




# Reinforced Adversarial Neural Computer for de Novo Molecular Design 논문 리뷰

J. Chem. Inf. Model. 2018, 58, 1194–1204

2020.09.04

# Reinforced Adversarial Neural Computer for *de Novo* Molecular Design

Evgeny Putin,<sup>\*,†,‡</sup>  Arip Asadulaev,<sup>‡</sup> Yan Ivanenkov,<sup>†,§,||</sup> Vladimir Aladinskiy,<sup>†,¶</sup>  
Benjamin Sanchez-Lengeling,<sup>||</sup> Alán Aspuru-Guzik,<sup>||,⊥</sup>  and Alex Zhavoronkov<sup>\*,†,#</sup> 



JOURNAL OF  
CHEMICAL INFORMATION  
AND MODELING

Chemical informatics, Molecular modeling

# Research Background

*In silico* modeling의 중요성

- Drug Design and Development (DDD)의 초기단계 중 하나
- 컴퓨터 기반 접근법은 잘 연구되어오고 있으나, 딥러닝과 관련된 연구의 적용은 초기 단계에 머무르고 있다

최근 머신러닝과 AI의 발달로, 새로운 molecule 생성에 데이터 기반 접근법이 적용

Data driven approach	Complete drug discovery
<ul style="list-style-type: none"><li>- 빠른 속도</li><li>- 많은 양의 sample 생성</li></ul>	<ul style="list-style-type: none"><li>- 느린 속도</li><li>- 비용 과다</li></ul>

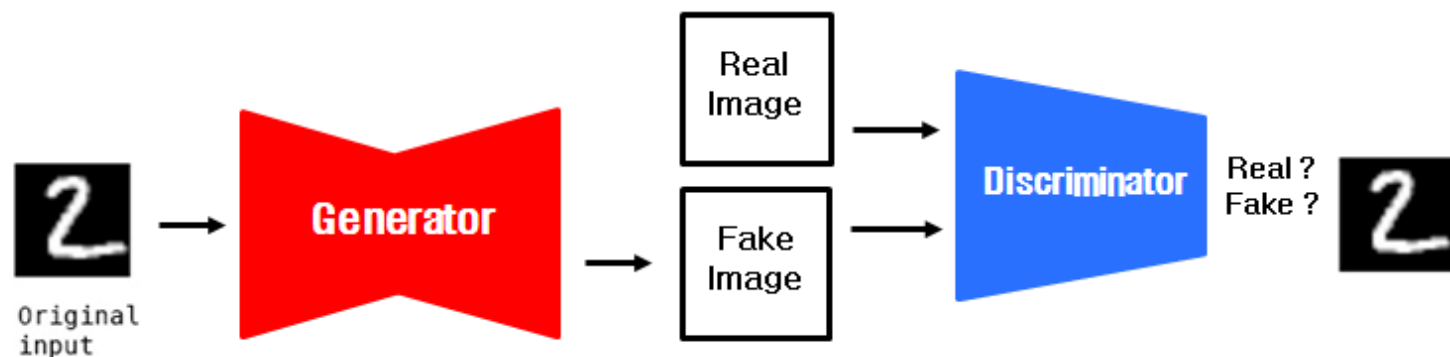
## Research Background

- Deep Neural Network(DNN)은 최근 DDD에서 엄청난 진전을 이루었다.
- 새로운 화학적 구조를 생성하는 데에 활발하게 사용됨
- 광범위한 화학적 공간 제공
- 최근 약물의 용도 변경 및 주요 특성(property)를 예측하는 데에 DNN 적용
- 하지만 이러한 네트워크들은 실제로 적용되기에는 여러 constraints에 부딪힐 것이기 때문에 시도는 계속해서 이루어지고 있지만 아직 해결되지는 않은 부분이 많다.

앞서 언급한 문제들을 해결하기 위한 가장 성공적인 시도 중 하나는 adversarial training과 강화학습(RL)을 결합한 ORGANIC architecture이다.

# Research Background

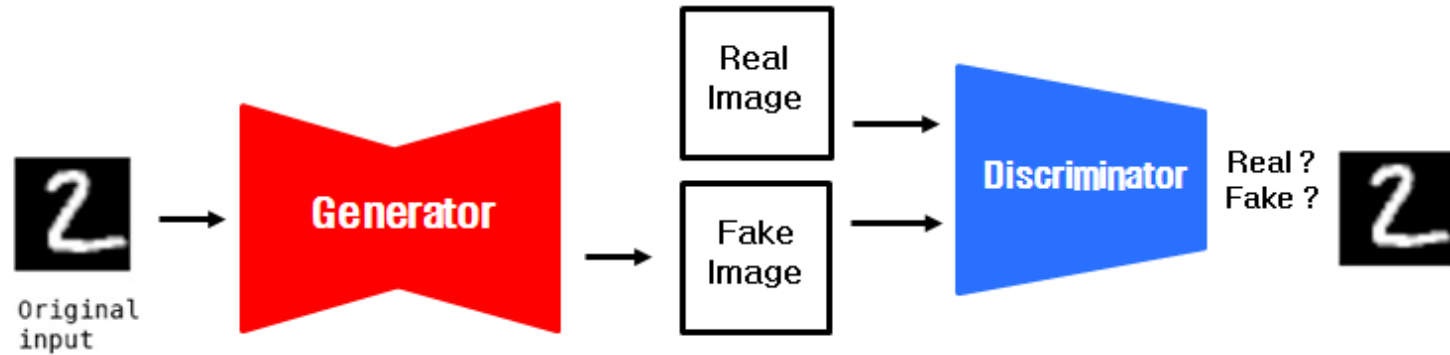
## ORGANIC



- Adversarial training(적대적 훈련)과 RL(강화학습)의 결합
- ORGANIC은 generator로 하여금 분자의 SMILES string 표현을 이용하여 다음을 생성하도록 한다.
  - discriminator 모델에 의해 penalty을 받거나 reward을 받는 분자 구조
  - drug-likeness등 원하는 특성을 계량화하는 reward function
- Discriminator는 데이터 분포에 기반하여 제안된 분자를 가짜 또는 진짜 분자로 분류하려고 시도하는데, 이는 본질적으로 generator가 실제적인 샘플을 만들도록 유인한다.
- Predefined desired property을 가진 분자를 생성할 수 있다.

# Research Background

## ORGANIC



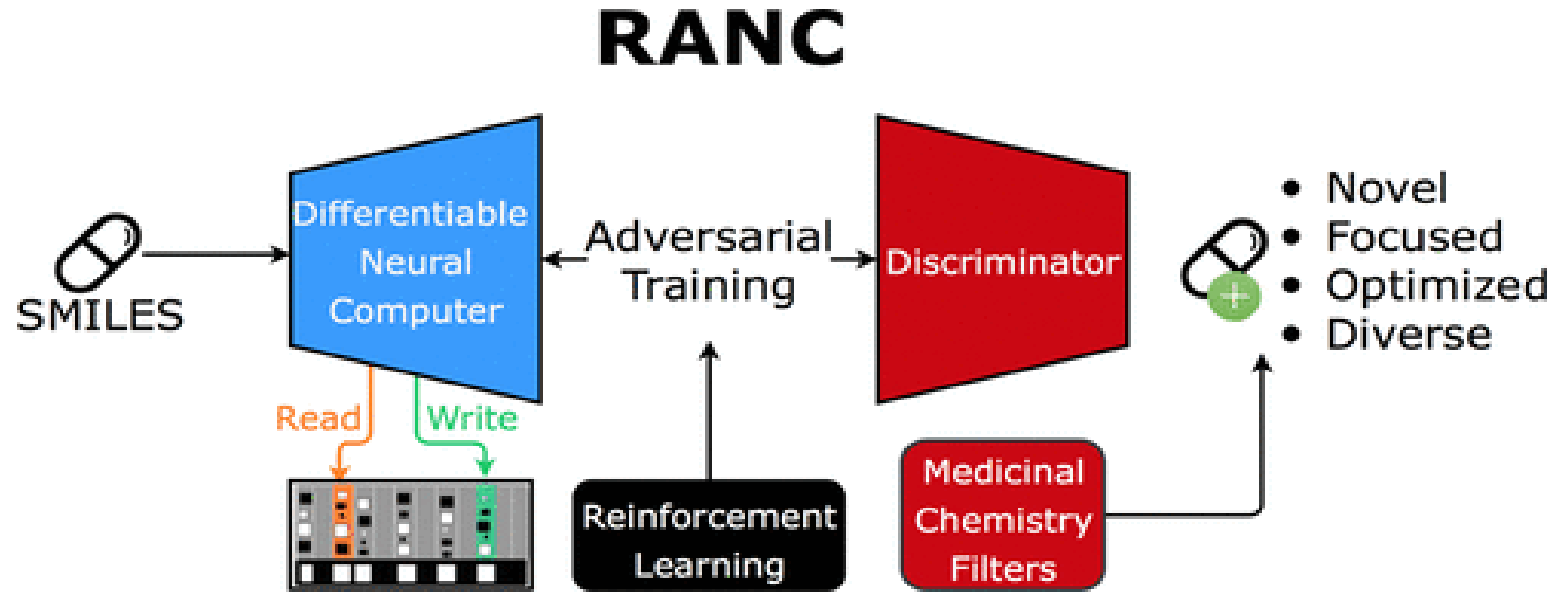
### 문제점

1. mode-collapse
2. perfect discriminator problem

결국 adversarial training과 강화학습(RL)을 결합한 형태임에도, adversarial training 과정에서 이런 문제가 발생하기 때문에 그냥 강화 학습(RL)에 불과하게 되는 것이다.

# RANC(Resinced Attorical Neural Computer)

DNC architecture과 RL(강화학습) 및 adversarial training을 결합한 최초의 접근방식



- 새로운 저분자 유기구조의 새로운 합성 설계(새로운 분자를 개발)를 위해 사용됨
- ORGANIC(DNN-based)의 성능을 여러 지표에서 능가
- 주요 chemical descriptor의 분포와 SMILES 문자열의 길이와 일치하는 구조를 생성
- 시간 절약, 넓은 화학적 공간 커버

# Contributions

1. DNC를 통해 외부 저장소가 있는 generator 구조를 제안함으로써 adversarial training 과정에서 문제가 발생하지 않고, RANC molecules가 training molecules의 길이와 주요 molecule descriptor 분포와 일치한다는 것을 보여주었다. 위의 것들은 모두 ORGANIC에 비해 개선된 점
2. ORGANIC 및 RANC 모델에 의해 생성된 분자 뿐만 아니라 이전에 보고되지 않았던 training molecule에 대한 광범위한 화학 분석을 수행한다. 이 분석을 통해 prefiltered training set이, 생성된 structures의 품질에 어떤 영향을 미치는지 평가하고 두 모델에서 생성된 분자를 의약화학 관점에서 비교한다.

# Method

## Proposal of DNC (Differentiable Neural Computer)

- RNN이나 LSTM을 사용하는 방식의 문제점으로는 exposure bias의 발생이 있다. 이에 이러한 문제점을 해결할 수 있는 새로운 아키텍처인 'DNC'가 제안되었다.
- 실제로 진행된 Graph experiment에서 200만개의 데이터를 학습시켰을 때 LSTM은 37%의 accuracy를 달성한 반면, 백만개만 학습한 시점에 DNC는 이미 98.8%의 accuracy를 달성했다.
- 이러한 DNC가 강화학습(RL)과 결합하면 Reward와 관련된 sequence 구조를 읽고 쓰는 것이 가능해진다.

## RANC는 LSTM 대신 DNC를 적용한다.

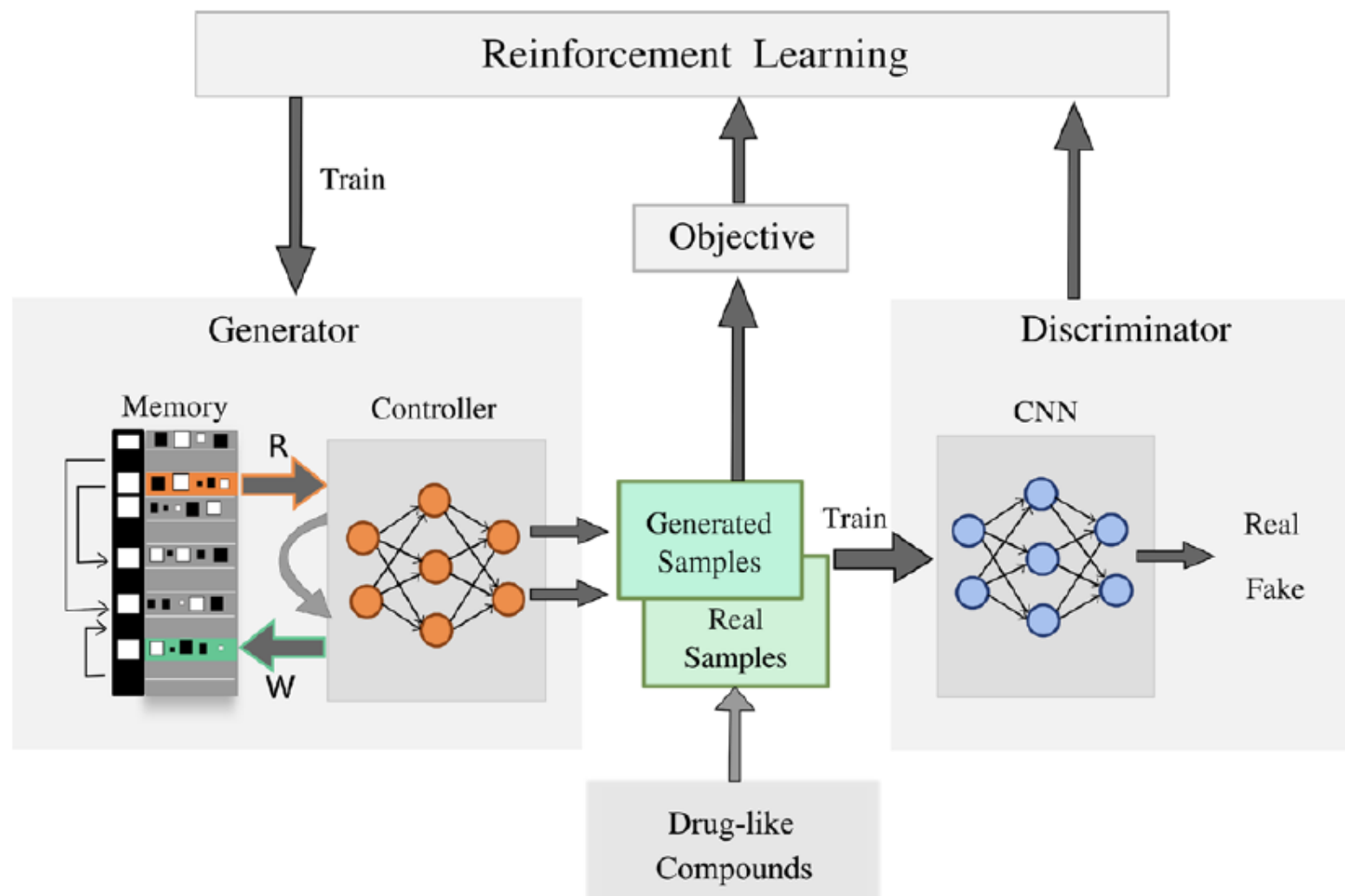
SMILES-based  $\theta$ -parametrized molecule generator  $G_\theta$ 로 DNC를 사용하는 것의 장점:

1. 매우 복잡한 시퀀스를 기억하는 DNC의 능력
2. LSTM에 비해 훨씬 긴 시퀀스  $Y_{1:T} = (y_1, \dots, y_T)$ 를 생성할 수 있다 ( $T$ =Sequence 길이)



# Method

RANC 모델의 개략도



# Method

## RANC 모델의 전반적인 훈련 절차

---

### Algorithm 1 Training procedure of the RANC model

---

**Require:** DNC generator policy  $G_\theta$  ; roll-out policy  $G_\beta$ ; Discriminator  $D_\phi$ ; objective  $O$ ; set  $S$  of sequences of length  $T$ ;  
Initialize  $G_\theta, D_\phi$  with random weights  $\theta, \phi$ ; Initialize  $\lambda, \zeta$ ;  
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  $S$ ;

- 1: **repeat**
- 2:   **for** g-steps **do**
- 3:     Generate a sequence  $Y_{1:T} = (y_1, \dots, y_T) \sim G_\theta$ ; additionally writing and reading from the external memory.
- 4:     **for**  $t$  in  $1 : T$  **do**
- 5:       Compute  $Q(s = Y_{1:t-1}, a = y_t)$ ; by Eq.(4) by calculating rewards using Eq.(5)
- 6:     **end for**
- 7:     Update generator parameters  $\theta$  via Eq.(6)
- 8:     **for** d-steps **do**
- 9:       Train  $D_\phi$  for  $k$  epochs by Eq.(2) via only verified generated negative samples using  $G_\theta$  and positive  $S$ ;
- 10:    **end for**
- 11: **end for**
- 12: **until** model coverage.

---

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

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

$$\text{MC}^{G_\theta}(Y_{1:t}; N) = \{Y_{1:T}^1, \dots, Y_{1:T}^N\} \quad (3)$$

$$Q(s = Y_{1:t-1}, a = y_t) = \begin{cases} \frac{1}{N} \sum_{n=1}^N R(Y_{1:T}^n), Y_{1:T}^n & \text{if } t < T \\ \in \text{MC}^{G_\theta}(Y_{1:T}; N), & \\ R(Y_{1:T}), & \text{if } t = T \end{cases} \quad (4)$$

$$R(Y_{1:T}) = \lambda \cdot D_\phi(Y_{1:T}) + (1 - \lambda) \cdot O(Y_{1:T}) \quad (5)$$

$$\nabla_\theta J(\theta) \simeq \frac{1}{T} \sum_{t=1}^T E_{y_t \sim G_\theta(y_t | Y_{1:t-1})} [\nabla_\theta \log G_\theta(y_t | Y_{1:t-1}) \cdot Q(Y_{1:t-1}, y_t)] \quad (6)$$

# Result

## Dataset

기준이 되는 분자(reference molecule)의 품질이 생성된 구조물의 특성에 어떤 영향을 미치는지 추정하기 위해 두 개의 training dataset에 대해 수행됨

- 1. Drugs data set: 15,000
- 2. ChemDiv(CD) data set: 15,000

Prefiltered training set가 생성된 structures의 품질에 어떤 영향을 미치는지 평가하는 것이 목적이다. Train DB의 화학적/통계적 특징 연구

Table 1. Key Chemical Statistical Parameters for the CD and Drugs Data Sets

data set	CD	Drugs
size	15000	15000
diversity	0.87	0.86
fragments <sup>a</sup>	8510	21773
clusters	793	223
singletons	3035	13554
heterocycles <sup>b</sup>	644	1052
cluster size <sup>c</sup>	18.9	67.2
MCF (%)	0.14	32.6
ROS (%)	77.9	89.7

- Tanimoto similarity를 이용해서 구조의 유사성에 의해 분자를 서로 다른 군집으로 분배
- 이때 군집 당 최소 구조물 수는 경험적으로 선택
- 모든 계산은 ChemoSoft 소프트웨어에 의해 수행됨

<sup>a</sup>Number of unique fragments considered to be diverse. <sup>b</sup>Number of unique heterocycles. <sup>c</sup>Average cluster size. Percentages are taken over the entire data set.

# Result

## Settings

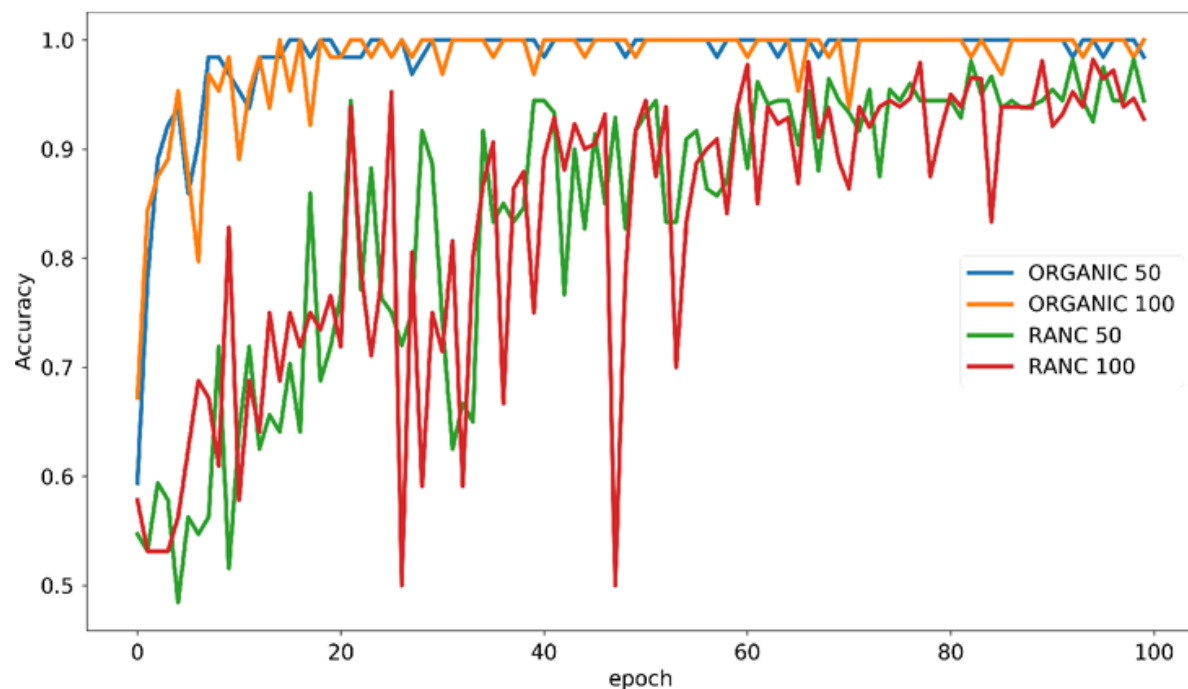
RANC		ORGANIC
동일한 hyperparameter와 설정으로 모든 실험을 수행 (training epoch 제외)		
Input	SMILES string encoded as a one-hot vector	
Training settings	NVIDIA Titan Pascal GPUs with 128 Gb of RAM	
Batch size	64	
Generator	<p>DNC generator <math>G_{\theta}</math></p> <ul style="list-style-type: none"><li>- LSTM controller with 256 neurons</li><li>- one read head <math>R</math></li><li>- memory size of <math>N \cdot W</math> (<math>N</math> = length of train data dictionary, <math>W</math> = max length of train data sequences)</li><li>- RMSprop (<math>1 \times 10^{-3}</math> leaning rate, 0.9 momentum)</li><li>- Coefficient <math>\lambda = 0.5</math></li><li>- coefficient <math>\zeta = 0.8</math></li></ul>	
Discriminator	Same	

# Result

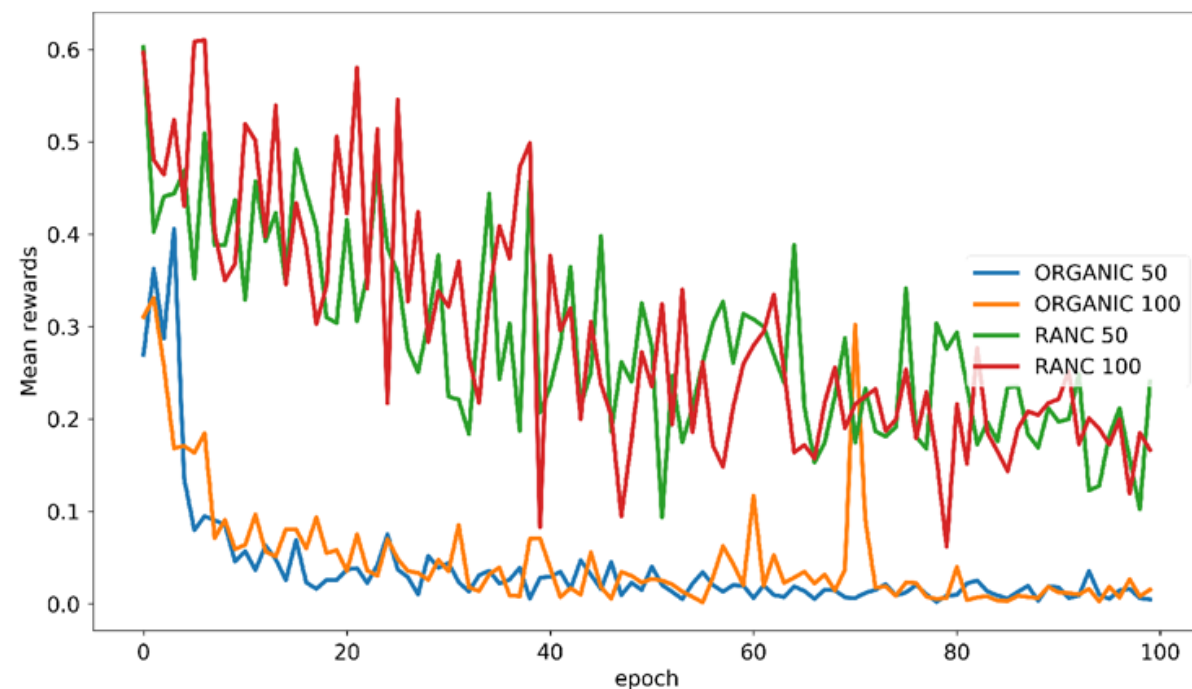
## Perfect Discriminator Problem

Generator과 discriminator는 각각 50, 100 epoch로 고정되었다.

Classification accuracy of discriminators



Mean values of the adversarial rewards of the generators



Reference 분자의 특성을 유지하는 분자 생성의 경우 generator에 대해 adversarial reward를 제공하는 것이 중요하다. 위의 분석으로부터 RANC 모델은 perfect discriminator problem이 발생하지 않는다는 것을 알 수 있다.

# Result

## Training and Evaluation

RANC와 ORGANIC 모델은 각각 CD, Drugs dataset에 train되었다. RO5를 객관적인 award function으로 사용

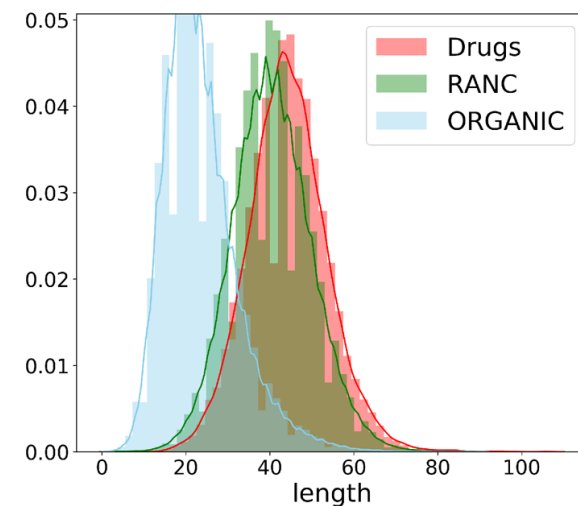
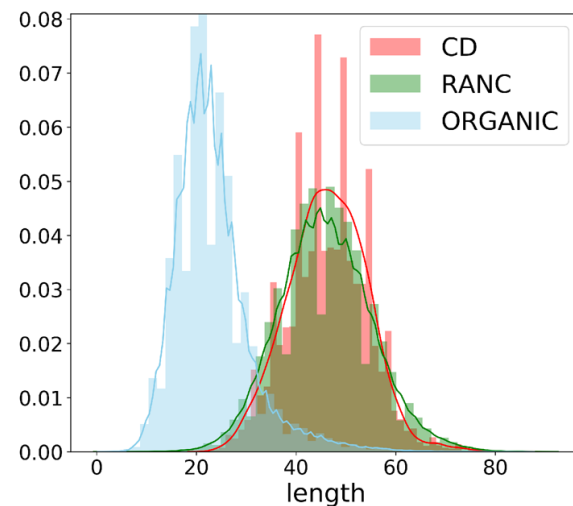
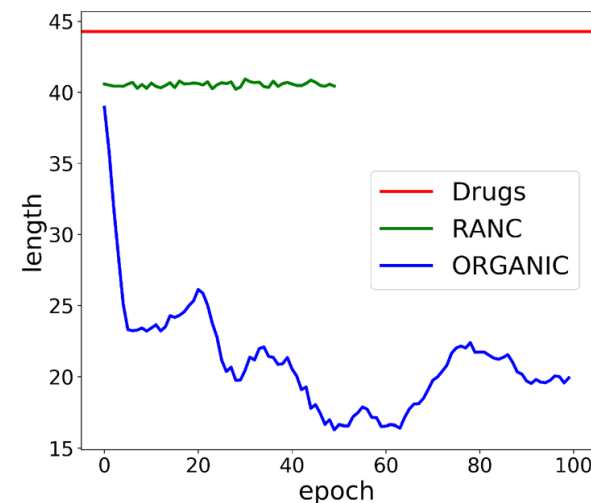
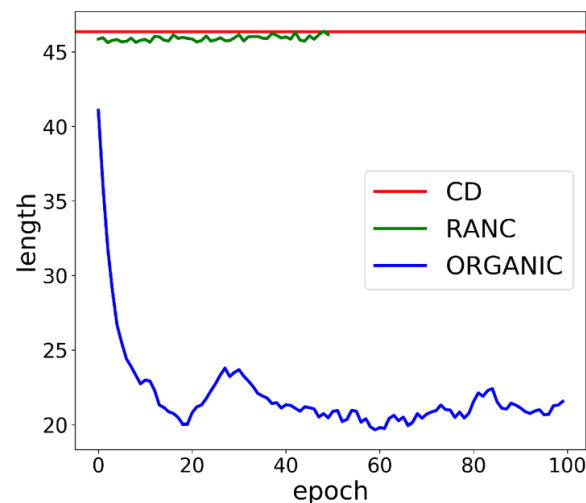
생성된 SMILES string에 대한 통계 분석

**Table 2. Statistics for the Generated SMILES Strings by RANC and ORGANIC Models on the CD and Drugs Data Sets**

	CD		Drugs	
data set approach	RANC	ORGANIC	RANC	ORGANIC
ave. length	46	23	40	23
valid (%)	58	87	76	92
unique (%)	48	18	76	24

ORGANIC은 두 데이터 세트 모두 유효 스마일 문자열의 생성율은 매우 좋지만, 긴 것보다 짧은 유효 SMILES 스트링의 발생이 훨씬 쉽다

각 train epoch에 생성된 SMILES 문자열의 평균 길이



# Result

## Training and Evaluation

RANC	ORGANIC
train molecule과 비슷한 길이의 SMILES 문자열을 안정적으로 생성	유효한 SMILES 생성율은 매우 높지만, RL epoch를 거치며 기존 molecule에 비해 훨씬 짧은 SMILES을 생성

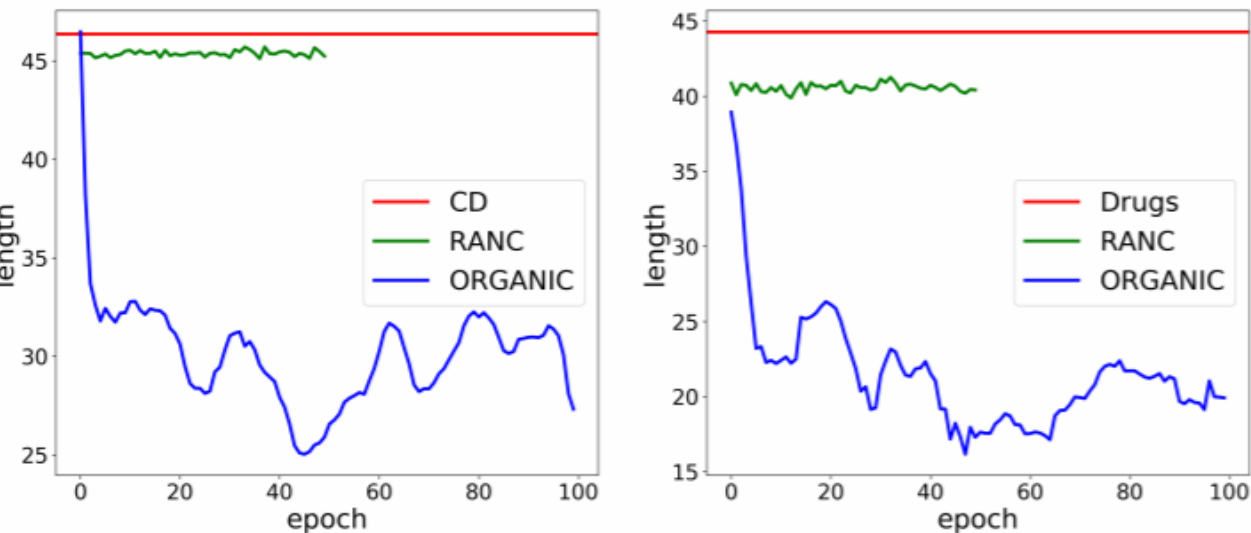
ORGANIC 모델에서 **perfect discriminator problem**이 발생한다는 또 하나의 증거가 된다.

이로 인해 adversarial training의 영향을 사라지고 강화학습(RL)의 만이 남아 generator의 학습 과정에 영향을 주게 되기 때문에 위와 같은 현상이 발생하는 것이다.

# Result

## Training and Evaluation

Muegge 기준을 객관적인 reward로 사용했을 때  
CD의 ORGANIC 및 RANC 모델에 의한 각 training epoch에 생성되는  
각 SMILES 문자열의 평균 길이



R05 및 Muegge 기준을 사용하여  
CD 및 약물 데이터 세트의 RANC와 유기농 모델에 의해 생성된  
unique하고 valid한 SMILES 문자열의 metric

Dataset	RANC	ORGANIC	RANC	ORGANIC
	ChemDiv	Lipinski	ChemDiv	Muegge
Total	320 k	896 k	320 k	896 k
Valid (%)	58	87	60	77
Unique (%)	48	18	51	33
Length	46	23	45	28

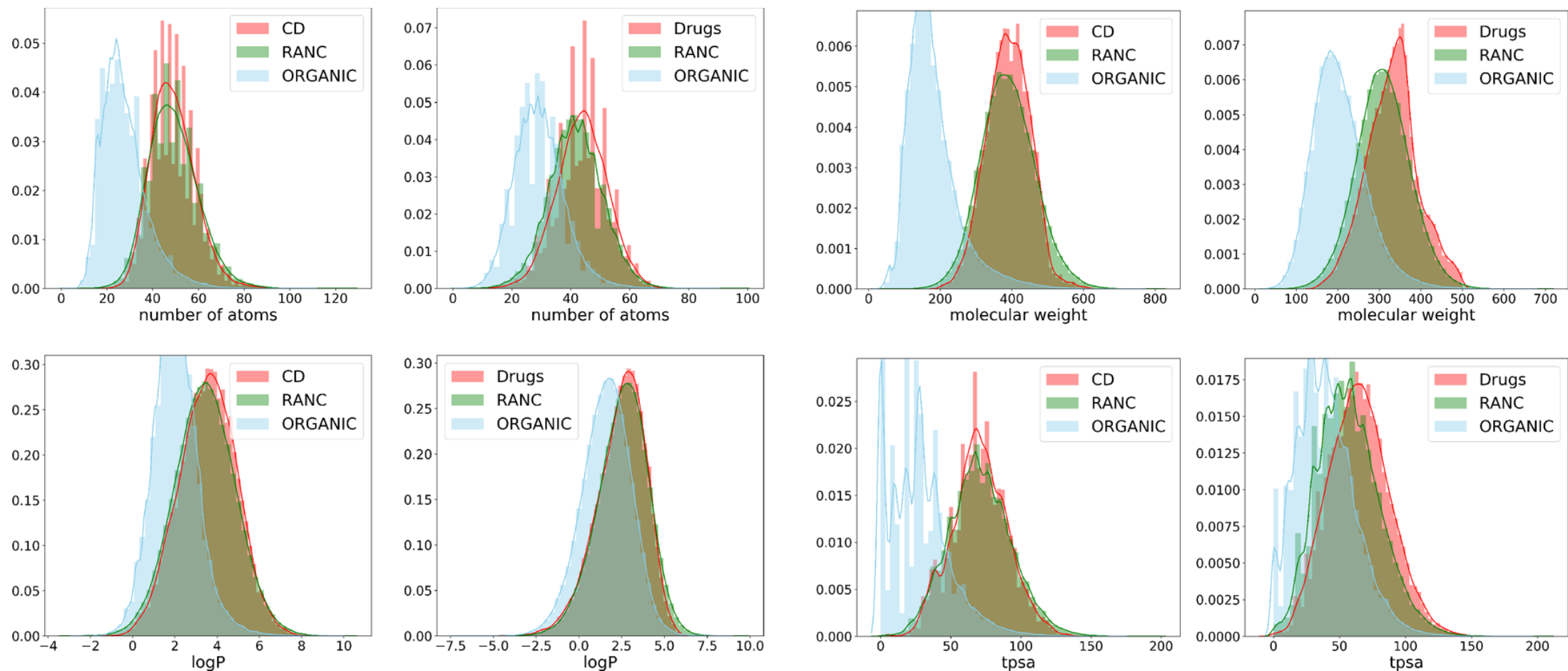
Dataset	RANC	ORGANIC	RANC	ORGANIC
	Drugs	Lipinski	Drugs	Muegge
Total	320 k	896 k	320 k	896 k
Valid (%)	76	92	71	85
Unique (%)	76	24	66	30
Length	40	23	41	23



# Result

## Training and Evaluation

생성된 분자에 대해 계산된 대표적인 molecular descriptor의 분포.



RANC는 SMILES 문자열 길이, 분자량, logP, TPSA 등 여러 핵심 descriptor에 있어서 training molecule 분포와 맞는 구조를 생성한다. 이는 현대 DDD에서 매우 중요한 요소

# Result

## Chemical Analysis of Generated Molecules

In Silico 분자 설계에서 RANC와 ORGANIC의 적용가능성을 알아보기 위해 두 모델에서 생성된 분자를 조사함  
RO5 대신 더 엄격한 기준인 Muegge drug-likeness filter를 충족한 것들 중 무작위로 20,000개 선택

CD 및 Drugs 데이터 세트의 RANC와 ORGANIC에 의해 생성된 분자에 대한 주요 화학통계

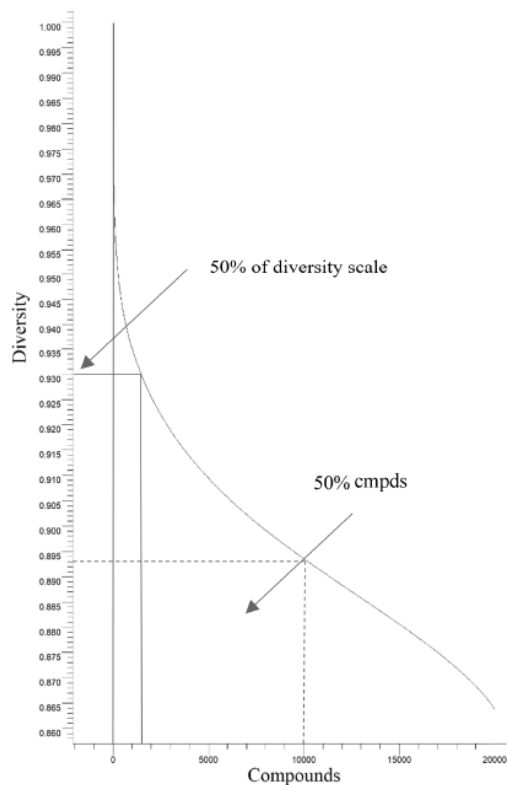
data set approach	CD		Drugs	
	RANC	ORGANIC	RANC	ORGANIC
size	20000	20000	20000	20000
diversity	0.86	0.89	0.85	0.86
fragments <sup>a</sup>	256000	428000	341000	324000
clusters	704	559	434	697
singletons	12797	14202	16330	11340
heterocycles <sup>b</sup>	2286	3805	1277	1215
cluster size <sup>c</sup>	10.2	10.4	8.5	12.4
MCF (%)	7.1	34.7	23.7	39.6

- RANC가 생성한 구조물은 MCF를 성공적으로 통과한 화합물의 수가 ORGANIC보다 각각 5배, 1.7배 높기 때문에 DDD 측면에서 더 매력적
- CD train dataset을 사용해 더 좋은 결과를 얻었다
- RANC에 의해 생성된 구조물은 적절한 singletons 수로 잘 군집 되어있다. (어떤 경우에는 ORGANIC이 singleton이나 군집 수를 더 많이 가지고 있다)

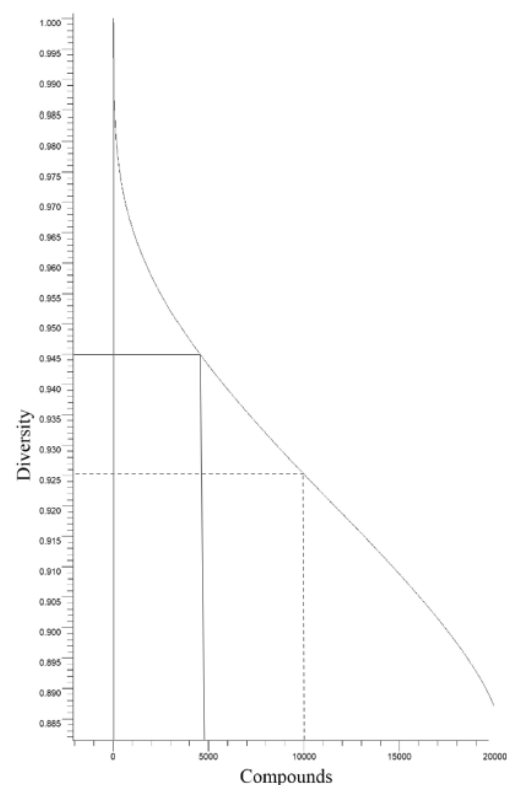
# Result

RANC, ORGANIC에 의해 생성된 분자의 common diversity and heterocyclic diversity

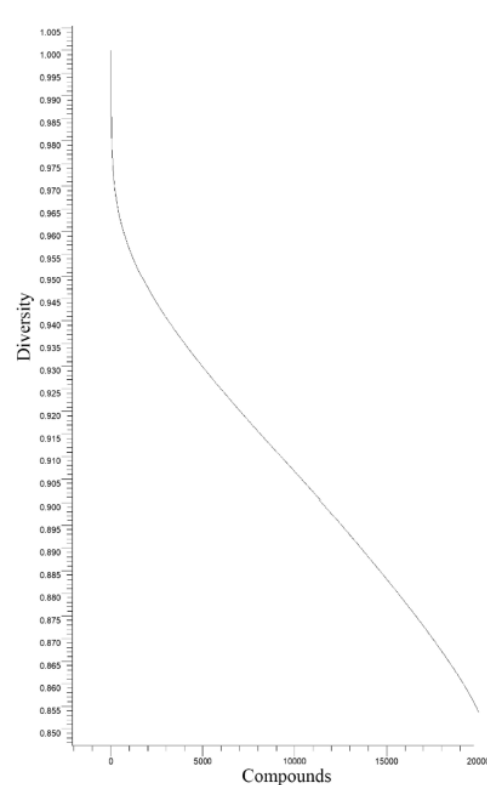
RANC, CD



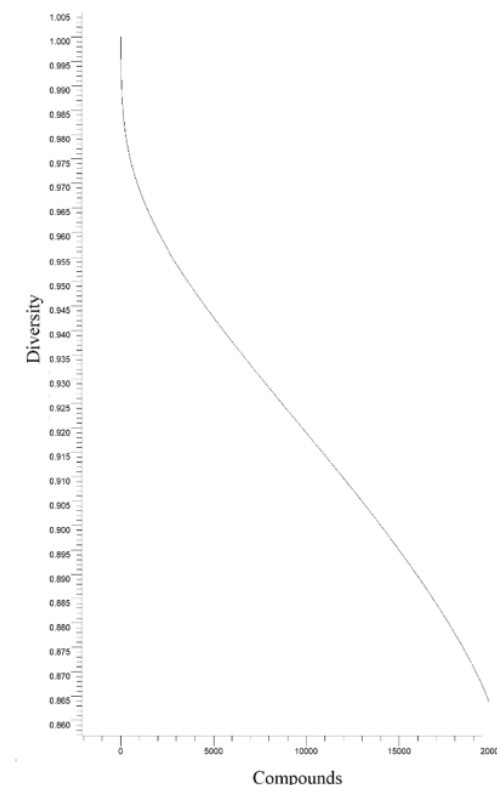
ORGANIC, CD



RANC, Drug



ORGANIC, Drug

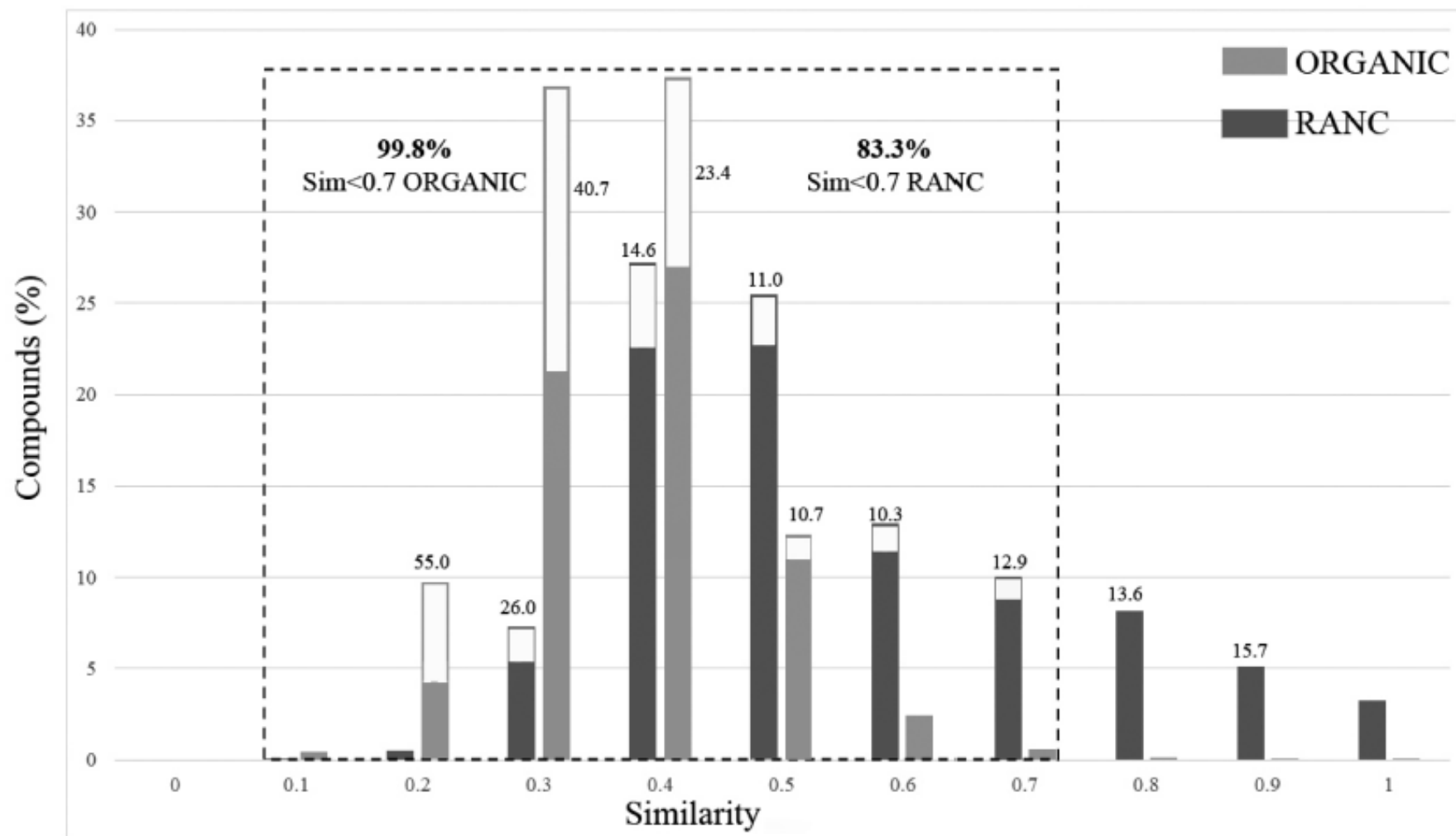


CD 데이터 세트를 사용하여 ORGANIC에 의해 생성된 구조는 RANC와 대조적으로 구조상 다양성이 더 높다는 결론  
하지만 높은 다양성은 MCF(6952개 화합물, 34.7%)를 통과하지 못한 화합물이 주로 달성한 것이라고 추측됨

# Result

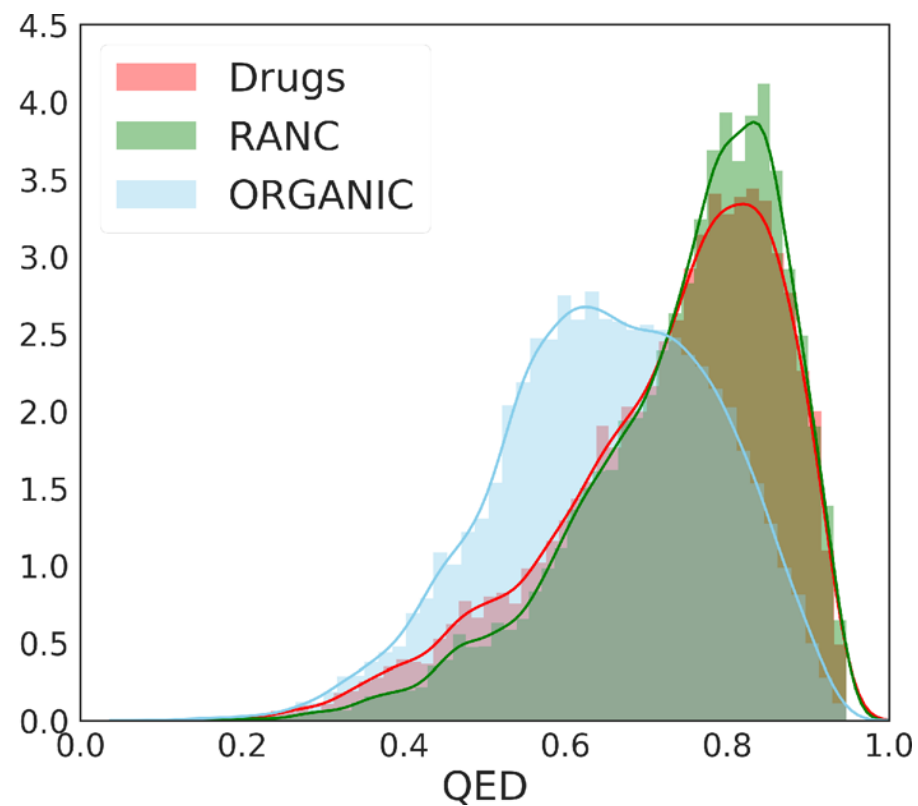
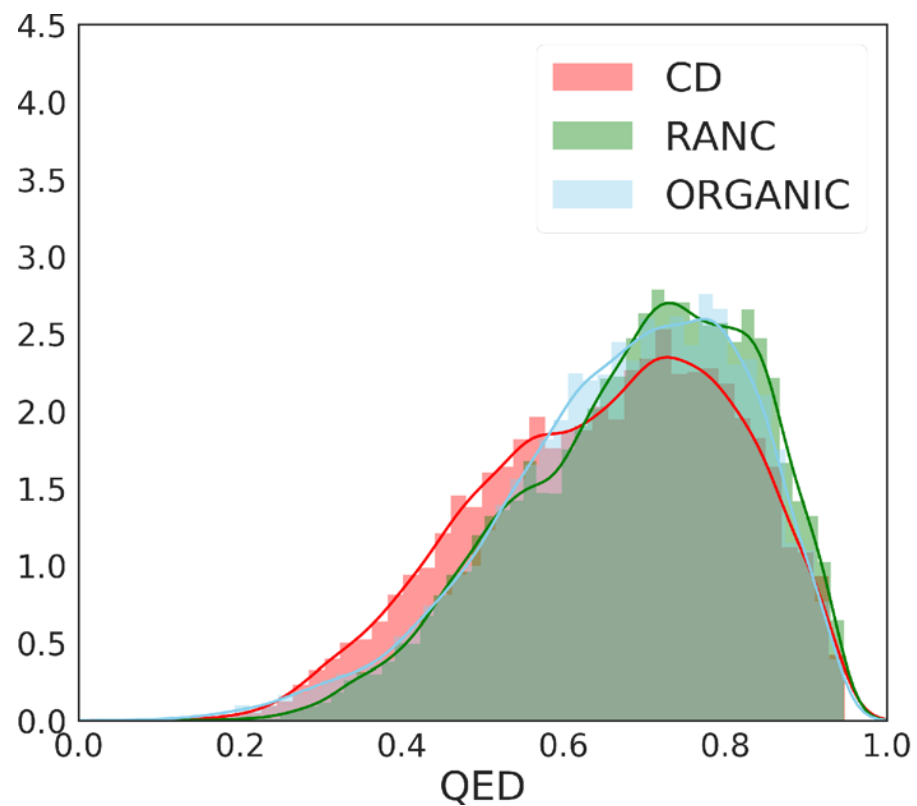
Train dataset의 화합물과 유사하게 생성된 구조물의 분포를 조사

CD train dataset: 생성된 structure와 화합물과의 유사성에 따른 분포 (빈 영역은 MCF를 통과하지 않은 화합물)



# Result

생성된 분자의 QED(Quantitative Estimate of Drug-likeness) score

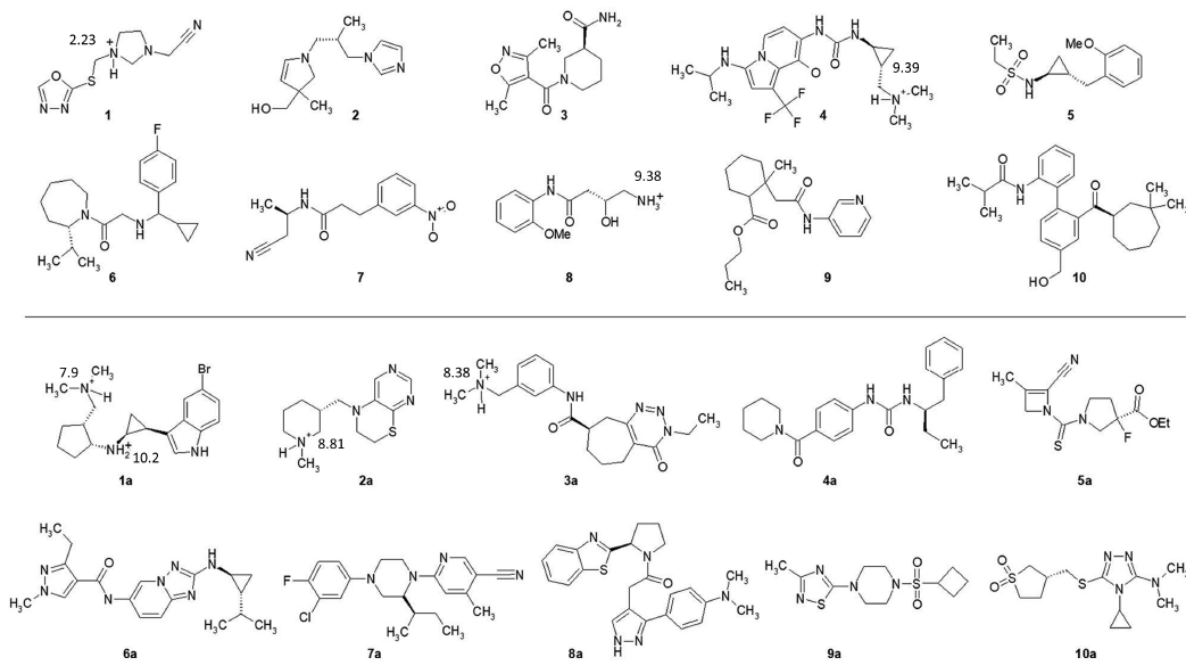


<i>CD Dataset</i>	<i>Drugs Dataset</i>
RANC와 ORGANIC 모두 비슷한 성능을 얻음	<ul style="list-style-type: none"><li>• RANC: 큰 값을 중심으로 dataset과 분포 동일</li><li>• ORGANIC: 분포에서 눈에 띄게 낮은 값으로 떨어짐</li></ul>

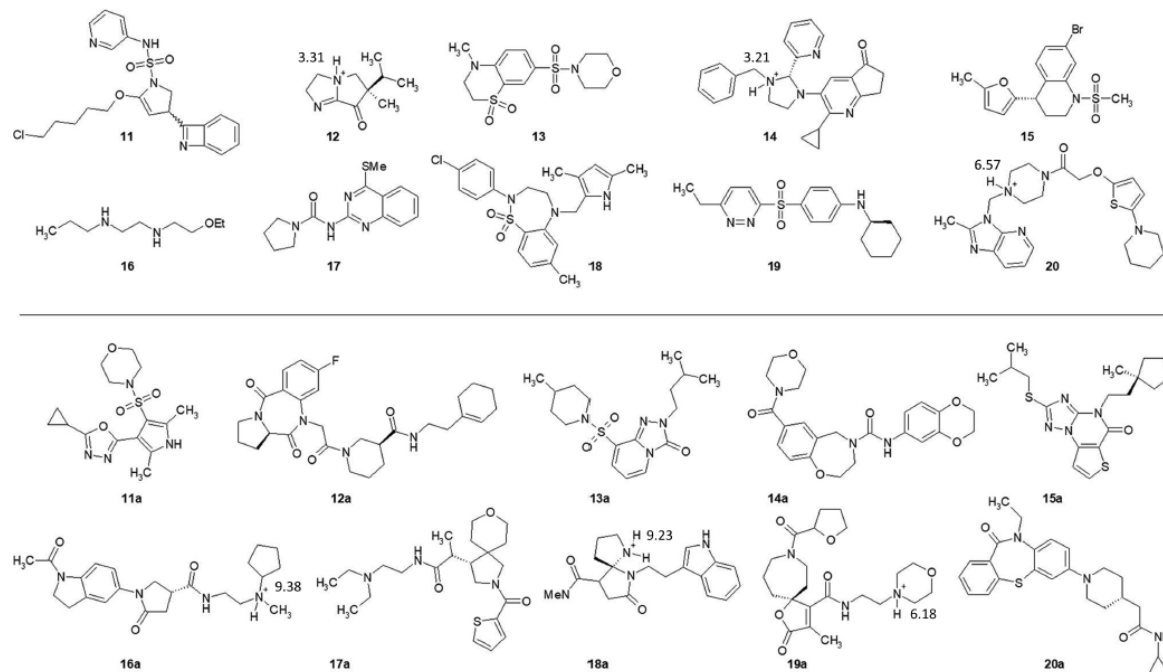
CD dataset과 drugs dataset 모두 RANC QED 점수가 높은 분자의 수가 더 많다

# Result

ORGANIC 모델 및 RANC 모델을 통해 Drugs Dataset  
을 사용하여 생성한 structure



ORGANIC 모델 및 RANC 모델을 통해 CD Dataset을 사용하여  
생성한 structure



# Conclusion

- 새로운 분자 설계를 위한 RANC 아키텍처 소개
- DNC를 generator 모델로 활용
- adversarial training 과정에서 generator와 discriminator사이의 균형
- 기존 architecture ORGANIC 개선
- MCF를 성공적으로 통과
- train molecule의 기초 화학적 property의 분포를 고려하면서 생성된 구조물의 길이에 큰 손실 없이 SMILES 문자열을 사용하여 매우 효율적이고 안정적인 학습을 진행
- Unique하고 적절한 구조를 생성 - unique하면서도 기존에 볼 수 없었던 특별한 substructure를 포함

## 추후 연구 방향

보다 복잡한 reward function을 다목적 제약 최적화 방식으로 통합하는 것

추가적인 다른 지표(metrics)에 대한 조사

adversarial training를 안정시키고 이해하는 데에 사용할 수 있는 다른 방법론들 시도

## 진행 중인 연구에 적용

Desired property에 대한 molecule을 생성하는 데에 SMILES strings로 표현된 데이터를 활용하여 알고리즘 적용

RANC에서 사용된 DNC라는 딥러닝 모델을 적용해서 대용량의 복잡한 데이터를 처리하는 데에 활용 가능