Electronic Supplementary Material (ESI) for Digital Discovery.

Link-INVENT: Generative Linker Design with Reinforcement Learning

Jeff Guo[1,⊥], Franziska Knuth[2,⊥], Christian Margreitter[1], Jon Paul Janet[3], Kostas Papadopoulos[1], Ola Engkvist[1,4], Atanas Patronov[1,*]

[1] Molecular AI, Discovery Sciences, R&D, AstraZeneca, Gothenburg, Sweden

[2] Department of Physics, Norwegian University of Science and Technology, Trondheim, Norway

[3] Medicinal Chemistry, Research and Early Development, Cardiovascular, Renal and Metabolism (CVRM), BioPharmaceuticals R&D, AstraZeneca, Gothenburg, Sweden

[4] Department of Computer Science and Engineering, Chalmers University of Technology, Gothenburg 41756, Sweden

* Corresponding author: patronov@gmail.com

⊥ These authors have contributed equally

Contents:

Physico-chemical Properties Score Transformations

Training Plots

Example Binding Poses

Binding Poses of Selected DeLinker and SyntaLinker Generated Molecules

Docking Scores Distribution of Link-INVENT, DeLinker, and SyntaLinker Generated Molecules

Experiment 2: Scaffold Hopping

Molecular Docking Score Transformation

Physico-chemical Properties Score Transformations

Training Plots

Example Binding Poses

Experiment 3: PROTACs

Physico-chemical Properties Score Transformations

Sub-Experiment 1: Scoring Function Components Transformations

Sub-Experiment 1: Controlling Linker Length Training Plots

Sub-Experiment 2: Scoring Function Components Transformations

Sub-Experiment 2: Controlling Linker Linearity Training Plots

Sub-Experiment 3: Scoring Function Component Transformation

Sub-Experiment 3: Controlling Linker Flexibility Training Plots

Data Preparation

In this section, detailed information regarding each step of the data preparation is presented starting from the raw ChEMBL data.[1]

ChEMBL Processing

1. **Initial Filtering of ChEMBL Data**

   Dataset size: 1,941,410 SMILES to 1,676,918 SMILES

   Initial filtering of the raw ChEMBL data was adapted from our previous Lib-INVENT work for library design by Fialkova et al.[2] with some modifications. The following criteria was used:

   1) Permitted Elements: {C, N, O, F, S, Cl, Br}

      The Prior model used in this work does not support phosphorus

   2) Number of Heavy Atoms ≤ 90
   3) Remove Salts
   4) Neutralize Charges
   5) Sanitize Molecules (according to RDKit)
   6) Number of Atoms ≥ 6

      Remove molecules that are small

   7) Molecular Weight ≤ 1500 Da
   8) Number of Rings ≤ 10
   9) Number of Aromatic Rings ≤ 8
   10) Heteroatom Ratio (Number of Carbon Atoms over Total Number of Atoms) > 0.5

2. **Reaction-based Slicing**

   Dataset size: 1,676,918 SMILES to 41,373,724 Tuples

   The filtered ChEMBL SMILES strings were sliced (maximum of 2 cuts permitted) to generate tuples with the form[3]: **(Linker, Warheads, Full Molecule)**

   The specific reaction SMIRKS used are provided as a Supplementary File and are identical to the ones used in our previous Lib-INVENT work.[2]

   At this step, the only criterion used to filter the resulting tuples from slicing was:

   Warhead Molecular Weight ≤ 500 Da

3. **Keep only Tuples with 2 Warheads**

   Dataset size: 41,373,724 Tuples to 29,915,654 Tuples

   Since we are interested in the task of linking two molecular subunits, the tuples generated by making a single cut were removed. The tuples remaining are those that have 3 components: 2 warheads and a linker.

4. **Remove Duplicates**

   Dataset size: 29,915,654 Tuples to 29,915,654 Tuples

   There were no duplicate tuples.

5. **Tuples Filtering**

Starting Dataset size: 29,915,654 Tuples

The purpose of this step was to remove unrealistic linkers and warheads. The final dataset adhered to the following criteria:

1) Linker Effective Length > 2

   Dataset size: 21,158,512 Tuples

   This criterion removes linkers where the attachment atoms are separated by a single atom.

2) Linker Length Ratio > 0.7

   Dataset size: 11,918,600 Tuples

   The 'Linker Length Ratio' is defined as the 'Linker Effective Length' (number of bonds between the attachment atoms) divided by the 'Linker Maximum Graph Length' (longest sequence of bonds in the linker). This criterion removes linkers with extensive branching.

3) Linker and Warheads Maximum Ring Size = 8

   Dataset size: 11,675,329 Tuples

   This criterion removes linkers and warheads with rings greater than 8 atoms.

4) Linker Molecular Weight ≤ 500 Da

   Dataset size: 10,615,401 Tuples

   This criterion removes excessively large linkers. We note that 500 Da linkers are still quite large, but it allows the Prior (trained using this data) to be successful in linking molecular subunits of even a few atoms (essentially constructing a full small molecule).

5) Linker Max Number of Rings <= 10

   Dataset size: 10,615,394 Tuples

   This criterion removes linkers with more than 10 rings. There were only 7 examples of these linkers as shown by the dataset size from criterion 4 to 5.

6. **Remove Rare SMILES Tokens**

   Dataset size: 10,615,394 Tuples to 10,613,969 Tuples

   The purpose of this filtering step was to remove SMILES tokens that were both rare, i.e., low frequency of occurrence, and/or represent undesirable atoms, e.g., carbanions which are highly reactive. The Prior generative model is more efficient after removal of rare tokens as it does not need to keep track of the conditional probabilities associated with them during the generative process.

   Linker Rare Tokens:

   [O], [O+], %10, [C-], [NH-], [c+]

   Warhead Rare Tokens:

   [O+], [C+]

Train/Validation Set Splitting

A validation set was generated by holding out unique Bemis-Murcko scaffolds *containing at least two rings*.[4] This was done by iterating over all tuples in the overall dataset such that *for each* tuple, all other tuples containing the same scaffold was removed. This process was repeated until at least 300,000 tuples were held out. See the Supplementary File for a list of all scaffolds in the SMILES format held out in the validation set. The final sizes of the datasets were:

Training: 10,313,109 Tuples (contains none of the scaffolds in the validation set)

Validation: 300,860 Tuples (contains 294 unique Bemis-Murcko scaffolds)

Prior Training

1. **Train/Validation Set SMILES Randomization**

   The SMILES in the train and validation sets were randomized (generating 20 training and validation SMILES files, each with randomized SMILES) such that resulting Prior is able to generalize better in 'chemical space'.[5]

2. **Train Prior**

   The Prior was trained for 20 epochs using the teacher forcing algorithm[6] following the same protocol as our Lib-INVENT work.[2] The Prior from epoch 4 was kept as it showed the best performance without overfitting. The vocabulary of the Prior is as follows:

   Warhead Vocabulary (37 Tokens):

   ['<pad>', '$', '^', '#', '(', ')', '*', '-', '1', '2', '3', '4', '5', '6', '7', '8', '=', 'Br', 'C', 'Cl', 'F', 'N', 'O', 'S', '[N+]', '[N-]', '[O-]', '[S+]', '[n+]', '[nH]', '[o+]', '[s+]', 'c', 'n', 'o', 's', '|']

   In this work, we are interested only in linking two molecular subunits. The '|' token separates the warheads in the tuple such that any given tuple has the form:

   (Linker, Warhead 1 | Warhead 2, Full Molecule)

   Linker Vocabulary (36 Tokens):

   ['<pad>', '$', '^', '#', '(', ')', '-', '1', '2', '3', '4', '5', '6', '7', '8', '=', 'Br', 'C', 'Cl', 'F', 'N', 'O', 'S', '[*]', '[N+]', '[N-]', '[O-]', '[S+]', '[n+]', '[nH]', '[o+]', '[s+]', 'c', 'n', 'o', 's']

Link-INVENT Scoring Function

The Scoring Function specifies the components to optimize via reinforcement learning (RL). The Scoring Function formulation is identical to the previously reported REINVENT platform and thus shares the same implementation.[7] The previously existing components operate on the full molecule as the first application of REINVENT was for small molecules.[7–9] In addition to these components, linker specific components have been implemented in Link-INVENT to provide control over the linker itself. This section lists all the supported components.

Linker Specific Components:

1. 'Linker Effective Length' (number of bonds between the attachment atoms)
2. 'Linker Maximum Graph Length' (longest sequence of bonds)
3. 'Linker Length Ratio' ('Linker Effective Length' over 'Linker Maximum Graph Length')
4. 'Linker Number of Rings'
5. 'Linker Number of Aromatic Rings'
6. 'Linker Number of Aliphatic Rings'
7. 'Linker Number of sp3 Atoms'
8. 'Linker Number of sp2 Atoms'
9. 'Linker Number of sp Atoms'
10. 'Linker Number of Hydrogen Bond Donors'
11. 'Linker Number of Hydrogen Bond Acceptors'
12. 'Linker Molecular Weight'
13. 'Linker Ratio of Rotatable Bonds' (number of rotatable bonds over total number of bonds)

Molecular Docking Protocol

This section contains all details related to performing molecular docking for Experiment 2: Fragment Linking[10,11] and Experiment 3: Scaffold Hopping.[12,13]

Target Preparation

The target structures for Experiment 1: Fragment Linking[10,11] and Experiment 2: Scaffold Hopping[12,13] were extracted from the corresponding PDB structures released (**5CU4** and **5CEO**, respectively). Schrodinger's Protein Preparation Wizard in Maestro (2019-4 release)[14] was used to pre-process the raw PDB structures using default parameters except for H-bond optimizer at pH 7.4. All waters beyond 3.0 het groups were removed.

Receptor Grid Generation

Schrodinger's Receptor Grid Generation in Maestro (2019-4 release)[14] was used with default settings to generate the receptor grids. Using the pre-processed proteins from the previous step, the reference ligand was used to define the molecular docking box.

LigPrep Ligand Embedding

All linked molecules proposed by Link-INVENT underwent LigPrep with default parameters except for the following: use Epik with pH 7.0 ± 1.0 and allow a maximum of 2 stereoisomers.[15] LigPrep was parallelized over 8 CPU cores on an Amazon Web Services (AWS) p3.8xlarge EC2 instance using 'DockStream'.[16]

Glide Molecular Docking

All ligands prepared by LigPrep were docked in the generated receptor grids using Glide with default parameters (Standard Precision).[17–20] Docking was parallelized over 48 CPU cores on an AWS p3.8xlarge EC2 instance using 'DockStream'.[16]

Core Constrained Glide Molecular Docking

This docking protocol corresponds to Experiment 1b: Fragment Linking in the main manuscript. All ligands prepared by LigPrep were docked in the generated receptor grid using Glide with Standard Precision. Constrained docking was applied via 'DockStream' [16] and the core definition was defined using the following SMARTS corresponding to the constituent fragment:

[H]n1cnc([H])c1-c2c([H])c([H])c(Br)c([H])c2[H]

The maximum allowed RMSD deviation from the reference ligand pose was 0.3 Å and in case a generated molecule cannot conform to this constraint, no fallback was allowed.[17–20]

Hardware Experiment Compute Times

All Link-INVENT experiments used a NVIDIA Tesla V100 Volta GPU with 32GB memory. Where molecular docking was used, the computation was parallelized over 8 and 48 CPU cores for LigPrep[15] and Glide docking[17–20], respectively, on an AWS p3.8xlarge EC2 instance using 'DockStream'.[16] The time discrepancies between replicate runs for the Fragment Linking and Scaffold Hopping experiments is attributed to the shared production environment used. The experiments were sent to a queue and CPU resources for docking sometimes experienced delay.

Experiment	Replicate 1	Replicate 2	Replicate 3
Illustrative	1 min 56 sec	1 min 48 sec	1 min 53 sec
Fragment Linking	32 hours 26 min 47 sec	26 hours 52 min 13 sec	23 hours 58 min 52 sec
Scaffold Hopping	22 hours 29 min 40 sec	10 hours 51 min 11 sec	18 hours 32 min 8 sec
PROTACs*	**Average Run Time**		
Sub-Experiment 1: Controlling Length	16 min 16 sec		
Sub-Experiment 2: Controlling Linearity	14 min 9 sec		
Sub-Experiment 3: Controlling Flexibility	21 min 26 sec		

Table S1: Experiment compute times. *For the PROTACs experiments, the average run time is reported. For Sub-Experiment 1, this is averaged over 15 replicates. For Sub-Experiments 2 and 3, this is averaged over 9 replicates.

Illustrative Example

The illustrative example demonstrates how Link-INVENT gradually learns to generate favourable linkers (and the corresponding full) molecules via RL. The scoring function in this experiment consists of the following components:

1. **Linker Number of Hydrogen Bond Donors:** The objective was to generate linkers that contained as few hydrogen bond donors in the linker as possible

2. **Linker Number of Rings:** The objective was to generate linkers with no rings

The Link-INVENT configuration JSON used is provided as an additional file. All specific parameters used are in the JSON.

Scoring Function Transformation

Transformation functions are applied to yield a score in the interval [0, 1] (for a given component) for agent feedback. The “linker_number_of_rings” component was formulated such that a score of 1 (perfect) was given if the linker contained no rings and 0 otherwise. Therefore, the scoring function transformation for this component is not continuous and is not shown. In contrast, the “linker_number_of_hydrogen_bond_donors” scoring function transformation was formulated to yield a (near) perfect score if no hydrogen bond donors are present in the linker. The transformation is shown below:

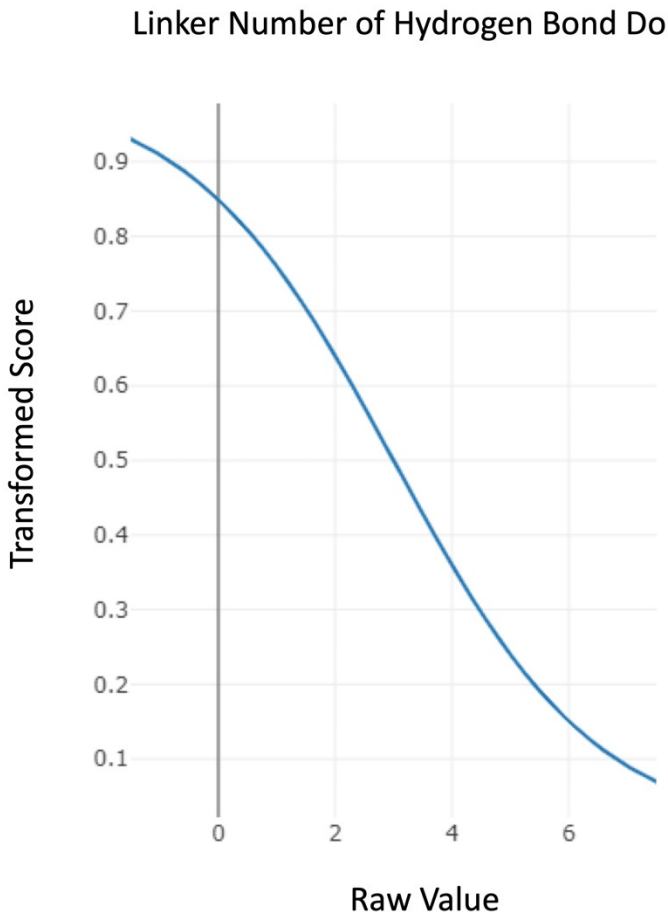


Fig. S1: Illustrative Example linker number of hydrogen bond donors scoring function transformation. The transformation ensures a score in the interval [0,1] is returned for agent feedback. The specific transformation parameters were chosen to encourage generation of linkers with no hydrogen bond donors.

Experiment 1a: Fragment Linking

This section shows all the training plots associated with Experiment 1a: Fragment Linking.[10,11] All Scoring Function components were assigned a transformed score in the interval [0,1], where 1 denotes perfect satisfaction. 'Average Score' is the aggregated score of each component in the Scoring Function. 'SMILES Found' refers to the number of valid SMILES generated. 'Valid SMILES' refers to the percentage of valid SMILES (RDKit successfully parses) generated at each epoch. The scoring function consists of the following components:

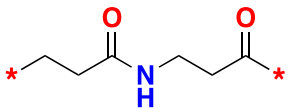
1. **DockStream: Molecular Docking**

2. **Linker Length Ratio ≥ 70**

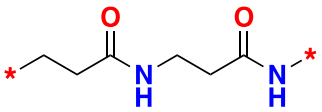
3. **Linker Molecular Weight ≤ 200 Da**

The Link-INVENT configuration JSON used is provided as an additional file. All specific parameters used are in the JSON.

Reference Linker ([*]C(CCNC(CC[*])=O)=O)



Most Similar Linker in Training Set (O=C(CC[*])NCCC(N[*])=O)



Tanimoto Similarity to Reference Linker = 0.694

Fragment Linking: Docking Score Transformation

The following transformation function was applied to the docking score to yield a score in the interval [0, 1].

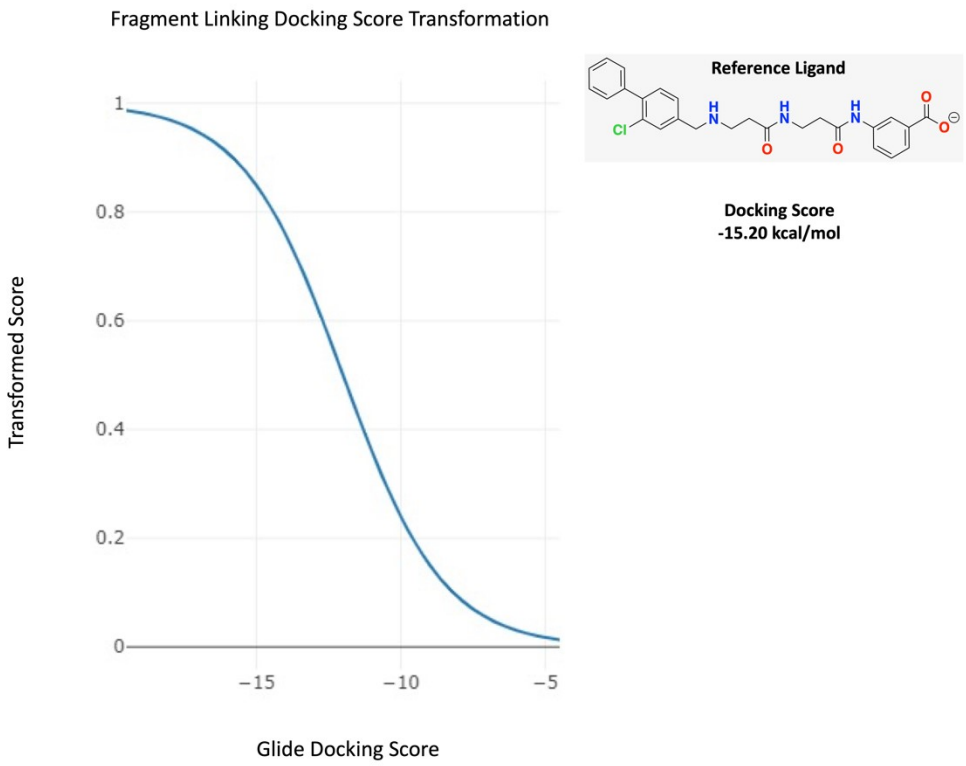


Fig. S2: Experiment 1a: Fragment Linking docking score transformation. The transformation ensures a score in the interval [0,1] is returned for agent feedback. The specific transformation parameters were chosen to encourage generation of molecules that possess a more favourable docking score than the reference ligand, shown in the figure.

Fragment Linking: Physico-chemical Properties Score Transformations

The following transformations functions were applied to the physico-chemical properties values to yield a score in the interval [0, 1].

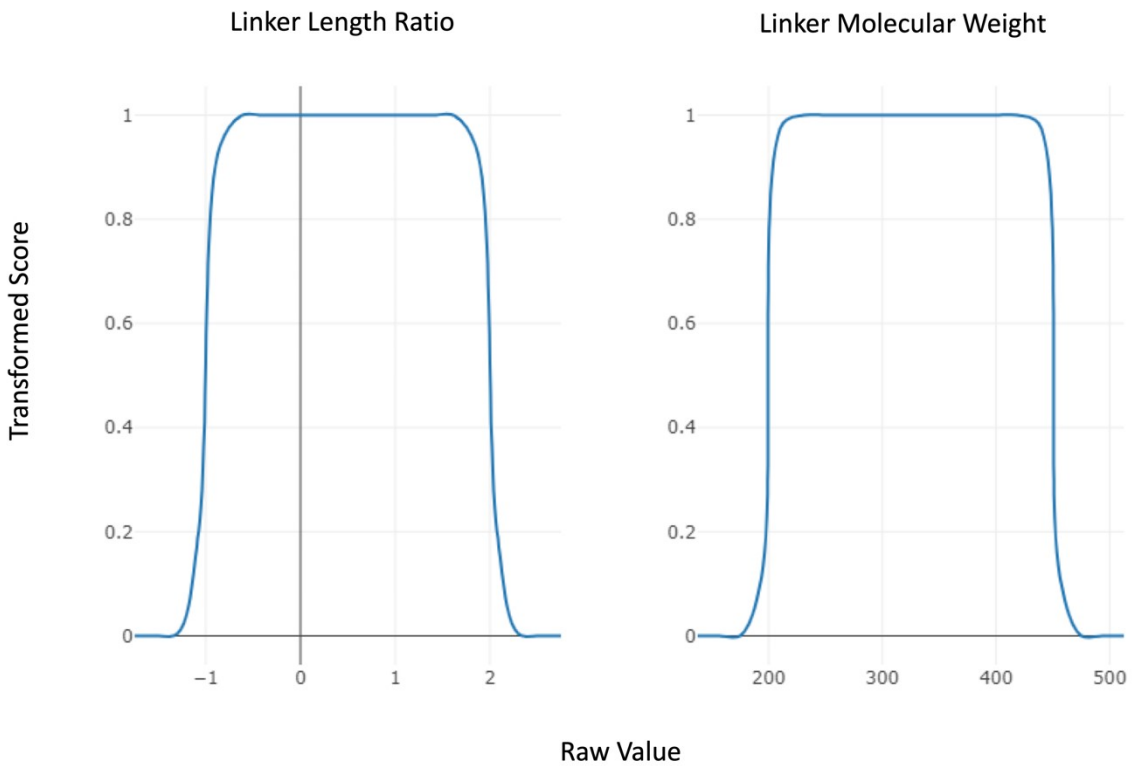


Fig. S3: Experiment 1a: Fragment Linking physico-chemical properties scoring function transformations. The transformation ensures a score in the interval [0,1] is returned for agent feedback.

Fragment Linking: Training Plots

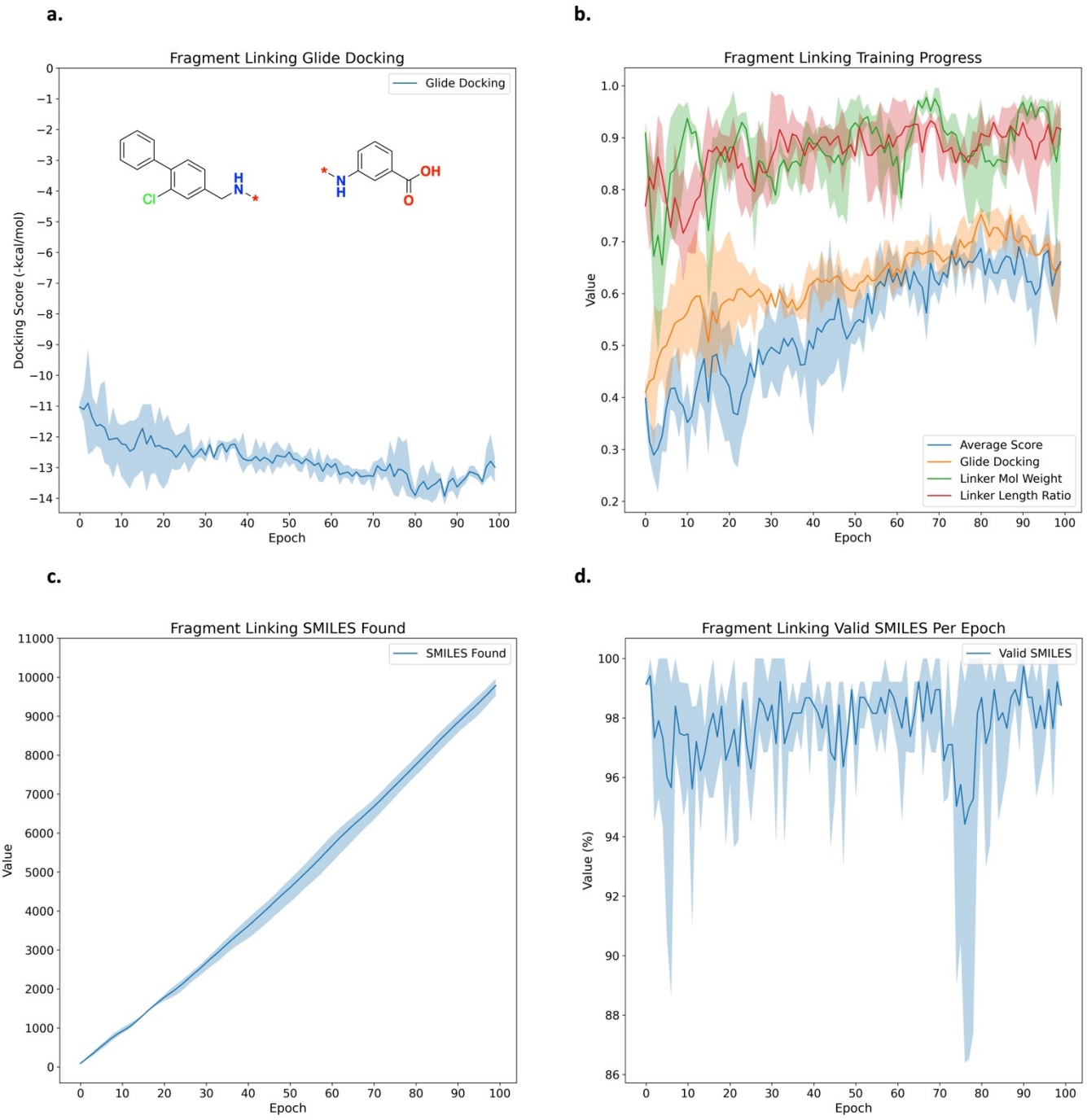


Fig. S4: Experiment 1a: Fragment Linking Training Plots. The experiment was run in triplicate. The curve shows the average score and the upper and lower bounds of the shaded region represent the maximum and minimum score, respectively. **a.** Raw docking score. **b.** Transformed score of all components in the scoring function. All components are optimized over the course of training. The transformed docking score achieves a moderate score due to the nature of the transformation applied. The maximum reward is only given if the generated molecule possesses a considerably better docking score than the reference ligand. **c.** SMILES Found. **d.** Percentage of valid SMILES.

Fragment Linking: Example Binding Poses

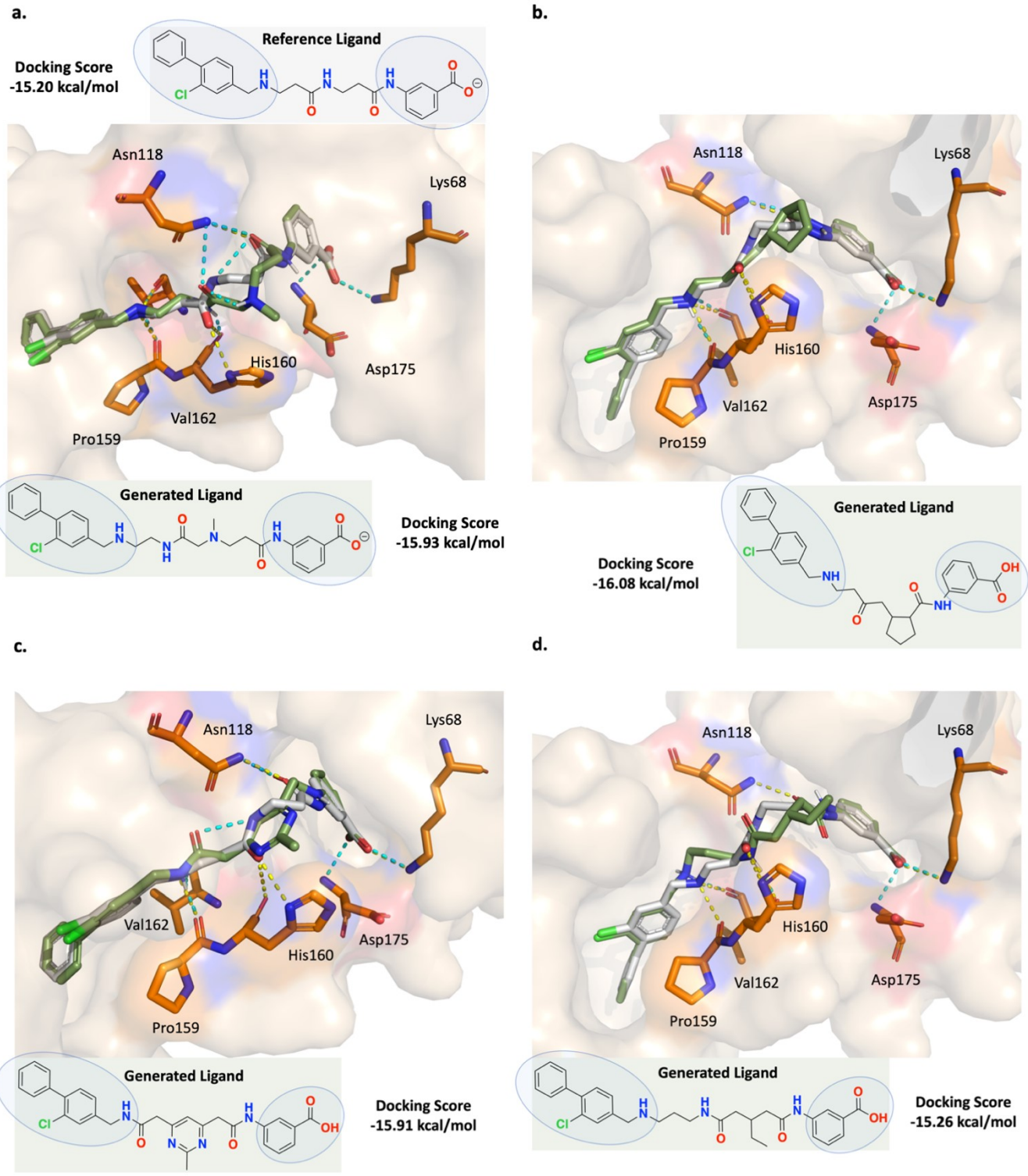


Fig. S5: Experiment 1a: Fragment Linking Example Binding Poses. **PDB ID: 5CU4**. The reference ligand and example generated ligand structures are shown in the figure. The interactions formed by the reference ligand and the generated ligands are shown as yellow and turquoise dotted lines, respectively. All generated ligands retain the interaction with Lys68, as enforced by the docking constraint. There is also extensive overlap between all the generated ligands and the reference ligand.

Experiment 1b: Comparison Fragment Linking

This section shows all the training plots associated with Experiment 1b: Fragment Linking.[21] All Scoring Function components were assigned a transformed score in the interval [0,1], where 1 denotes perfect satisfaction. 'Average Score' is the aggregated score of each component in the Scoring Function. 'SMILES Found' refers to the number of valid SMILES generated. 'Valid SMILES' refers to the percentage of valid SMILES (RDKit successfully parses) generated at each epoch. The scoring function consists of the following components:

1. **DockStream: Molecular Docking**

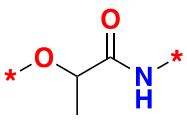
2. **3 <= Linker Effective Length <= 5**

3. **Linker Length Ratio ≥ 70**

4. **Linker Molecular Weight ≤ 150 Da**

The Link-INVENT configuration JSON used is provided as an additional file. All specific parameters used are in the JSON.

Reference Linker ([*]OC(C(N[*])=O)C)



The training set contained this linker.

Comparison Fragment Linking: Docking Score Transformation

The following transformation function was applied to the docking score to yield a score in the interval [0, 1].

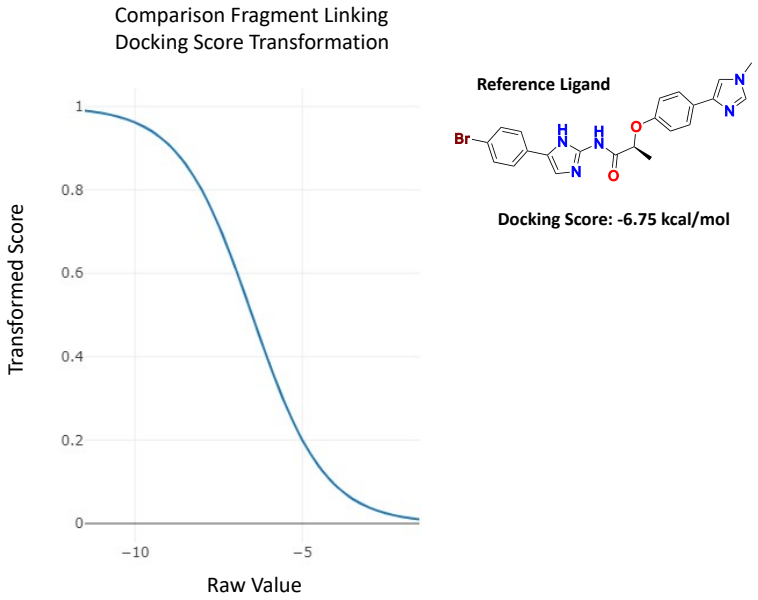


Fig. S6: Experiment 1b: Fragment Linking docking score transformation. The transformation ensures a score in the interval [0,1] is returned for agent feedback. The specific transformation parameters were chosen to encourage generation of molecules that possess a more favourable docking score than the reference ligand, shown in the figure.

Comparison Fragment Linking: Physico-chemical Properties Score Transformations

The following transformations functions were applied to the physico-chemical properties values to yield a score in the interval [0, 1].

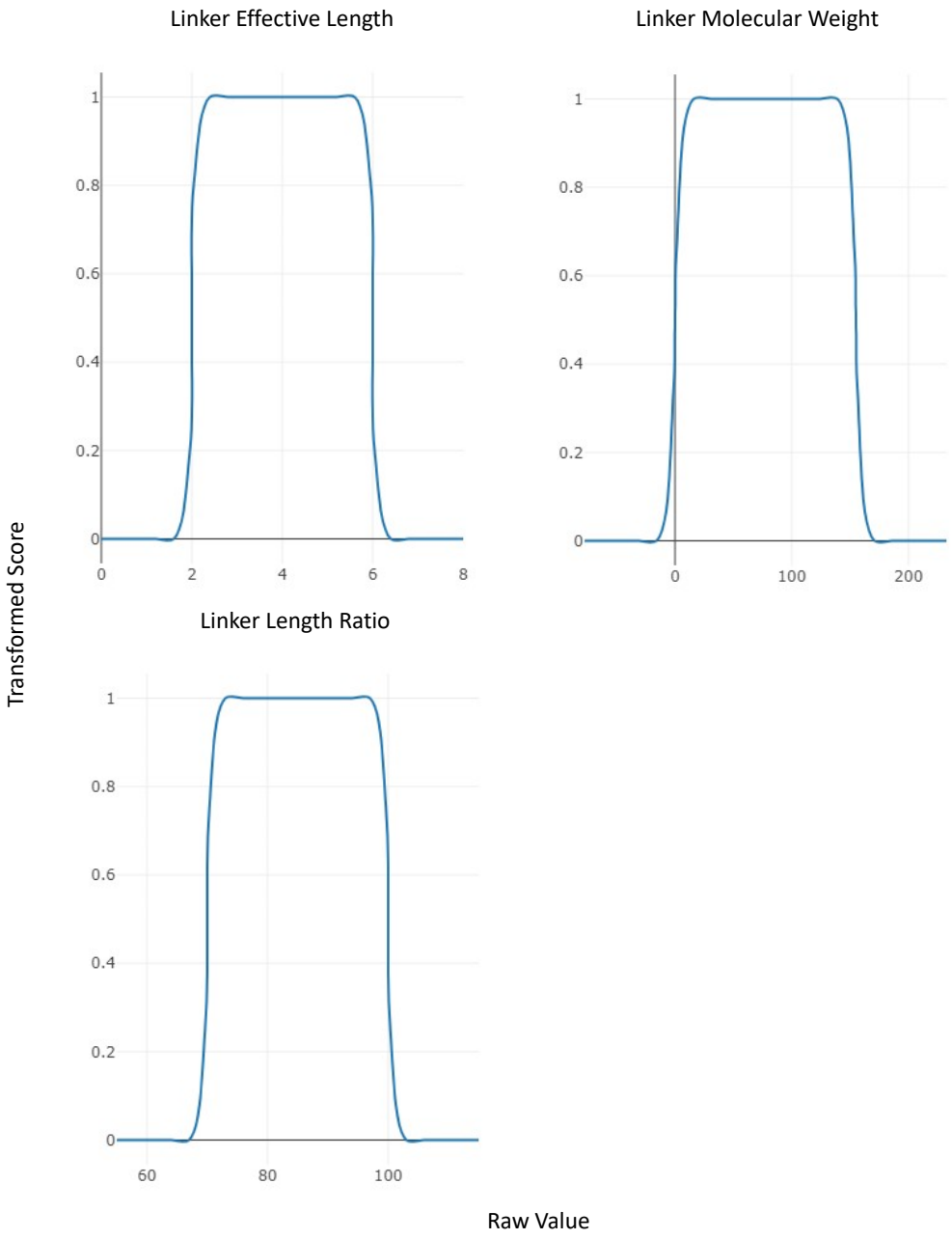
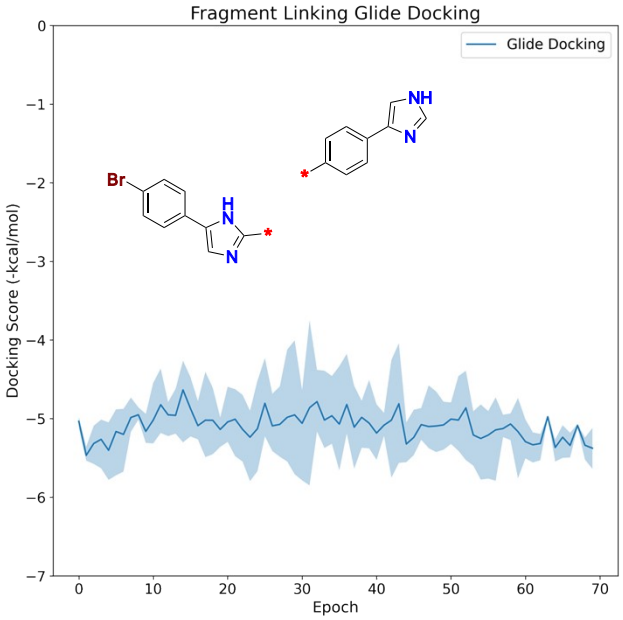


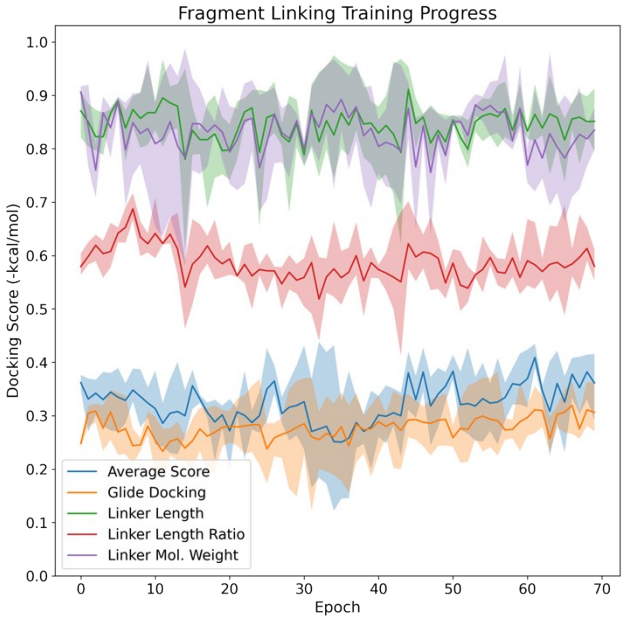
Fig. S7: Experiment 1b: Fragment Linking physico-chemical properties scoring function transformations. The transformation ensures a score in the interval [0,1] is returned for agent feedback.

Comparison Fragment Linking: Training Plots

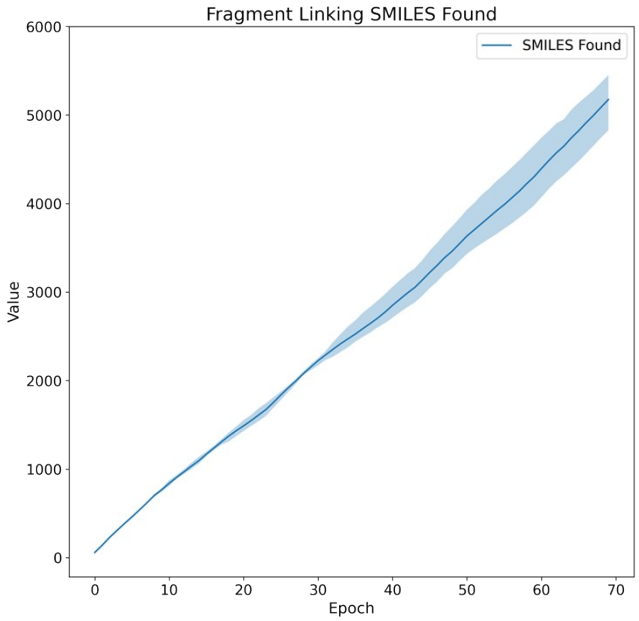
a.



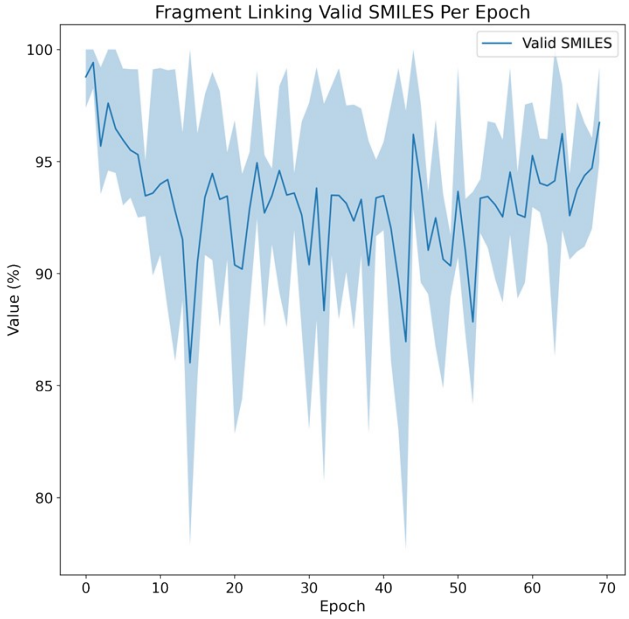
b.



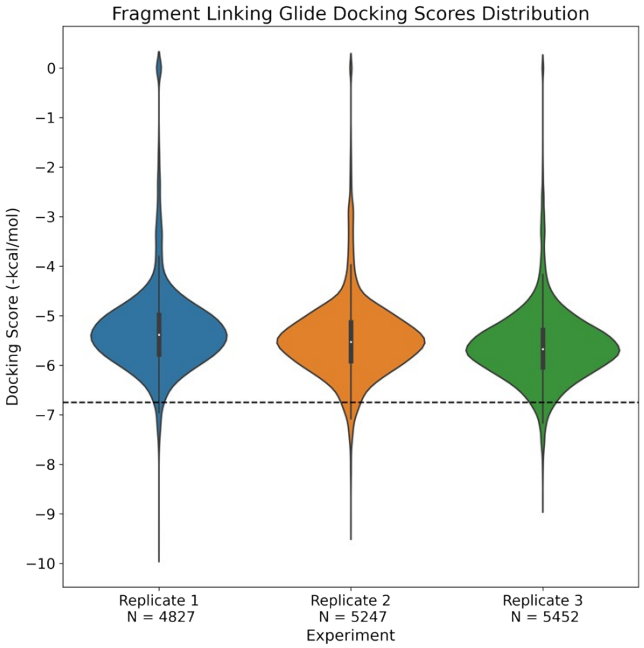
c.



d.



e.



f.

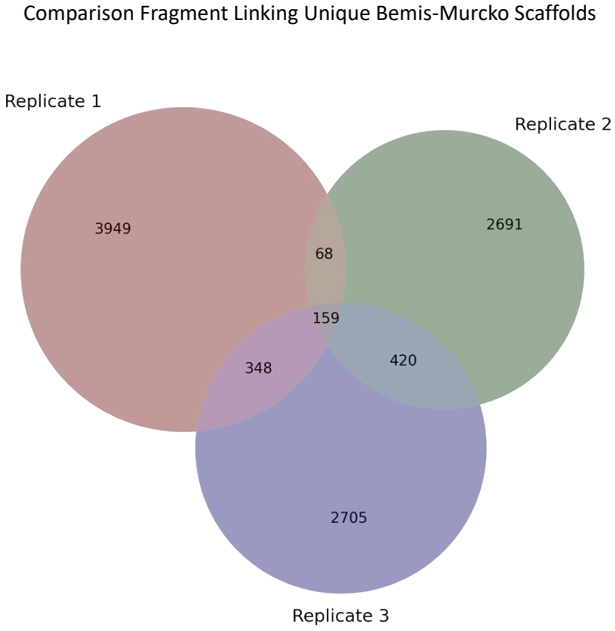


Fig. S8: Experiment 1b: Fragment Linking Training Plots. The experiment was run in triplicate. The curve shows the average score and the upper and lower bounds of the shaded region represent the maximum and minimum score, respectively. **a.** Raw docking score. **b.** Transformed score of all components in the scoring function. All components are optimized over the course of training. The transformed docking score achieves a moderate score due to the nature of the transformation applied. The maximum reward is only given if the generated molecule possesses a considerably better docking score than the reference ligand. **c.** SMILES Found. **d.** Percentage of valid SMILES. **e.** Docking scores distribution of all molecules. The black dotted line at -6.75 is the docking score of the reference ligand and some generated molecules possess a more favourable docking score. The majority of generated molecules possess a comparable docking score. **f.** The number of unique Bemis-Murcko scaffolds generated. There is minimal overlap across the triplicate runs, demonstrating that Link-INVENT explores generally different chemical space. The absolute number of Bemis-Murcko scaffold shows that Link-INVENT can generate a diverse number of linker ideas.

Comparison Fragment Linking: Example Binding Poses

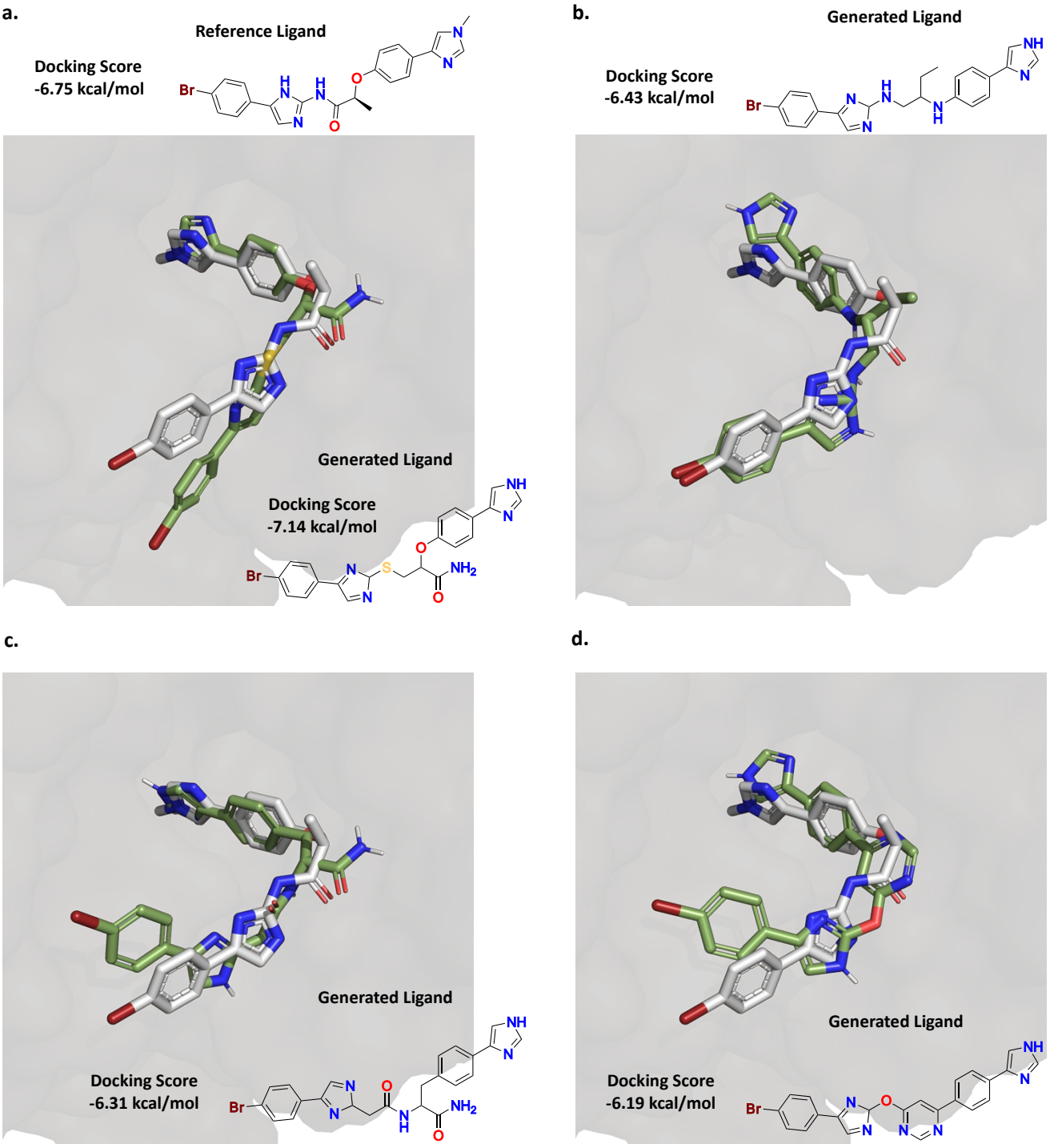


Fig. S9: Experiment 1b: Fragment Linking Example Binding Poses. **PDB ID: 5OU3**. The reference ligand and example generated ligand structures are shown in the figure. At least one constituent fragment of the generated molecules overlaps extensively with the reference fragment, as enforced by the constrained docking protocol.

Binding Poses of Selected DeLinker and SyntaLinker Generated Molecules

In the DeLinker and SyntaLinker works, twenty and three example molecules were released.[22,23] The same constrained docking protocol used in our work was applied to these molecules and those with the best binding pose agreement to the reference fragments are shown.

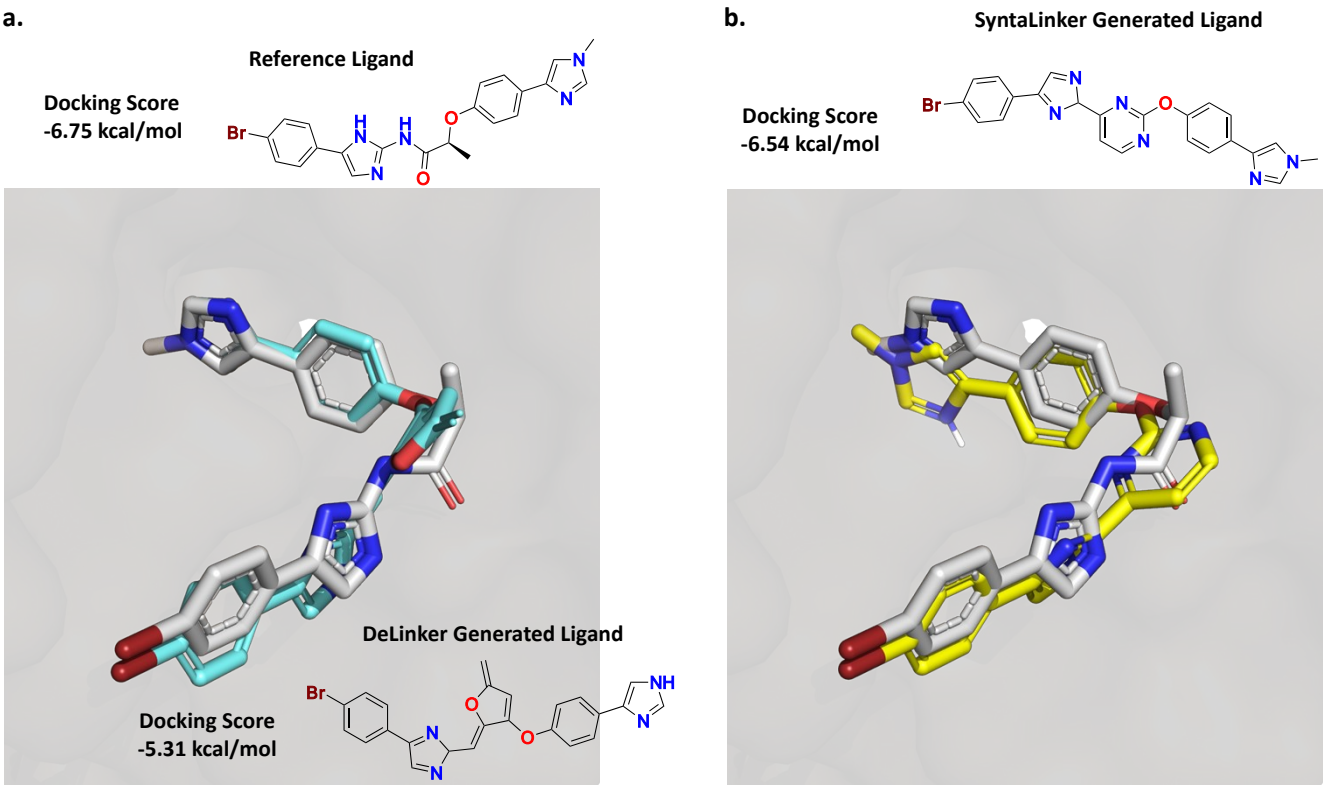


Fig. S10: Experiment 1b: Fragment Linking Example Binding Poses of DeLinker and SyntaLinker Generated Molecules.[22,23] **PDB ID: 5OU3**. The reference ligand and example generated ligand structures are shown in the figure. The molecules with the best binding pose agreement to the reference fragments are shown. The constituent fragments are circled. SyntaLinker introduced bias by providing their model with information about the methyl substituent and ether linkage which are present in the reference ligand and only known after Trapero et al. performed fragment linking.[21] Neither the DeLinker nor SyntaLinker generated molecule possesses a docking score better than the reference ligand.

Docking Scores Distribution of Link-INVENT, DeLinker, and SyntaLinker Generated Molecules

The table in this section shows the number of generated molecules by Link-INVENT, DeLinker, and SyntaLinker below different docking score thresholds.[22,23] The total number of molecules generated is also shown. In the case of DeLinker and SyntaLinker, only 20 and 3 molecules, respectively, were provided in their work. We further note that SyntaLinker recovers the reference ligand but in their experimental design, they introduced bias by providing their model with information only known after successful fragment linking by Trapero et al.[21] Specifically, in one of the fragments, Trapero et al. introduce a methyl substituent on the imidazole ring due to synthetic accessibility and the linker with the greatest potency was an ether linkage. In the SyntaLinker work, this methyl substituent and the ether linkage information was provided to their model.[23] We also acknowledge that it is possible that some DeLinker and SyntaLinker proposed molecules possess favourable docking scores and the analysis that follows is only based on what the authors have provided. Finally, while Link-INVENT is tasked with generating 8960 SMILES over 70 epochs, all linked molecules which do not satisfy the core constrained docking were discarded.

Experiment	<= Reference Ligand (-6.75 kcal/mol)	<= 6 kcal/mol	<= 5.5 kcal/mol	<= 5 kcal/mol
Link-INVENT Replicate 1	79/4827	688/4827	2024/4827	3577/4827
Link-INVENT Replicate 2	176/5247	1091/5247	2728/5247	4279/5247
Link-INVENT Replicate 3	220/5452	1491/5452	3396/5452	4732/5452
DeLinker	0/20	5/20	11/20	14/20
SyntaLinker	1/3*	2/3	3/3	3/3

Table S2: Experiment 1b: Fragment Linking comparison of docking scores of generated molecules by Link-INVENT, DeLinker, and SyntaLinker. *SyntaLinker recovers the reference ligand.[22,23]

Experiment 2: Scaffold Hopping

This section shows all the training plots associated with Experiment 2: Scaffold Hopping.[12,13] All Scoring Function components were assigned a transformed score in the interval [0,1], where 1 denotes perfect satisfaction. 'Average Score' is the aggregated score of each component in the Scoring Function. 'SMILES Found' refers to the number of valid SMILES generated. 'Valid SMILES' refers to the percentage of valid SMILES (RDKit successfully parses) generated at each epoch. The scoring function consists of the following components:

1. **DockStream: Molecular Docking. The transformation function applied to the raw docking scores is shown below.**

2. **Molecular Weight < 450 Da**

3. **Number of Hydrogen Bond Donors < 2**

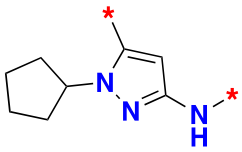
4. **3 ≤ logP ≤ 4**

5. **Topological Polar Surface Area < 90 Å^2**

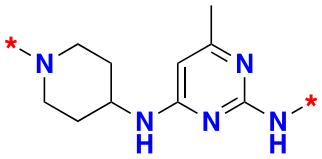
6. **1 ≤ Linker Number of Aromatic Rings ≤ 2**

The Link-INVENT configuration JSON used is provided as an additional file. All specific parameters used are in the JSON.

Reference Linker ([*]NC1=NN(C([*])=C1)C2CCCC2)



Most Similar Linker in Training Set ([*]N1CCC(CC1)NC2=CC(C)=NC(N[*])=N2)



Tanimoto Similarity to Reference Linker = 0.288

Scaffold Hopping: Docking Score Transformation

The following transformation function was applied to the docking score to yield a score in the interval [0, 1].

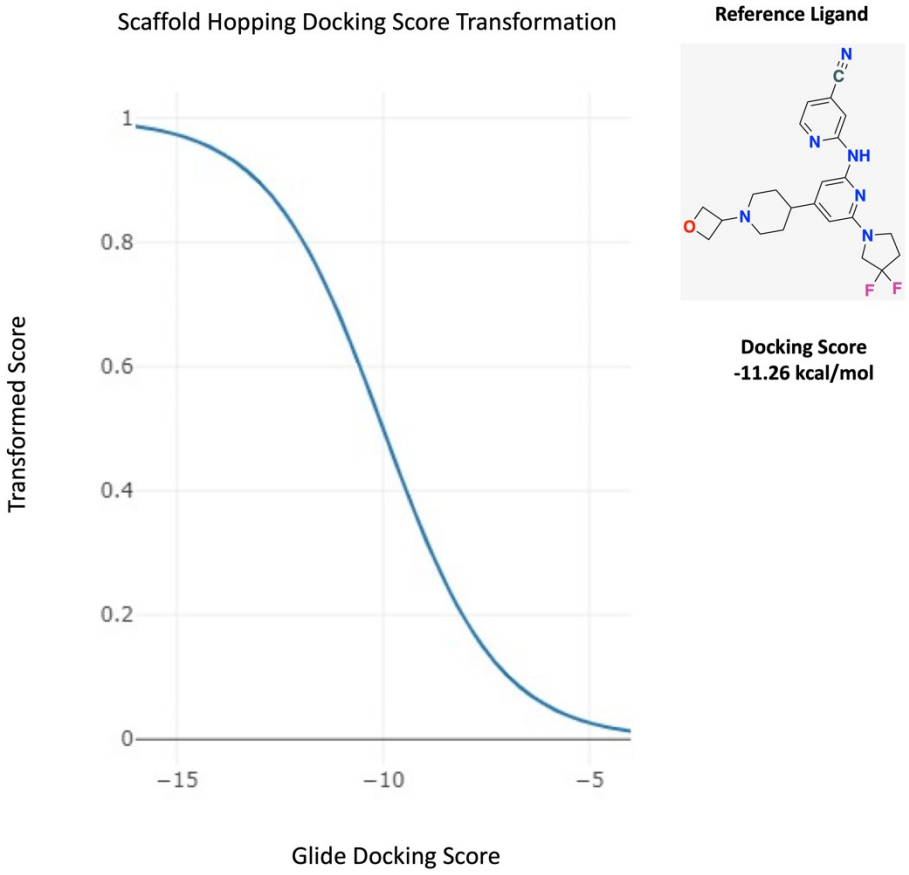


Fig. S11: Experiment 2: Scaffold Docking Score Transformation. The transformation ensures a score in the interval [0,1] is returned for agent feedback. The specific transformation parameters were chosen to encourage generation of molecules that possess a more favourable docking score than the reference ligand, shown in the figure.

Scaffold Hopping: Physico-chemical Properties Score Transformations

The following transformations functions were applied to the physico-chemical properties values to yield a score in the interval [0, 1].

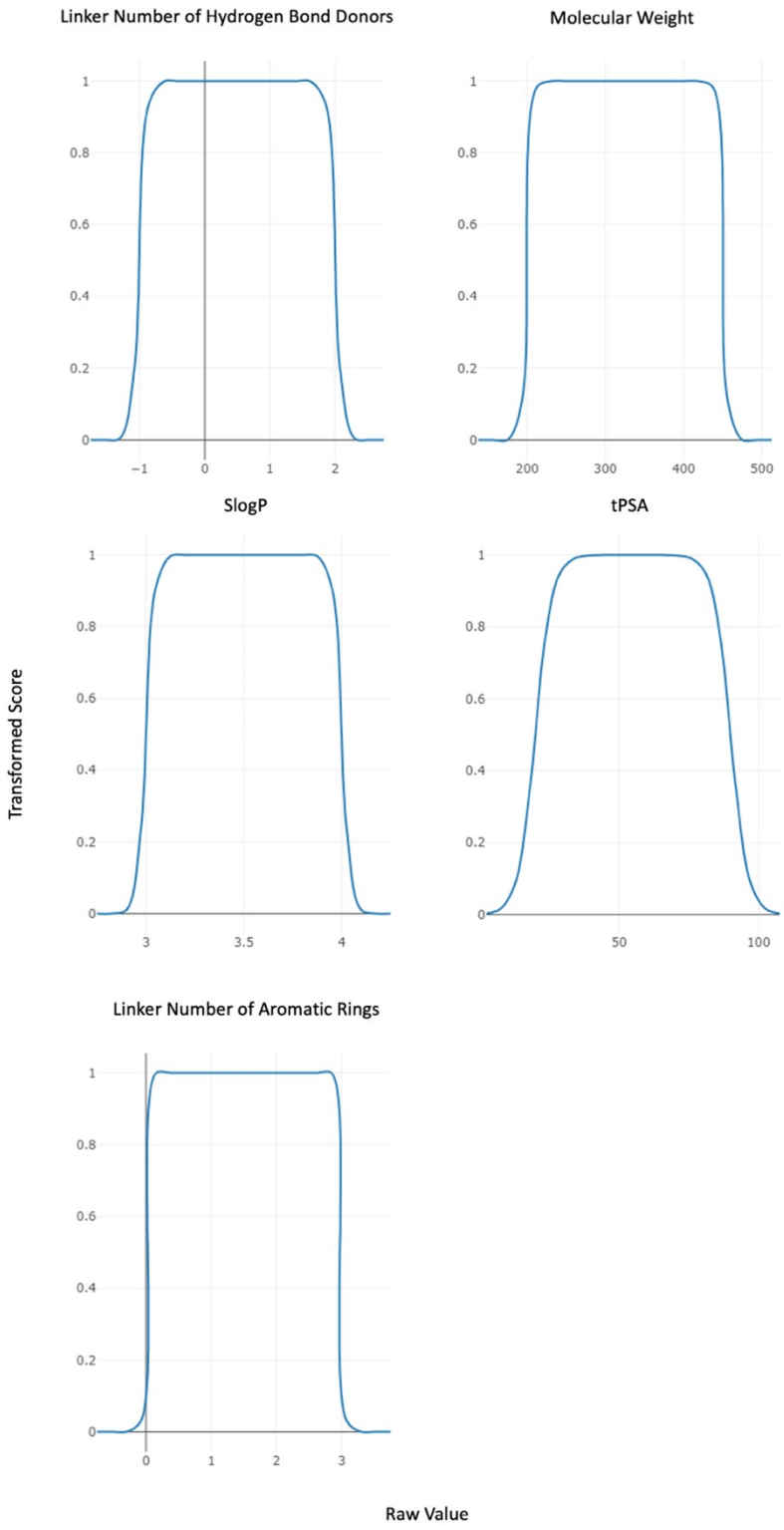


Fig. S12: Experiment 2: Scaffold Hopping physico-chemical properties scoring function transformations. The transformation ensures a score in the interval [0,1] is returned for agent feedback.

Scaffold Hopping: Training Plots

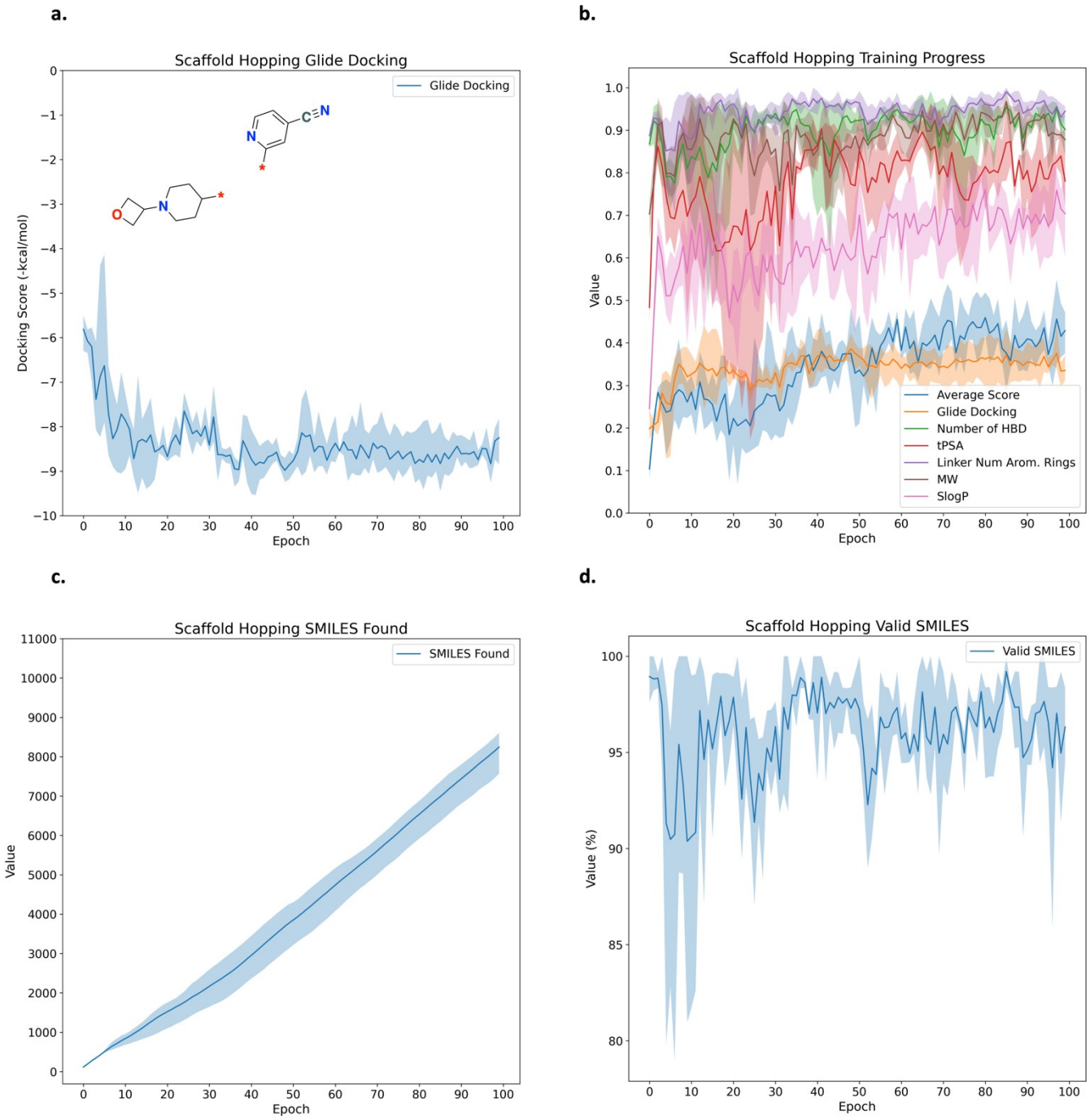


Fig. S13: Experiment 2: Scaffold Hopping Training Plots. The experiment was run in triplicate. The curve shows the average score and the upper and lower bounds of the shaded region represent the maximum and minimum score, respectively. **a.** Raw docking score. **b.** Transformed score of all components in the scoring function. All components are optimized over the course of training. The transformed docking score achieves a moderate score due to the nature of the transformation applied. The maximum reward is only given if the generated molecule possesses a considerably better docking score than the reference ligand. **c.** SMILES Found. **d.** Percentage of valid SMILES.

Scaffold Hopping: Example Binding Poses

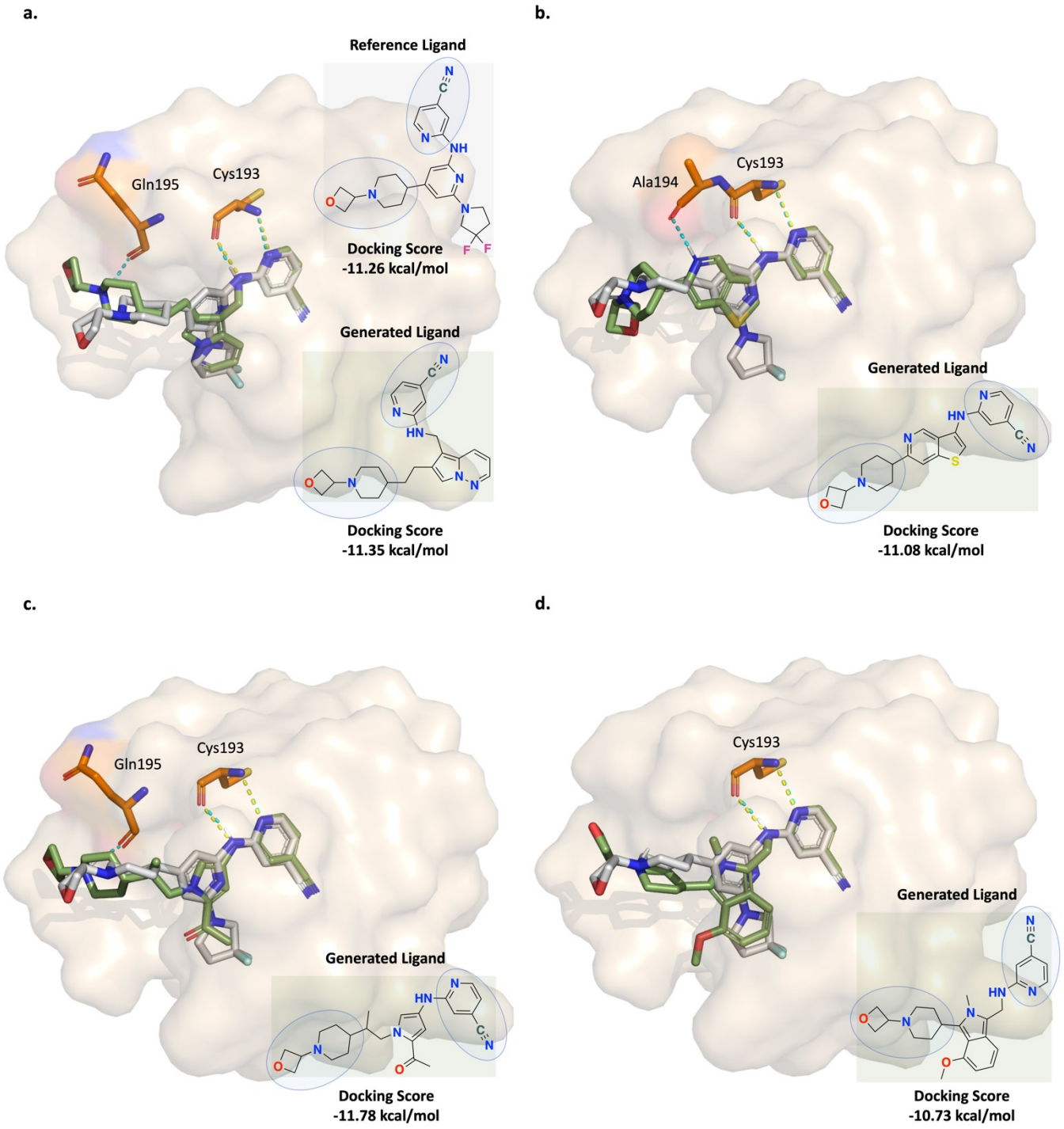


Fig. S14: Experiment 2: Scaffold Hopping Example Binding Poses. **PDB ID: 5CEO**. The reference ligand and example generated ligand structures are shown in the figure. The interactions formed by the reference ligand and the generated ligands are shown as yellow and turquoise dotted lines, respectively. All generated ligands retain the two hydrogen-bonding interactions with Cys193, as enforced by the docking constraint. There is also extensive overlap between all the generated ligands and the reference ligand. Finally, all generated ligands satisfy the desired CNS physico-chemical properties.

Experiment 3: PROTACs

This section contains all the training plots for the PROTACs Sub-Experiments.[24] 'SMILES Found' refers to the number of valid SMILES generated. 'Valid SMILES' refers to the percentage of valid SMILES (RDKit successfully parses) generated at each epoch. The Link-INVENT configuration JSONs used for all Sub-Experiments in this section are provided as an additional file. All specific parameters used are in the JSON. Sub-Experiments 1 and 2 enforced a set of physico-chemical properties within specific intervals derived from experimentally observed PROTACs:[25,26]

1. **tPSA ≤ 250 Å^2**

2. **3.5 ≤ logP ≤ 6.0**

3. **Hydrogen Bond Acceptors ≤ 16**

4. **3 ≤ Hydrogen Bond Donors ≤ 6**

5. **Rotatable Bonds < 25**

PROTACs: Physico-chemical Properties Score Transformations

The following transformations functions were applied to the physico-chemical properties values to yield a score in the interval [0, 1].

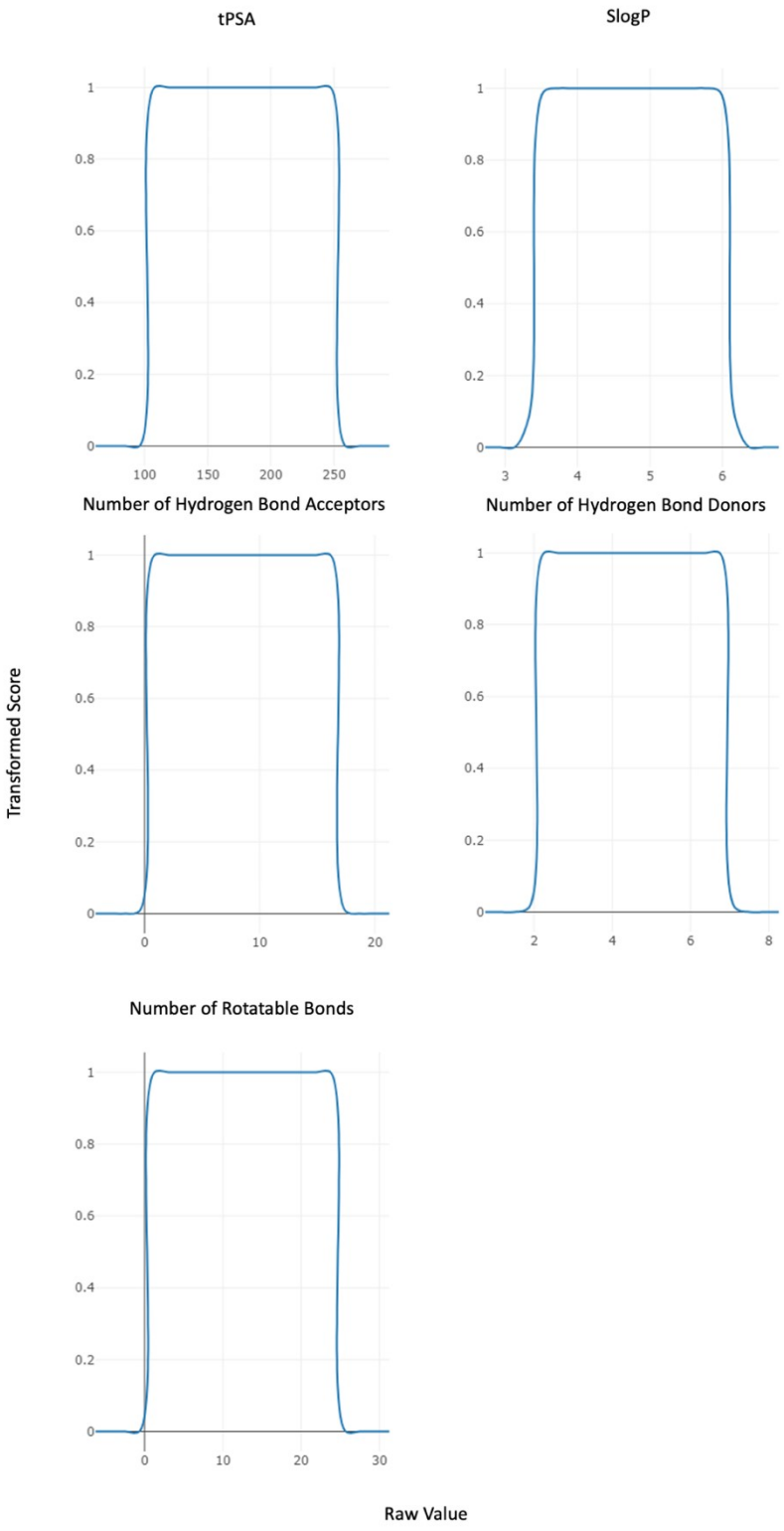


Fig. S15: Experiment 3: PROTACs physico-chemical properties scoring function transformations for Sub-Experiments 1 and 2. The transformation ensures a score in the interval [0,1] is returned for agent feedback.

Sub-Experiment 1: Controlling Linker Length

This Sub-Experiment controls for linker lengths while keeping physico-chemical properties fixed within a specified interval. Linkers were generated within the following length intervals: [4,6], [7,9], [10,12], [13,15], and 'Baseline' which does not enforce length. The corresponding training plots are labelled according to their length intervals. In addition to enforcing physico-chemical properties shown above, the Scoring Function contains the following additional components:

1) **Linker Effective Length = [4,6], [7,9], [10,12], or [13,15]:** this component enforces linkers to possess an effective length within the specified intervals

2) **Linker Length Ratio = 100** this component prevents linker branching

The "Linker Length Ratio" was enforced to be 100, giving as perfect score (1) if achieved and 0 otherwise. Therefore, no score transformation is shown as the outcome is binary. In contrast, a transformation was applied to the "Linker Effective Length" to enforce generated linkers to be within specified length intervals. The transformation is shown below:

Sub-Experiment 1: Scoring Function Components Transformations

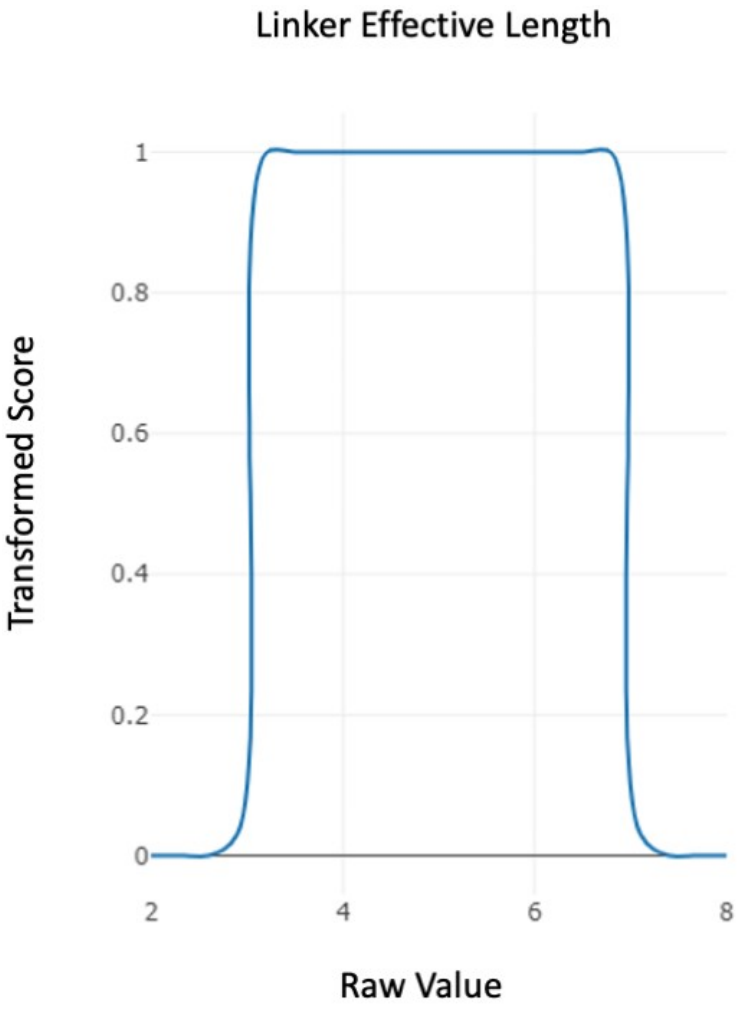


Fig. S16: Experiment 3: PROTACs Sub-Experiment 1 Linker Effective Length Scoring Function transformation. The transformation ensures a score in the interval [0,1] is returned for agent feedback. In the figure, linkers are enforced to possess a length in the interval [4-6]. For the experiments involving other length intervals, the transformation is identical, except shifted to the lengths of interest.

Sub-Experiment 1: Controlling Linker Length Training Plots

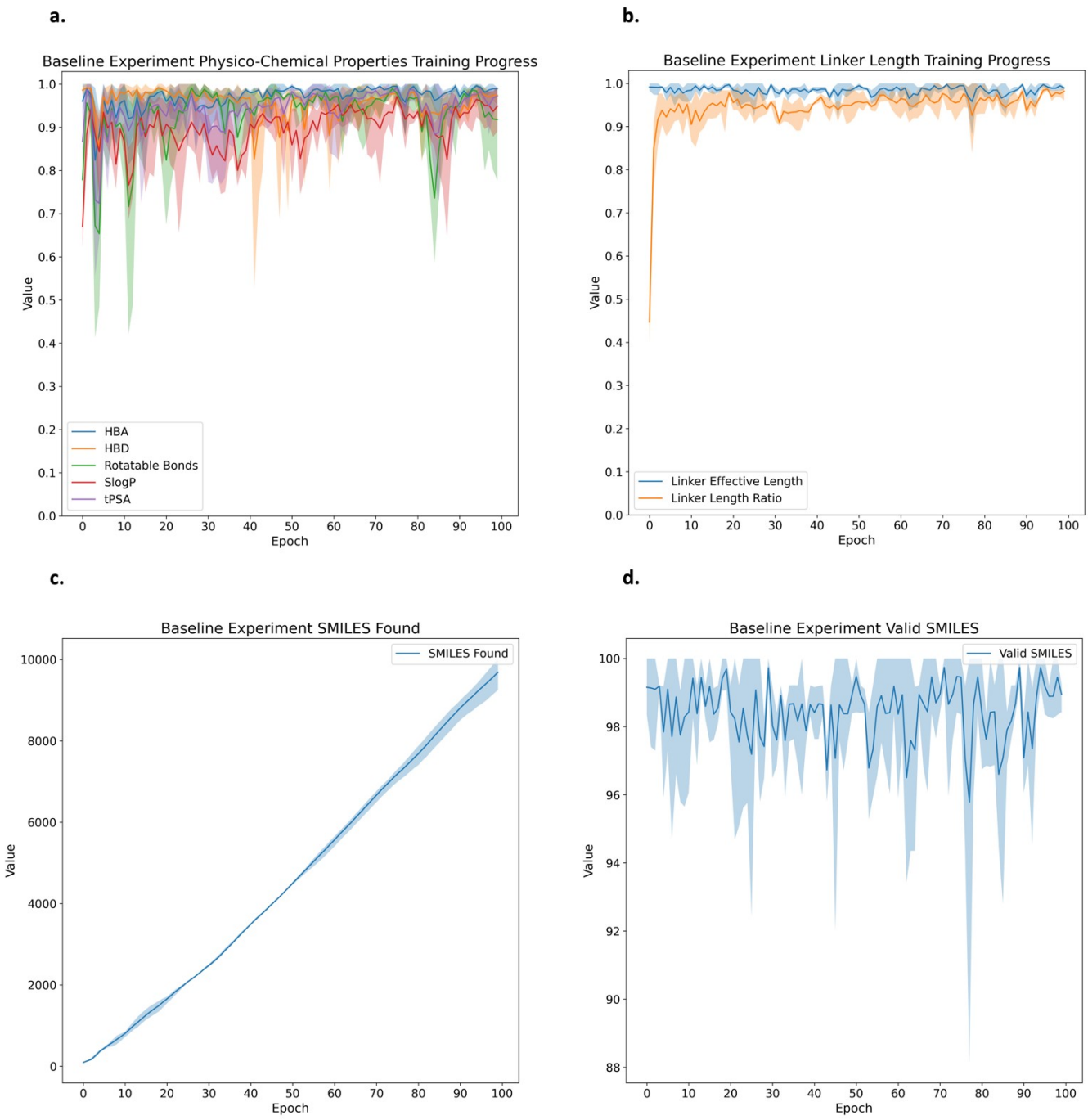


Fig. S17: Experiment 3: PROTACs Sub-Experiment 1: Controlling Linker Length **Baseline** Training Plots. The experiment was run in triplicate. The curve shows the average score and the upper and lower bounds of the shaded region represent the maximum and minimum score, respectively. **a.** Transformed score of all physico-chemical properties components in the scoring function. All components are optimized over the course of training. **b.** Transformed score of all linker specific properties. Note that the "Linker Effective Length" component is always satisfied as the **Baseline** experiment does not enforce any specific length. **c.** SMILES Found. **d.** Percentage of valid SMILES.

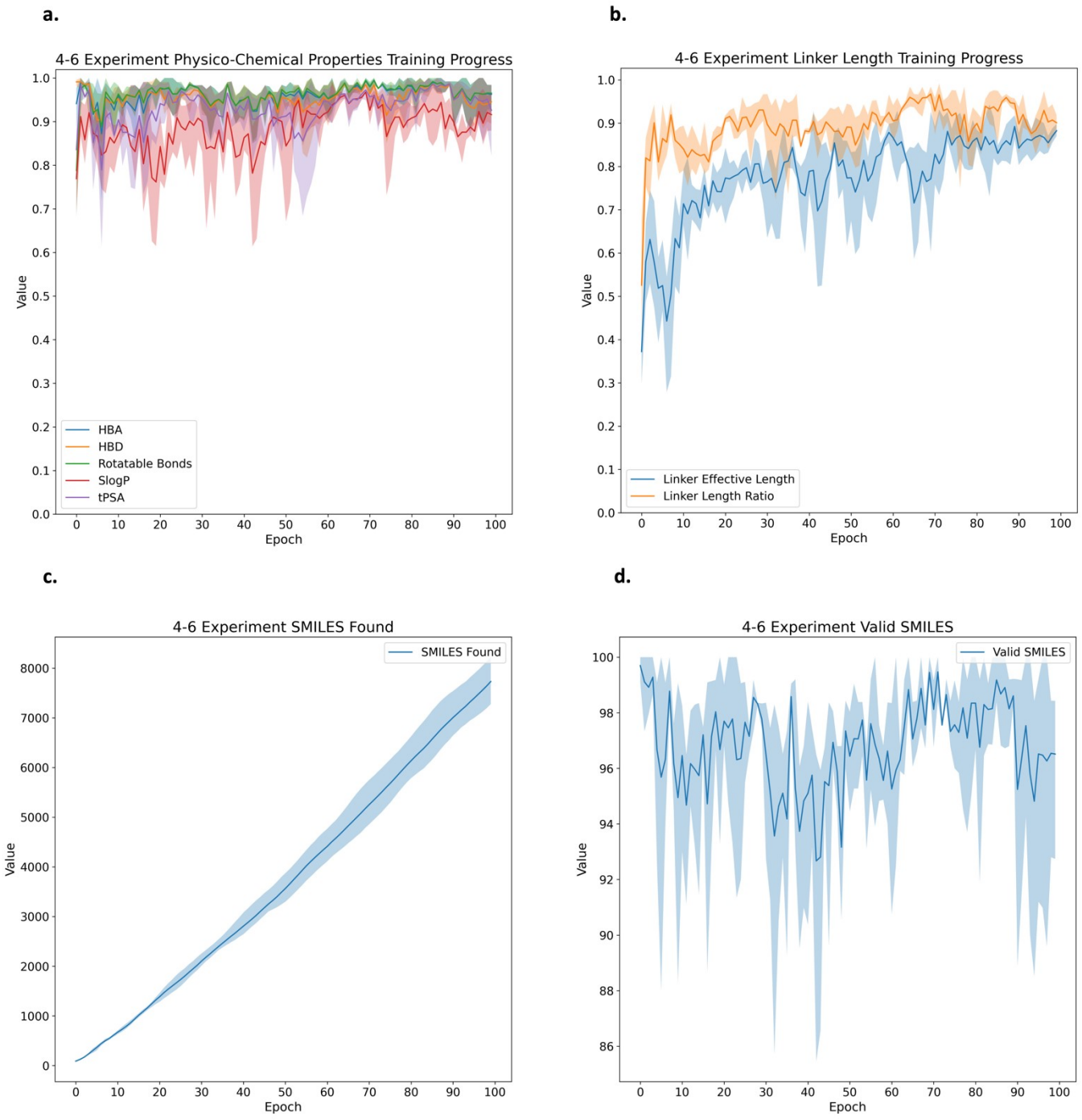


Fig. S18: Experiment 3: PROTACs Sub-Experiment 1: Controlling Linker Length **4-6** Training Plots. The experiment was run in triplicate. The curve shows the average score and the upper and lower bounds of the shaded region represent the maximum and minimum score, respectively. **a.** Transformed score of all physico-chemical properties components in the scoring function. All components are optimized over the course of training. **b.** Transformed score of all linker specific properties. All components are optimized over training as the agent learns to generate linkers within the specified **4-6** interval. **c.** SMILES Found. **d.** Percentage of valid SMILES.

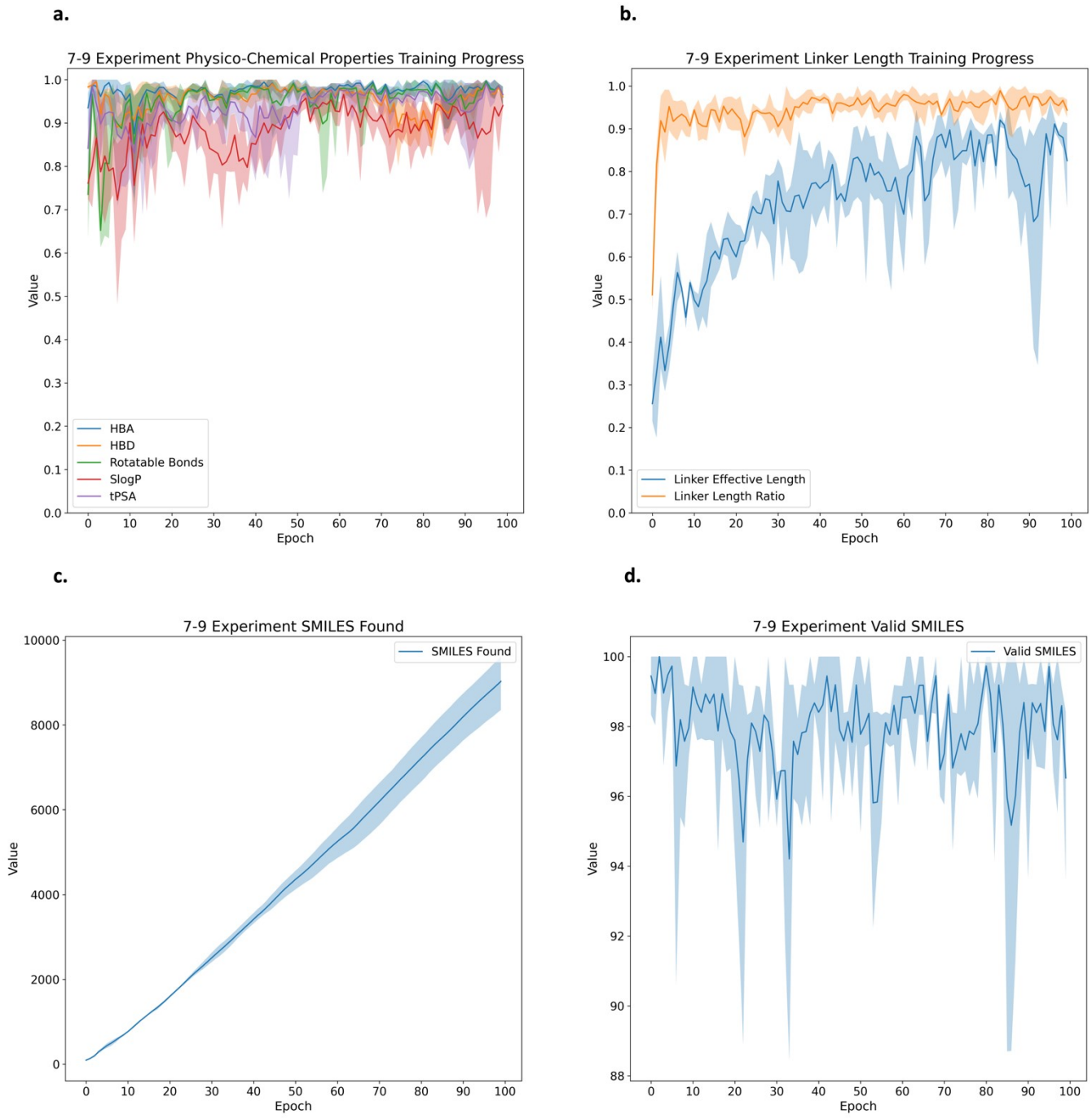


Fig. S19: Experiment 3: PROTACs Sub-Experiment 1: Controlling Linker Length **7-9** Training Plots. The experiment was run in triplicate. The curve shows the average score and the upper and lower bounds of the shaded region represent the maximum and minimum score, respectively. **a.** Transformed score of all physico-chemical properties components in the scoring function. All components are optimized over the course of training. **b.** Transformed score of all linker specific properties. All components are optimized over training as the agent learns to generate linkers within the specified **7-9** interval. **c.** SMILES Found. **d.** Percentage of valid SMILES.

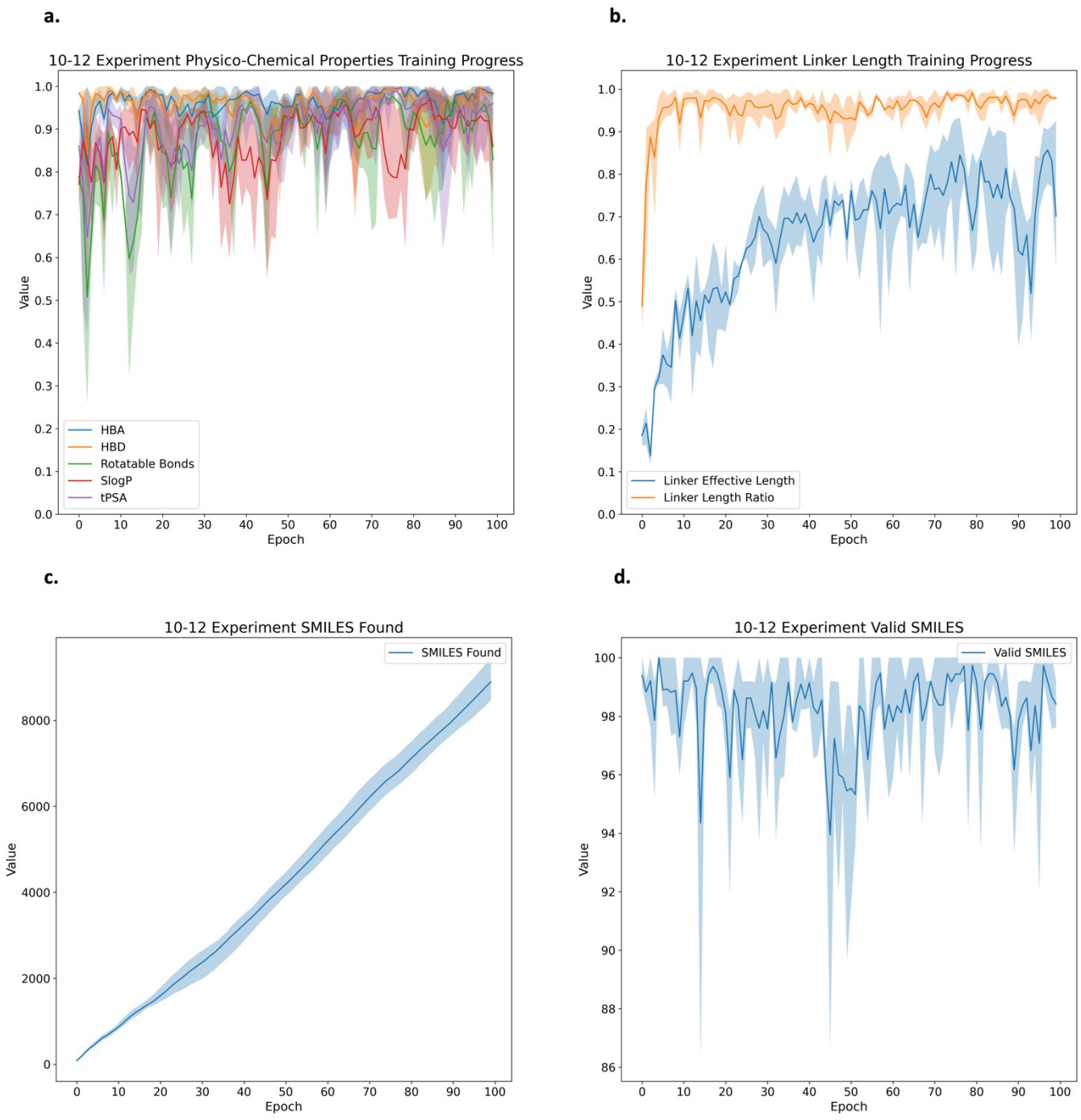


Fig. S20: Experiment 3: PROTACs Sub-Experiment 1: Controlling Linker Length **10-12** Training Plots. The experiment was run in triplicate. The curve shows the average score and the upper and lower bounds of the shaded region represent the maximum and minimum score, respectively. **a.** Transformed score of all physico-chemical properties components in the scoring function. All components are optimized over the course of training. **b.** Transformed score of all linker specific properties. All components are optimized over training as the agent learns to generate linkers within the specified **10-12** interval. **c.** SMILES Found. **d.** Percentage of valid SMILES.

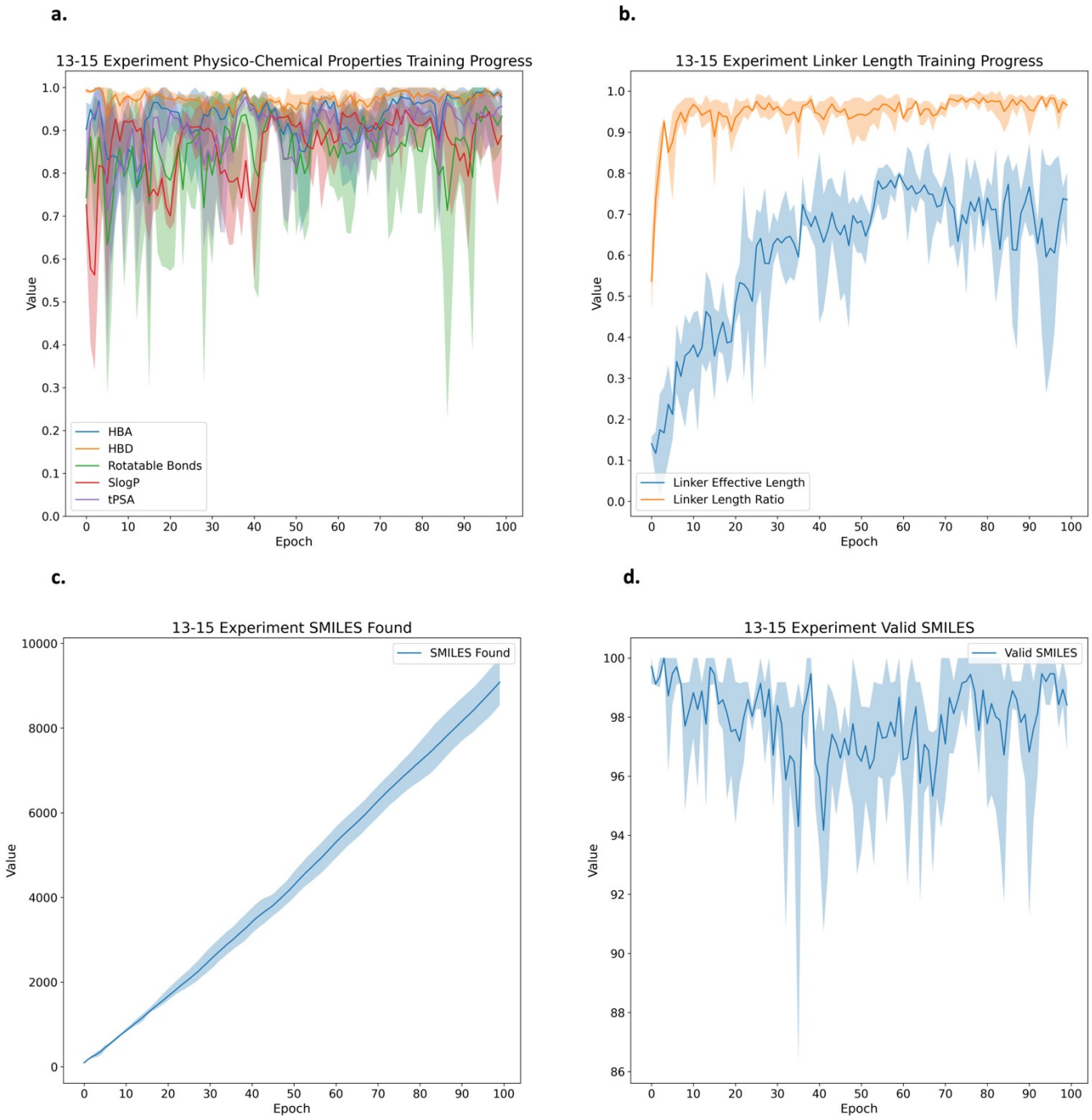


Fig. S21: Experiment 3: PROTACs Sub-Experiment 1: Controlling Linker Length **13-15** Training Plots. The experiment was run in triplicate. The curve shows the average score and the upper and lower bounds of the shaded region represent the maximum and minimum score, respectively. **a.** Transformed score of all physico-chemical properties components in the scoring function. All components are optimized over the course of training. **b.** Transformed score of all linker specific properties. All components are optimized over training as the agent learns to generate linkers within the specified **13-15** interval. **c.** SMILES Found. **d.** Percentage of valid SMILES.

Sub-Experiment 2: Controlling Linker Linearity

This Sub-Experiment controls for linker linearity while keeping physico-chemical properties fixed within a specified interval and linker length fixed in the interval [7,9]. The 'Linear' experiment enforces linear linkers, the 'Cyclic' experiment enforces linkers with at least one ring, and the 'Baseline' experiment enforces neither. In addition to enforcing physico-chemical properties shown in the beginning of this section, the Scoring Function contains the following additional components:

1) **Linker Effective Length = [7,9]:** this component enforces linkers to possess an effective length within the specified interval of [7,9]

2) **Linker Length Ratio = 100** this component prevents linker branching

3) **Linker Number of Rings = 0** this component enforces linkers to possess no rings, i.e., the linker is linear. In the experiment where we want to generate linkers with rings, we simply omit this component in the Scoring Function

The "Linker Length Ratio" was enforced to be 100, giving as perfect score (1) if achieved and 0 otherwise. Therefore, no score transformation is shown as the outcome is binary. In contrast, a transformation was applied to the "Linker Effective Length" to enforce generated linkers to be within specified length intervals. A score transformation was also applied to "Linker Number of Rings" in the experiment where we enforce the generation of linkers without rings. The transformation allows the Agent to gradually learn, and we observe it gives smoother training than a binary transformation like the one applied for the "Linker Length Ratio". The transformations are shown below:

Sub-Experiment 2: Scoring Function Components Transformations

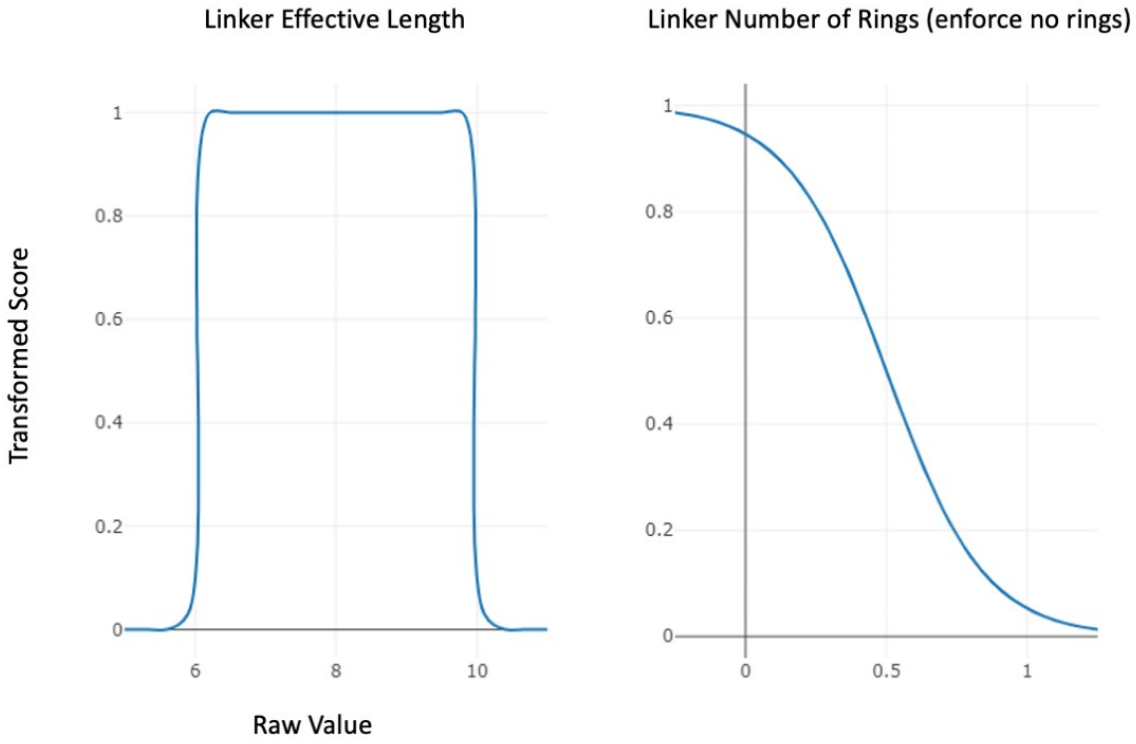
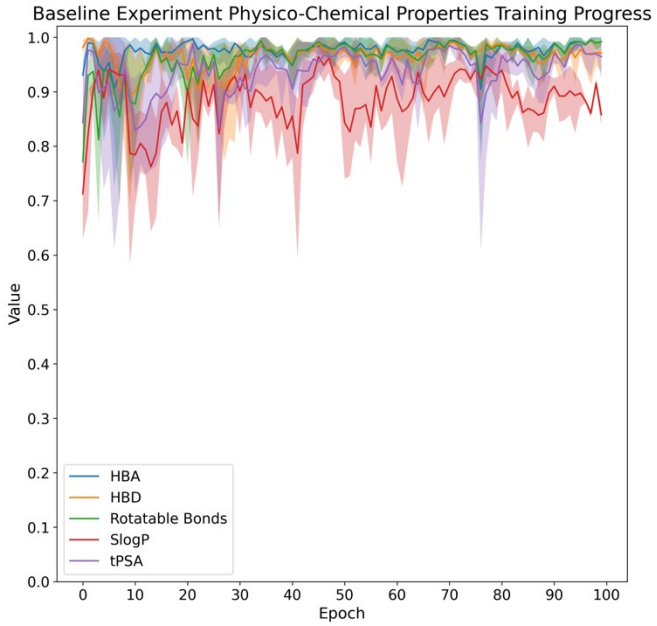


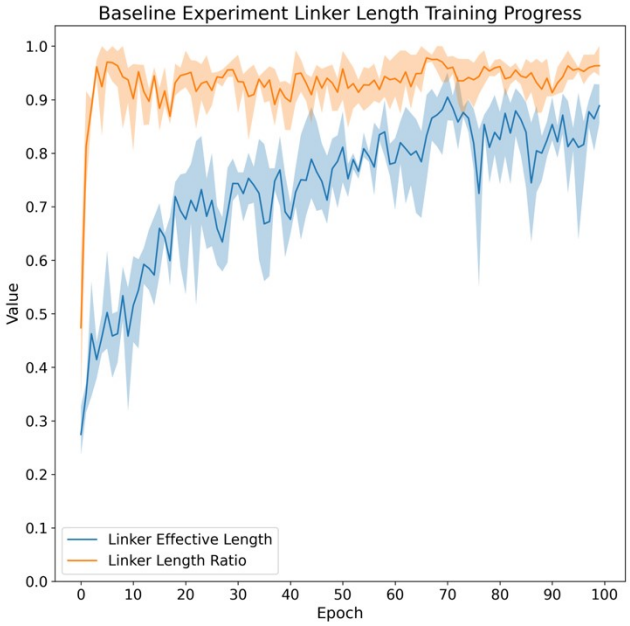
Fig. S22: Experiment 3: PROTACs Sub-Experiment 2 additional Scoring Function transformations. The transformation ensures a score in the interval [0,1] is returned for agent feedback. The linkers are enforced to possess a length in the interval [7-9]. In the experiment where we enforce linkers to contain no rings, a gradient transformation is applied for smoother Agent training (a linker with 1 ring is not automatically given a score of 0. Instead, it achieves a low score).

Sub-Experiment 2: Controlling Linker Linearity Training Plots

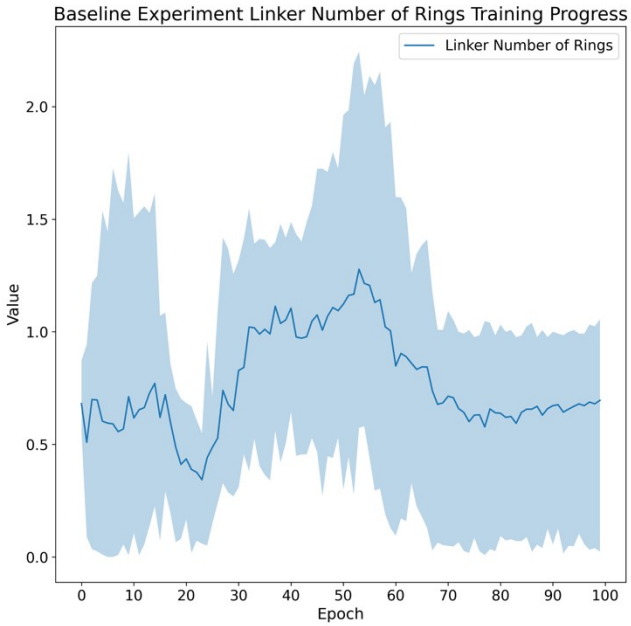
a.



b.



c.



d.

e.

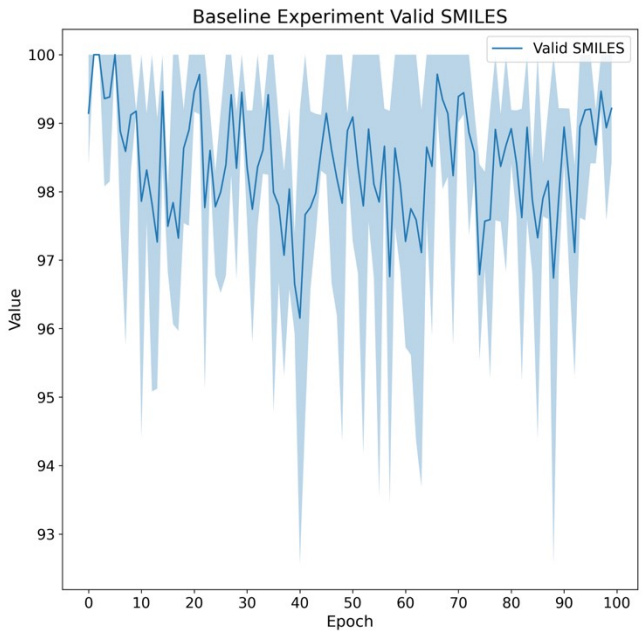
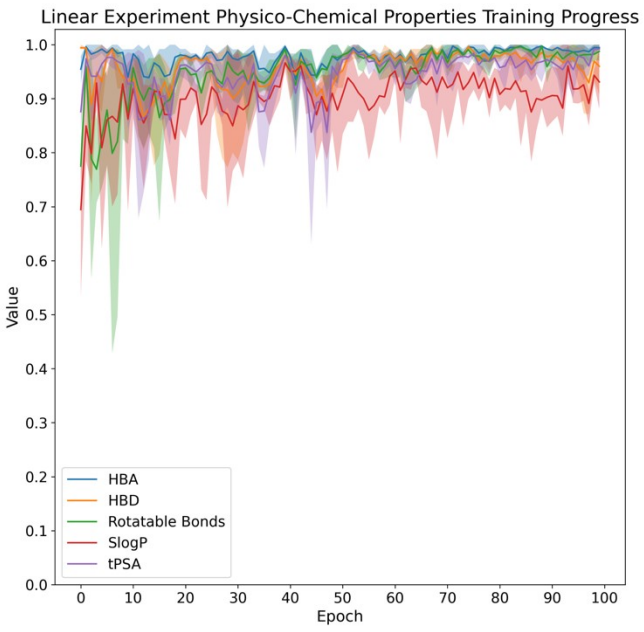
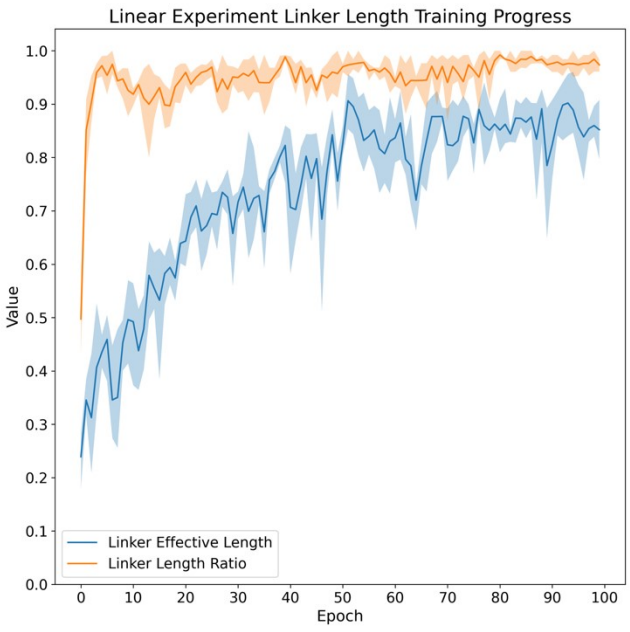


Fig. S23: Experiment 3: PROTACs Sub-Experiment 2: Controlling Linker Linearity **Baseline** Training Plots. The experiment was run in triplicate. The curve shows the average score and the upper and lower bounds of the shaded region represent the maximum and minimum score, respectively. **a.** Transformed score of all physico-chemical properties components in the scoring function. All components are optimized over the course of training**. b.** Transformed score of all linker specific properties. All components are optimized over training as the agent learns to generate linkers within the specified **7-9** interval. **c.** Raw linker number of rings. In the **Baseline** experiment, the agent generated linkers with and without rings as neither is enforced. **d.** SMILES Found. **e.** Percentage of valid SMILES.

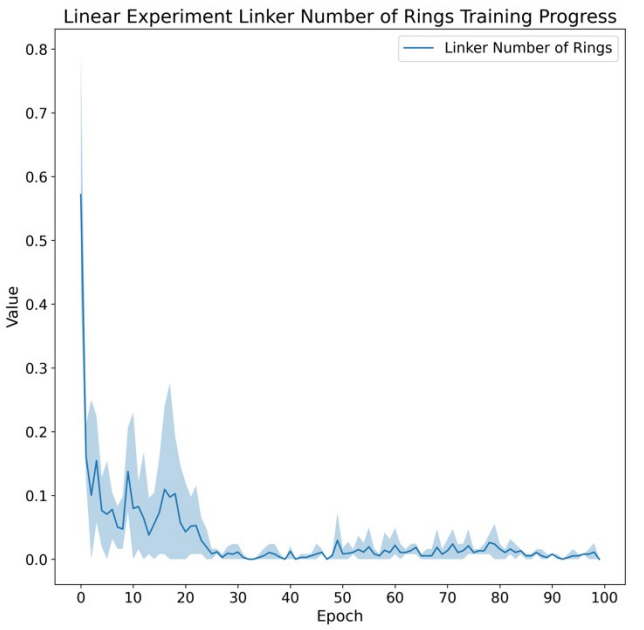
a.


Linear Experiment Physico-Chemical Properties Training Progress
Value
Epoch
HBA
HBD
Rotatable Bonds
SlogP
tPSA

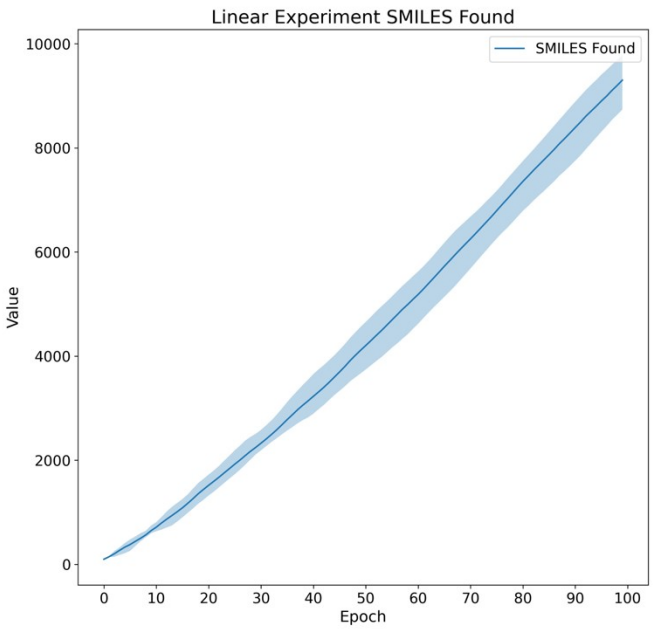
b.


Linear Experiment Linker Length Training Progress
Value
Epoch
Linker Effective Length
Linker Length Ratio

c.


Linear Experiment Linker Number of Rings Training Progress
Value
Epoch
Linker Number of Rings

d.


Linear Experiment SMILES Found
Value
Epoch
SMILES Found

e.

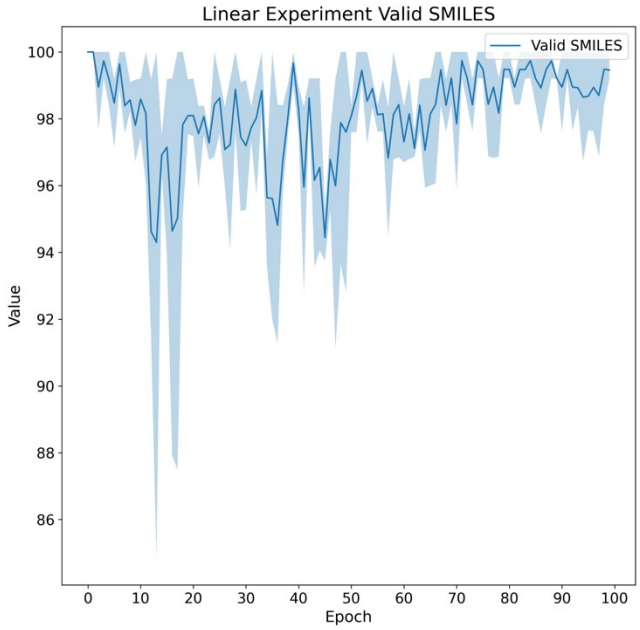
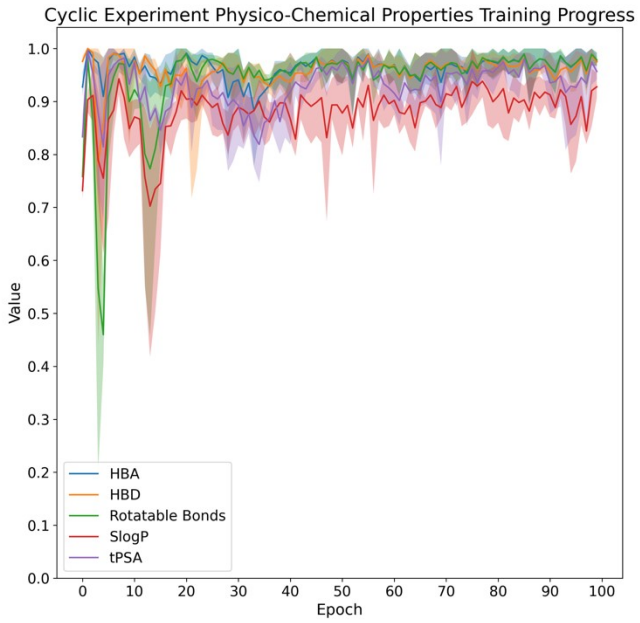
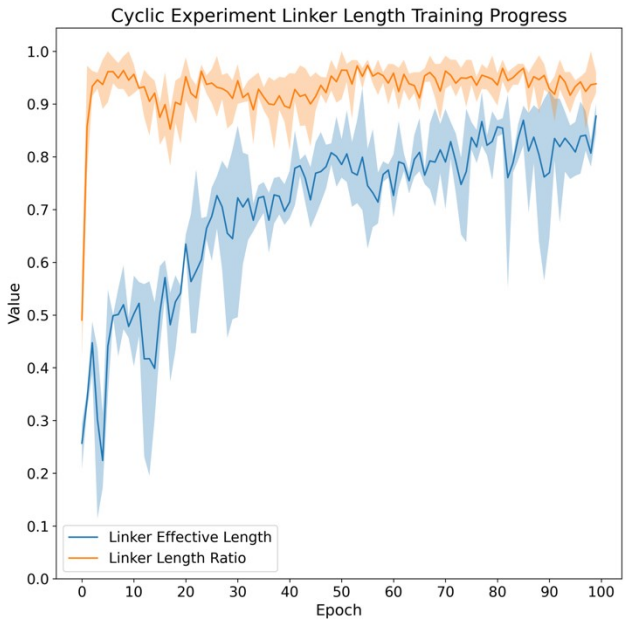


Fig. S24: Experiment 3: PROTACs Sub-Experiment 2: Controlling Linker Linearity **Linear** Training Plots. The experiment was run in triplicate. The curve shows the average score and the upper and lower bounds of the shaded region represent the maximum and minimum score, respectively. **a.** Transformed score of all physico-chemical properties components in the scoring function. All components are optimized over the course of training. **b.** Transformed score of all linker specific properties. All components are optimized over training as the agent learns to generate linkers within the specified **7-9** interval. **c.** Raw linker number of rings. In the **Linear** experiment, the agent learns to generate linkers without rings. **d.** SMILES Found. **e.** Percentage of valid SMILES.

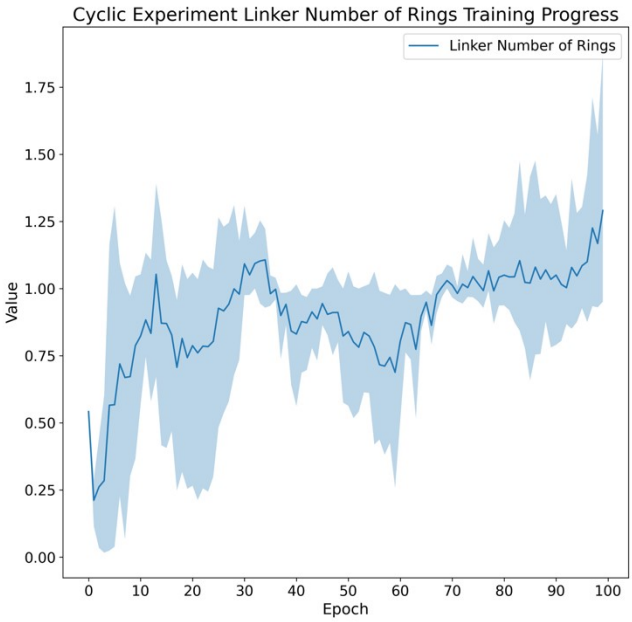
a.


Cyclic Experiment Physico-Chemical Properties Training Progress
Value
Epoch
HBA
HBD
Rotatable Bonds
SlogP
tPSA

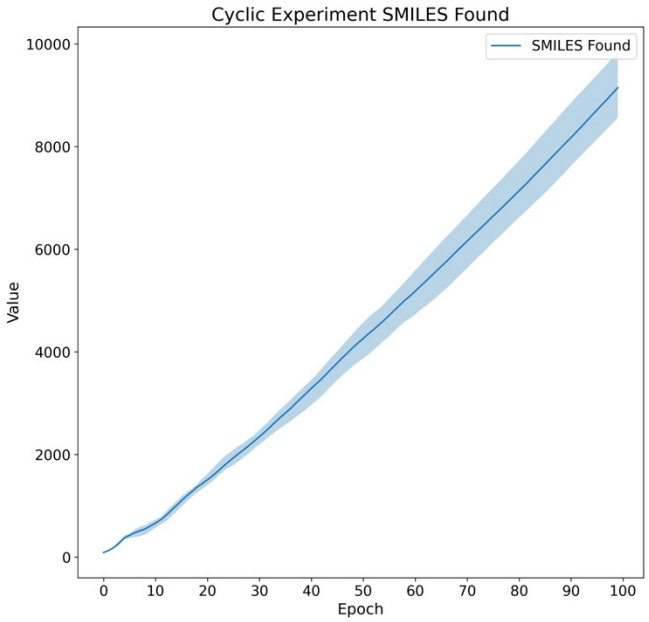
b.


Cyclic Experiment Linker Length Training Progress
Value
Epoch
Linker Effective Length
Linker Length Ratio

c.


Cyclic Experiment Linker Number of Rings Training Progress
Value
Epoch
Linker Number of Rings

d.


Cyclic Experiment SMILES Found
Value
Epoch
SMILES Found

e.

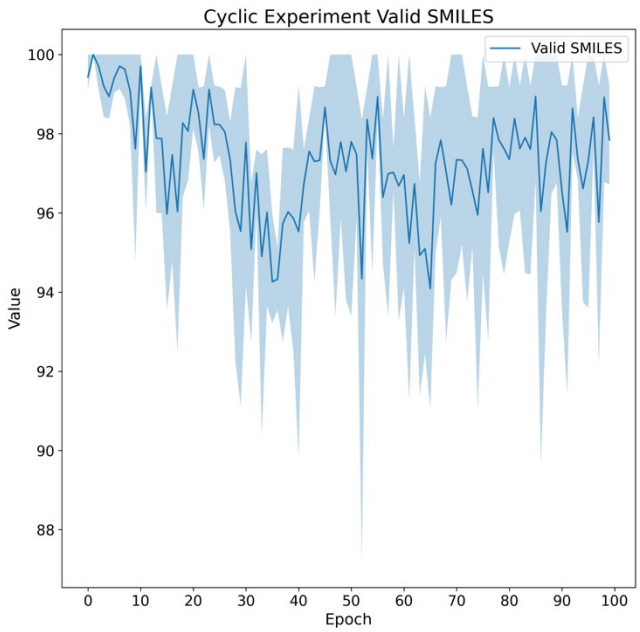


Fig. S25: Experiment 3: PROTACs Sub-Experiment 2: Controlling Linker Linearity **Cyclic**Training Plots. The experiment was run in triplicate. The curve shows the average score and the upper and lower bounds of the shaded region represent the maximum and minimum score, respectively. **a.** Transformed score of all physico-chemical properties components in the scoring function. All components are optimized over the course of training. **b.** Transformed score of all linker specific properties. All components are optimized over training as the agent learns to generate linkers within the specified **7-9** interval. **c.** Raw linker number of rings. In the **Cyclic** experiment, the agent learns to generate linkers with at least one ring. **d.** SMILES Found. **e.** Percentage of valid SMILES.

Sub-Experiment 3: Controlling Linker Flexibility

This Sub-Experiment controls for linker flexibility. Linkers were generated within the following 'linker ratio of rotatable bonds' intervals: "Low": [0,30], "Moderate": [40,60], and "High": [70,100]. No physico-chemical properties were enforced in this Sub-Experiment. Instead, the Scoring Function contains only one component:

1) **Linker Ratio of Rotatable Bonds = [0,30], [40,60], [70,100]:** the defined intervals correspond to "Low", "Moderate", and "High" flexibility

The Scoring Function transformation is shown below:

Sub-Experiment 3: Scoring Function Component Transformation

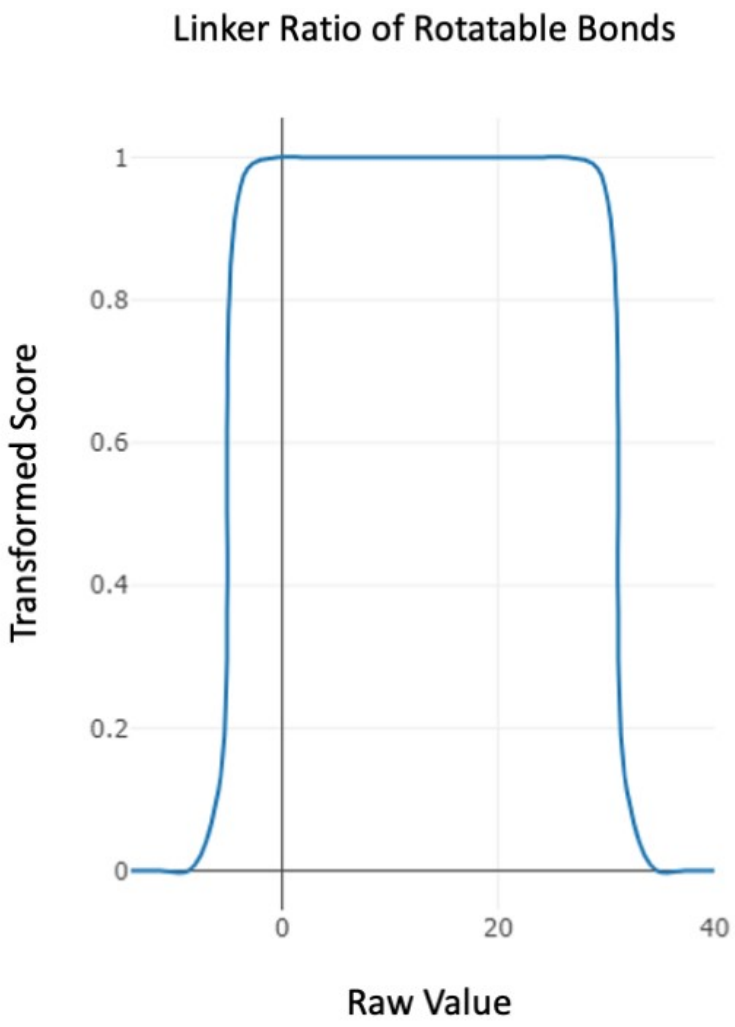


Fig. S26: Experiment 3: PROTACs Sub-Experiment 3 Linker Ratio of Rotatable Bonds Scoring Function transformation. The transformation ensures a score in the interval [0,1] is returned for agent feedback. In the figure, linkers are enforced to possess a Linker Length Ratio in the interval [0,30]. For the experiments involving other intervals, the transformation is identical, except shifted to the values of interest.

Sub-Experiment 3: Controlling Linker Flexibility Training Plots

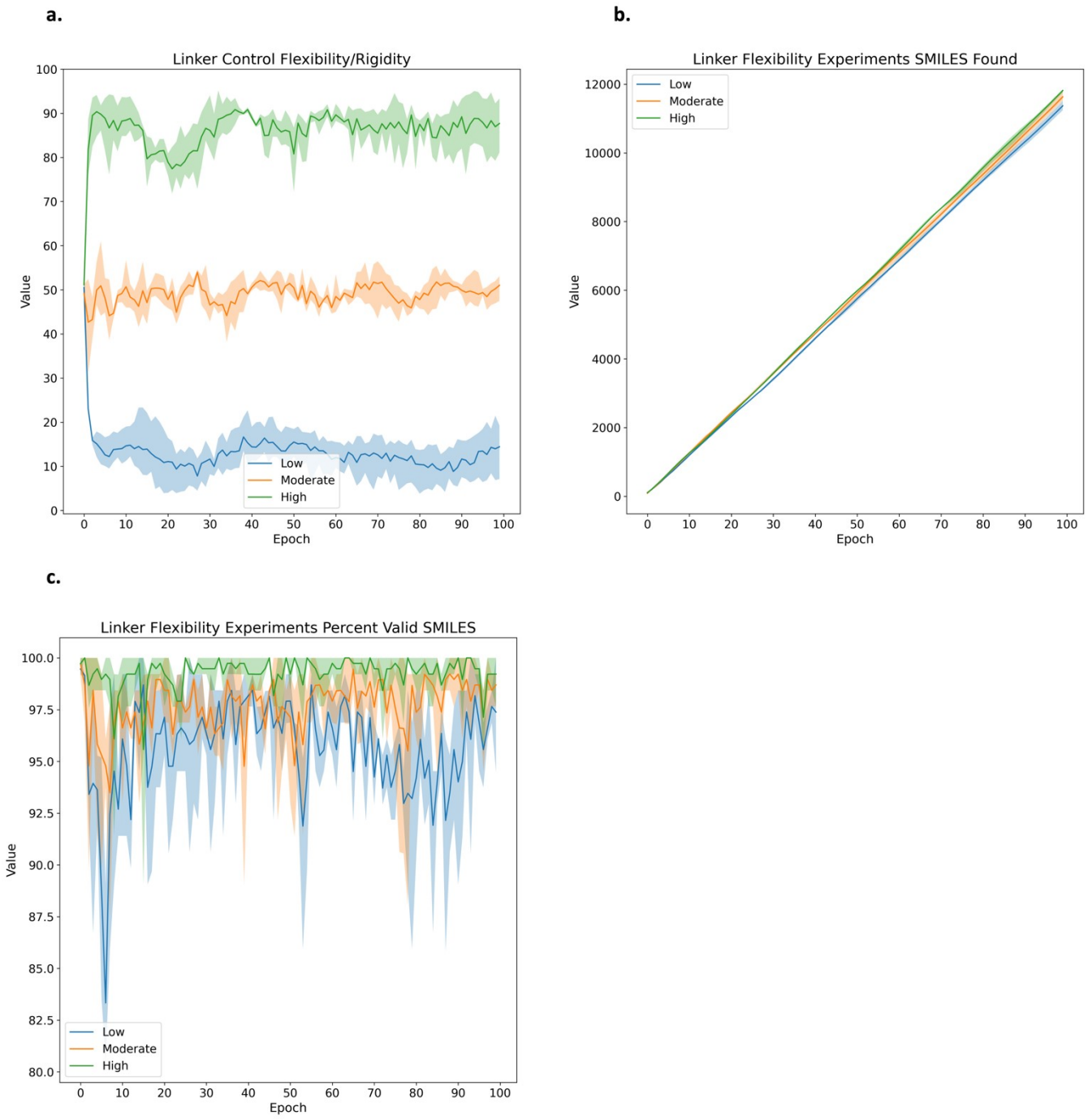


Fig. S27: Experiment 3: PROTACs Sub-Experiment 3: Controlling Linker Flexibility Training Plots. The experiment was run in triplicate. The curve shows the average score and the upper and lower bounds of the shaded region represent the maximum and minimum score, respectively. **a.** Raw 'linker ratio of rotatable bonds' values for the "Low", "Moderate" and "High" experiments which enforces a value within the intervals, [0,30], [40-60], and [70-100], respectively. In each experiment, the agent learns to generate linkers with 'linker ratio of rotatable bonds' values within the specified intervals. **b.** SMILES Found. **c.** Percentage of valid SMILES.

References

(1) Gaulton, A.; Hersey, A.; Nowotka, M.; Bento, A. P.; Chambers, J.; Mendez, D.; Mutowo, P.; Atkinson, F.; Bellis, L. J.; Cibrián-Uhalte, E.; Davies, M.; Dedman, N.; Karlsson, A.; Magariños, M. P.; Overington, J. P.; Papadatos, G.; Smit, I.; Leach, A. R. The ChEMBL Database in 2017. *Nucleic Acids Research* **2017**, *45* (D1), D945–D954. https://doi.org/10.1093/nar/gkw1074.

(2) *LibINVENT: Reaction-based Generative Scaffold Decoration for in Silico Library Design | Journal of Chemical Information and Modeling*. https://pubs.acs.org/doi/10.1021/acs.jcim.1c00469 (accessed 2022-01-01).

(3) Weininger, D. SMILES, a Chemical Language and Information System. 1. Introduction to Methodology and Encoding Rules. *J. Chem. Inf. Comput. Sci.* **1988**, *28* (1), 31–36. https://doi.org/10.1021/ci00057a005.

(4) Bemis, G. W.; Murcko, M. A. The Properties of Known Drugs. 1. Molecular Frameworks. *J. Med. Chem.* **1996**, *39* (15), 2887–2893. https://doi.org/10.1021/jm9602928.

(5) Arús-Pous, J.; Johansson, S. V.; Prykhodko, O.; Bjerrum, E. J.; Tyrchan, C.; Reymond, J.-L.; Chen, H.; Engkvist, O. Randomized SMILES Strings Improve the Quality of Molecular Generative Models. *Journal of Cheminformatics* **2019**, *11* (1), 71. https://doi.org/10.1186/s13321-019-0393-0.

(6) Williams, R. J.; Zipser, D. A Learning Algorithm for Continually Running Fully Recurrent Neural Networks. *Neural Computation* **1989**, *1* (2), 270–280. https://doi.org/10.1162/neco.1989.1.2.270.

(7) Blaschke, T.; Arús-Pous, J.; Chen, H.; Margreitter, C.; Tyrchan, C.; Engkvist, O.; Papadopoulos, K.; Patronov, A. REINVENT 2.0: An AI Tool for De Novo Drug Design. *J. Chem. Inf. Model.* **2020**, *60* (12), 5918–5922. https://doi.org/10.1021/acs.jcim.0c00915.

(8) Olivecrona, M.; Blaschke, T.; Engkvist, O.; Chen, H. Molecular De-Novo Design through Deep Reinforcement Learning. *Journal of Cheminformatics* **2017**, *9* (1), 48. https://doi.org/10.1186/s13321-017-0235-x.

(9) Segler, M. H. S.; Kogej, T.; Tyrchan, C.; Waller, M. P. Generating Focused Molecule Libraries for Drug Discovery with Recurrent Neural Networks. *ACS Cent. Sci.* **2018**, *4* (1), 120–131. https://doi.org/10.1021/acscentsci.7b00512.

(10) De Fusco, C.; Brear, P.; Iegre, J.; Georgiou, K. H.; Sore, H. F.; Hyvönen, M.; Spring, D. R. A Fragment-Based Approach Leading to the Discovery of a Novel Binding Site and the Selective CK2 Inhibitor CAM4066. *Bioorg Med Chem* **2017**, *25* (13), 3471–3482. https://doi.org/10.1016/j.bmc.2017.04.037.

(11) Brear, P.; De Fusco, C.; Hadje Georgiou, K.; Francis-Newton, N. J.; Stubbs, C. J.; Sore, H. F.; Venkitaraman, A. R.; Abell, C.; Spring, D. R.; Hyvönen, M. Specific Inhibition of CK2α from an Anchor Outside the Active Site. *Chem. Sci.* **2016**, *7* (11), 6839–6845. https://doi.org/10.1039/C6SC02335E.

(12) Patel, S.; Cohen, F.; Dean, B. J.; De La Torre, K.; Deshmukh, G.; Estrada, A. A.; Ghosh, A. S.; Gibbons, P.; Gustafson, A.; Huestis, M. P.; Le Pichon, C. E.; Lin, H.; Liu, W.; Liu, X.; Liu, Y.; Ly, C. Q.; Lyssikatos, J. P.; Ma, C.; Scearce-Levie, K.; Shin, Y. G.; Solanoy, H.; Stark, K. L.; Wang, J.; Wang, B.; Zhao, X.; Lewcock, J. W.; Siu, M. Discovery of Dual Leucine Zipper Kinase (DLK, MAP3K12) Inhibitors with Activity in Neurodegeneration Models. *J. Med. Chem.* **2015**, *58* (1), 401–418. https://doi.org/10.1021/jm5013984.

(13) Patel, S.; Harris, S. F.; Gibbons, P.; Deshmukh, G.; Gustafson, A.; Kellar, T.; Lin, H.; Liu, X.; Liu, Y.; Liu, Y.; Ma, C.; Scearce-Levie, K.; Ghosh, A. S.; Shin, Y. G.; Solanoy, H.; Wang, J.; Wang, B.; Yin, J.; Siu, M.; Lewcock, J. W. Scaffold-Hopping and Structure-Based Discovery of Potent, Selective, And Brain Penetrant N-(1H-Pyrazol-3-Yl)Pyridin-2-Amine Inhibitors of Dual Leucine Zipper Kinase

(DLK, MAP3K12). *J. Med. Chem.* **2015**, *58* (20), 8182–8199. https://doi.org/10.1021/acs.jmedchem.5b01072.
(14) *Schrödinger Release 2019-4: Maestro, Schrödinger, LLC, New York, NY, 2019.*
(15) *Schrödinger Release 2019-4: LigPrep, Schrödinger, LLC, New York, NY, 2019.*
(16) Guo, J.; Janet, J. P.; Bauer, M. R.; Nittinger, E.; Giblin, K. A.; Papadopoulos, K.; Voronov, A.; Patronov, A.; Engkvist, O.; Margreitter, C. DockStream: A Docking Wrapper to Enhance de Novo Molecular Design. *Journal of Cheminformatics* **2021**, *13* (1), 89. https://doi.org/10.1186/s13321-021-00563-7.
(17) Friesner, R. A.; Banks, J. L.; Murphy, R. B.; Halgren, T. A.; Klicic, J. J.; Mainz, D. T.; Repasky, M. P.; Knoll, E. H.; Shelley, M.; Perry, J. K.; Shaw, D. E.; Francis, P.; Shenkin, P. S. Glide: A New Approach for Rapid, Accurate Docking and Scoring. 1. Method and Assessment of Docking Accuracy. *J. Med. Chem.* **2004**, *47* (7), 1739–1749. https://doi.org/10.1021/jm0306430.
(18) Halgren, T. A.; Murphy, R. B.; Friesner, R. A.; Beard, H. S.; Frye, L. L.; Pollard, W. T.; Banks, J. L. Glide: A New Approach for Rapid, Accurate Docking and Scoring. 2. Enrichment Factors in Database Screening. *J. Med. Chem.* **2004**, *47* (7), 1750–1759. https://doi.org/10.1021/jm030644s.
(19) Friesner, R. A.; Murphy, R. B.; Repasky, M. P.; Frye, L. L.; Greenwood, J. R.; Halgren, T. A.; Sanschagrin, P. C.; Mainz, D. T. Extra Precision Glide: Docking and Scoring Incorporating a Model of Hydrophobic Enclosure for Protein-Ligand Complexes. *J. Med. Chem.* **2006**, *49* (21), 6177–6196. https://doi.org/10.1021/jm051256o.
(20) *Schrödinger Release 2019-4: Glide, Schrödinger, LLC, New York, NY, 2019.*
(21) Trapero, A.; Pacitto, A.; Singh, V.; Sabbah, M.; Coyne, A. G.; Mizrahi, V.; Blundell, T. L.; Ascher, D. B.; Abell, C. Fragment-Based Approach to Targeting Inosine-5'-Monophosphate Dehydrogenase (IMPDH) from Mycobacterium Tuberculosis. *J. Med. Chem.* **2018**, *61* (7), 2806–2822. https://doi.org/10.1021/acs.jmedchem.7b01622.
(22) Imrie, F.; Bradley, A. R.; van der Schaar, M.; Deane, C. M. Deep Generative Models for 3D Linker Design. *J. Chem. Inf. Model.* **2020**, *60* (4), 1983–1995. https://doi.org/10.1021/acs.jcim.9b01120.
(23) Yang, Y.; Zheng, S.; Su, S.; Zhao, C.; Xu, J.; Chen, H. SyntaLinker: Automatic Fragment Linking with Deep Conditional Transformer Neural Networks. *Chem. Sci.* **2020**, *11* (31), 8312–8322. https://doi.org/10.1039/D0SC03126G.
(24) Wang, Z.; He, N.; Guo, Z.; Niu, C.; Song, T.; Guo, Y.; Cao, K.; Wang, A.; Zhu, J.; Zhang, X.; Zhang, Z. Proteolysis Targeting Chimeras for the Selective Degradation of Mcl-1/Bcl-2 Derived from Nonselective Target Binding Ligands. *J. Med. Chem.* **2019**, *62* (17), 8152–8163. https://doi.org/10.1021/acs.jmedchem.9b00919.
(25) Maple, H. J.; Clayden, N.; Baron, A.; Stacey, C.; Felix, R. Developing Degraders: Principles and Perspectives on Design and Chemical Space. *Med. Chem. Commun.* **2019**, *10* (10), 1755–1764. https://doi.org/10.1039/C9MD00272C.
(26) Weng, G.; Shen, C.; Cao, D.; Gao, J.; Dong, X.; He, Q.; Yang, B.; Li, D.; Wu, J.; Hou, T. PROTAC-DB: An Online Database of PROTACs. *Nucleic Acids Research* **2021**, *49* (D1), D1381–D1387. https://doi.org/10.1093/nar/gkaa807.