ACS OMEGA

http://pubs.acs.org/journal/acsodf

Article

Predictive Modeling of HMG-CoA Reductase Inhibitory Activity and Design of New HMG-CoA Reductase Inhibitors

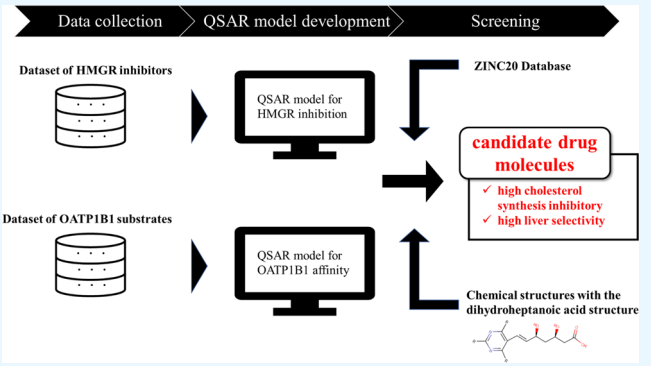
Shigeyoshi Samizo and Hiromasa Kaneko*

Cite This: *ACS Omega* 2023, 8, 27247−27255

Read Online

ACCESS | Metrics & More | Article Recommendations

ABSTRACT: As blood cholesterol increases, it accumulates in the intima of blood vessels, elevating the risk of atherosclerosis and coronary artery disease. Drugs that inhibit enzymes essential for cholesterol synthesis are effective in improving blood cholesterol levels. Statins are used to treat hypercholesterolemia as they inhibit 3-hydroxyl-3-methylglutaryl coenzyme A (HMG-CoA) reductase (HMGR), the rate-limiting enzyme in cholesterol synthesis. Statins are known to exert their effects by translocating to the liver, where they are taken up by the organic anion transporting polypeptide 1B1 (OATP1B1). Therefore, we hypothesized that a compound with high HMGR inhibitory activity and high affinity for OATP1B1 would be an excellent new therapeutic agent for hypercholesterolemia with increased liver selectivity and fewer side effects. In this study, we developed two models for predicting HMGR inhibitory activity and OATP1B1 affinity to propose the chemical structure of a new therapeutic agent for hypercholesterolemia with both high inhibitory activity and high liver selectivity. HMGR inhibitory activity and OATP1B1 affinity prediction models were constructed with high prediction accuracy for the test data: r^2 = 0.772 and 0.768, respectively. New chemical structures were then input into these models to search for candidate compounds. We found compounds with higher HMGR inhibitory activity and OATP1B1 affinity than rosuvastatin, the most recently developed statin drug, and compounds that did not have a common structure of statins with high HMGR inhibitory activity.



1. INTRODUCTION

Cholesterol is essential for living organisms and is a component of cell membranes, bile acids, and steroid hormones. However, as blood cholesterol increases, cholesterol accumulates in the intima of blood vessels, increasing the risk of atherosclerosis and coronary artery disease.[1] Drugs that inhibit enzymes essential for cholesterol synthesis can be used to improve blood cholesterol levels.[2] Statins, which are inhibitors of 3-hydroxyl-3-methylglutaryl coenzyme A (HMG-CoA) reductase (HMGR), are used to treat hypercholesterolemia because they inhibit HMGR or the rate-limiting enzyme in cholesterol synthesis.[2]

Although statins are generally considered safe drugs, serious side effects exist, including skeletal muscle disorders such as rhabdomyolysis, myopathy, and renal impairment;[3] thus, statins are required to have fewer side effects. The cause of these side effects is believed to be the action of statins on tissues other than the liver. To mitigate these side effects, a strategy is being implemented to minimize drug distribution to tissues other than the liver.[4] Drug uptake into the liver is known to involve transporters, and many statins are known to be taken up by organic anion transporting polypeptide 1B1 (OATP1B1).[4,5] Therefore, compounds with high HMGR inhibitory activity and high affinity for OATP1B1 would be an excellent new treatment for hypercholesterolemia with increased selectivity for the liver and fewer side effects.

As the development of a new drug generally takes 8−15 years and incurs huge costs, many studies have been conducted to improve its efficiency.[6] In silico drug discovery is attracting attention as a way to dramatically increase the efficiency of drug development.[7] Quantitative structure−activity relationship (QSAR)[8] model, constructed between the activity and molecular descriptors of chemical structures with compounds using machine learning, can predict the activity of new chemical structures.

The purpose of this study was to propose the chemical structures of a new treatment for hypercholesterolemia that has both high HMGR inhibitory activity and liver selectivity by constructing QSAR models to predict HMGR inhibitory activity and OATP1B1 affinity for efficient drug development.

Received: April 14, 2023
Accepted: June 30, 2023
Published: July 18, 2023

https://doi.org/10.1021/acsomega.3c02567
ACS Omega 2023, 8, 27247−27255

High-activity chemical structures were obtained using both models, and data from publicly available databases and published experimental data were used to construct the two QSAR models. The molecular descriptors calculated from the compound structures were used as explanatory variables to construct models for predicting the objective variables. The objective variable for the QSAR model predicting the HMGR inhibitory activity was pIC_{50}, which is the negative logarithm of the half-maximal inhibitory concentration (IC_{50}). For the QSAR model predicting the OATP1B1 affinity, pK_m, which is the negative logarithm of the Michaelis constant (K_m), was used as the objective variable. Using these two models for virtual screening, we searched for the chemical structure of a new treatment for hypercholesterolemia with high HMGR inhibition activity and liver selectivity.

2. METHOD

2.1. Data Collection. An HMGR inhibitor dataset was downloaded from the ChemBL database (ChEMBL3247). Compounds with IC_{50} data were selected, and their IC_{50} was converted to pIC_{50} using a negative log transformation. Duplicate samples within a dataset were averaged to obtain the pIC_{50} values. After processing, 833 compounds were retained in the samples. These compounds were descriptorized using the RDKit software.[9] These compounds were then split into training and test sets at a ratio of 70:30 and subjected to machine learning QSAR modeling.

The dataset for the OATP1B1 substrates was based on ChEMBL data (CHEMBL1697668) and additional data from the literature.[10] Compounds with Michaelis constants (Km) were selected; Km was transformed to pK_m using a negative log transformation. Duplicate samples in the dataset were averaged to obtain the pK_m values. After processing, 41 compounds were retained in the samples. These compounds were descriptorized using the RDKit software.

2.2. QSAR Model Development. The regression analysis methods used were ordinary least squares (OLS), partial least squares regression (PLS), ridge regression (RR), least absolute shrinkage and selection operator (LASSO), elastic net (EN), linear support vector regression (LSVR), nonlinear support vector regression (NLSVR), decision tree (DT), random forest (RF), Gaussian process regression (GPR), gradient boosting decision tree (GBDT), XGBoost (XGB), and the light gradient boosting model (LGBM). In the regression analysis, the hyperparameters were selected by maximizing the determination coefficients computed after a fivefold cross-validation procedure.

The dataset was divided into training and test data, and the predictive ability of models constructed with training data was validated with test data. The coefficient of determination r^2 (eq 1), root-mean-square error (RMSE) (eq 2), and mean absolute error (MAE) (eq 3) were used as indicators to evaluate the estimation performance of the model and are given as follows:

$$r^2 = 1 - \frac{\sum_{i=1}^{n} (y^{(i)} - y_{EST}^{(i)})^2}{\sum_{i=1}^{n} (y^{(i)} - y_{mean})^2} \quad (1)$$

$$MAE = \frac{\sum_{i=1}^{n} |y^{(i)} - y_{EST}^{(i)}|}{n} \quad (2)$$

$$RMSE = \sqrt{\frac{\sum_{i=1}^{n} (y^{(i)} - y_{EST}^{(i)})^2}{n}} \quad (3)$$

where r^2 is a measure of variability; if it is close to 1, the plots are clustered on the diagonal of the scatterplot. RMSE and MAE are measures of the overall error magnitude; the closer they are to zero, the smaller the error.

For the HMGR inhibition and OATP1B1 affinity prediction models, the model with the highest prediction accuracy was compared with and without feature selection using the Boruta algorithm. Boruta is a method for selecting important features according to the importance of variables calculated using RF. The percentile was set at 80.

A conformer-based 3D-QSAR (C3D-QSAR)[11] was used to construct an HMGR inhibition prediction model. C3D-QSAR is a method for generating conformers and constructing a QSAR model using representative descriptors (maximum, minimum, median, and average values) of each three-dimensional structure descriptor. Samples for inverse analysis by C3D-QSAR were subjected to inverse analysis by 2D-QSAR, and compounds predicted to have high inhibitory activity were used.

Descriptors were calculated using RDKit software for the 2D-QSAR model of HMGR inhibition. A total of 130 descriptors were selected by eliminating those for which the fractions of samples with the same values were 95% or higher. A regression model $[y = f(X)]$ was established to describe the relationship between the descriptors (X) and pIC_{50} values (y).

Mordred[12] was used to calculate the descriptors for C3D-QSAR analyses of HMGR inhibition. A regression model $[y = f(X)]$ was established to describe the relationship between the descriptors (X) and pIC_{50} values (y). The regression analysis method was the same as that used for the 2D-QSAR model for HMGR inhibition. The 180 descriptors were selected using the Boruta algorithm after eliminating the descriptors for which the fractions of samples with the same values were 95% or higher.

RDKit was used to calculate the descriptors for the 2D-QSAR analyses of OATP1B1 affinity. A regression model $[y = f(X)]$ was established to describe the relationship between the descriptors (X) and the pIC_{50} values (y). The prediction accuracies of different methods were compared by double cross-validation (DCV).[13] For the internal and external cross-validations (CVs), five-fold and leave-one-out CV techniques were used. The regression analysis method was the same as that used for the 2D-QSAR model for HMGR inhibition. A total of 141 descriptors were selected by eliminating those for which the fractions of samples with the same values were 95% or higher.

2.3. Inverse Analysis. Two datasets were used as compounds for inverse analysis. The first dataset used compounds from ZINC20.[14] In general, compounds with high water solubility are difficult to transfer into cells by passive diffusion. This reduces side effects because they are less likely to migrate to organs other than the liver. Therefore, in this study, we used compounds with a log P of 2.5, among the compounds included in ZINC20. This dataset is hereafter referred to as the ZINC20 dataset. For the second dataset, we used compounds that generated chemical structures with a dihydroheptanoic acid structure, a common statin structure. The generation of chemical structures is described in ref 15. This dataset is hereinafter referred to as the SGBR dataset. The

https://doi.org/10.1021/acsomega.3c02567
ACS Omega 2023, 8, 27247−27255

ZINC20 and SGBR datasets were inputted into the 2D-QSAR model for HMGR inhibition. Compounds with a predicted pIC_{50} greater than 8.0 were extracted as active compounds. The extracted compounds were used as samples for inverse analysis of the C3D-QSAR model for HMGR inhibition and the 2D-QSAR model for OATP1B1 affinity. As the newest statin currently on the market is rosuvastatin, the activity value of rosuvastatin was used as a reference. Since the pIC_{50} of rosuvastatin is 8.43 when calculated using the C3D-QSAR model, compounds with predicted values of around 8.43 were selected as compounds with high inhibitory activity. Similarly, the pK_m for rosuvastatin was calculated to be 1.71 using the 2D-QSAR model for OATP1B1 affinity, and thus, compounds with predicted values around 1.71 were selected as high-affinity compounds.

We established an applicability domain (AD)[16] for various prediction models because the performance of models is unreliable when predicting new samples in extrapolated regions. In this study, the distance calculated using the *k*-nearest-neighbor algorithm[17] was used as a measure of AD. When a new sample was obtained, the average of the Euclidean distances of the *k*-nearest neighbor samples from the new sample was calculated. If the average value was less than the threshold value, the sample was considered to be inside the AD. The threshold was set to 99.7% of the mean of the ascending training data. This threshold is derived from the 3-sigma rule,[18] which gives a probability of 0.997 for a sample inside an AD.

3. RESULTS AND DISCUSSION

3.1. Model Construction. *3.1.1. 2D-QSAR Model for HMGR Inhibition.* The statistics representing the estimation errors of the test data predicted by the different algorithms are listed in Table 1. As a result, LGBM was selected as the algorithm with the highest prediction accuracy because its r^2 was close to 1, and its MAE and RMSE were small. A plot of the predicted and measured values predicted by the LBGM based on the test data is shown in Figure 1. High prediction accuracy was confirmed by a good diagonal fit.

Table 1. Summary of the Statistical Parameters Is Derived from the 2D-QSAR Model for HMGR Inhibition in Test Data

	r^2	MAE	RMSE
OLS	−1.134	2.193	0.876
PLS	0.213	1.332	1.045
RR	0.583	0.97	0.775
LASSO	0.536	1.023	0.809
EN	0.582	0.971	0.772
LSVR	0.467	1.096	0.875
NLSVR	0.68	0.85	0.649
DT	0.186	1.355	1.097
RF	0.724	0.788	0.61
GPR	0.713	0.805	0.618
GBDT	0.705	0.815	0.643
XGB	0.717	0.798	0.613
LGBM	0.765	0.728	0.566

3.1.2. C3D-QSAR Model for HMGR Inhibition. The statistics representing the estimation error in predicting the test data using different algorithms are presented in Table 2. As a result, the algorithm with the highest prediction accuracy was selected because the r^2 of GRP is close to 1, and the MAE and RMSE are small. A plot of the predicted and measured values predicted by GPR on the test data is shown in Figure 2. High prediction accuracy was confirmed by a good diagonal fit (Table 2).

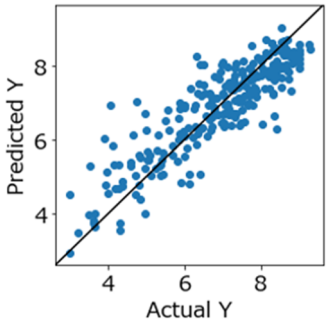


Figure 1. Scatterplot between the observed and the pIC_{50} is predicted with LGBM-based 2D-QSAR model for HMGR inhibition in test data.

Table 2. Summary of the Statistical Parameters Derived from the C3D-QSAR Model for HMGR Inhibition

	r^2	MAE	RMSE
OLS	0.548	1.009	0.711
PLS	0.617	0.929	0.71
RR	0.616	0.931	0.678
LASSO	0.584	0.968	0.733
EN	0.617	0.929	0.695
LSVR	0.404	1.159	0.901
NLSVR	0.766	0.726	0.535
DT	0.265	1.287	1.017
RF	0.736	0.772	0.596
GPR	0.772	0.716	0.542
GBDT	0.74	0.765	0.575
XGB	0.744	0.76	0.567
LGBM	0.763	0.73	0.544

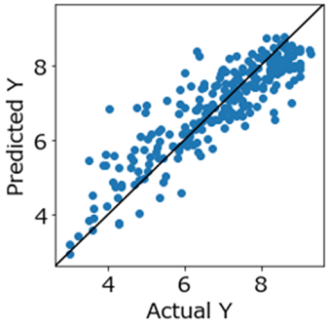


Figure 2. Scatterplot between the observed and the pIC_{50} is predicted with GPR-based C3D-QSAR model for HMGR inhibition in test data.

3.1.3. 2D-QSAR Model for OATP1B1 Affinity. The results of the comparison of the prediction accuracy of each method using the DCV are listed in Table 3. LASSO was selected as the algorithm with the highest prediction accuracy because its r^2 was close to 1, and its MAE and RMSE were small. A plot of the predicted and measured values predicted by LASSO for the test data is shown in Figure 3. High prediction accuracy was confirmed by a good diagonal fit.

https://doi.org/10.1021/acsomega.3c02567
ACS Omega 2023, 8, 27247−27255

Table 3. Summary of the Statistical Parameters Is Derived from 2D-QSAR Model for 2D-QSAR Model for OATP1B1 Affinity

	r^2	MAE	RMSE
OLS	−2.152	1.681	1.128
PLS	0.579	0.615	0.478
RR	0.674	0.541	0.429
LASSO	0.768	0.456	0.357
EN	0.669	0.545	0.436
LSVR	0.534	0.646	0.491
NLSVR	0.083	0.907	0.6
DT	0.369	0.752	0.545
RF	0.605	0.595	0.419
GPR	0.348	0.765	0.516
GBDT	0.188	0.853	0.582
XGB	0.327	0.777	0.598
LGBM	−2.152	1.681	1.128

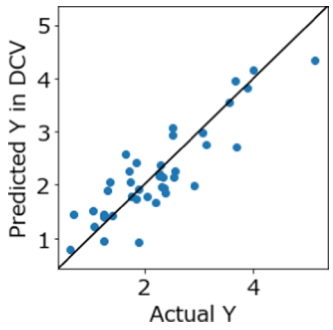


Figure 3. Scatterplot between the observed and the pK_m predicted with LASSO-based 2D-QSAR model for OATP1B1 affinity in test data.

3.2. Inverse Analysis. *3.2.1. 2D-QSAR Model for HMGR Inhibition.* The number of compounds with a predicted pIC_{50} over 8.0 was 65868 in the ZINC20 dataset and 1753 in the SGBR dataset. The chemical structures that were particularly active are shown in Figures 4 and 5. Numerous non-statin-like compounds were extracted. Various studies have reported the potential of non-statin-like compounds as HMGR inhibitors.[19]

3.2.2. C3D-QSAR Model for HMGR Inhibition. Compounds with a predictive value of around 8.43 were selected as highly active compounds. There were no compounds in the ZINC dataset with predicted values greater than 8.43. However, 37 compounds were predicted to be higher than the pIC_{50} of pravastatin (7.4), a first-generation statin drug. Some of these compounds, such as those shown in Figure 6, do not share a common structure with statins. As shown in Figure 7, there were two compounds in the SGBR dataset with predicted values greater than 8.43.

3.2.3. 2D-QSAR Model for OATP1B1 Affinity. Compounds with a predicted pK_m value exceeding 1.71 were selected as high-affinity compounds. Consequently, 42683 compounds in the ZINC20 dataset exceeded the predicted value of 1.71. Figure 8 shows the chemical structures of compounds with pK_m predictions exceeding 1.71 and pIC_{50} exceeding 7.4, as predicted in Section 3.2.2 C3D-QSAR Model for HMGR Inhibition. In the SGBR dataset, 911 compounds had predicted values exceeding 1.71. The active chemical structures are shown in Figure 9.

3.2.4. Chemical Structures of New HMGR Inhibitor Candidates. The chemical structure shown in Figure 10 exhibited higher predicted pIC_{50} and pK_m values than those of rosuvastatin.

When the HMGR inhibitory activity was predicted, the chemical structure with an isopropyl group had a higher predicted HMGR inhibitory activity than the chemical

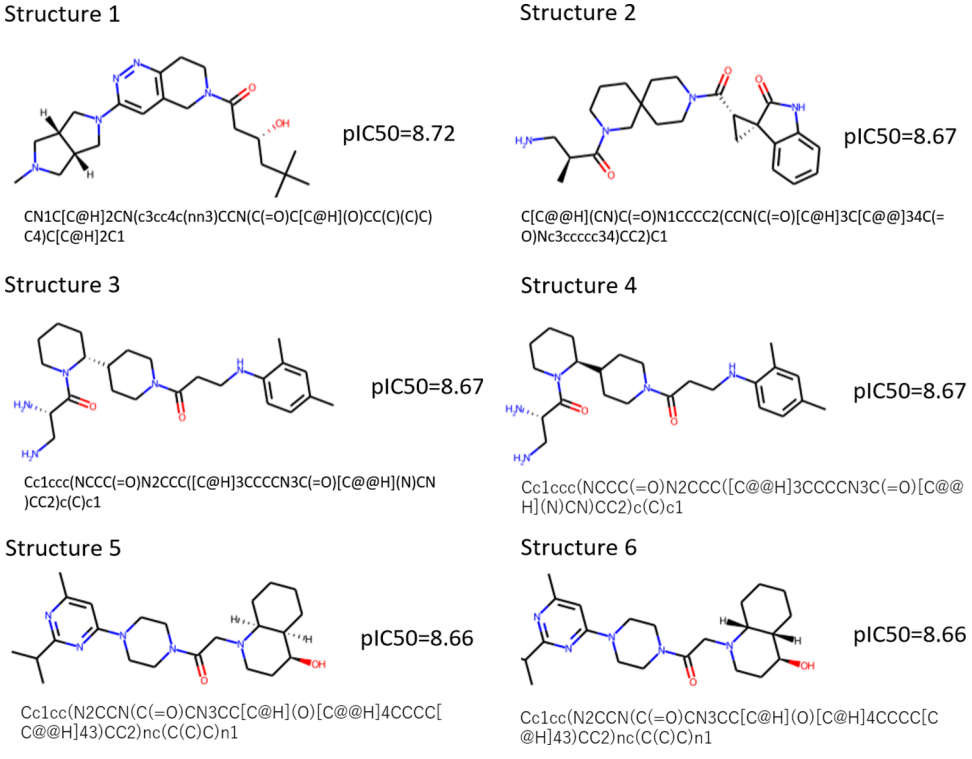


Figure 4. Chemical structures with pIC_{50} predictions over 8.0 by 2D-QSAR on the ZINC20 dataset.

https://doi.org/10.1021/acsomega.3c02567
ACS Omega 2023, 8, 27247−27255

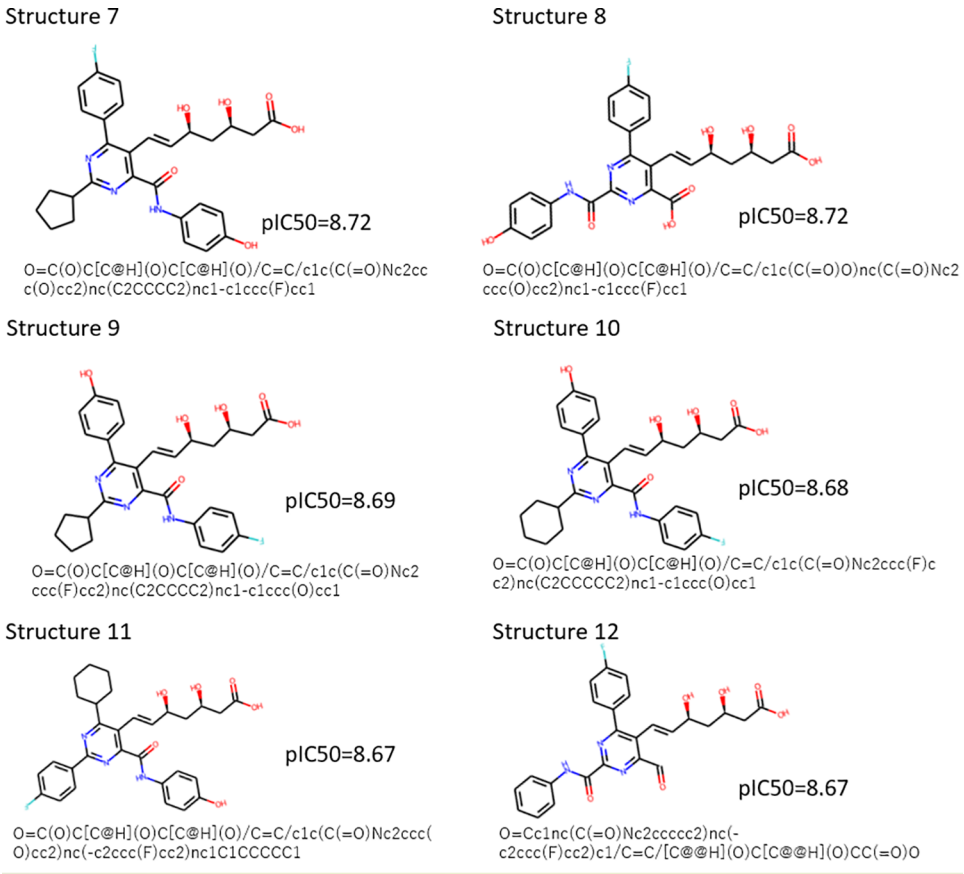


Figure 5. Chemical structures with pIC_{50} predictions over 8.0 by 2D-QSAR on the SGBR dataset.

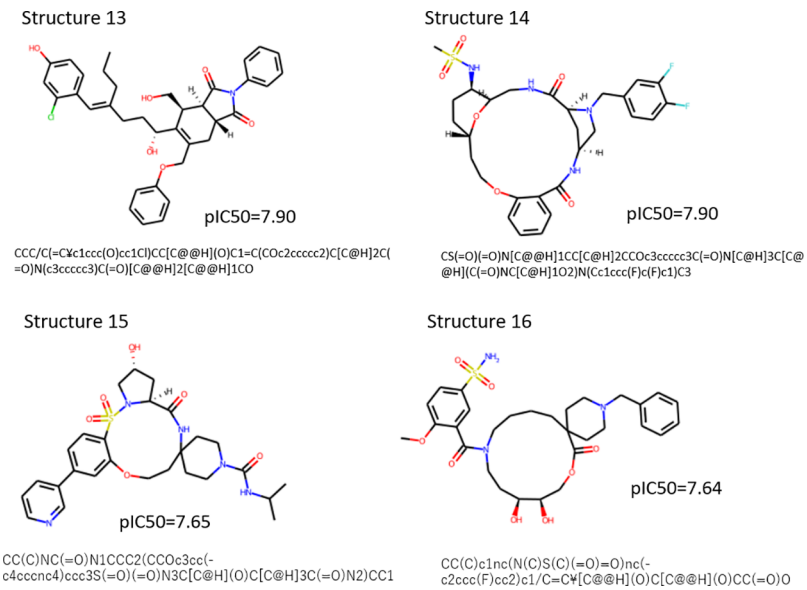


Figure 6. Chemical structures with pIC_{50} predictions over 7.4 by C3D-QSAR on the ZINC dataset.

structure with a methyl group. This is consistent with the results reported by Beck et al.[20] Wang et al. studied the interaction of atorvastatin derivatives with HMGR using MD simulations and reported that the three rings attached to the pyrrole ring interact hydrophobically with HMGR.[21] The phenyl group in the chemical structure shown in Figure 10 may also exhibit high inhibitory activity due to similar hydrophobic interactions. Istvan et al. reported a hydrogen bond between the carbonyl oxygen of atorvastatin and Ser 565 of HMGR.[22] Since the chemical structure shown in Figure 10 is similar to that of atorvastatin, it may interact with Ser 565 of HMGR and inhibit the activity of HMGR in the same way as atorvastatin. In addition, the compound shown in Figure 10 has a hydroxyl group attached to the phenyl group. These hydroxyl groups may interact with HMGR by binding to hydrophilic amino acid residues. Gui et al. suggested that a

https://doi.org/10.1021/acsomega.3c02567
ACS Omega 2023, 8, 27247−27255

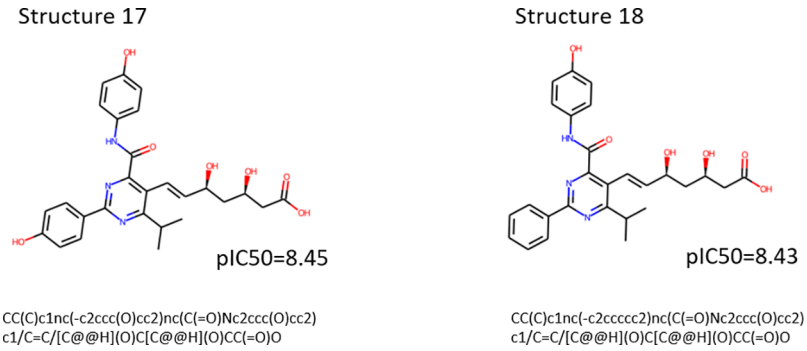


Figure 7. Chemical structures with pIC_{50} predictions over 8.43 by C3D-QSAR on the SGBR dataset.

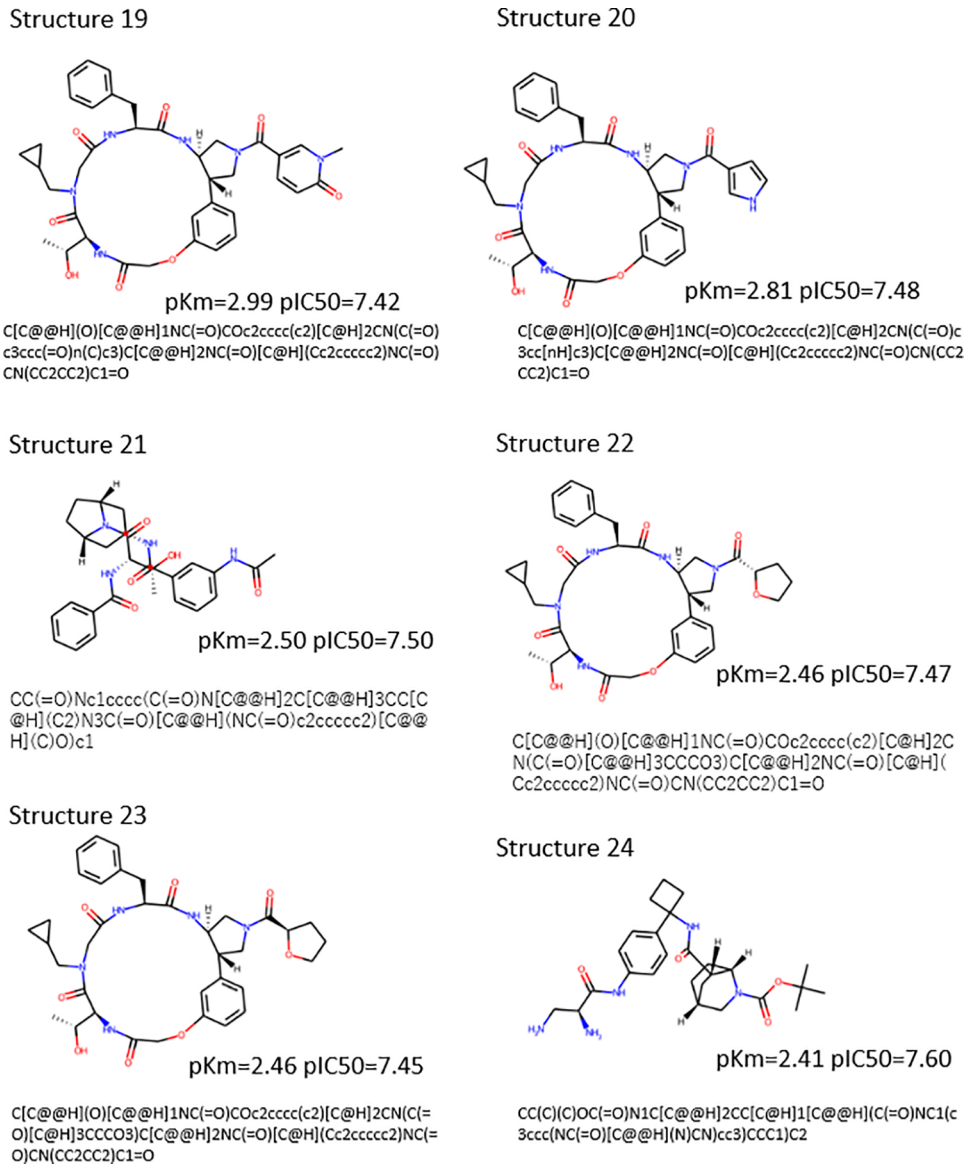


Figure 8. Chemical structures with pK_m predictions over 1.71 by 2D-QSAR and pIC_{50} predictions over 7.4 by C3D-QSAR on the ZINC dataset.

sterically large hydrophobic region in the center with either a negative charge or polar groups at either end could interact with OATP1B1.[23] In the compound in Figure 10, the structure is similar, so the results are consistent with those reported by Gui et al. Based on these observations, it is possible that the compound with the chemical structure shown in Figure 10 is a better statin drug than rosuvastatin because of its high HMGR inhibitory activity and high OATP1B1 affinity.

None of the compounds in the ZINC dataset were predicted to have higher HMGR inhibitory activity than rosuvastatin. However, we found that several compounds, such as those shown in Figure 8, were predicted to have higher inhibitory

https://doi.org/10.1021/acsomega.3c02567
ACS Omega 2023, 8, 27247−27255

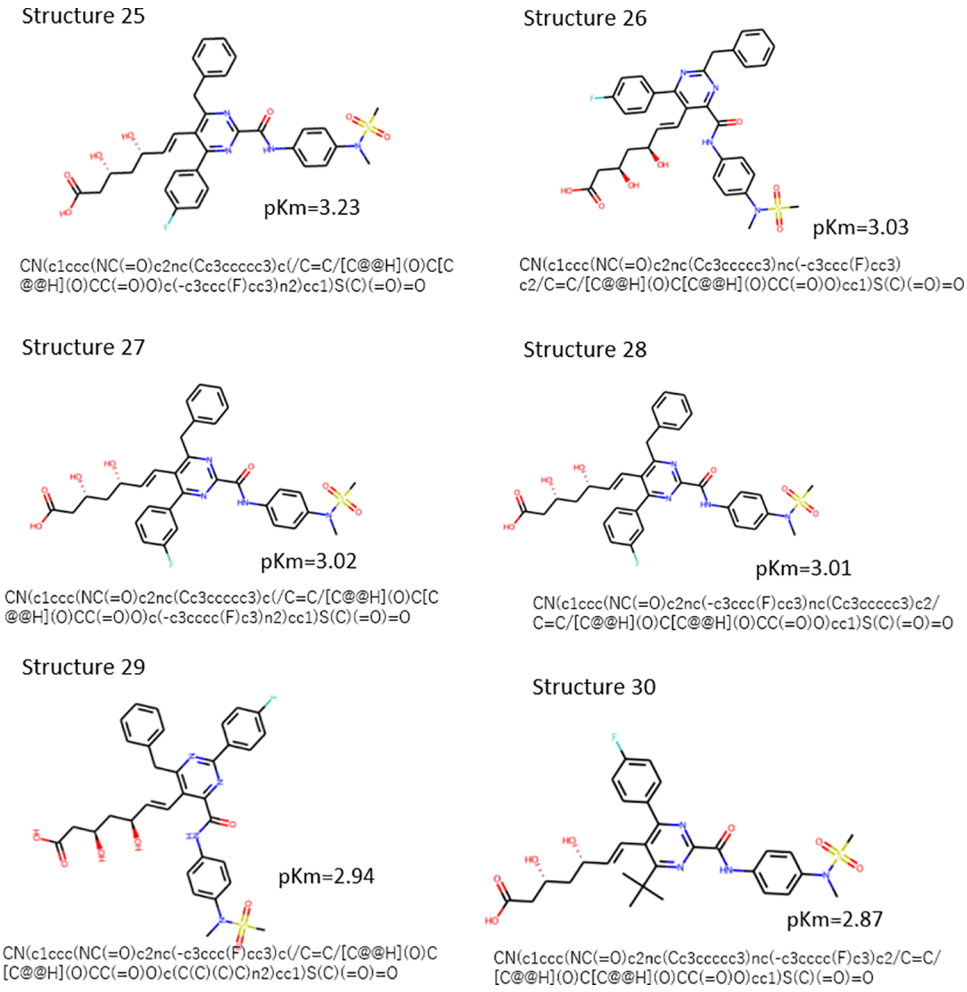


Figure 9. Chemical structures with pK_m predictions over 1.71 by 2D-QSAR on the SGBR dataset.

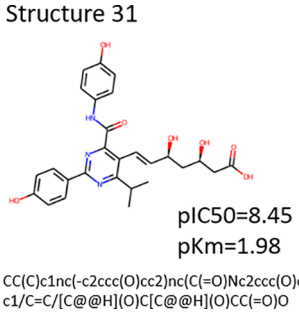


Figure 10. Chemical structures of new statin drug candidates.

activity than pravastatin, some of which had predicted pK_m values greater than those of rosuvastatin. They are potentially new HMGR inhibitors because these compounds are more likely to be transferred to the liver than rosuvastatin.

4. CONCLUSIONS

In this study, we assumed that compounds with high HMGR inhibitory activity and high affinity for OATP1B1 would be suitable as hyperlipidemic drugs, and constructed HMGR inhibition prediction models and OATP1B1 affinity prediction models. Two datasets, ZINC20 and SGBR dataset, were used as screening data. We identified compounds with higher HMGR inhibitory activity and OATP1B1 affinity than those of rosuvastatin in the SGBR dataset. Although there were no compounds in the ZINC20 dataset with HMGR inhibitory activity higher than that of rosuvastatin, there were compounds with HMGR inhibitory activity higher than that of pravastatin that may be potential new HMGR inhibitors.

In this study, three regression models were constructed and used to propose new statin chemical structures. All constructed models had high prediction accuracy. In addition, the proposed chemical structures had higher pIC_{50} or pK_m values than known statins, suggesting that they may be excellent statin drugs. However, their study was limited to predictions using QSAR. In future studies, it will be necessary to synthesize these compounds and measure their activity experimentally. Additionally, other transporters such as OATP2B1 and toxicity effects should be considered in the future.

■ ASSOCIATED CONTENT

Data Availability Statement

The software that supports the findings of this study are available at https://github.com/hkaneko1985/dcekit.

■ AUTHOR INFORMATION

Corresponding Author

Hiromasa Kaneko − *Department of Applied Chemistry, School of Science and Technology, Meiji University, Kawasaki, Kanagawa 214-8571, Japan;* orcid.org/0000-

https://doi.org/10.1021/acsomega.3c02567
ACS Omega 2023, 8, 27247−27255

0001-8367-6476; Phone: +81-44-934-7197;
Email: hkaneko@meiji.ac.jp

Author

Shigeyoshi Samizo − *Department of Applied Chemistry, School of Science and Technology, Meiji University, Kawasaki, Kanagawa 214-8571, Japan*

Complete contact information is available at:
https://pubs.acs.org/10.1021/acsomega.3c02567

Notes

The authors declare no competing financial interest.

■ ACKNOWLEDGMENTS

This study was supported by a Grant-in-Aid for Scientific Research (KAKENHI) (grant number 19K15352) from the Japan Society for the Promotion of Science.

■ REFERENCES

(1) Grundy, S. M. Cholesterol and coronary heart-disease - a new era. *JAMA* **1986**, *256*, 2849−2858.
(2) Friesen, J. A.; Rodwell, V. W. The 3-hydroxy-3-methylglutaryl coenzyme-A (HMG-CoA) reductases. *Genome Biol.* **2004**, *5*, No. 248.
(3) Franc, S.; Dejager, S.; Bruckert, E.; Chauvenet, M.; Giral, P.; Turpin, E. A comprehensive description of muscle symptoms associated with lipid-lowering drugs. *Cardiovasc. Drugs Ther.* **2003**, *17*, 459−465.
(4) Vildhede, A.; Karlgren, M.; Svedberg, E. K.; Wisniewski, J. R.; Lai, Y.; Noren, A.; Artursson, P. Hepatic uptake of atorvastatin: influence of variability in transporter expression on uptake clearance and drug-drug interactions (vol 42, pg 1210, 2014). *Drug Metab. Dispos.* **2015**, *43*, 786−787.
(5) (a) Ahmad, S.; Madsen, C. S.; Stein, P. D.; Janovitz, E.; Huang, C.; Ngu, K.; Bisaha, S. R.; Kennedy, L. J.; Chen, B. C.; Zhao, R. L.; et al. (3R,5S,E)-7-(4-(4-fluorophenyl)-6-isopropyl-2-(methyl(1-methyl-1H-1,2,4-triazo1-5-yl)amino)pyrimidin-5-yl)-3,5-dihydroxyhept-6-enoic acid (BMS-644950): a rationally designed orally efficacious 3-hydroxy-3-methylglutaryl coenzyme-A reductase inhibitor with reduced myotoxicity potential. *J. Med. Chem.* **2008**, *51*, 2722−2733. (b) Ho, R. H.; Tirona, R. G.; Leake, B. F.; Glaeser, H.; Lee, W.; Lemke, C. J.; Wang, Y.; Kim, R. B. Drug and bile acid transporters in rosuvastatin hepatic uptake: function, expression, and pharmacogenetics. *Gastroenterology* **2006**, *130*, 1793−1806.
(6) (a) Davood, A.; Nematollahi, A.; Iman, M.; Shafiee, A. Computational studies of new 1,4-dihydropyridines containing a 4-(5)-chloro-2-ethyl-5-(4)-imidazolyl substituent: QSAR and docking. *Med. Chem. Res.* **2010**, *19*, 58−70. (b) Guccione, S.; Doweyko, A. M.; Chen, H. M.; Barretta, G. U.; Balzano, F. 3D-QSAR using 'multiconformer' alignment: The use of HASL in the analysis of 5-HT1A thienopyrimidinone ligands. *J. Comput.-Aided Mol. Des.* **2000**, *14*, 647−657.
(7) (a) Kitano, H. Computational systems biology. *Nature* **2002**, *420*, 206−210. (b) van de Waterbeemd, H.; Gifford, E. ADMET in silico modelling: Towards prediction paradise? *Nat. Rev. Drug Discovery* **2003**, *2*, 192−204.
(8) Tropsha, A. Best practices for QSAR Model development, validation, and exploitation. *Mol. Inf.* **2010**, *29*, 476−488.
(9) https://www.rdkit.org/ (accessed February 18, 2023).
(10) (a) Bednarczyk, D.; Boiselle, C. Organic anion-transporting polypeptide (OATP)-mediated transport of coproporphyrins I and III. *Xenobiotica* **2016**, *46*, 457−466. (b) van de Steeg, E.; Greupink, R.; Schreurs, M.; Nooijen, I. H. G.; Verhoeckx, K. C. M.; Hanemaaijer, R.; Ripken, D.; Monshouwer, M.; Vlaming, M. L. H.; DeGroot, J.; et al. Drug-drug interactions between rosuvastatin and oral antidiabetic drugs occur at the OATP1B1 level. *Drug Metab. Dispos.* **2013**, *41*, 592−601. (c) Brenner, S.; Riha, J.; Giessrigl, B.; Thalhammer, T.; Grusch, M.; Krupitza, G.; Stieger, B.; Jager, W. Effects of organic anion-transporting polypeptides 1B1, 1B3, and 2B1 on the antitumor activity of flavopiridol in breast cancer cells. *Int. J. Oncol.* **2015**, *46*, 324−332. (d) Zhang, Z. Y.; Li, Q. S.; Si, D. Y.; Yi, X. L.; Liu, C. X. Application of LC-MS/MS method for evaluating rosuvastatin affinity with OATP1B1 and OATP2B1 in vitro. *Asian J. Chem.* **2014**, *26*, 335−341. (e) Izumi, S.; Nozaki, Y.; Komori, T.; Takenaka, O.; Maeda, K.; Kusuhara, H.; Sugiyama, Y. Investigation of fluorescein derivatives as substrates of organic anion-transporting polypeptide (OATP) 1B1 to develop sensitive fluorescence-based OATP1B1 inhibition assays. *Mol. Pharmaceutics* **2016**, *13*, 438−448. (f) Cao, L.; Zhou, J.; Wen, J. H. Transport of salvianolic acid B via the human organic anion transporter 1B1 in the liver. *Phytother. Res.* **2019**, *33*, 197−204. (g) Suga, T.; Yamaguchi, H.; Sato, T.; Maekawa, M.; Goto, J.; Mano, N. Preference of conjugated bile acids over unconjugated bile acids as substrates for OATP1B1 and OATP1B3. *PLoS One* **2017**, *12*, No. e0169719. (h) Jeong, H. U.; Kwon, M.; Lee, Y.; Yoo, J. S.; Shin, D. H.; Song, I. S.; Lee, H. S. Organic anion transporter 3-and organic anion transporting polypeptides 1B1-and 1B3-mediated transport of catalposides. *Drug Des. Devel. Ther.* **2015**, *9*, 643−653. (i) Meier-Abt, F.; Faulstich, H.; Hagenbuch, B. Identification of phalloidin uptake systems in rat and human livers. *Biochim. Biophys. Acta, Biomembr.* **2004**, *1664*, 64−69. (j) Wang, C.; Huo, X. K.; Wang, C. Y.; Meng, Q.; Liu, Z. H.; Sun, P. Y.; Cang, J.; Sun, H. J.; Liu, K. Organic anion transporting polypeptide and efflux transporter-mediated hepatic uptake and biliary excretion of cilostazol and its metabolites in rats and humans. *J. Pharm. Sci.* **2017**, *106*, 2515−2523. (k) Wang, H. P.; Sun, P. Y.; Wang, C. Y.; Meng, Q.; Liu, Z. H.; Huo, X. K.; Sun, H. J.; Ma, X. D.; Peng, J. Y.; Liu, K. X. Liver uptake of cefditoren is mediated by OATP1B1 and OATP2B1 in humans, and Oatp1a1, Oatp1a4, and Oatp1b2 in rats. *RSC Adv.* **2017**, *7*, 30038−30048.
(11) Nitta, F.; Kaneko, H. Two- and three-dimensional quantitative structure-activity relationship models based on conformer structures. *Mol. Inf.* **2021**, *40*, No. 2000123.
(12) http://mordred-descriptor.github.io/documentation/master/ (accessed February 18, 2023).
(13) Filzmoser, P.; Liebmann, B.; Varmuza, K. Repeated double cross validation. *J. Chemometr.* **2009**, *23*, 160−171.
(14) Irwin, J. J.; Tang, K. G.; Young, J.; Dandarchuluun, C.; Wong, B. R.; Khurelbaatar, M.; Moroz, Y. S.; Mayfield, J.; Sayle, R. A. ZINC20-A free ultralarge-scale chemical database for ligand discovery. *J. Chem. Inf. Model.* **2020**, *60*, 6065−6073.
(15) https://github.com/hkaneko1985/structure_generator_based_on_r_group. (accessed February 18, 2023).
(16) Dragos, H.; Marcou, G.; Alexandre, V. Predicting the predictability: a unified approach to the applicability domain problem of QSAR models. *J Chem. Inf. Model.* **2009**, *49*, 1762−1776.
(17) Kanno, Y.; Kaneko, H. Improvement of predictive accuracy in semi-supervised regression analysis by selecting unlabeled chemical structures. *Chemom. Intell. Lab. Syst.* **2019**, *191*, 82−87.
(18) https://en.wikipedia.org/wiki/68%E2%80%9395%E2%80%9399.7_rule. (accessed February 18, 2023).
(19) (a) Son, M.; Baek, A.; Sakkiah, S.; Park, C.; John, S.; Lee, K. W. Exploration of virtual candidates for human HMG-CoA reductase inhibitors using pharmacophore modeling and molecular dynamics simulations. *PLoS One* **2013**, *8*, No. e83496. (b) Teli, M. K.; K, R. G. Pharmacophore generation and atom-based 3D-QSAR of N-isopropyl pyrrole-based derivatives as HMG-CoA reductase inhibitors. *Org. Med. Chem. Lett.* **2012**, *2*, No. 25. (c) Zhang, Q. Y.; Wan, J.; Xu, X.; Yang, G. F.; Ren, Y. L.; Liu, J. J.; Wang, H.; Guo, Y. Structure-based rational quest for potential novel inhibitors of human HMG-CoA reductase by combining CoMFA 3D QSAR modeling and virtual screening. *J. Comb. Chem.* **2007**, *9*, 131−138.
(20) Beck, G.; Kesseler, K.; Baader, E.; Bartmann, W.; Bergmann, A.; Granzer, E.; Jendralla, H.; Vonkerekjarto, B.; Krause, R.; Paulus, E.; et al. Synthesis and biological-activity of new HMG-COA reductase inhibitors. 1. Lactones of pyridine-substituted and pyrimidine-

substituted 3,5-dihydroxy-6-heptenoic(-heptanoic) acids. *J. Med. Chem.* **1990**, *33*, 52−60.
(21) Wang, Z.; Cheng, L. P.; Kai, Z. P.; Wu, F. H.; Liu, Z. Y.; Cai, M. F. Molecular modeling studies of atorvastatin analogues as HMGR inhibitors using 3D-QSAR, molecular docking and molecular dynamics simulations. *Bioorg. Med. Chem. Lett.* **2014**, *24*, 3869−3876.
(22) Istvan, E. S.; Deisenhofer, J. Structural mechanism for statin inhibition of HMG-CoA reductase. *Science* **2001**, *292*, 1160−1164.
(23) Gui, C.; Wahlgren, B.; Lushington, G. H.; Hagenbuch, B. Identification, Ki determination and CoMFA analysis of nuclear receptor ligands as competitive inhibitors of OATP1B1-mediated estradiol-17beta-glucuronide transport. *Pharmacol. Res.* **2009**, *60*, 50−56.

https://doi.org/10.1021/acsomega.3c02567
ACS Omega 2023, 8, 27247−27255