorithms for connectionist reinforcement learning. *Machine learning* **1992**, *8*, 229–256.
- (55) Guimaraes, G. L.; Sanchez-Lengeling, B.; Farias, P. L. C.; Aspuru-Guzik, A. Objective-Reinforced Generative Adversarial Networks (ORGAN) for Sequence Generation Models. *arXiv preprint arXiv:1705.10843*, **2017**.
- (56) Benhenda, M. ChemGAN challenge for drug discovery: can AI reproduce natural diversity? *arXiv:1708.08227v3*, **2017**, [stat.ML].
- (57) Olivecrona, M.; Blaschke, T.; Engkvist, O.; Chen, H. Molecular de-novo design through deep reinforcement learning. *J. Cheminf.* **2017**, *9*, 48.
- (58) Sanchez-Lengeling, B.; Outeiral, C.; Guimaraes, G. L.; Aspuru-Guzik, A. Optimizing distributions over molecular space. An Objective-Reinforced Generative Adversarial Network for Inverse-design Chemistry (ORGANIC). *ChemRxiv preprint: 5309668*, **2017**.
- (59) Irwin, J. J.; Sterling, T.; Mysinger, M. M.; Bolstad, E. S.; Coleman, R. G. ZINC: a free tool to discover chemistry for biology. *J. Chem. Inf. Model.* **2012**, *52*, 1757–1768.
- (60) ORGANIC is exactly the ORGAN model, and further in the text we will refer only as ORGANIC because it is more related to molecular design than ORGAN.
- (61) Graves, A.; et al. Hybrid computing using a neural network with dynamic external memory. *Nature* **2016**, *538*, 471.
- (62) Sutton, R. S.; Barto, A. G. *Reinforcement Learning: An Introduction*; MIT Press: Cambridge, 1998; Vol. 1.
- (63) Sutton, R. S.; McAllester, D. A.; Singh, S. P.; Mansour, Y. Policy gradient methods for reinforcement learning with function approximation. *Advances in Neural Information Processing Systems* **2000**, 1057–1063.
- (64) Hochreiter, S.; Schmidhuber, J. Long short-term memory. *Neural computation* **1997**, *9*, 1735–1780.
- (65) Rogers, D.; Hahn, M. Extended-connectivity fingerprints. *J. Chem. Inf. Model.* **2010**, *50*, 742–754.
- (66) Zhang, B.; Vogt, M.; Maggiora, G. M.; Bajorath, J. Design of chemical space networks using a Tanimoto similarity variant based

upon maximum common substructures. *J. Comput.-Aided Mol. Des.* **2015**, *29*, 937–950.

(67) Abadi, M.; Barham, P.; Chen, J.; Chen, Z.; Davis, A.; Dean, J.; Devin, M.; Ghemawat, S.; Irving, G.; Isard, M.; et al. TensorFlow: A System for Large-Scale Machine Learning. *OSDI* **2016**, 265–283.

(68) Further in the text we will refer to the models and their objective reward functions using abbreviation such as ATNC-IDC, ORGANIC-MU, etc.

(69) Landrum, G. RDKit: Open-source cheminformatics. <http://www.rdkit.org>.

(70) Muegge, I. Selection criteria for drug-like compounds. *Med. Res. Rev.* **2003**, *23*, 302–321.

(71) Egan, W. J.; Merz, K. M.; Baldwin, J. J. Prediction of drug absorption using multivariate statistics. *J. Med. Chem.* **2000**, *43*, 3867–3877.

(72) Veber, D. F.; Johnson, S. R.; Cheng, H.-Y.; Smith, B. R.; Ward, K. W.; Kopple, K. D. Molecular properties that influence the oral bioavailability of drug candidates. *J. Med. Chem.* **2002**, *45*, 2615–2623.

(73) Ghose, A. K.; Viswanadhan, V. N.; Wendoloski, J. J. A knowledge-based approach in designing combinatorial or medicinal chemistry libraries for drug discovery. 1. A qualitative and quantitative characterization of known drug databases. *J. Comb. Chem.* **1999**, *1*, 55–68.

(74) Von Ahsen, O.; Bömer, U. High-Throughput Screening for Kinase Inhibitors. *ChemBioChem* **2005**, *6*, 481–490.

(75) Kaller, M. R.; Zhong, W.; Henley, C.; Magal, E.; Nguyen, T.; Powers, D.; Rzasa, R. M.; Wang, W.; Xiong, X.; Norman, M. H. Design and synthesis of 6-oxo-1, 6-dihydropyridines as CDK5 inhibitors. *Bioorg. Med. Chem. Lett.* **2009**, *19*, 6591–6594.

(76) Cheng, C. C.; Huang, X.; Shipps, G. W.; Wang, Y.-S.; Wyss, D. F.; Soucy, K. A.; Jiang, C.-k.; Agrawal, S.; Ferrari, E.; He, Z.; Huang, H.-C. Pyridine carboxamides: Potent palm site inhibitors of HCV NS5B Polymerase. *ACS Med. Chem. Lett.* **2010**, *1*, 466–471.

(77) Ishihara, T.; Koga, Y.; Iwatsuki, Y.; Hirayama, F. Identification of potent orally active factor Xa inhibitors based on conjugation strategy and application of predictable fragment recommender system. *Bioorg. Med. Chem.* **2015**, *23*, 277–289.

(78) Wai, J.; Fisher, T. E.; Duggan, M. E.; Hartman, G. D.; Perkins, J. J. Fibrinogen receptor antagonists. US Patent US 5,780,480, 1998.

(79) Mossalam, H.; Assael, Y. M.; Roijers, D. M.; Whiteson, S. Multi-objective deep reinforcement learning. *arXiv preprint arXiv:1610.02707*, **2016**.

■ NOTE ADDED AFTER ASAP PUBLICATION

After this paper was published ASAP March 30, 2018, a new reference was added to the reference list as ref 56. The revised version was reposted on May 21, 2018.