In Silico Screening of Zinc (II) Enzyme Inhibitors Using ILP

Tadasuke Ito,[1] Shotaro Togami,[2] Shin Aoki,[2] and Hayato Ohwada[1]

1. Faculty of Science and Technology, Tokyo University of Science, Chiba, Japan
7414601@ed.tus.ac.jp, ohwada@rs.tus.ac.jp
2. Faculty of Pharmaceutical Sciences, Tokyo University of Science, Chiba, Japan
shinaoki@rs.tus.ac.jp

Abstract. In silico screening is a powerful drug-discovery tool. However, the application of traditional structure-based and mechanism-based drug design is hampered by the limited availability of three-dimensional structures of target enzymes or proteins. Thus, we propose a new method of screening good inhibitors of target enzymes without using their precise structures, based on machine learning. With this method, the data of ligands and decoys are collected from the inhibitor's Database of Useful Decoys: Enhanced (DUD-E). We evaluated the accuracy of Inductive-Logic Programing (ILP) by applying a classification model that learned ligands and decoys from DUD-E to ligand candidates that are not included in DUD-E. In the present study, this technique is applied to the screening of inhibitors of carbonic anhydrase. ILP exhibited high classification performance. Furthermore, we visualized the rules from ILP and obtained a clear classification model.

Keywords: Carbonic anhydrase, In silico screening, Machine learning

1. INTRODUCTION

In-silico drug-screening discovery is a powerful, low-cost method of finding strong binders for proteins and enzymes from a large number of compounds [1, 2]. For example, Structure-Based Virtual Screening (SBVS) is a method of preparing three-dimensional data of given enzymes and receptors, and calculating the fitting of their binding sites with ligand molecules [3]. Although the structures of many enzymes have been disclosed by X-ray crystal structure analysis, enzymes whose structures are unknown cannot be applied to SBVS. In addition, calculation of the force field of ligands and enzymes in SBVS is not always correct [4]. It is especially difficult to estimate the affinity of ligands with metalloenzymes, which have metal cations in their active center [5]. Although the use of quantum mechanics calculation of metal-ligand coordination bonds has been reported, it is not considered suitable for screening many ligand candidates, due to the length of time required [6, 7].

An alternative way to screen good ligands is Ligand-Based Virtual Screening (LBVS) [8, 9], which predicts ligand candidates by superimposing their structures with those of known ligands [10-12]. The present study focuses on ligand screening using machine learning. We decided to conduct Inductive-Logic Programming (ILP), which is a machine-learning method using the dataset of both of good ligands and decoys obtained from the Database of Useful Decoys: Enhanced (DUD-E), which includes information on good and poor ligands for various enzymes and is now open to the public ("decoys" indicate randomly collected compounds and/or ligands that bind to other enzymes, which may lower the reliability of the calculation output) [13]. A discrimination

model obtained using machine learning to determine the attribute value is a black box used by drug-discovery researchers. However, ILP can visually express the common rule of ligands, because ligand discrimination rules are represented by a set of logical expressions. Obtaining a clear classification model is important for drug-discovery researchers.

In this study, machine learning is applied to screen inhibitors of carbonic anhydrase (CA), which is a zinc (II) enzyme that catalyzes the reversible conversion between carbon dioxide and bicarbonate [14-17]. In addition, the CA mechanism has been extensively studied, and several CA inhibitors (e.g., acetazolamide and dorzolamide) have been clinically used to treat epilepsy and glaucoma [18]. CAH2, one of the many subtypes of CA in the human body, has been extensively studied, and most CA ligands listed in DUD-E are CAH2 ligands, possibly because CAH2 has been recognized as an important target.

We evaluated the accuracy of ILP by applying the classification model that learned ligands and decoys from DUD-E to the ligand candidates that are not included in DUD-E. This study was conducted to determine the validity of the prediction of ligands by machine learning using DUD-E.

2. MATERIALS AND METHODS

Figure 1 illustrates the process of this method, which includes three parts. First, we extract the structure of a chemical compound from DUD-E. Second, we produce a model that discriminates between the ligand and the decoy, using the structure data. Finally, we predict whether an inhibitor candidate is a ligand or a decoy using the classification model.

2.1 Dataset from DUD-E

DUD-E contains data on ligands that are extracted from the biological activity database site ChEMBL and that have Ki values of <1μM [19]. Decoys are defined as compounds that have physical parameters similar to those of ligands but have different structures and lower affinity, and these data were obtained from the ZINC database [20]. The file "actives_final.mol2" contains information on the three-dimensional structures of good ligands, and the file "decoys_final.mol2" contains information on those of decoys. Specific numbers were assigned to each ligand according to ChEMBL and to each decoy according to ZINC. The number of CA compounds is indica-

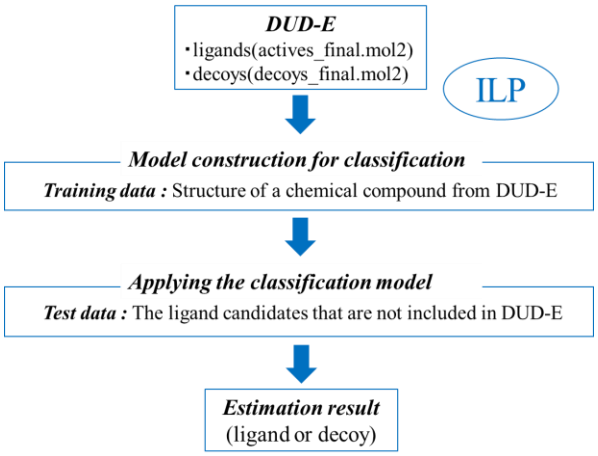


Fig. 1. Overview of the proposed method

Table 1. Number of CA inhibitors

	Ligand	Decoy
Total	835	31710
Total without almost identical compounds	492	31133
The number of the compounds used for the machine learning	492	3000

ted in Table 1. The ligands have been clustered by ChEMBL ID, as well as the decoys. In each cluster, the ligands have the same structure but different charges. Because almost all the properties of such ligands are the same, careful treatment is needed to use the cluster in performance evaluation. Therefore, we performed our experiment using this cluster. In addition, we reduced the number of decoys in order to avoid over-fitting.

2.2 Inductive-Logic Programming

Inductive-Logic Programming (ILP) is a framework for performing inductive logic according to logic programming [21]. ILP is used to learn the relation expressions that a set of attribute values cannot represent. The greatest feature of ILP is its ability to learn background knowledge that is written in first-order predicate logic. For this reason, ILP can learn complex rules.

Background knowledge. Mol2 files describe atoms included in the compound and interatomic bond strength. It is necessary to convert the input format of ILP because it cannot learn in this state. For example, a sentence defines the structure of an atom, such as a bond (C1,a1,a2,2). This sentence indicates that compound C1 contains a bond between atoms a1 and a2; the fourth argument specifies that this bond is a double bond. Background knowledge (input data) is composed of a set of such logical expressions.

Our method uses the following statements, known as clauses:

- bond(compound, atomid, atomid, bondtype)
- atom(compound, atomid, atomtype)
- ring(compound, ringid, atomid, ringsize).

Bond() is a description of the bonding between atoms. Atom() is a description of types of atoms. Ring() is a description of the atoms contained in a cyclic compound. ILP creates a rule (hypothesis) that applies to only ligands using a combination of these clauses (e.g., bond(A, B, C, 2), atom(A, B, cl), and ring(A, D, B, 6)). These rules indicate that compound A has atoms B and C that are connected by a double bond. In addition, B is a chlorine atom and is included in hexagonal cyclic compound D.

GKS. In this study, we use GKS as our ILP system software [22]. GKS takes input parameter values for depth, positive, negative, and clause_size. Depth is the allowable depth of a variable (e.g., A, B, and C). Positive is the minimum number of ligands covered by a rule. Negative is the maximum number of decoys covered by a rule. Clause_size is the maximum number of clauses in the rule.

3. EXPERIMENT

3.1 Experiment Parameters

GKS takes input parameter values for depth, positive, negative, and clause_size. In this study, we set the following values: depth = 10, negative = 10, positive = 10, and clause_size = 6.

3.2 Evaluation experiment

We evaluate the accuracy of machine-learning methods by applying the classification model that learned ligands and decoys from DUD-E to ligand candidates that are not included in DUD-E. We prepare 22 inhibitor candidates that are not included in DUD-E. Most of these compounds are obtained from the literature, or are designed and synthesized.

The measures for evaluating a classification are Accuracy, Precision, Recall, and F1 score. Accuracy is the rate of correct classification for all compounds. Precision is the rate of ligands in the compounds that have been classified as ligands. Recall is the rate of ligands that have been correctly classified as ligands. The F1 score is the harmonic mean of Precision and Recall. It is difficult to detect ligands in this study because the number of decoys in the training data (DUD-E) is very large. Moreover, the screening method should select more ligands than decoys. For these reasons, when Precision and Recall are high and the F1 score is high, screening performance is high.

3.3 Results and Discussion

Results. Table 2 lists the hypotheses that were derived by ILP using the training data from DUD-E. The values given as scores are the covering numbers of the rules in the training data or test data (training data: ligands = 492, decoys = 3000 | test data : ligands = 14, decoys = 8); p is the number of positives (ligands), and n is that of negatives (decoys) covered by the rule. ILP determined the ligand by applying these rules to the test data.

The classification results in Table 3 indicate that ILP can classify ligands and decoys with excellent accuracy. The F1 score of ILP was almost 0.9; therefore, the discrimination of ILP was very high. High F1 values suggest that correlation properties of most ligands are correctly predicted. The dataset from DUD-E is imbalanced, and the number of decoys is very high; precision is high and the number of incorrectly predicted compounds is small. Therefore, it is possible to

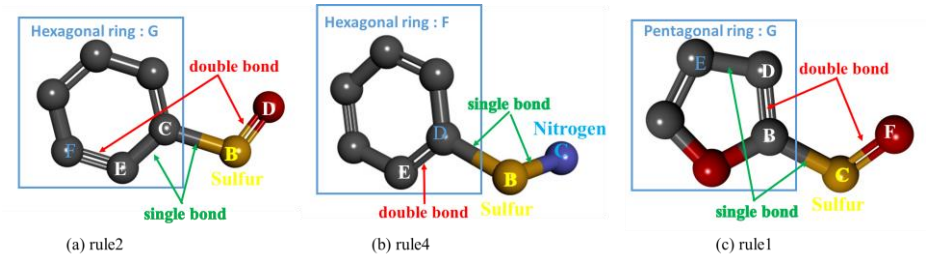
Table 2. Hypotheses derived by ILP

Number	Rule	Score(trainning data)		Score(test data)	
		p	n	p	n
1	dock(A) :- bond(A, B, C, 1), atom(A, C, s), bond(A, D, B, 2), bond(A, E, D, 1), bond(A, C, F, 2), ring(A, G, E, 5)	118	7	1	0
2	dock(A) :- atom(A, B, s), bond(A, C, B, 1), bond(A, B, D, 2), bond(A, C, E, 1), bond(A, E, F, 2), ring(A, G, F, 6)	125	8	12	2
3	dock(A) :- atom(A, B, n), atom(A, C, s), bond(A, D, B, 2), bond(A, E, D, 1), bond(A, E, F, 2), bond(A, C, G, 2)	14	5	3	0
4	dock(A) :- bond(A, B, C, 1), atom(A, B, s), atom(A, C, n), bond(A, D, B, 1), bond(A, D, E, 2), ring(A, F, D, 6)	191	10	1	0
5	dock(A) :- atom(A, B, o), atom(A, C, s), bond(A, B, D, 1), bond(A, C, E, 2), ring(A, F, D, 5)	21	6	1	0
6	dock(A) :- bond(A, B, C, 2), atom(A, B, s), bond(A, D, B, 1), bond(A, E, D, 1), bond(A, F, E, 1), bond(A, F, G, 2)	22	3	1	0
7	dock(A) :- atom(A, B, o), atom(A, C, s), bond(A, B, C, 1)	25	6	1	0
8	dock(A) :- atom(A, B, s), atom(A, C, o), bond(A, D, C, 1), bond(A, E, D, 1), bond(A, B, F, 2), bond(A, E, G, 2)	36	8	2	0
9	dock(A) :- bond(A, B, C, 1), atom(A, C, s), bond(A, D, B, 2), bond(A, D, E, 1), bond(A, C, F, 2), bond(A, E, G, 2)	41	6	0	0
10	dock(A) :- bond(A, B, C, 1), atom(A, B, s), atom(A, C, n), bond(A, D, B, 1), bond(A, E, D, 1), ring(A, F, E, 5)	50	6	0	0
11	dock(A) :- atom(A, B, s), atom(A, C, f), bond(A, D, C, 1), bond(A, E, D, 1), bond(A, B, F, 2)	58	9	0	0

Table 3. Classification result obtained by ILP

Method	tp	fn	tn	fp	Accuracy	Recall	Precision	F1
ILP	13	1	6	2	0.864	0.929	0.867	0.897

tp(true positive) is actual ligands that were correctly classified as ligands.
fn(false negative) is ligands that were incorrectly marked as decoys.
tn(true negative) is all the remaining compounds, correctly classified as decoys.
fp(false positive) is decoys that were incorrectly labeled as ligands.



(a) rule2 (b) rule4 (c) rule1

Fig. 2. Visualization of the rules

screen ligands from a large number of compounds including decoys. Thus, our machine-learning method is useful.

Insight from rules. Positive (ligand) and negative (decoy) examples in the test data covered by each rule are presented in Table 3. Focusing on positive examples covered by the rule, most ligands are detected by rules 2 and 4. These rules have excellent discrimination ability, as they get great ligand coverage in the training data. One advantage of learning the structure in ILP is visualization of the classification rules. The classification model of machine-learning methods is a black box. Figure 2(a), depicts rule 2, and Fig. 2(b) depicts rule 4. These are actually observed forms in the test data.

Figure 2(c) illustrates rule 1. Rule 1 is very similar to rule 2. This rule has high ligand coverage in the training data. However, few ligands are covered by rule 1 in the test data because of the small number of pentagonal rings in the test data. It may be possible to detect more ligands with more test data. Rules 5 and 10 are also similar.

Other rules express just the bond without information on the ring. In addition, these rules classified ligands that had been determined by rules with the ring description. We do not regard these rules as important.

The ligand of CAH2 is affected by the sulfur atom and cyclic compound. It is possible to design the compound as depicted in Fig. 2. Obtaining a clear classification model and new knowledge is important for drug discovery researchers.

4. CONCLUSION

In summary, we have reported on the prediction method for ligand screening using machine learning. Our method uses both ligands and decoys; therefore, it can predict ligand high performance. The dataset from DUD-E has higher reproducibility and reliability than the traditional dataset. It is possible to screen many inhibitor candidates using our method. In addition, machine learning can predict proteins and enzymes without their three-dimensional structures. Therefore, machine learning can be applied to angiotensin-converting enzyme, histone deacetylase, metallo-B-lactamase, and other zinc enzymes. It is possible to classify many new inhibitor candidates and

to detect new ligands of enzymes with high performance by collecting ligands and decoys same as DUD-E. In addition, ILP provides a clear classification model and new knowledge for drug-discovery researchers.

REFERENCES

1. S. Kortagere, In Silico Models for Drug Discovery (Methods in Molecular Biology), Humana Press, 2013.
2. H. Gohike, Protein-Ligand Interactions (Methods and Principles in Medicinal Chemistry), Wiley-VCH, 2012.
3. Lyne, P. D., Drug Discovery Today, 2002, 7, 1047-1055.
4. Kosuge, T.; Mashima, J.; Kodama, Y.; Fujisawa, T.; Kaminuma, E.; Ogasawara, O.; Okubo, K.; Takagi, T.; Nakamura, Y., Nucleic Acids Res., 2014, 42, 44-49.
5. Hanaya, K.; Suetsugu, M.; Saijo, S.; Yamato, I.; Aoki, S., J. Biol. Inorg. Chem., 2012, 17,517-529.
6. Khandelwal, A.; Lukacova, V.; Comez, D.; Kroll, D. M.; Raha, S.; Balaz, S., J. Med. Chem., 2005, 48, 5437-5447.
7. Taufer, M.; Crowley, M.; Price, D.; Chien, A. A.; Brooks III, C. L., Concurrency and Computation: Practice & Experience, 2004, 17, 1627 – 1641.
8. Koeppen, H.; Kriegl, J.; Lessel, U.; Tautermann, C. S.; Wellenzohn, B., Wiley-VCH, 2011, 48.
9. Aparoy, P.; Reddy, K. K.; Reddanna, P., Current Medicinal Chemistry, 2012, 19, 3763-3778.
10. Schneider, G.; Neidhart, W.; Giller, T.; Schmid, G., Angew. Chem., Int. Ed., 1999, 38, 2894-2896.
11. Renner, S.; Noeske, T.; Parsons, C. G.; Schneider, P.; Weil, T.; Schneider, G., Chem-BioChem, 2005, 6, 620-625.
12. Schuffenhauer, A.; Floersheim, P.; Acklin, P.; Jacoby, E., J. Chem. Inf. Comput. Sci., 2003, 43, 391-405.
13. Mysinger, M. M.; Carchia, M.; Irwin, J. J.; Shoichet B. K., J. Med. Chem., 2012, 55, 6582-6594.
14. Supuran, C. T.; Scozzafava, A.; Casini, A., Medicinal Research Reviews, 2003, 23, 146-189.
15. Supuran, C. T.; Scozzafava, A., Bioorg Med Chem., 2007, 15, 4336-4350.
16. C. T. Supuran, Carbonic anhydrases: novel therapeutic applications for inhibitors and activators, Nat Rev Drug Discov., 2008, Vol. 7, Issue 2, pp. 168-181.
17. Aoki, S.; Kimura, E., Comprehensive Coordination Chemistry II, 8, 2003, 601-640.
18. Weston, J., Chem Rev. 2005, 105, 2151-2174
19. Gaulton, A.; Bellis, L. J.; Bento, A. P.; Chambers, J.; Davies, M.; Hersey, A.; Light, Y.; McGlinchey, S.; Michalovich, D.; Al-Lazikani, B.; Overington J. P., Nucleic Acids Research. 2012, 40, 1100-1107.
20. Irwin J. J., Shoichet B. K., ZINC-a free database of commercially available compounds for virtual screening, J. Chem. Inf. Model., Vol. 45, pp. 177-182, 2005.
21. Muggleton S., Inverse entailment and progol, NewGeneration Computing, Vol.13, pp 245-286, 1995.
22. Mizoguchi,F., Owada,H., Constrained Relative Least General Generalization for Inducing Constraint Logic Programs, New Generation Computing, Vol.13, pp.335-368, 1995.