Deep Reinforcement Learning for *De novo* Synthesis of Eye Drops

Anastasia Vepreva[1], **Nina Gubina**[1], **Andrei Dmitrenko**[1,2], **Vladimir Vinogradov**[1]

[1]Center for AI in Chemistry, SCAMT institute, ITMO University, Russian Federation
[2]D ONE AG, Zurich, Switzerland
dmitrenko@scamt-itmo.ru

Abstract

When developing eye drops, it is necessary to take into account not only the biological activity of the drug, but also a number of additional properties affecting its effectiveness. In this work, we present a novel pipeline based on deep reinforcement learning for *de novo* design of ophthalmic drugs with desired properties. Using the pipeline, we generate 2000 new molecules with improved corneal permeability, melanin binding and irritation as key properties characterizing active substances of eye drops. The resulting molecules are validated against already known drugs by a set of properties commonly used to evaluate drug-like compounds.

Code — https://github.com/ai-chem/ophthalmic_drugs

Introduction

Eye drops are the preferred method of treatment for various ophthalmic diseases due to their convenience, cost-effectiveness, and safety. However, their efficacy is limited by the challenges of delivering drugs through the anterior segment of the eye, which has unique physiology and anatomy restricting bioavailability (Gause et al. 2016). Development of eye drops is no different to any other drug and remains a long and resource-intensive task. In the early stages of rational drug design, new molecules can be created by merging fragments of existing compounds or using optimization techniques like genetic algorithms (**?**). Recently, deep learning methods have gained traction in drug discovery, with autoencoder-based models (Gómez-Bombarelli et al. 2018), variational autoencoders (VAEs) and adversarial autoencoders (AAEs) (Dai et al. 2018), generative adversarial networks (GANs) (De Cao and Kipf 2018). The models discussed do not account for protein-ligand interactions, which are crucial for therapeutic effects (Śledź and Caflisch 2018). A promising strategy for generating protein-conditioned molecules involves reinforcement learning (RL), where an agent explores chemical space and receives rewards based on the properties of generated molecules (Danel et al. 2023). The binding affinity of a molecule to a target protein serves as an ideal reward. Among generative RL models for de novo drug design, REINVENT and MolDQN are notable (Popova, Isayev, and Tropsha 2018; Zhou et al. 2019). A recent study comparing the improved FREED++ model with these on docking score optimization for the USP7 target showed that FREED++ achieved higher scores, likely due to its focus on fragment generation, which narrows the search space and enhances the RL agent's learning efficiency (Telepov et al. 2024).

As mentioned earlier, development of eye drops is fraught with difficulties, since this type of drug administration has low bioavailability. First, eye drops need to pass a natural barrier - the cornea. Prediction of corneal permeability using machine learning methods has been performed by various scientific groups (Agatonovic-Kustrin, Evans, and Alany 2003; Ghorbanzad'e et al. 2011; Kidron et al. 2010; Ramsay et al. 2018; Dargó et al. 2019). Classical machine learning models were used to predict the speed of drug passage through a rabbit's cornea. Corneal permeability is not the only factor for increasing the bioavailability of eye drops, however. The eye tissues have pigmented cells containing melanin. Binding of the drug to this biopolymer can lead to longer retention of the drug in the tissues of the eye. Several research groups(Jakubiak et al. 2018; Lowrey et al. 1997; Radwa et al. 1995; Reilly et al. 2015) have developed an in silico approach to predict drug binding to melanin.

In this work, we present a novel pipeline for generating potential active substances for eye drops based on protein-ligand interactions with predefined properties. We trained classical machine learning (ML) models to predict corneal permeability, drug binding to melanin, and eye irritation. These models were integrated into a fragment-based generative model based on FREED++. Finally, we performed the generation procedure using our pipeline. The generated molecules utilized the predictions from our ML models in the reward function. We evaluate these structures against our trained models and compared their properties with those produced by the FREED++ model. Our approach yielded molecules with enhanced potential for eye drops, demonstrating improved target characteristics. Additionally, we validated the resulting compounds against key drug-like properties, including lipophilicity and synthetic accessibility.

Results and discussion

Pipeline for generative design of ophthalmic drugs

We present a novel pipeline for generative design of ophthalmic drugs with predefined properties (Fig.1).

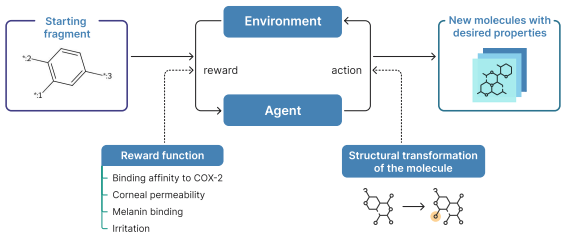

Figure 1: Novel pipeline for *de novo* design of ophthalmic drugs based on RL. We constructed the reward function to explicitly take into account the desired properties of eye drops and binding affinity to COX-2 protein.

We used RL-based sequential generation method based on FREED++ model (Telepov et al. 2024). The agent takes the current state (starting fragment) and selects an action which is a molecular fragment appended to the current state. The transition dynamic is straightforward: the new state is assembled from the previous state by attaching a new fragment to it. In the context of the RL cycle, machine learning models are used as reward to predict specific properties of eye drops, namely, corneal permeability, melanin binding, and irritation. In our implementation, we have adopted an approach where the weights for all predicted properties are set to 1. This means that each property contributes equally to the overall reward calculation during the training process. The rationale behind this decision is to ensure that no single property dominates the learning process, thereby allowing the model to develop a balanced understanding of all properties being predicted. Thus, by repeating several RL cycles, we fine-tune the model to produce molecules with the desired properties. To implement the reinforcement learning approach, we used a high-performance computing setup with single NVIDIA A6000 GPU and 20 GB of RAM. In the following sections, we describe the results of the individual components of the pipeline in more detail.

Data

We manually collected three datasets describing the target properties of eye drops - corneal permeability, melanin binding, and irritation (Fig. 2A). For corneal permeability, we selected studies that tested the rate of drug permeability through the rabbit cornea *in vivo*, since it is the most common method for measuring corneal permeability (Agatonovic-Kustrin, Evans, and Alany 2003; Ghorbanzad'e et al. 2011; Kidron et al. 2010; Ramsay et al. 2018). For melanin binding, we used the data obtained from *in vitro* studies, since their number was noticeably larger. We used a dataset from a study where drug binding to melanin was assessed by analytically quantifying the unbound fraction after centrifugation (Jakubiak et al. 2018). For irritation, we used *in vivo* data obtained from the Draize test, a toxicity test that involves applying a substance to the cornea of a live, immobilized rabbit, then washing the substance off after a certain time and recording the effects (Wang et al. 2017).

SMILES (Simplified Molecular-Input Line-Entry System) notation is the most widely used way of representing molecules. Therefore, we used the SMILES representations to describe the composition and structure of chemical molecules with short strings. To convert strings into the vector form, we used MACCS (Molecular ACCess System) fingerprints, RDKit descriptors and Morgan fingerprints. For each dataset, we employed three variants of representing molecules (i.e., RDKit descriptors, MACCS and Morgan fingerprints) and compared their effectiveness in terms of performance metrics of the corresponding machine learning models. Additionally, we experimented with multiple feature engineering and selection techniques (such as PCA for dimension reduction, correlation and feature importance analysis) striving for the best predictive performance (Fig. 2B).

Prediction of corneal permeability, melanin binding and irritation

We experimented with different ways to represent molecules and compared their influence on the model performance. The best results across all datasets were achieved using gradient boosting models. Specifically, the XGBRegression model attained an R^2 score of 0.67 for predicting corneal permeability, while the XGBClassifier achieved an F1-score of 0.86 for melanin binding prediction. The LGBM model excelled with an F1-score of 0.95 for predicting eye irritation. The relatively low performance in predicting corneal permeability can be attributed to the limited training data, as many studies report small sample sizes, such as the 32 compounds analyzed by Ramsay et al. (Ramsay et al. 2018). Despite this, our extensive data collection on drug permeability through rabbit corneas allows us to consider these performance metrics as state-of-the-art. To interpret the machine learning models, we conducted a feature importance analysis using MACCS fingerprints, which provide a binary representation of molecular substructures (Yang et al. 2022). The RDKit library was employed to link bit numbers to their corresponding substructures in SMARTS format. The analysis of most important features revealed that hydrophobic groups, such as alkyl chains and aromatic rings, are crucial for estimating corneal permeability due to the lipophilic nature of the cornea. For melanin binding predictions, aromatic rings significantly impact model performance, likely due to their ability to interact with melanin through π-π stacking. In predicting eye irritation, 6-membered rings and nitrogen-carbon bond fragments were identified as the most important features.

KAN for more reliable melanin binding prediction

Tree-based gradient boosting models are recognized as leading methods for tabular data. Despite achieving high performance in predicting melanin binding, we noted significant overfitting, with accuracy differing notably between training and test sets. To address this, we explored the potential

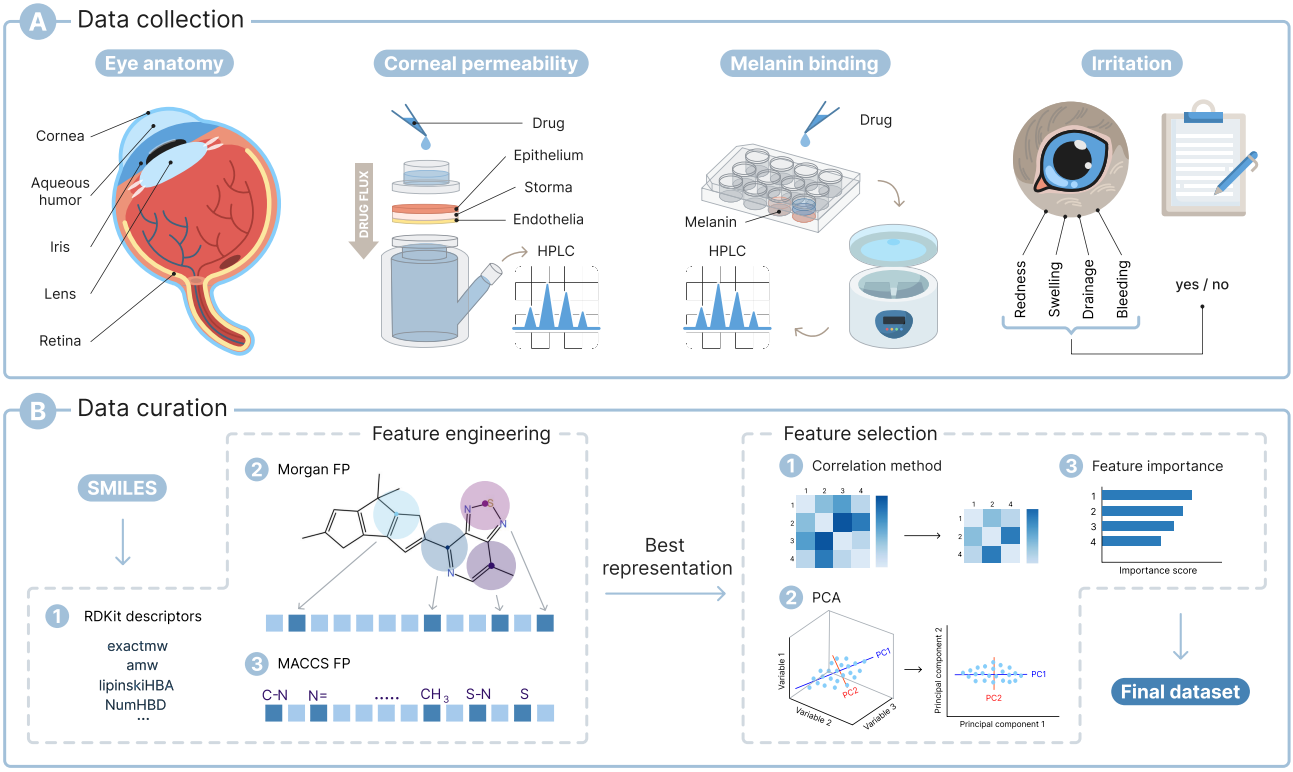

Figure 2: Data collection and curation. A) Schematic representation of the measurement of properties required of eye drops. B) Data preprocessing. Feature engineering and dimensionality reduction.

of Kolmogorov-Arnold Networks (KANs) (Liu et al. 2024) alongside traditional models like gradient boosting (GB) and multilayer perceptrons (MLP).

Table 1: Comparison of train and test accuracy between GB, MLP and KAN models.

Model	Train accuracy	Test accuracy	Δ
GB	0.937 ± 0.010	0.799 ± 0.028	0.138
MLP	0.994 ± 0.010	0.781 ± 0.029	0.213
KAN	0.786 ± 0.005	0.759 ± 0.025	0.027

Our experiments showed that KAN achieved accuracies of 0.76 on the test set and 0.78 on the training set. While KAN's test accuracy was 2-4% lower than that of GB and MLP, it exhibited a significantly reduced overfitting, with the delta (the absolute difference between training and test accuracy) being five times lower for KAN compared to GB and even more so against MLP (Table 1). These findings indicate that KAN may serve as a more reliable predictive model in drug discovery, where managing false positives and negatives is critical for the safety and efficacy of new medications.

Case study: predicted inhibitors of COX-2

To demonstrate the effectiveness of our pipeline for *de novo* synthesis of active components in eye drops, we generated 6000 molecules with high inhibitory capacity against the COX-2 protein, which plays a key role in synthesizing prostanoids involved in inflammation and pain relief (Zarghi and Arfaei 2011). We also utilized FREED++ as base model to assess the impact of our modified reward function. In Table 2, we compare the predicted properties of the generated molecules, including average corneal permeability (logarithm of permeability through rabbit cornea), melanin binding probability, and irritation (toxicity). We also included docking scores (DS), estimated using QVina02 software, to characterize ligand-protein interactions (Alhossary et al. 2015). While melanin binding and docking scores showed only moderate changes between the base and modified models, we observed significant improvements in corneal permeability and irritation. Specifically, the logarithm of corneal permeability increased by 1.07, indicating a tenfold improvement in average permeability rate. Additionally, the probability of irritation decreased by over 40%. These findings underscore the advantages of our reward function and validate the effectiveness of our approach.

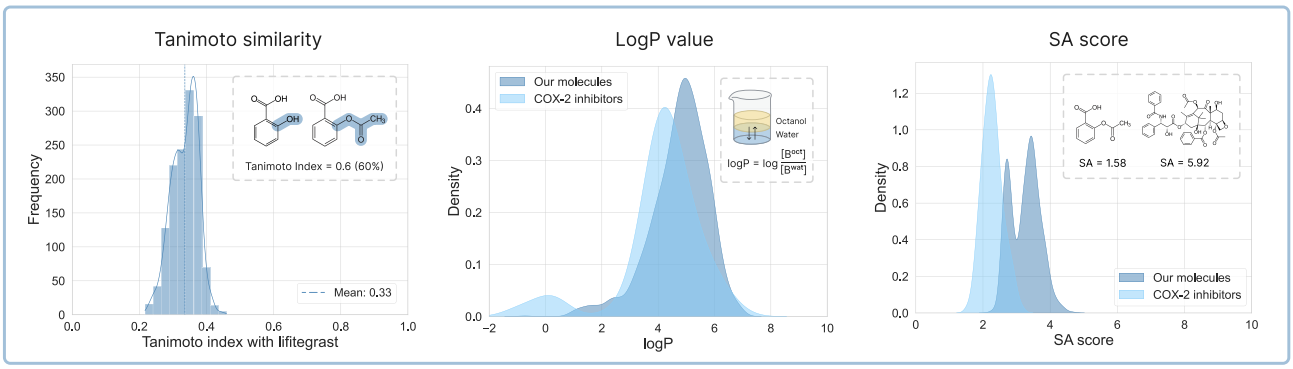

Figure 3: Validation of pipeline. Tanimoto index with lifitegrast. The Tanimoto index (or Tanimoto coefficient) is a measure of similarity between two sets, commonly used in cheminformatics to compare the similarity of chemical compounds. LogP and SA score for our molecules and for known inhibitors of COX-2. LogP, or the logarithm of the partition coefficient (P), is a measure of a compound's hydrophobicity (lipophilicity). The SA score, or Synthetic Accessibility score, is a metric used in cheminformatics to evaluate how easy or difficult it is to synthesize a particular chemical compound.

Table 2: Property comparison of molecules generated by FREED++ and our model. Arrows in brackets indicate better values.

Property	Base model	Our model	Δ
CP (↑)	3.44 ± 0.19	4.51 ± 0.11	+1.07 (+31.10%)
MB (↑)	0.75 ± 0.01	0.76 ± 0.02	+0.01 (+1.01%)
EI (↓)	0.55 ± 0.06	0.33 ± 0.09	-0.22 (-40.00%)
DS (↓)	-7.07 ± 0.16	-6.49 ± 0.25	+0.58 (+8.20%)

Validation

The validation process employed key metrics to assess the drug-like properties of the generated molecules (Fig. 3). We calculated Tanimoto similarity scores to compare our compounds with the known ophthalmic anti-inflammatory agent lifitegrast, yielding an average index of 0.33, indicating minimal shared features. The LogP value for our compounds was 4.05, compared to 4.25 for known COX-2 inhibitors, suggesting they are relatively lipophilic and may penetrate the cornea effectively. Additionally, the Synthetic Accessibility (SA) score of 3.66 indicates that these compounds are reasonably accessible for laboratory synthesis. In summary, the moderate Tanimoto similarity, high LogP values, and satisfactory SA score suggest promising drug-like properties, warranting further investigation as potential ophthalmic anti-inflammatory agents.

Limitations

The results presented above are promising for the practical application of our pipeline. However, experimental validation, including the synthesis of generated molecules, as well as *in vivo* and *in vitro* tests, is necessary to confirm its effectiveness. Based on our results, we can highlight the following limitations:

1. Due to the small amount of data used to train predictive models, the diversity of generated molecules is limited.
2. As described earlier, FREED++ is a fragment-based model, and a starting fragment is necessary in the training phase to achieve the target properties of the generated molecules.

Conclusions

We developed a novel pipeline for designing eye drops with specific properties such as corneal permeability, melanin binding, and eye irritation. This involved creating three datasets from scientific articles, which were processed to include RDKit descriptors, MACCS and Morgan fingerprints. We trained classical machine learning models based on gradient boosting to predict the desired properties, achieving state-of-the-art performance with R2 = 0.67, F1-scores of 0.86 and 0.95 for irritation and melanin binding, respectively. Utilizing a dataset with melanin, we applied the new KAN architecture, significantly reducing overfitting compared to classical models. Our models were integrated into the FREED++ framework for targeted molecule generation, demonstrating improved properties over standard FREED++ outputs. We also assessed lipophilicity, synthetic accessibility, and Tanimoto index of the generated molecules against known COX-2 inhibitors. Our findings indicate that this pipeline effectively predicts and optimizes essential eye drop formulation properties, enhancing the potential for developing therapeutic agents. Our research marks just the beginning of optimizing ocular drug development. We also plan to compare our results with other RL architectures, such as REINVENT and MOLDQN, in the near future

References

Agatonovic-Kustrin, S.; Evans, A.; and Alany, R. G. 2003. Prediction of corneal permeability using artificial neural networks. *Die Pharmazie*, 58(10): 725–729.

Alhossary, A.; Handoko, S. D.; Mu, Y.; and Kwoh, C.-K. 2015. Fast, accurate, and reliable molecular docking with QuickVina 2. *Bioinformatics*, 31(13): 2214–2216.

Dai, H.; Tian, Y.; Dai, B.; Skiena, S.; and Song, L. 2018. Syntax-Directed Variational Autoencoder for Structured Data. Version Number: 1.

Danel, T.; Łeski, J.; Podlewska, S.; and Podolak, I. T. 2023. Docking-based generative approaches in the search for new drug candidates. *Drug Discovery Today*, 28(2): 103439.

Dargó, G.; Vincze, A.; Müller, J.; Kiss, H. J.; Nagy, Z. Z.; and Balogh, G. T. 2019. Corneal-PAMPA: A novel, non-cell-based assay for prediction of corneal drug permeability. *European Journal of Pharmaceutical Sciences*, 128: 232–239.

De Cao, N.; and Kipf, T. 2018. MolGAN: An implicit generative model for small molecular graphs. Publisher: arXiv Version Number: 2.

Gause, S.; Hsu, K.-H.; Shafor, C.; Dixon, P.; Powell, K. C.; and Chauhan, A. 2016. Mechanistic modeling of ophthalmic drug delivery to the anterior chamber by eye drops and contact lenses. *Advances in Colloid and Interface Science*, 233: 139–154.

Ghorbanzad'e, M.; Fatemi, M. H.; Karimpour, M.; and Andersson, P. L. 2011. Quantitative and qualitative prediction of corneal permeability for drug-like compounds. *Talanta*, 85(5): 2686–2694.

Gómez-Bombarelli, R.; Wei, J. N.; Duvenaud, D.; Hernández-Lobato, J. M.; Sánchez-Lengeling, B.; Sheberla, D.; Aguilera-Iparraguirre, J.; Hirzel, T. D.; Adams, R. P.; and Aspuru-Guzik, A. 2018. Automatic Chemical Design Using a Data-Driven Continuous Representation of Molecules. *ACS Central Science*, 4(2): 268–276.

Jakubiak, P.; Reutlinger, M.; Mattei, P.; Schuler, F.; Urtti, A.; and Alvarez-Sánchez, R. 2018. Understanding Molecular Drivers of Melanin Binding To Support Rational Design of Small Molecule Ophthalmic Drugs. *Journal of Medicinal Chemistry*, 61(22): 10106–10115.

Kidron, H.; Vellonen, K.-S.; Del Amo, E. M.; Tissari, A.; and Urtti, A. 2010. Prediction of the Corneal Permeability of Drug-Like Compounds. *Pharmaceutical Research*, 27(7): 1398–1407.

Liu, Z.; Wang, Y.; Vaidya, S.; Ruehle, F.; Halverson, J.; Soljačić, M.; Hou, T. Y.; and Tegmark, M. 2024. KAN: Kolmogorov-Arnold Networks. ArXiv:2404.19756 [cond-mat, stat].

Lowrey, A. H.; Fameini, G. R.; Loumbev, V.; Wilson, L. Y.; and Tosk, J. M. 1997. Modeling Drug-Melanin Interaction With Theoretical Linear Solvation Energy Relationships. *Pigment Cell Research*, 10(5): 251–256.

Popova, M.; Isayev, O.; and Tropsha, A. 2018. Deep reinforcement learning for de novo drug design. *Science Advances*, 4(7): eaap7885.

Radwa, A.; Frackowiak, T.; Ibrahim, H.; Aubry, A.; and Kaliszan, R. 1995. Chromatographic modelling of interactions between melanin and phenothiazine and dibenzazepine drugs. *Biomedical Chromatography*, 9(5): 233–237.

Ramsay, E.; Del Amo, E. M.; Toropainen, E.; Tengvall-Unadike, U.; Ranta, V.-P.; Urtti, A.; and Ruponen, M. 2018. Corneal and conjunctival drug permeability: Systematic comparison and pharmacokinetic impact in the eye. *European Journal of Pharmaceutical Sciences*, 119: 83–89.

Reilly, J.; Williams, S. L.; Forster, C. J.; Kansara, V.; End, P.; and Serrano-Wu, M. H. 2015. High-Throughput Melanin-Binding Affinity and In Silico Methods to Aid in the Prediction of Drug Exposure in Ocular Tissue. *Journal of Pharmaceutical Sciences*, 104(12): 3997–4001.

Telepov, A.; Tsypin, A.; Khrabrov, K.; Yakukhnov, S.; Strashnov, P.; Zhilyaev, P.; Rumiantsev, E.; Ezhov, D.; Avetisian, M.; Popova, O.; and Kadurin, A. 2024. FREED++: Improving RL Agents for Fragment-Based Molecule Generation by Thorough Reproduction. Version Number: 1.

Wang, Q.; Li, X.; Yang, H.; Cai, Y.; Wang, Y.; Wang, Z.; Li, W.; Tang, Y.; and Liu, G. 2017. In silico prediction of serious eye irritation or corrosion potential of chemicals. *RSC Advances*, 7(11): 6697–6703.

Yang, J.; Cai, Y.; Zhao, K.; Xie, H.; and Chen, X. 2022. Concepts and applications of chemical fingerprint for hit and lead screening. *Drug Discovery Today*, 27(11): 103356.

Zarghi, A.; and Arfaei, S. 2011. Selective COX-2 Inhibitors: A Review of Their Structure-Activity Relationships. *Iranian journal of pharmaceutical research: IJPR*, 10(4): 655–683.

Zhou, Z.; Kearnes, S.; Li, L.; Zare, R. N.; and Riley, P. 2019. Optimization of Molecules via Deep Reinforcement Learning. *Scientific Reports*, 9(1): 10752.

Śledź, P.; and Caflisch, A. 2018. Protein structure-based drug design: from docking to molecular dynamics. *Current Opinion in Structural Biology*, 48: 93–102.