JOURNAL OF CHEMICAL INFORMATION AND MODELING

Article

pubs.acs.org/jcim

Modeling Compound–Target Interaction Network of Traditional Chinese Medicines for Type II Diabetes Mellitus: Insight for Polypharmacology and Drug Design

Sheng Tian,[†] Youyong Li,[†] Dan Li,[‡] Xiaojie Xu,[§] Junmei Wang,[⊥] Qian Zhang,[†] and Tingjun Hou*[,†,‡]

[†]Institute of Functional Nano & Soft Materials (FUNSOM) and Jiangsu Key Laboratory for Carbon-Based Functional Materials & Devices, Soochow University, Suzhou, Jiangsu 215123, China

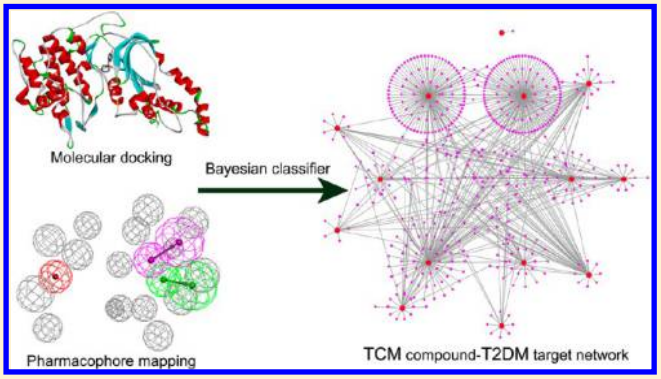
[‡]College of Pharmaceutical Sciences, Zhejiang University, Hangzhou, Zhejiang 310058, China

[§]College of Chemistry and Molecular Engineering, Peking University, Beijing 100871, China

[⊥]Department of Biochemistry, The University of Texas Southwestern Medical Center, 5323 Harry Hines Boulevard, Dallas, Texas 75390, United States

Supporting Information

ABSTRACT: In this study, in order to elucidate the action mechanism of traditional Chinese medicines (TCMs) that exhibit clinical efficacy for type II diabetes mellitus (T2DM), an integrated protocol that combines molecular docking and pharmacophore mapping was employed to find the potential inhibitors from TCM for the T2DM-related targets and establish the compound–target interaction network. First, the prediction capabilities of molecular docking and pharmacophore mapping to distinguish inhibitors from noninhibitors for the selected T2DM-related targets were evaluated. The results show that molecular docking or pharmacophore mapping can give satisfactory predictions for most targets but the validations are still quite necessary because the prediction accuracies of these two methods are variable across different targets. Then, the Bayesian classifiers by integrating the predictions from molecular docking and pharmacophore mapping were developed, and the well-validated Bayesian classifiers for 15 targets were utilized to find the potential inhibitors from TCM and establish the compound–target interaction network. The analysis of the compound–target network demonstrates that a small portion (18.6%) of the predicted inhibitors can interact with multitargets. The pharmacological activities for some potential inhibitors have been experimentally confirmed, highlighting the reliability of the Bayesian classifiers. Besides, it is interesting to find that a considerable number of the predicted multitarget inhibitors have free radical scavenging/antioxidant activities, which are closely related to T2DM. It appears that the pharmacological effect of the TCM formulas is determined not only by the compounds that interact directly with one or more T2DM-related targets, but also by the compounds with other supplementary bioactivities important for relieving T2DM, such as free radical scavenging/antioxidant effects. The mechanism uncovered by this study may offer a deep insight for understanding the theory of the classical TCM formulas for combating T2DM. Moreover, the predicted inhibitors for the T2DM-related targets may provide a good source to find new lead compounds against T2DM.



INTRODUCTION

Diabetes mellitus, or simply diabetes, is a group of metabolic disorder diseases. The diabetes patients have high blood glucose, because either the body system cannot produce enough insulin (absolute insulin deficiency) or the related cells are not sensitive to or respond to insulin (insulin resistance or relative insulin deficiency). Diabetes mellitus has become a serious health problem around the world. Diabetes mellitus is classified into three categories: type 1, type 2, and gestational diabetes. Type 1 diabetes mellitus (T1DM) is characterized by the loss of the insulin-producing beta cells of the islets of langerhans in the pancreas, leading to absolute insulin deficiency. Type 2 diabetes mellitus (T2DM) is characterized by insulin resistance, which may be combined with relatively reduced insulin secretion. The last one, gestational diabetes mellitus (GDM), which occurs in about 2–5% of all pregnancies, resembles T2DM in several respects, involving a combination of relatively inadequate secretion and responsiveness. Globally, it is estimated that 285 million people have diabetes mellitus, with T2DM making up about 90% of the cases in 2010.[1] Its incidence is increasing rapidly, and by 2030, this number is estimated to be almost double.[2]

Received: March 9, 2013
Published: June 16, 2013

T2DM is a chronic disease, associated with a ten-year shorter life expectancy, caused by a number of complications including cardiovascular, eye, kidney, and nervous system diseases.[3−6] Most T2DM patients need medication to reduce their blood glucose to near-normal level.[7] To control high-level blood glucose and prevent diabetic complications, many drugs, such as alpha-glucosidase inhibitors, sulfonylureas, metformin, thiazolidinediones (TZDs), and insulin injections, have been used to treat T2DM patients.[7] To our disappointment, all of these therapeutic agents have limited efficacy and are associated with undesired side effects, including hypoglycemia, weight gain, gastrointestinal disturbances, edema, anemia, lactic acidosis, etc.[8,9] Due to the deficiencies of marketed drugs for T2DM, the philosophy of rational drug design, or more specifically, the "one gene, one drug, one disease" paradigm for treating T2DM was met with controversy.

Over the past decade, the number of new molecular entities (NME) approved by US Food and Drug Administration (FDA) did not change significantly. The reasons that have been proposed for this discouraging situation may not be technological, environmental, or even scientific but philosophical.[10] Also, it has been proven that many effective drugs act on multiple rather than one specific disease-related target.[11] It seems that the "magic bullets" cannot cure some specific diseases as expected.

The new drug candidates with low efficacy and undesired ADME/T properties are not the only factor to debate the traditional paradigm for drug design/discovery.[12,13] The phenotype usually cannot be disturbed by single-gene knockouts[14] due to redundant functions and alternative compensatory signaling pathways.[10,15] The intrinsic robustness of living systems against various perturbations is an essential factor that hinders such magic bullet drugs' success in the drug market.[16]

Compared with the traditional single-target−drug paradigm, the new philosophy for drug design known as polypharmacology or network pharmacology has attracted more attentions.[10,17−19] Therefore, it has been realized that designing drug candidates to target a disease-associated network rather than a single target may be a better strategy to find better drug candidates for some complicated diseases, such as T2DM.

It is estimated that approximated 34% of new drugs were directly or indirectly from natural products from 1981 to 2010.[20] As an important source of natural products in clinical use, traditional Chinese medicines (TCMs) have played an important role in the healthcare system for Chinese people for more than several thousand years. Because of satisfactory in vivo efficacy and safety of TCM in targeting complicated chronic diseases,[21] finding potential drug candidates from TCM may be a good practice.[22,23] The main philosophical theories of TCM are based on Yin−Yang, five elements, human body meridian systems, and Zang Fu theory.[24] It is well-known that the basic form of TCM for combating diseases is TCM formula (or prescription), which is a mixture of special herbs with suitable dosages. A TCM formula consists of a large number of chemical compounds, which may interact with multitargets. Therefore, at the molecular level, the mechanisms of TCM formulas may be explained by polypharmacology or network pharmacology.

Compared with the time-consuming and costly experimental methods to uncover the action mechanism of TCM for combating diseases, the use of computational approaches may provide an easier and more practical way to understand the mysterious power of TCM.[25−27] Previous studies suggest that molecular modeling techniques, especially molecular docking and pharmacophore mapping, not only can find potentially active compounds from TCM for targeting disease-related proteins, but also can identify potential targets for compounds from TCM.[28−34] However, it seems that most previous studies just simply used computational approaches to find potential inhibitors from TCM or even establish the compound−target networks to explore the polypharmacology of TCM. It should be noted that the validations of the computational approaches are quite critical to guarantee the reliability of the predictions for TCM given by the computational approaches.

Through thousands of years of medical use for various diseases in the Chinese medical system, a large amount of knowledge has also been accumulated for diabetes therapy that is termed as "XiaoKeZheng".[24,35] Similar to the combination therapy of multicomponent drugs, the multicomponents from TCM can interact with multitargets through an integrated way to lowing blood glucose or relieving diabetes.[21,36] However, we still know little about how the active compounds in the TCM formulas modulate the network for combating diabetes. Therefore, in this study, we tried to establish the compound−target interaction network by theoretical approaches and uncover the underlying action mechanisms of TCM for T2DM. To make our predictions more reliable, the capabilities of molecular docking and pharmacophore mapping to discriminate the known inhibitors from the noninhibitors for each T2DM-related target were assessed. To integrate the advantages of molecular docking and pharmacophore modeling, the naïve Bayesian classifiers were developed by combining the predictions from molecular docking and pharmacophore mapping together. The combined models show better performance than those only based on the predictions from molecular docking or pharmacophore mapping. Then, the well-validated Bayesian classifiers were utilized to find potentially active compounds in TCM that can interact with the important T2DM-related targets. Finally, the compound−target interaction network was established to uncover the action mechanisms of the classical TCM formulas for relieving T2DM.

METHODS AND MATERIALS

Collection of the T2DM-Related Targets. By searching the public data sources, including Therapeutic Targets Database (TTD),[37] Drugbank,[38] Kyoto Encyclopedia of Genes and Genomes (KEGG) pathway database,[39] and previous reported literature,[31,40] the important targets related to T2DM were identified and the available crystal structures of the protein−ligand complexes for these targets were retrieved from the RCSB protein data bank.[41] For each target, the crystal structure with the highest resolution was reserved if multiple structures were available, and finally, a nonredundant set of 48 crystal complexes for the T2DM-related targets was obtained. It is necessary to point out that almost all of the T2DM-related targets may have several synonyms because different standards might be used in different target databases.[37] For example, the corticosteroid 11-beta-dehydrogenase isozyme 1 also can be named as 11-Beta hydroxysteroid dehydrogenase 1, 11-beta-HSD1, 11HSD1, and so on.[37] The specific target with different names should be check carefully. The names, PDB entries, and the corresponding crystal resolutions for the selected protein−ligand complexes are summarized in Table S1 in the Supporting Information.

Collection of the Compounds in the T2DM-Related TCM Formulas. By extensive data profiling, 52 classical TCM

formulas with clinical efficacy for T2DM were compiled. There are 95 Chinese herbs in those 52 TCM formulas. The frequency of each herb was counted, and only the 32 Chinese herbs with a frequency of 5 or higher were considered in our study (Table S2 in the Supporting Information). The Latin names for the 32 studied Chinese herbs were collected from the latest version of TCMCD developed in our group[42−45] and TCM-Database@Taiwan.[46] It should be noted that multiple Latin names might be available for the same Chinese herb. The reason is that different standards may be used in different TCM databases. For example, the herb "Cang Zhu" (Chinese name) has the Latin name *Atractylodes lancea* in TCMCD and *Atractylodes lancea (Thunb.) DC.* in TCM-Database@Taiwan (Table S2 in the Supporting Information). A total of 2479 nonduplicated compounds isolated from these 32 Chinese herbs were found from the TCMCD and TCM-Database@Taiwan databases. These compounds were standardized and optimized by using Discovery Studio 3.1 (DS3.1) molecular simulation package.[47]

Preparation of the Validation Data Sets. The validation data set for each T2DM-related target was prepared to evaluate the prediction accuracies of molecular docking or pharmacophore mapping to distinguish known inhibitors from noninhibitors. The known inhibitors were obtained from the BindingDB database.[48] The number of the inhibitors for each target is listed in Table S3 in the Supporting Information. To ensure the statistical significance of the validation results, the targets with the number of the known inhibitors less than 10 are excluded. Besides, in order to achieve acceptable computational efficiency, when the number of the known inhibitors for a target is higher than 500, only 500 randomly selected inhibitors were included in the validation data set. Then, because the experimentally validated noninhibitors are quite limited, the compounds were randomly chosen from the Chembridge database to serve as noninhibitors for each target.[49] The inhibitors and noninhibitors were compared and the duplicates were removed from the noninhibitor subset. To mimic the unbalanced nature of inhibitors versus noninhibitors, we chose to set the ratio of noninhibitors versus inhibitors to 20. By applying the processing protocol mentioned above, the validation data sets for the 33 T2DM-related targets were generated (Table S3 in the Supporting Information).

Validation of Molecular Docking-Based Virtual Screening. The crystal structures of the protein−ligand complexes for the 33 T2DM-related targets were used for the docking calculations performed with *Glide*[50] in Schrodinger 9.0.[51] For each crystal structure, the crystallographic water molecules were removed, the missing hydrogen atoms were added, and the protonated states and partial charges were assigned with the OPLS-2005 force field.[52] Minimizations were performed until the average root-mean-square deviation of nonhydrogen atoms reached 0.3 Å. Preparation and refinement were accomplished with the *protein preparation wizard* in Schrodinger 9.0.[51]

The validation data sets were processed using the *LigPrep* module in Schrodinger.[51] For each small molecule, the possible ionization states were generated by using *Epik* at pH = 7.0. Because the noninhibitors randomly chosen from Chembridge do not have 3D structural information, all possible stereoisomers for each molecule were generated. All of the other parameters were kept as the default settings in *LigPrep*.

Then, the bounding box of size 10 Å × 10 Å × 10 Å was defined for each target and centered on the mass center of the cocrystal ligand to confine the mass center of the docked ligand by using the *Receptor Grid Generation* function of *Glide*.[51] According to our docking results, the size of the bounding box is big enough for almost all of the ligands to be docked into the binding site successfully for each target. The default *Glide* settings were used for grid generation. All of the compounds were docked into the binding site of the studied targets and scored by the standard precision (SP) and extra precision (XP) scoring modes of *Glide*.[51] Five thousand poses per ligand were generated during the initial phase of the docking calculation, out of which the best 400 poses were chosen for energy minimization. A dielectric constant of 2.0 and 100 steps of conjugate gradient minimizations were used in energy minimization stage. The performances of the SP and XP scoring modes of Glide on the 33 T2DM-related targets were evaluated and compared.

Validation of Pharmacophore-Based Virtual Screening. The low-energy conformations of the molecules in the validation data sets were generated by using the *Conformation Generation protocol* in DS3.1[47] with the FAST setting and the number of the maximum conformations *per* compound was set to 100. Then, the receptor−ligand interaction patterns (or pharmacophore features) for each target was generated by using the *Receptor−Ligand Pharmacophore Generation (RLPG)* protocol in DS3.1.[47] A set of predefined pharmacophore features that include hydrogen bond acceptor/donor (HBA/HBD), hydrophobic (HYD), negative/positive ionizable (PI/NI), and ring aromatic (RA) of the ligand are identified for each receptor−ligand complex first and then pruned when not satisfying the protein−ligand interactions using adjustable topological rules.[53]

The *RLPG* protocol in DS3.1 was applied to the receptor−ligand complexes for the 33 T2DM-related targets. For each pharmacophore model, the minimum number of the pharmacophore features was set to 3 and the maximum number of the pharmacophore features was the same as the number of the total features that can match the receptor−ligand interactions. Up to ten pharmacophore models per receptor−ligand complex were built and ranked by the selectivity as the fault settings in DS3.1. The selectivity of a pharmacophore model was evaluated by a rule-based score function based on the genetic function approximation (GFA) technique.[54] The details of the GFA model for pharmacophore modeling were described in the previous study.[55] Besides, the discrimination power of a pharmacophore model was determined by the capability to distinguish known inhibitors from noninhibitors. The discrimination capability of a pharmacophore model is more important than the selectivity for virtual screening. An ideal pharmacophore model used in virtual screening must have the ability not only to find the most promising active compounds but also to abandon the inactive compounds. The discrimination capability of each pharmacophore model for the corresponding validation set was evaluated.

Generation of the Molecule−Target Interaction Network. As mentioned above, the performance of molecular docking or pharmacophore modeling for each target was validated. The validation process aiming to distinguish the known inhibitors from noninhibitors for each target is a binary classification problem. Recently, the naïve Bayesian classification technique was successfully applied to different classification problems by our group.[45,56,57] Naïve Bayesian classification can process large amounts of data, learn fast, and is tolerant of random noise. It only requires a small amount of training data

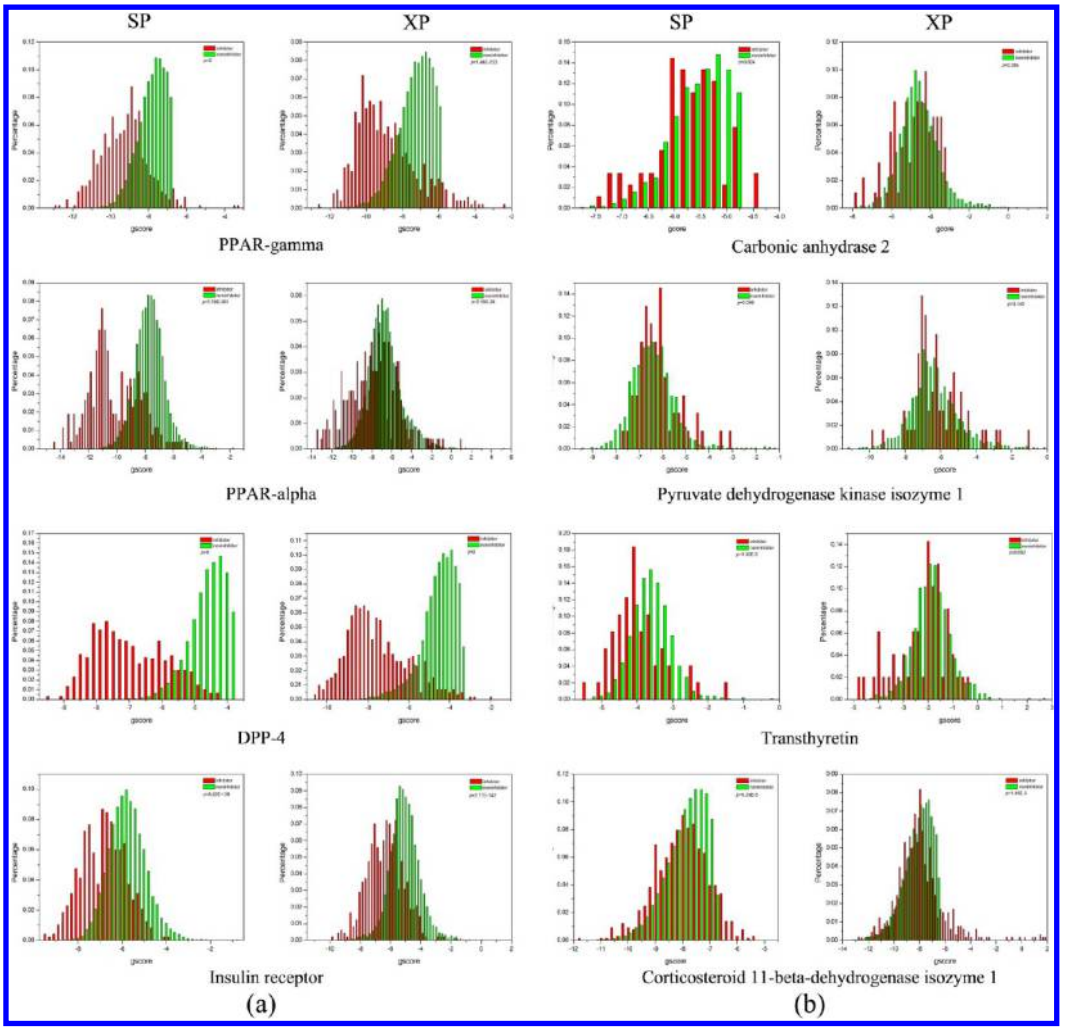


Figure 1. Distributions of the docking scores of the validation data sets for (a) four targets (PPAR-gamma, PPAR-alpha, DPP-4, and insulin receptor) with low *p*-values and (b) four targets (carbonic anhydrase 2, pyruvate dehydrogenase [lipoamide] kinase isozyme 1, transthyretin and corticosteroid 11-beta-dehydrogenase, isozyme 1) with high *p*-values.

to estimate the parameters (means and variances of the variables) necessary for classification. Moreover, in the scheme of naïve Bayesian classification, the belonging of a compound in which class can be qualitatively evaluated by the posterior probability. The mathematical procedure to train a naive Bayesian classifier was described previously.[45,56,57] To integrate the advantages of molecular docking and pharmacophore mapping, each compound in the validation data sets was docked into the binding site first and then mapped onto the pharmacophore model for each target. The docking scores given by molecular docking and the fit values given by pharmacophore mapping were used as the independent variables to train the naïve Bayesian classifiers.

After rigorous validation, the naïve Bayesian classifiers with reliable predictions for 15 targets were utilized to virtually screen the TCM compound data set. First, The 2479 nonduplicate compounds from 32 Chinese herbs of the TCM formulas for T2DM (Table S2 in the Supporting Information) were processed using the *LigPrep* module in Schrodinger.[51] Then, those compounds were docked into the binding sites of the 15 T2DM-related targets by using Glide in Schrodinger 9.0 and scored by the SP and XP scoring modes. Next, each compound was mapped onto the pharmacophore model with highest discrimination capability for each target. On the basis of the calculations from molecular docking and pharmacophore mapping, each compound was predicted by the reliable naïve Bayesian classifiers for 15 T2DM-related targets.

By applying the Bayesian classifiers, the potential inhibitors for each target were determined. In order to balance the prediction accuracy and the complexity of interaction network, only the top 10% of the predicted inhibitors ranked by decreasing the Bayesian scores were extracted to generate the compound−target interaction networks by using the Osprey network visualization program (version 1.2.0).[58]

RESULTS AND DISCUSSION

Performance of Molecular Docking-Based Ligand Profiling. Before applying molecular docking to screen the TCM compound data set (Table S1 in the Supporting Information), the performance of molecular docking to distinguish inhibitors from noninhibitors for 33 T2DM-related targets (Table S2 in the Supporting Information) was evaluated.

First, as a representative measurement for the accuracy and reliability of molecular docking, the "docking power" was examined.[59] The cocrystallized ligands were extracted and then redocked into the binding sites. The root-mean-squared deviation (RMSD) of the docked pose from the original

dx.doi.org/10.1021/ci400146u | J. Chem. Inf. Model. 2013, 53, 1787−1803

position for each complex was computed to evaluate whether the Glide docking is reliable for virtual screening. The RMSD was used as the criterion for the success of the prediction. If the RMSD of the docked pose was less than or equal to 2.0 Å from the experimentally observed conformation, we considered it to be a successful prediction. As shown in Table S3 in the Supporting Information, the SP and XP scoring modes can successfully recognize the near-native conformations for 25 and 26 of the entire 33 complexes, i.e. the successful rate are 75.8% and 78.8%, respectively. In addition, only four complexes that do not meet the criterion of RMSD $\leq$ 2.0 Å by using both SP and XP modes (Table S3 in the Supporting Information). From this data, it is clear that for most complexes the Glide docking can successfully recognize the correct binding conformations.

Then, we evaluated the "discrimination power" of molecular docking to distinguish known inhibitors from noninhibitors for each target. The student's *t*-test was employed to evaluate the significance of the difference between the means of the two distributions of the Glide SP or XP scores for the known inhibitor and noninhibitor classes (Table S3 in the Supporting Information). On the whole, the molecular docking calculations for most targets give satisfactory results, indicated by the quite low *P*-values associated with the difference in the mean SP or XP scores of the inhibitors versus those of the noninhibitors. The *P*-values for 25 of 33 targets (76%) are lower than 10^{-20}, suggesting that molecular docking can successfully discriminate the known inhibitors from noninhibitors for most targets. In other words, the Glide docking can be used as a reliable tool to identify known inhibitors from diverse chemical space for most targets. The distributions of the Glide SP or XP scores for four targets with good predictions and four targets with poor predictions are illustrated in Figures 1a and b as examples.

Generally speaking, the XP scoring (extra precision) is believed to be better than the SP scoring (standard precision). However, it is interesting to observe that the XP scoring of *Glide* does not always show better predictions than the SP scoring. As shown in Table S3 in the Supporting Information, the XP scoring only shows better or equal discrimination power for 15 (45.5%) targets according to the measurement of the *P*-values. The results suggest that the XP scoring does not always yield better performance than SP. Therefore, the comparison study of the performance between the SP and XP scoring modes of *Glide* is quite necessary for a specific target. For each target, based on the comparison results, more reliable scoring mode was chosen for the following virtual screening of the TCM compound data set.

In summary, the *Glide* docking based on the SP or XP scoring mode can successfully recognize the near-native conformations for 25 or 26 of the entire 33 complexes. It is indicated that the *Glide* docking is suitable for our studies. Moreover, it can successfully discriminate the known inhibitors from noninhibitors for 25 of 33 targets, proving the reliability of molecular docking for most targets. Ideally, if the results of docking can meet both of the two requirements: RMSD $\leq$ 2.0 Å and *P*-value $\leq 10^{-20}$, the docking-based virtual screening is reliable.

The reason of the failure of the docking calculations for some targets lies in the fact that the performance of scoring function used by *Glide* is often varied across different protein families.[59,60] For example, for HMG-CoA reductase, six crystal complexes can be retrieved from PDB database (Table S4 in the Supporting Information). According to our analysis, there is only one complex (1DQA) can meet the demand of RMSD ($\leq$2.0 Å) by using the SP mode of *Glide*. Meanwhile, the capability of distinguishing the known inhibitors from non-inhibitors for 1DQA complex is not good, indicated by the high *P*-value (1.13×10^{-6}) associated with the difference in the mean SP scores of the known inhibitors versus that of the noninhibitors (Figure S1 in the Supporting Information). That is to say, all of the docking calculations by using the SP and XP scoring modes for HMG-CoA reductase are questionable either due to unacceptable RMSD or poor discrimination capacity. In most previous application of molecular docking on the TCM studies, the capabilities of molecular docking on the studied systems were not evaluated, and therefore, the reported results may be questionable. According to our calculations, it is strongly recommended that the validation process is necessary before any docking study for TCM.

Performance of Pharmacophore-Based Ligand Profiling. On the basis of the 33 receptor−ligand complexes of the T2DM-related targets (Table S5 in the Supporting Information), 28 complex-based pharmacophore models were generated (Table S5), and no pharmacophore model can be generated for five complexes. There are two possible reasons for explaining the failures for these five complexes: (1) the number of extracted pharmacophore features is quite limited (less than three), and (2) some pharmacophore features extracted from the receptor−ligand complex do not comply with the minimum feature distance criterion (<1.0 Å).

For each complex, the generated pharmacophore models were ranked by the selectivity scores (Table S5 in the Supporting Information). Then, the prediction power of each pharamcophore model to distinguish the known inhibitors from noninhibitor in the validation data set was evaluated. The discrimination power of a pharmacophore model was quantitatively characterized by the AUC area under a receiver operating characteristics (ROC) curve. All pharmacophore features derived from the receptor−ligand interactions (total features), the pharmacophore features for the most selective pharmacophore model (feature set A), and those for the pharmacophore model with the highest AUC value (feature set B) are listed in Table S5. The comparison of the numbers of the pharmacophore features for 25 targets among total features, feature set A, and feature set B are shown in Figure 2a and listed in Table S6 in the Supporting Information. We can observe that the numbers of total features are generally more than those of feature set A, and the numbers of feature set A are also more than those of feature set B for the same complex (Table S5) for most cases. Besides, the percentage of the number of the pharmacophore features that are more than five for total features, feature set A, and feature set B are 71.4% (20/28), 64.3% (18/28), and 32.1% (9/28), respectively.

It is expected that a pharmacophore model with more pharmacophore features should be more selective than that with less pharmacophore features. According to our analysis, the numbers of feature set A for most selective pharmacophore models are not always equal to those of total features. If the number of total features derived from a complex is *n*, the numbers of feature set A for most selective pharmacophore models determined by GFA are usually equal to *n* (14/28) or *n* − 1 (11/28), except *n* − 2 for PPAR-γ, PTP1B, and estrogen receptor. The results demonstrate that some pharmacophore features derived directly from the protein−ligand interactions have no effect on the identification of the known inhibitors from chemical space. Meanwhile, the numbers of feature set B of the pharmacophore models that have the highest

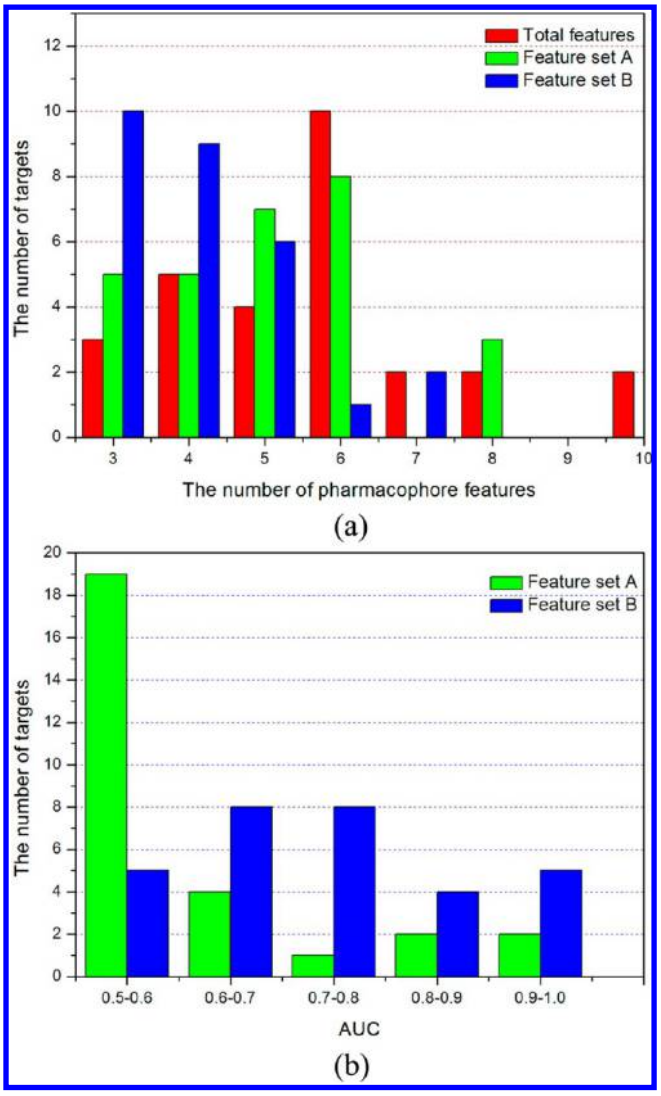


Figure 2. (a) Comparison of the numbers of the pharmacophore features in total features, feature set A, and feature set B and (b) comparison of the discrimination capacity indicated by the AUC values of the pharmacophore models based on feature sets A and B.

discrimination capacity determined by the AUC values are usually $n - 1$ (11/28) or $n - 2$ (11/28), except n for angiotensin-converting enzyme, insulinlike growth factor 1 receptor and mitogen-activated protein kinase 1, $n - 3$ for PPAR-γ, PTP1B, and estrogen receptor. By comparing the pharmacophore features between feature sets A and B, it was found that the numbers of the pharmacophore models with feature set B equaling to, one less than, and two less than those in feature set A are 21.4% (6/28), 64.3% (18/28), and 14.3% (4/28), respectively.

In summary, the pharmacophore models with feature set B determined by the validation data sets are more reasonable to be used than those with feature set A determined by GFA (Table S5 in the Supporting Information and Figure 2b). If AUC higher than 0.70 was used as the criterion to judge a successful prediction, 5 (18%) of the most selective pharmacophore models yield acceptable predictions while 15 (55%) of the pharmacophore models with the best discrimination capabilities yield acceptable predictions. The most selective pharmacophore model was optimized by GFA, but it should be cautiously used in virtual screening when there are not enough known inhibitors for the validation. The aim of virtual screening is to find more promising compounds (true positives) and avoid false positives from the huge chemical space. The most distinguishing pharmacophore models, which are validated by the validation data set, are more suitable to be used in structure-based virtual screening. Obviously, the validation process before any pharmacophore modeling study for TCM is quite necessary.

Bayesian Classifier by Combining the Predictions from Molecular Docking and Pharmacophore Mapping. As mentioned above, molecular docking gives good predictions for 21 targets, and pharmacophore mapping gives acceptable predictions for 15 targets. Both of molecular docking (RMSD $\leq$ 2.0 Å and P-value $\leq 10^{-20}$) and pharmacophore mapping (AUC $\geq$ 0.7) give good predictions for eight targets, and five targets cannot be well-predicted by either molecular docking or pharmacophore mapping. However, molecular docking or pharmacophore mapping can give relatively reliable predictions for 28 T2DM-related targets (28/33 = 84.8%). It must keep in mind that the virtual screening based on molecular docking or pharmacophore mapping is a binary classification problem. Then, we may raise the following question: could the combination of these two structure-based methods enhance the prediction? For each molecule, the docking score and fit value are the quantitative measurements for molecular docking and pharmacophore mapping, respectively. Therefore, the naïve Bayesian classifiers were built by using the docking scores and fit values as the independent variables for 24 targets with reliable predictions of molecular docking and/or phamacophore mapping. Three criteria should be satisfied for the chosen targets: (1) the number of known inhibitors should be equal to or higher than 10; (2) the nativelike conformation of the ligand should be reproduced by molecular docking (RMSD $\leq$ 2.0 Å); (3) a reasonable pharmacophore model should be generated. In order to compare the prediction capacity among molecular docking, pharmacophore modeling and the combination of the predictions from both, the AUC values with 5-fold cross validation were employed to measure the performance of the naïve Bayesian classifiers (Table 1 and Figure 3). According to statistical results shown in Table 1, we observe that the AUC values of the Bayesian classifiers based on the docking scores and fit values are higher than those only based on the docking scores or fit values for almost all targets (except PTP1B). For example, for insulin-like growth factor 1 receptor (IGF-1R), the AUC values for the Bayesian classifier based on the fit values and that based on the docking scores are 0.729 and 0.788, respectively, while the AUC value for the Bayesian classifier based on both the fit values and docking scores is 0.904 (Figure 4). The reason why the naïve Bayesian classifier based on both fit values and docking scores for PTP1B cannot get the improvement was also investigated. We found that only six known inhibitors could be docked into the binding site of PTP1B and mapped onto the pharmacophore model of the PTP1B complex, and a limited number of the PTP1B inhibitors could not guarantee the statistical reliability of the validation results.

Here, the AUC values in the range of 0.5−0.6, 0.6−0.7, 0.7−0.8, 0.8−0.9, and 0.9−1.0 represent fail, poor, fair, good, and excellent predictions for the Bayesian classifiers (Table S7 in the Supporting Information and Figure 3). The definition of AUC for indicating the performance of prediction models is consistent with the definition used by the *Discovery Studio* simulation package. The Bayesian classifiers built based on

Table 1. Performance of the Naïve Bayesian Classifiers by Using Fit Values, Docking Scores, and the Combination of Fit Values and Docking Scores for 24 Targets

				AUC		
no.	target name	N_{inh}	N_{non}	fit value	score	combined
1	carbonic anhydrase 1	52	1040	0.908	0.841	0.942
2	carbonic anhydrase 2	91	1820	0.803	0.500	0.867
3	growth factor receptor-bound protein 2	16	320	0.687	0.597	0.741
4	peroxisome proliferator-activated receptor gamma	500	10000	0.532	0.795	0.837
5	glycogen synthase kinase-3 β	69	1380	0.833	0.916	0.971
6	aldose reductase	145	2900	0.717	0.836	0.892
7	pancreatic α-amylase	22	440	0.795	0.743	0.865
8	peroxisome proliferator-activated receptor delta (PPAR-δ)	224	4480	0.624	0.781	0.830
9	pyruvate dehydrogenase [lipoamide] kinase isozyme 1	59	1180	0.641	0.573	0.672
10	insulin receptor	370	7400	0.614	0.741	0.814
11	insulin-like growth factor 1 receptor	177	2540	0.788	0.729	0.904
12	mRNA of protein-tyrosine phosphatase, nonreceptor type 1 (PTP1B)	34	680	0.604	0.759	0.596
13	peroxisome proliferator-activated receptor α (PPAR-α)	263	5260	0.533	0.809	0.847
14	glucokinase (hexokinase D)	314	6280	0.624	0.856	0.896
15	protein kinase C, alpha type	500	10000	0.572	0.600	0.683
16	mitogen-activated protein kinase 10	500	10000	0.573	0.655	0.720
17	corticosteroid 11-beta-dehydrogenase, isozyme 1 (11-beta hydroxysteroid dehydrogenase 1)	500	10000	0.526	0.578	0.604
18	mitogen-activated protein kinase 8	500	10000	0.551	0.601	0.661
19	mitogen-activated protein kinase 1	130	2600	0.562	0.668	0.722
20	receptor protein-tyrosine kinase erbB-2	13	260	0.669	0.937	0.957
21	estrogen receptor	500	10000	0.676	0.897	0.940
22	glucocorticoid receptor	115	2300	0.628	0.736	0.788
23	fructose-1,6-bisphosphatase 1 (FBPase)	89	1780	0.740	0.821	0.856
24	phosphatidylinositol-4,5-bisphosphate 3-kinase catalytic subunit, delta isoform (phosphoinositide 3-kinase delta)	16	320	0.840	0.898	0.934

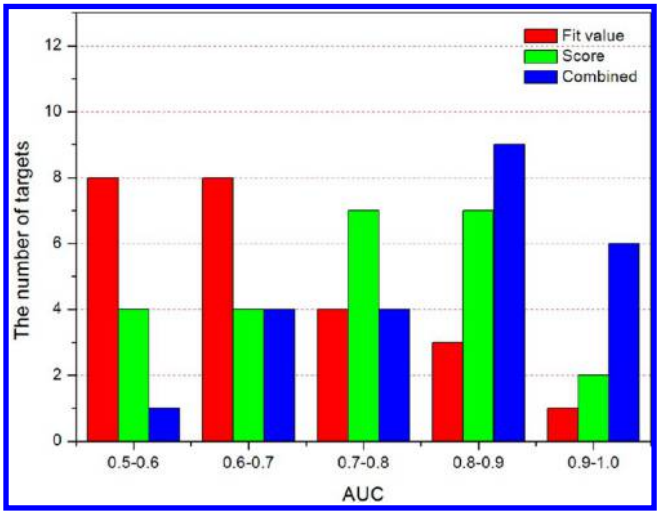


Figure 3. Comparison of the discrimination capability indicated by the AUC values for the naïve Bayesian classifiers based on fit values, docking scores, and the combination of fit values and docking scores.

docking scores or fit values can only give good or excellent predictions for 4 or 9 targets, while the combination of these two measurements can give good or excellent predictions for 15 targets. Clearly, the combination of docking scores and fit values can enhance the discrimination capacity.

Deep Analysis of the Potential Active Compounds for the T2DM-Related Targets. As we discussed in the above section, for 15 targets (15/24 = 62.5%), the Bayesian classifiers based on both docking scores and fit values give reliable predictions (AUC $\geq$ 0.8), and therefore, only the Bayesian classifier for these 15 targets were employed to virtually screen the TCM compound data set.

The numbers of the predicted inhibitors are listed in Table S8 in the Supporting Information. For protein-tyrosine kinase erbB-2, no potential inhibitor was found by the Bayesian classifier. For protein-tyrosine kinase erbB-2, there are only ten inhibitors for building and validating the Bayesian classifier. As shown in Figure S2 in the Supporting Information, these ten inhibitors almost have the same molecular framework (Murcko framework).[61] Therefore, the conservation of the target–inhibitor interaction patterns may be unfavorable to generate a general prediction model for finding other dissimilar compounds from the TCM compound data set. Then, we investigated the correlation between the number of the potential inhibitors and the volume and hydrophobic property of the binding pocket for each target (Table S8 in the Supporting Information) and did not find an obvious relationship between the number of the potential inhibitors for each target and the volume or hydrophobic property of the binding pocket.

It is well-known that two similar compounds or two compounds that share similar framework usually have similar pharmacological effects.[62] Therefore, the similarities between the potential inhibitors from TCM predicted by the Bayesian classifiers and the known inhibitors for each target (Table S8 in the Supporting Information) were evaluated by the 2D similarities (*Tanimoto Coefficient*) based on the MDLPubicKeys fingerprints[63] supported in DS3.1. The results are summarized in Table S9 in the Supporting Information. It is encouraging to find that a number of the potential inhibitors are similar to the known inhibitors for 12 of 15 targets when

dx.doi.org/10.1021/ci400146u | J. Chem. Inf. Model. 2013, 53, 1787–1803

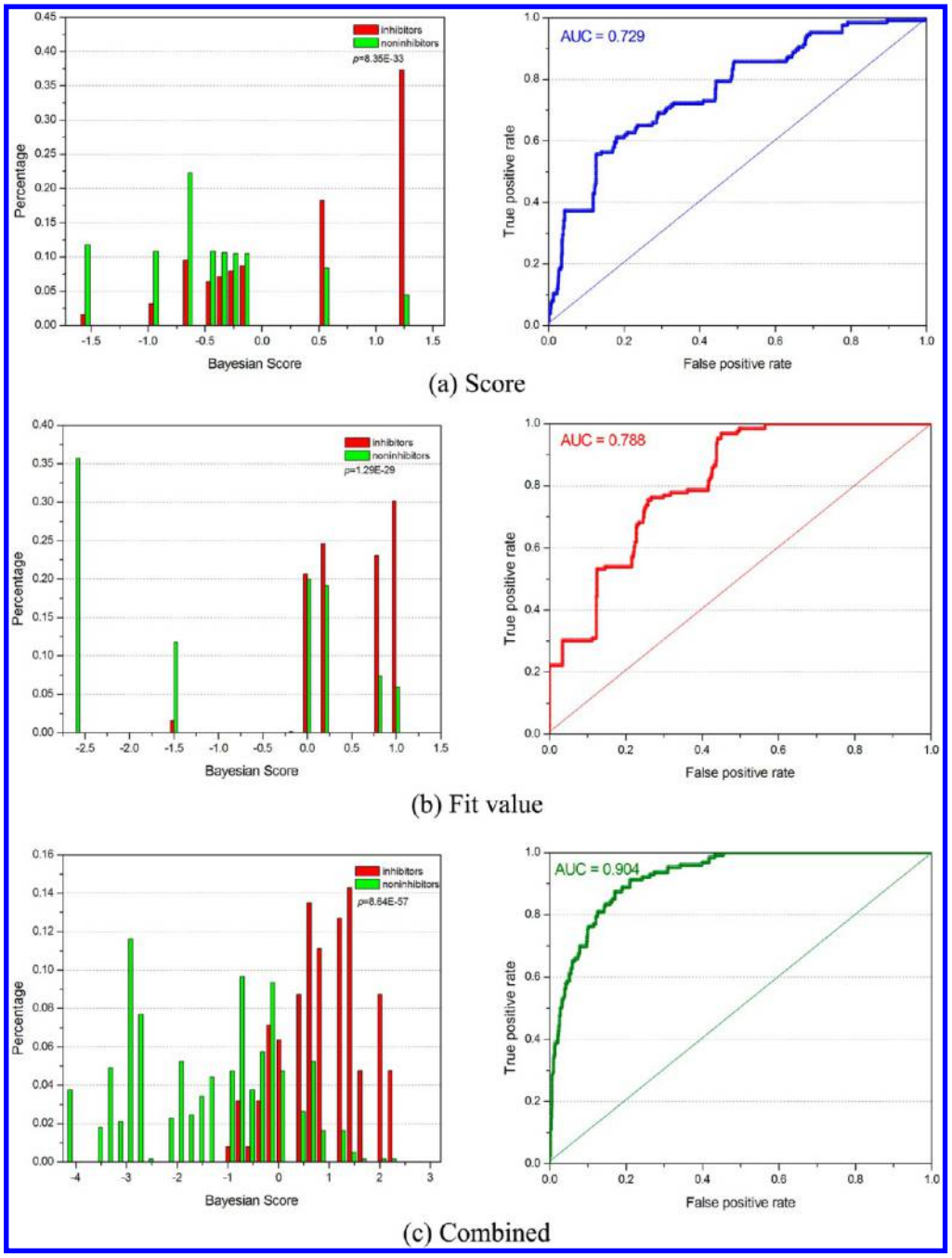


Figure 4. Distributions of Bayesian scores of the inhibitors and noninhibitors and the AUC value of the naïve Bayesian classifiers based on (a) docking scores, (b) fit values, and (c) the combination of docking scores and fit values for insulin-like growth factor 1 receptor (IGF-1R).

the similarity coefficient higher than 0.6 was used as the threshold. In addition, for insulin receptor and estrogen receptor, 7 and 11 potential inhibitors have the similarity coefficient of 1.0 to the known inhibitors. For estrogen receptor, 9 of the 11 potential inhibitors are the known inhibitors (Figure 5a), and the other 2 are quite similar to the known inhibitors of estrogen receptor (Figure 5b).

It is interesting to find that all of the eight compounds shown in Figure 5a were directly isolated from natural products. Be specific, compounds 2 (Emodin), 6 (Genistein), and 7 (Daidzein) were isolated from *Polygonum cuspidatum (polygonaceae)*,[64] compound 8 (Broussonin A) was from *Broussonetia Kazinoki*,[65] compound 4 (8-PN) was from *Humulus lupulus L.*,[66] compound 3 (Coumestrol) was from plant-based natural molecules or derivatives,[67] and compounds 1 (Resveratrol)[68] and 5 (Apigenin)[69] were also from natural plants. It should be noted that all of the eight compounds are not only found in the natural plants mentioned above but also found in the Chinese herbs in the classical TCM formulas for treating T2DM. Furthermore, for insulin receptor, we can also observe that 21 predicted inhibitors have the similarity coefficients higher than 0.9 to the known inhibitors of insulin receptor (Table S9 in the Supporting Information). The structural analysis shows that 18 of the 21 molecules are quite similar to the known inhibitor, *myricetin*, of insulin receptor. The *myricetin* and the 18 potential inhibitors similar to *myricetin* are shown in Figure 6. Obviously, these 18 molecules share the same molecular framework

dx.doi.org/10.1021/ci400146u | J. Chem. Inf. Model. 2013, 53, 1787−1803

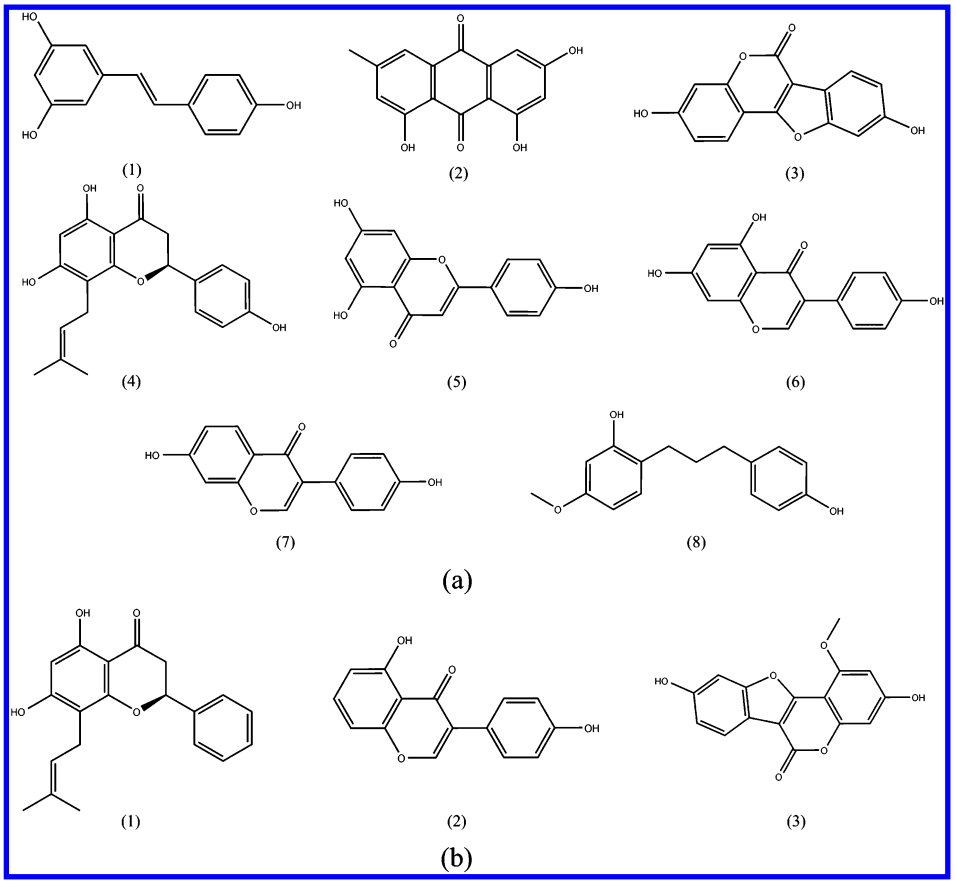


Figure 5. (a) Eight potential inhibitors predicted by the Bayesian classifier are identical to the known inhibitors of estrogen receptor and (b) three potential inhibitors predicted by the Bayesian classifier are similar to the known inhibitors of estrogen receptor.

(Murcko framework)[61] with *myricetin* and can be seen as the derivatives of *myricetin*.

In summary, our results imply that the Bayesian classifiers by combining the predictions from molecular docking and pharmacophore mapping have exciting capability to find promising inhibitors from TCM for the T2DM-related targets. Many potential inhibitors predicted by the Bayesian classifiers are quite similar or even identical to the known inhibitors. Moreover, by the 2D similarity analysis, we can observe that some potential inhibitors are dissimilar (low similarity) to the known inhibitors for most targets and they may be novel lead compounds for the studied targets. Our observation can provide a deep insight for understanding the action mechanism of TCM formulas for T2DM. The main mechanism of the TCM formulas for treating T2DM may be similar to western drugs: the active compounds or inhibitors in TCM can interact with the important disease-related targets, forming the leading force for combating diabetes.

Analysis of the Compound–Target Interaction Network. A TCM formula usually consists of various herbs with numerous compounds to exert potency toward diseases. The action mechanism of TCM is also termed as "cocktails of drugs", which represents multiple components interacting with multiple targets. Two possible mechanisms may be proposed for the pharmacological effect of the TCM formulas for treating T2DM: (1) multiactive-components interacting with different T2DM-related targets to achieve synergic effect and (2) limited individual components interacting with multiple targets, most of them having activities against single specific T2DM-related targets.

After employing the Bayesian classifiers to virtually screen the TCM compound data set, 1996 nonduplicated compounds were determined to be potential inhibitors for 14 targets. Among these 1996 compounds, 1590 (74.38%) of them were predicted to be drug-like by using the drug-likeness model developed in our group.[45] The distributions of eight important molecular physicochemical properties, including MW, PSA, Alog*P*, log*D*, log*S*, the number of hydrogen bond acceptors/donors, and the number of rotatable bonds for these compounds are displayed in Figure S3 in the Supporting Information. In order to achieve more reliable predictions, only the top 10% of the compounds (693) ranked by decreasing the Bayesian scores for each target were chosen for the following analysis (Table S8 in the Supporting Information). The numbers of the predicted potential inhibitors for each target are listed in Figure 7.

It is interesting to find that there are a number of potential inhibitors that interact with multiple targets (Figure 8 and Table S10 in the Supporting Information). There are 519 nonduplicated predicted inhibitors against a single target, 94 against two targets, 28 against three targets, 8 against four targets, and only one against five targets. The compound–target interaction networks are shown in Figure 9. As can be seen from Figure 8, the number of the potential inhibitors decreases rapidly with the increase of the number of targets. Apparently, most compounds in Chinese herbs have potential activities against limited targets. As can be seen in Figure 9a, the whole compound–target network has 519 nodes (potential inhibitors) and 696 edges (interactions) with an average of 1.34 interactions per node. About 81.39% (564/693) of the

dx.doi.org/10.1021/ci400146u | *J. Chem. Inf. Model.* 2013, 53, 1787–1803

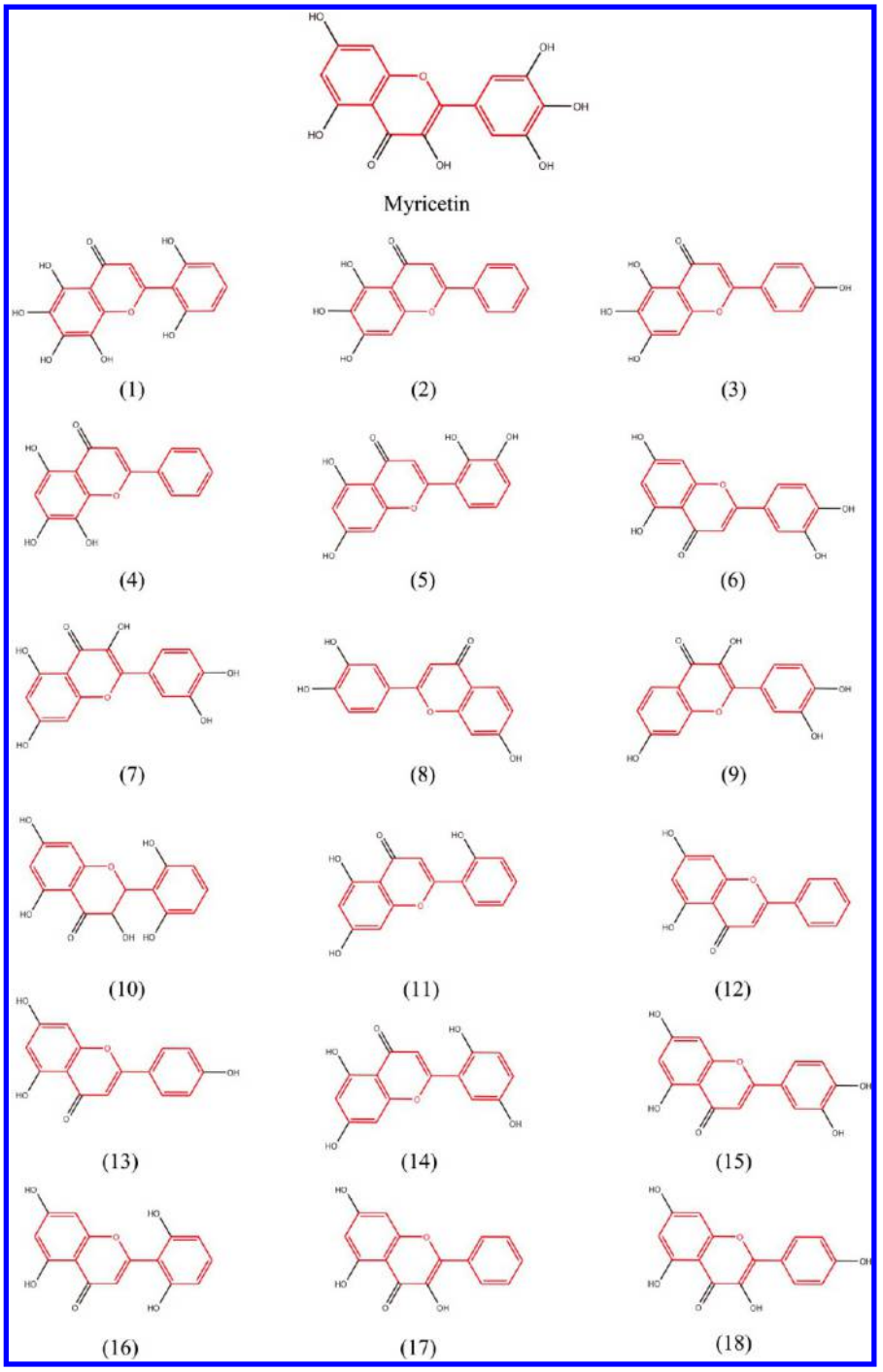


Figure 6. Eighteen potential inhibitors from the TCM formulas predicted by the Bayesian classifier are similar to the known inhibitor, *myricetin*, of insulin receptor.

potential inhibitors only interact with one target. According to our analysis, we can conclude that the mechanism of the TCM formulas may belong to the second hypothesis: limited individual components interact with multitargets and the others only interact with the single specific T2DM-related targets.

We then investigated the potential multitarget inhibitors predicted by the Bayesian classifiers. The similarities between the potential multitarget inhibitors and the known inhibitors for each target were calculated. When the minimum similarity coefficient based on the MDLPubicKeys fingerprint was set to 0.6, no similar known inhibitor was found for these potential inhibitors for all of four and five targets, but similar known inhibitors were successfully found for some potential inhibitors for all of two and three targets.

For the only potential inhibitor (mirificin) against five targets (carbonic anhydrase 1, carbonic anhydrase 2, FBPase, glycogen synthase kinase-3 beta, and insulin receptor), similar known inhibitors for three of five targets (carbonic anhydrase 1, carbonic anhydrase 2, and insulin receptor) could be found. For the eight potential inhibitors against four targets, similar known inhibitors for two or three of four targets could be found. Besides, for the predicted inhibitors against two and three targets, similar known inhibitors for all of two or three targets

dx.doi.org/10.1021/ci400146u | J. Chem. Inf. Model. 2013, 53, 1787−1803

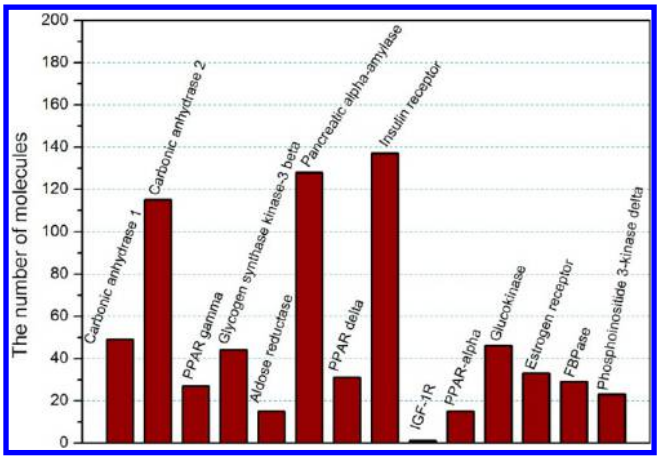


Figure 7. Number of the potential inhibitors ranked by the Bayesian scores for establishing the compounds−targets network.

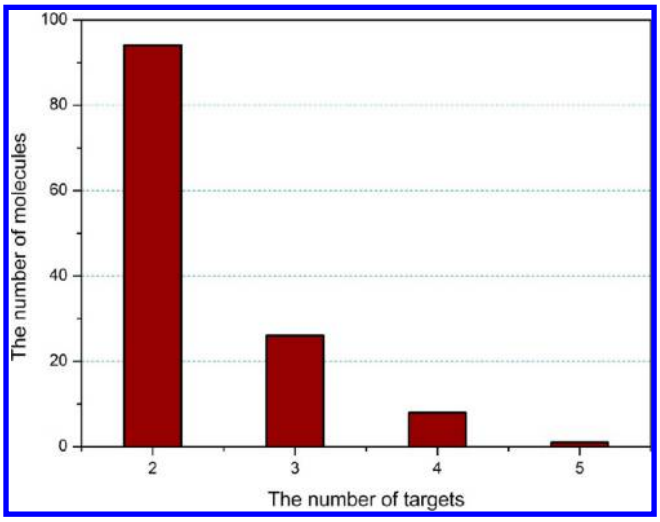


Figure 8. Distribution of the number of the potential multitarget inhibitors.

could be found. For example, the compound (Pureoside A) that was predicted to be a potential inhibitor of carbonic anhydrase 1, PPAR delta, and PPAR gamma and similar known inhibitors for each target were found and shown in Figure 10; two compounds (Epicatechin and (2*RS*)-2-(4-*O*-*β*-D-glucopyranosylphenyl) propionyl *β*-D-glucopyranoside) that were predicted to be potential inhibitors for two targets and similar known inhibitors were found and shown in Figure 11.

In summary, our predictions suggest that most of the compounds extracted from Chinese herbs in TCM formulas can only interact with a single specific target. Although the theory of TCM formulas is multiple compounds for multiple targets, according to our analysis, it is obvious that only limited compounds from the TCM formulas for T2DM show activities against multiple targets. The 2D-similarity analysis shows that some known inhibitors are similar to the predicted potential inhibitors against multitargets. Furthermore, the low similarities between some of the predicted potential inhibitors, and the known inhibitors for each target suggest that some novel potential inhibitors might be found from the predictions given by the Bayesian classifiers. Those novel potential inhibitors may provide good source for pharmaceutical chemists to find promising leads for treating T2DM.

Possible Mechanisms of the Classical TCM Formulas for Treating T2DM. As discussed in the above section, it is demonstrated that the 519 nonduplicated potential inhibitors can interact with 14 T2DM-related targets. More than 80% of the potential inhibitors only interact with one target, and the other potential inhibitors interact with two or more targets. The multiple compounds from one herb or different herbs of TCM formulas may form synergistic interactions for combating T2DM. Then, the pharmacological activities of those predicted inhibitors that can interact with two or more targets were examined by extensive literature searching. First, we analyzed the nine compounds that interact with at least four targets (Table S10 in the Supporting Information). The only compound, mirificin, which was predicted to be an inhibitor of five targets (carbonic anhydrase 1, carbonic anhydrase 2, FBPase, glycogen synthase kinase-3 beta, and insulin receptor) has free radical scavenging activity.[70] Moreover, the eight compounds that are predicted to be the inhibitors of four targets have reported activities. Interestingly, the reported activities for three compounds (Apiopaeonoside, Kakkalide, and tectorigenin 7-*O*-xylosylglucoside) are not related to the studied targets while related to aldose reductase.[71,72] Previous studies have shown the close relationship between aldose reductase and free radical scavenging/antioxidant.[73] Aldose reductase might induce the oxidative damage resulted from the competition between aldose reductase and glutathione reductase for NADPH. Furthermore, the inhibitors of aldose reductase not only show activities for aldose reductase, but also have antioxidant activity.[74] By checking the top 10% of the predicted inhibitors of aldose reductase (Table S8 in the Supporting Information), we also found that 6 of the 15 compounds have free radical scavenging/antioxidant activities. According to those observations, the linkage between the three potential inhibitors for four targets and the one potential inhibitor for five targets can be established. Although the predicted potential inhibitors for more than four targets do not have reported activities as predicted by us, all of those four compounds have the reported activities of free radical scavenging or antioxidant. Besides, the previous studies have shown that there were direct relationship between free radical scavenging/antioxidant and T2DM.[75] The level of free radical increases with the T2DM disease progressing,[76] and, in addition, some complications of T2DM, including obesity, retinopathy and nephropathy, and some T2DM-related diseases, such as coronary arteriosclerotic heart disease and arteriosclerosis, are also related to the out of control of free radical.[77−81] At last, it is necessary to emphasize that some potential inhibitors that can interact only one, two, or three targets also possess the reported activities of free radical scavenging/antioxidant, for example, the compound, apigenin-7-glucoside, a potential inhibitor of two targets (insulin receptor and pancreatic alpha-amylase). Actually, the activity of apigenin-7-glucoside against pancreatic alpha-amylase has been reported;[82] in addition, this compound also has aldose reductase and free radical scavenging/antioxidant activities.[83−85] The compound 4′-*O*-*β*-glucopyranosyl-3′,5,7-trihydroxyflavone with predicted activities for carbonic anhydrase 1 and pancreatic alpha-amylase targets has aldose reductase and free radical scavenging activities.[84,86] For the compound taxifolin, a predicted inhibitor against three targets (carbonic anhydrase 2, estrogen receptor, and pancreatic alpha-amylase), the experimental activity against estrogen receptor has been reported,[87] and this compound also shows free radical

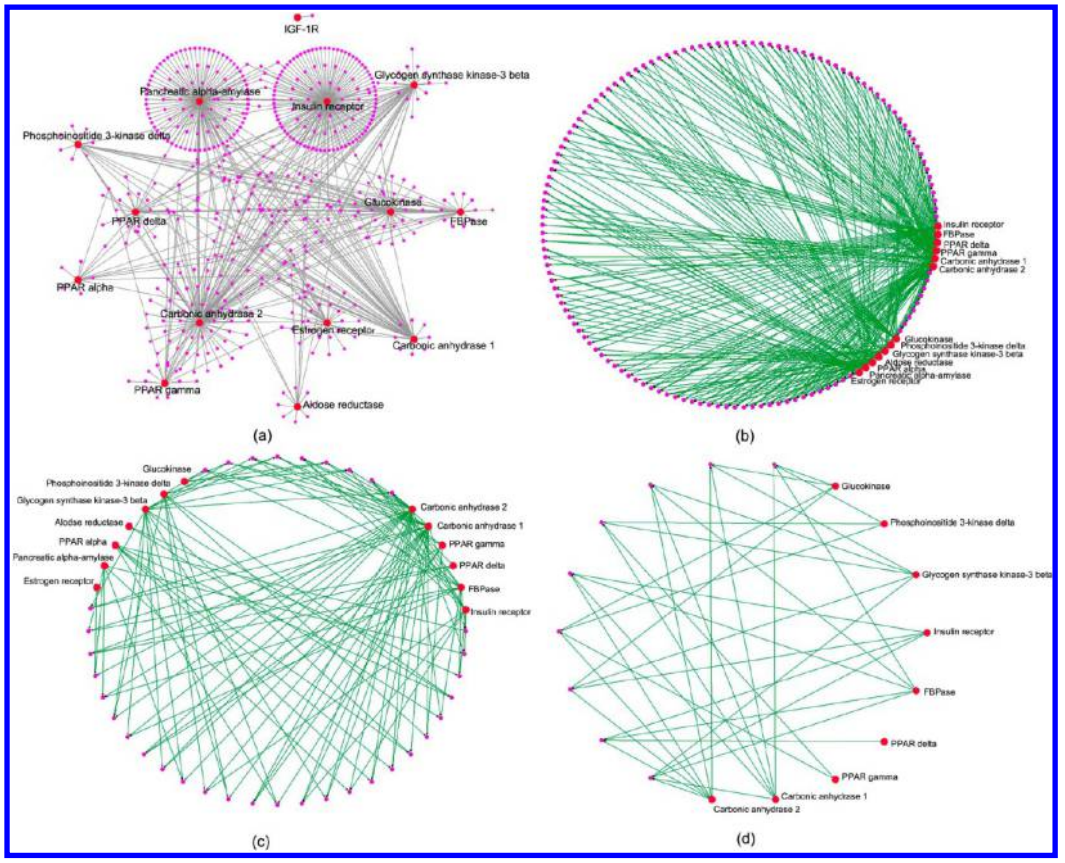


Figure 9. (a) Compound−target network for all potential inhibitors, (b) for those against no less than two targets, (c) for those against no less than three targets, and (d) for those against no less than four targets.

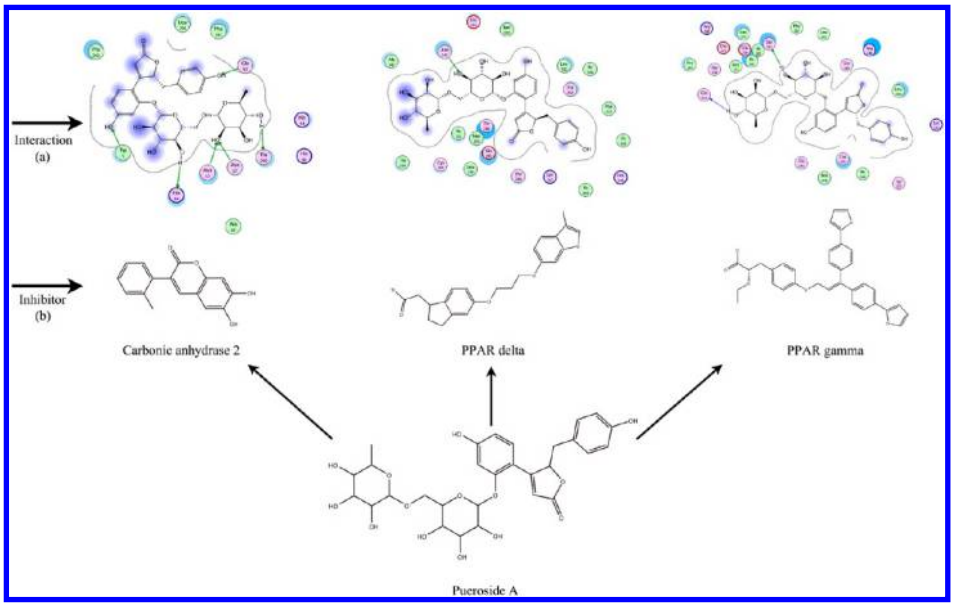


Figure 10. Molecule Pueroside A was predicted to an inhibitor of three targets (carbonic anhydrase, PPAR delta and PPAR gamma). (a) Schematic representation of the interaction between Pueroside A and three targets and (b) the known inhibitors that have the maximum similarity to Pueroside A based on the MDLPulicKeys fingerprints for three targets, respectively.

scavenging/antioxidant,[88,89] PI3K inhibition,[90] and PPAR gamma inhibition activities.[91] Moreover, we also found that some potential multitarget inhibitors did not have any documented activities for the T2DM-related targets. For example, the compound lobetyolin was predicted to be a potential inhibitor of carbonic anhydrase 2, insulin receptor, and pancreatic alpha-amylase targets, but the reported activity of lobetyolin is antibacterial.[92] It was interesting to find that the known drug of T2DM, glibenclamide, also has antibacterial activity. The antibacterial activity is also found to be valuable in relieving diabetes and the associated secondary disorders.[93] According to our observations, we can find that some of the

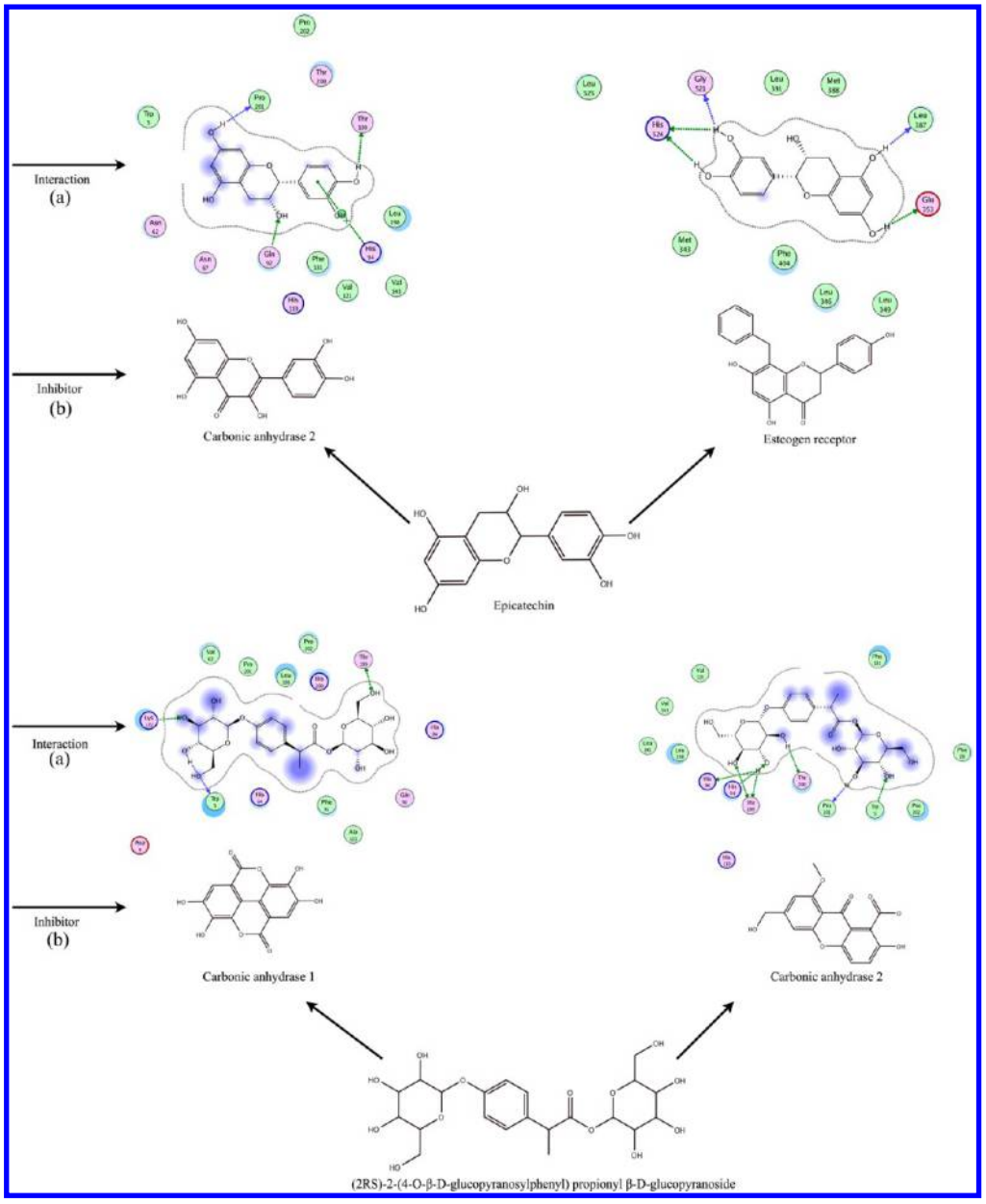


Figure 11. Molecules, Epicatechin and (2*RS*)-2-(4-*O*-*β*-D-glucopyranosylphenyl) propionyl *β*-D-glucopyranoside were predicted to be the inhibitors of two targets. (a) Schematic representation of the interactions between Epicatechin and carbonic anhydrase 2 and estrogen receptor and those between (2*RS*)-2-(4-*O*-*β*-D-glucopyranosylphenyl) propionyl *β*-D-glucopyranoside and carbonic anhydrase 1 and carbonic anhydrase 2 targets and (b) the known inhibitors of carbonic anhydrase 2 and estrogen receptor that have the maximum similarity to Epicatechin; the known inhibitors of carbonic anhydrase 1 and carbonic anhydrase that have the maximum similarity to (2*RS*)-2-(4-*O*-*β*-D-glucopyranosylphenyl) propionyl *β*-D-glucopyranoside based on the MDLPublicKeys fingerprints.

predicted potential inhibitors have reported pharmacological activities related to T2DM for more than one target.

In summary, we have examined the pharmacological activities of the multitargets compounds extensively. It was found that some pharmacological effects of the predicted inhibitors of multitargets have been reported previously, highlighting the reliable predictions of the Bayesian classifiers. Furthermore, some potential inhibitors predicted by our study need experimental verifications. Then, during the analysis of the pharmacological activities of potential multitarget inhibitors, it was found that most of these potential compounds have free radical scavenging/antioxidant activities that have been proved to be effective to relieve T2DM and the related diseases of T2DM.

The theory of TCM formulas is multicompounds versus multitargets. The hundreds or even thousands of compounds extracted from the herbs of TCM formulas interact with multitargets to achieve synergistic effects. How the compounds from the herbs of TCM formulas affect the T2DM-related targets is also unclear and needs to be uncovered. According to our analysis, it was found that most of the predicted potential inhibitors can only interact with a single target. However, a few of them still can interact with more than one target. Some pharmacological effects, such as radical scavenging/antioxidant and antibacterial activities, which are not directly related to the studied targets, were found. Most compounds from the herbs in the TCM formulas only interact with an individual target, forming the leading fighting force to combat T2DM. Then,

dx.doi.org/10.1021/ci400146u | *J. Chem. Inf. Model.* 2013, 53, 1787−1803

those potential multitarget compounds may influence the T2DM-related targets, forming additional forces to enhance the therapeutic effects. At last, a portion of the compounds are responsible for remedying the other related symptoms that are proved to be related to T2DM, such as abnormal status of free radical and oxidation in the human body. All of these observations found in this study can be seen as a proper way to reveal the classical theory "Monarch, Minster, Assistant, and Guide" in TCM prescriptions at the molecular level.

CONCLUSION

In this study, we employed structure-based and complex-based virtual screening approaches to find potential inhibitors from the Chinese herbs of the classical TCM formulas for T2DM against the important T2DM-related targets. In order to achieve more reliable predictions, the performance of molecular docking and pharmacophore mapping for the 33 T2DM-related targets was evaluated. By combining the advantages of molecular docking and pharmacophore mapping together, the Bayesian classifiers were developed by integrating the predictions from molecular docking and pharmacophore mapping. The validation results show that the integrated models are superior to that based on the prediction from molecular docking or pharmacophore mapping only.

The Bayesian classifiers with satisfactory discrimination capabilities for 15 targets were employed to screen the 2479 compounds from the Chinese herbs of the TCM formulas for T2DM. The pharmaceutical activities of some compounds predicted by the Bayesian classifiers are confirmed by the experimental studies. In order to uncover the underlying mechanism of the TCM formulas for combating T2DM, the top 10% of the potential compounds ranked by the Bayesian scores for each target were extracted to build the compound–target interaction network. The analysis of the compound–target network shows that ~19% of the potential inhibitors interact with more than one target. The examination of the pharmacological activities demonstrates that the biological activities for a small portion of these potential multitarget inhibitors have been reported, but the underlying activities for most potential multitarget inhibitors need experimental verification. Besides, it is important for us to find that a variety of the potential multitarget inhibitors have free radical scavenging/antioxidant activities, which are closely related to T2DM. We can conclude that the pharmacological effects of the TCM formulas for T2DM are not only determined by the compounds that interact directly with one or more T2DM-related targets, but also by the compounds with other supplementary bioactivities important for treating T2DM, such as free radical scavenging/antioxidant activities. The holistic way to express join forces for treating T2DM may be the essence of the TCM formulas to break the robustness of the phenotype of T2DM and restore the human body to normal physiological condition.

ASSOCIATED CONTENT

Supporting Information

Table S1: names, PDB entries, and crystal resolutions for the 48 T2DM-related targets. Table S2: 32 Chinese herbs in the TCM classical formulas for T2DM with the highest frequency. Table S3: docking power and discrimination power of the *Glide* docking for 33 T2DM-related targets. Table S4: docking power of the *Glide* docking based on the six crystal structures of the complexes of HMG-CoA reductase. Table S5: numbers of total features, feature set A, and feature set B and the selectivity and discrimination capabilities of the pharmacophore models based on feature sets A and B. Table S6: comparisons of the number of pharmacophore features in total features, feature set A, and feature set B. Table S7: comparison of the performance of the naïve Bayesian classifiers built by using docking scores, fit values, and the combination of docking scores and fit values. Table S8: volume and hydrophobic properties of the binding pockets and the numbers of the potential inhibitors predicted by the naïve Bayesian classifiers for 15 targets. Table S9: number of the potential inhibitors from the TCM formulas that are similar to the known inhibitors for 15 targets using different thresholds of the similarity based on the MDLPublicKeys fingerprints. Table S10: number of the predicted inhibitors that can interact with two, three, four, and five targets. Figure S1: distributions of the docking scores of the known inhibitors and noninhibitors by using the (a) SP and (b) XP scoring modes of *Glide* for 1DQA of HMG-CoA reductase. Figure S2: ten known inhibitors of receptor protein-tyrosine kinase erbB-2 target with similar Murcko framework (colored in red). Figure S3: distributions of the eight important physicochemical properties for the 1996 nonduplicated potential inhibitors for 15 targets. This material is available free of charge via the Internet at http://pubs.acs.org.

AUTHOR INFORMATION

Corresponding Author

*E-mail: tingjunhou@hotmail.com or tingjunhou@zju.edu.cn. Phone: +86-512-65882039.

Notes

The authors declare no competing financial interest.

ACKNOWLEDGMENTS

This study was supported by the National Science Foundation of China (21173156), the National Basic Research Program of China (973 program, 2012CB932600), and the Priority Academic Program Development of Jiangsu Higher Education Institutions (PAPD)

REFERENCES

(1) Smyth, S.; Heron, A. Diabetes and obesity: the twin epidemics. *Nature Med.* **2006**, *12*, 75–80.
(2) Wild, S.; Roglic, G.; Green, A.; Sicree, R.; King, H. Global prevalence of diabetes - Estimates for the year 2000 and projections for 2030. *Diabetes Care* **2004**, *27*, 1047–1053.
(3) Florez, H. J.; Sanchez, A. A.; Marks, J. B. Type 2 Diabetes. *Diabetes and the Brain* **2010**, 33–53.
(4) Pasquier, F. Diabetes and cognitive impairment: how to evaluate the cognitive status? *Diabetes Metab.* **2010**, *36*, S100–S105.
(5) Reaven, G. M. Pathophysiology of insulin-resistance in human-disease. *Physiol. Rev.* **1995**, *75*, 473–486.
(6) Ripsin, C. M.; Kang, H.; Urban, R. J. Management of Blood Glucose in Type 2 Diabetes Mellitus. *Am. Fam. Physician* **2009**, *79*, 29–36.
(7) Morral, N. Novel targets and therapeutic strategies for type 2 diabetes. *Trends Endocrin. Met.* **2003**, *14*, 169–175.
(8) Marcus, A. O. Safety of drugs commonly used to treat hypertension, dyslipidemia, and type 2 diabetes (the metabolic syndrome): part 1. *Diabetes Technol. Ther.* **2000**, *2*, 101–10.
(9) Marcus, A. O. Safety of drugs commonly used to treat hypertension, dyslipidemia, and Type 2 diabetes (the metabolic syndrome): part 2. *Diabetes Technol. Ther.* **2000**, *2*, 275–81.
(10) Hopkins, A. L. Network pharmacology: the next paradigm in drug discovery. *Nature Chem. Biol.* **2008**, *4*, 682–690.

(11) Roth, B. L.; Sheffler, D. J.; Kroeze, W. K. Magic shotguns versus magic bullets: selectively non-selective drugs for mood disorders and schizophrenia. *Nat. Rev. Drug Discovery* **2004**, *3*, 353−359.
(12) Hou, T.; Wang, J. Structure - ADME relationship: still a long way to go? *Expert Opin. Drug Metab. Toxicol.* **2008**, *4*, 759−770.
(13) Hou, T. J.; Xu, X. J. Recent development and application of virtual screening in drug discovery: An overview. *Curr. Pharm. Des.* **2004**, *10*, 1011−1033.
(14) Winzeler, E. A.; Shoemaker, D. D.; Astromoff, A.; Liang, H.; Anderson, K.; Andre, B.; Bangham, R.; Benito, R.; Boeke, J. D.; Bussey, H.; Chu, A. M.; Connelly, C.; Davis, K.; Dietrich, F.; Dow, S. W.; El Bakkoury, M.; Foury, F.; Friend, S. H.; Gentalen, E.; Giaever, G.; Hegemann, J. H.; Jones, T.; Laub, M.; Liao, H.; Liebundguth, N.; Lockhart, D. J.; Lucau-Danila, A.; Lussier, M.; M'Rabet, N.; Menard, P.; Mittmann, M.; Pai, C.; Rebischung, C.; Revuelta, J. L.; Riles, L.; Roberts, C. J.; Ross-MacDonald, P.; Scherens, B.; Snyder, M.; Sookhai-Mahadeo, S.; Storms, R. K.; Veronneau, S.; Voet, M.; Volckaert, G.; Ward, T. R.; Wysocki, R.; Yen, G. S.; Yu, K. X.; Zimmermann, K.; Philippsen, P.; Johnston, M.; Davis, R. W. Functional characterization of the S-cerevisiae genome by gene deletion and parallel analysis. *Science* **1999**, *285*, 901−906.
(15) Kitano, H. Towards a theory of biological robustness. *Mol. Syst. Biol.* **2007**, *3*.
(16) Kitano, H. Innovation - A robustness-based approach to systems-oriented drug design. *Nat. Rev. Drug Discov.* **2007**, *6*, 202−210.
(17) Berger, S. I.; Iyengar, R. Network analyses in systems pharmacology. *Bioinformatics* **2009**, *25*, 2466−2472.
(18) Goh, K.-I.; Cusick, M. E.; Valle, D.; Childs, B.; Vidal, M.; Barabasi, A.-L. The human disease network. *Proc. Natl. Acad. Sci. U.S.A.* **2007**, *104*, 8685−8690.
(19) Yildirim, M. A.; Goh, K.-I.; Cusick, M. E.; Barabasi, A.-L.; Vidal, M. Drug-target network. *Nat. Biotechnol.* **2007**, *25*, 1119−1126.
(20) Newman, D. J.; Cragg, G. M. Natural Products As Sources of New Drugs over the 30 Years from 1981 to 2010. *J. Nat. Prod.* **2012**, *75*, 311−335.
(21) Zhao, J.; Jiang, P.; Zhang, W. Molecular networks for the study of TCM Pharmacology. *Brief. Bioinform.* **2009**, *11*, 417−430.
(22) Ji, H.-F.; Li, X.-J.; Zhang, H.-Y. Natural products and drug discovery Can thousands of years of ancient medical knowledge lead us to new and powerful drug combinations in the fight against cancer and dementia? *EMBO Rep.* **2009**, *10*, 194−200.
(23) Newman, D. J.; Cragg, G. M.; Snader, K. M. The influence of natural products upon drug discovery. *Nat. Prod. Rep.* **2000**, *17*, 215−234.
(24) Lukman, S.; He, Y.; Hui, S.-C. Computational methods for Traditional Chinese Medicine: A survey. *Comput. Methods Programs Biomed.* **2007**, *88*, 283−294.
(25) Chang, S.-S.; Huang, H.-J.; Chen, C.Y.-C. Two Birds with One Stone? Possible Dual-Targeting H1N1 Inhibitors from Traditional Chinese Medicine. *PLoS Comput. Biol.* **2011**, *7*.
(26) Chen, K.-C.; Chang, S.-S.; Tsai, F.-J.; Chen, C.Y.-C. Han ethnicity-specific type 2 diabetic treatment from traditional Chinese medicine? *J. Biomol. Struct. Dyn.* **2012**, 1−17.
(27) Tou, W. I.; Chang, S.-S.; Lee, C.-C.; Chen, C.Y.-C. Drug Design for Neuropathic Pain Regulation from Traditional Chinese Medicine. *Sci. Rep.* **2013**, 3.
(28) Chen, K.-C.; Chang, K.-W.; Chen, H.-Y.; Chen, C.Y.-C. Traditional Chinese medicine, a solution for reducing dual stroke risk factors at once? *Mol. Biosyst.* **2011**, *7*, 2711−2719.
(29) Chen, X.; Ung, C. Y.; Chen, Y. Z. Can an in silico drug-target search method be used to probe potential mechanisms of medicinal plant ingredients? *Nat. Prod. Rep.* **2003**, *20*, 432−444.
(30) Ehrman, T. M.; Barlow, D. J.; Hylands, P. J. In silico search for multi-target anti-inflammatories in Chinese herbs and formulas. *Bioorg. Med. Chem.* **2010**, *18*, 2204−2218.
(31) Gu, J.; Zhang, H.; Chen, L.; Xu, S.; Yuan, G.; Xu, X. Drug-target network and polypharmacology studies of a Traditional Chinese Medicine for type II diabetes mellitus. *Comput. Biol. Chem.* **2011**, *35*, 293−297.
(32) Harvey, A. L.; Clark, R. L.; Mackay, S. P.; Johnston, B. F. Current strategies for drug discovery through natural products. *Expert Opin. Drug Discov.* **2010**, *5*, 559−568.
(33) Schuster, D.; Wolber, G. Identification of Bioactive Natural Products by Pharmacophore-Based Virtual Screening. *Curr. Pharm. Des.* **2010**, *16*, 1666−1681.
(34) Tou, W. I.; Chen, C.Y.-C. In Silico Investigation of Potential Src Kinase Ligands from Traditional Chinese Medicine. *PLoS One* **2012**, *7*.
(35) Li, W. L.; Zheng, H. C.; Bukuru, J.; De Kimpe, N. Natural medicines used in the traditional Chinese medical system for therapy of diabetes mellitus. *J. Ethnopharmacol.* **2004**, *92*, 1−21.
(36) Herrick, T. M.; Million, R. P. Tapping the potential of fixed-dose combinations. *Nat. Rev. Drug Discovery* **2007**, *6*, 513−514.
(37) Zhu, F.; Shi, Z.; Qin, C.; Tao, L.; Liu, X.; Xu, F.; Zhang, L.; Song, Y.; Liu, X.; Zhang, J.; Han, B.; Zhang, P.; Chen, Y. Therapeutic target database update 2012: a resource for facilitating target-oriented drug discovery. *Nucleic Acids Res.* **2012**, *40*, D1128−D1136.
(38) Wishart, D. S.; Knox, C.; Guo, A. C.; Cheng, D.; Shrivastava, S.; Tzur, D.; Gautam, B.; Hassanali, M. DrugBank: a knowledgebase for drugs, drug actions and drug targets. *Nucleic Acids Res.* **2008**, *36*, D901−D906.
(39) Kanehisa, M.; Goto, S.; Sato, Y.; Furumichi, M.; Tanabe, M. KEGG for integration and interpretation of large-scale molecular data sets. *Nucleic Acids Res.* **2012**, *40*, D109−D114.
(40) Moller, D. E. New drug targets for type 2 diabetes and the metabolic syndrome. *Nature* **2001**, *414*, 821−827.
(41) Berman, H. M.; Westbrook, J.; Feng, Z.; Gilliland, G.; Bhat, T. N.; Weissig, H.; Shindyalov, I. N.; Bourne, P. E. The Protein Data Bank. *Nucleic Acids Res.* **2000**, *28*, 235−242.
(42) Qiao, X. B.; Hou, T. J.; Zhang, W.; Guo, S. L.; Xu, S. J. A 3D structure database of components from Chinese traditional medicinal herbs. *J. Chem. Inf. Comput. Sci.* **2002**, *42*, 481−489.
(43) Shen, M.; Tian, S.; Li, Y.; Li, Q.; Xu, X.; Wang, J.; Hou, T. Drug-likeness analysis of traditional Chinese medicines: 1. property distributions of drug-like compounds, non-drug-like compounds and natural compounds from traditional Chinese medicines. *J. Cheminform.* **2012**, *4*, 1−13.
(44) Tian, S.; Li, Y.; Wang, J.; Xu, X.; Xu, L.; Wang, X.; Chen, L.; Hou, T. Drug-likeness analysis of traditional Chinese medicines: 2. Characterization of scaffold architectures for drug-like compounds, non-drug-like compounds, and natural compounds from traditional Chinese medicines. *J. Cheminform.* **2013**, *5*, 1−14.
(45) Tian, S.; Wang, J.; Li, Y.; Xu, X.; Hou, T. Drug-likeness Analysis of Traditional Chinese Medicines: Prediction of Drug-likeness Using Machine Learning Approaches. *Mol. Pharm.* **2012**, *9*, 2875−86.
(46) Chen, C.Y.-C. TCM Database@Taiwan: The World's Largest Traditional Chinese Medicine Database for Drug Screening In Silico. *PloS One* **2011**, *6*.
(47) *Discovery Studio 3.1 Guide*; Accelrys Inc., San Diego, 2012; http://www.accelrys.com.
(48) Liu, T.; Lin, Y.; Wen, X.; Jorissen, R. N.; Gilson, M. K. BindingDB: a web-accessible database of experimentally determined protein-ligand binding affinities. *Nucleic Acids Res.* **2007**, *35*, D198−D201.
(49) Krueger, D. M.; Evers, A. Comparison of Structure- and Ligand-Based Virtual Screening Protocols Considering Hit List Complementarity and Enrichment Factors. *Chemmedchem* **2010**, *5*, 148−158.
(50) 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*, 6177−6196.
(51) *Schrödinger*, version 9.0; Schrödinger, LLC, New York, 2009; http://www.schrodinger.com.
(52) Kaminski, G. A.; Friesner, R. A.; Tirado-Rives, J.; Jorgensen, W. L. Evaluation and reparametrization of the OPLS-AA force field for

proteins via comparison with accurate quantum chemical calculations on peptides. *J. Phys. Chem. B* **2001**, *105*, 6474−6487.
(53) Sutter, J.; Li, J.; Maynard, A. J.; Goupil, A.; Luu, T.; Nadassy, K. New Features that Improve the Pharmacophore Tools from Accelrys. *Curr. Comput.-Aided Drug Des.* **2011**, *7*, 173−180.
(54) Rogers, D.; Hopfinger, A. J. Application of genetic function approximation to quantitative structure-activity-relationships and quantitative structure-property relationships. *J. Chem. Inf. Comput. Sci.* **1994**, *34*, 854−866.
(55) Meslamani, J.; Li, J.; Sutter, J.; Stevens, A.; Bertrand, H.-O.; Rognan, D. Protein-Ligand-Based Pharmacophores: Generation and Utility Assessment in Computational Ligand Profiling. *J. Chem. Inf. Model.* **2012**, *52*, 943−955.
(56) Chen, L.; Li, Y.; Zhao, Q.; Peng, H.; Hou, T. ADME Evaluation in Drug Discovery. 10. Predictions of P-Glycoprotein Inhibitors Using Recursive Partitioning and Naive Bayesian Classification Techniques. *Mol. Pharm.* **2011**, *8*, 889−900.
(57) Wang, S.; Li, Y.; Wang, J.; Chen, L.; Zhang, L.; Yu, H.; Hou, T. ADMET Evaluation in Drug Discovery. 12. Development of Binary Classification Models for Prediction of hERG Potassium Channel Blockage. *Mol. Pharm.* **2012**, *9*, 996−1010.
(58) Breitkreutz, B. J.; Stark, C.; Tyers, M. Osprey: a network visualization system. *Genome Biol.* **2003**, DOI: 10.1186/gb-2003-4-3-r22.
(59) Cheng, T.; Li, X.; Li, Y.; Liu, Z.; Wang, R. Comparative Assessment of Scoring Functions on a Diverse Test Set. *J. Chem. Inf. Model.* **2009**, *49*, 1079−1093.
(60) Cross, J. B.; Thompson, D. C.; Rai, B. K.; Baber, J. C.; Fan, K. Y.; Hu, Y.; Humblet, C. Comparison of Several Molecular Docking Programs: Pose Prediction and Virtual Screening Accuracy. *J. Chem. Inf. Model.* **2009**, *49*, 1455−1474.
(61) Bemis, G. W.; Murcko, M. A. The properties of known drugs 0.1. Molecular frameworks. *J. Med. Chem.* **1996**, *39*, 2887−2893.
(62) Yang, L.; Wang, K.; Chen, J.; Jegga, A. G.; Luo, H.; Shi, L.; Wan, C.; Guo, X.; Qin, S.; He, G.; Feng, G.; He, L. Exploring Off-Targets and Off-Systems for Adverse Drug Reactions via Chemical-Protein Interactome - Clozapine-Induced Agranulocytosis as a Case Study. *PLoS Comput. Biol.* **2011**, *7*.
(63) Durant, J. L.; Leland, B. A.; Henry, D. R.; Nourse, J. G. Reoptimization of MDL keys for use in drug discovery. *J. Chem. Inf. Comput. Sci.* **2002**, *42*, 1273−1280.
(64) Matsuda, H.; Shimoda, H.; Morikawa, T.; Yoshikawa, M. Phytoestrogens from the roots of Polygonum cuspidatum (Polygonaceae): Structure-requirement of hydroxyanthraquinones for estrogenic activity. *Bioorg. Med. Chem. Lett.* **2001**, *11*, 1839−1842.
(65) Lee, D. Y.; Kim, D. H.; Lee, H. J.; Lee, Y.; Ryu, K. H.; Jung, B.-I.; Song, Y. S.; Ryu, J.-H. New estrogenic compounds isolated from Broussonetia kazinoki. *Bioorg. Med. Chem. Lett.* **2010**, *20*, 3764−3767.
(66) Roelens, F.; Heldring, N.; Dhooge, W.; Bengtsson, M.; Comhaire, F.; Gustafsson, J.-A.; Treuter, E.; De Keukeleire, D. Subtle side-chain modifications of the hop phytoestrogen 8-prenylnaringenin result in distinct agonist/antagonist activity profiles for estrogen receptors alpha and beta. *J. Med. Chem.* **2006**, *49*, 7357−7365.
(67) Zhao, L. Q.; Brinton, R. D. Structure-based virtual screening for plant-based ER beta-selective ligands as potential preventative therapy against age-related neurodegenerative diseases. *J. Med. Chem.* **2005**, *48*, 3463−3466.
(68) Lyons, M. M.; Yu, C. W.; Toma, R. B.; Cho, S. Y.; Reiboldt, W.; Lee, J.; Van Breemen, R. B. Resveratrol in raw and baked blueberries and bilberries. *J. Agric. Food Chem.* **2003**, *51*, 5867−5870.
(69) Ferreira, C. V.; Justo, G. Z.; Souza, A. C. S.; Queiroz, K. C. S.; Zambuzzi, W. F.; Aoyama, H.; Peppelenbosch, M. P. Natural compounds as a source of protein tyrosine phosphatase inhibitors: Application to the rational design of small-molecule derivatives. *Biochimie* **2006**, *88*, 1859−1873.
(70) Wang, F.-R.; Yang, X.-W.; Zhang, Y.; Liu, J.-X.; Yang, X.-B.; Liu, Y.; Shi, R.-B. Three new isoflavone glycosides from Tongmai granules. *J. Asian Nat. Prod. Res.* **2011**, *13*, 319−329.
(71) Ha, D. T.; Tran Minh, N.; Lee, I.; Lee, Y. M.; Kim, J. S.; Jung, H.; Lee, S.; Na, M.; Bae, K. Inhibitors of Aldose Reductase and Formation of Advanced Glycation End-Products in Moutan Cortex (Paeonia suffruticosa). *J. Nat. Prod.* **2009**, *72*, 1465−1470.
(72) Matsuda, H.; Morikawa, T.; Toguchida, I.; Yoshikawa, M. Structural requirements of flavonoids and related compounds for aldose reductase inhibitory activity. *Chem. Pharm. Bull.* **2002**, *50*, 788−795.
(73) Srivastava, S. K.; Ansari, N. H. Prevention of sugar-induced cataractogenesis in rats by butylated hydroxytoluene. *Diabetes* **1988**, *37*, 1505−1508.
(74) Wolff, S. P.; Crabbe, M. J. C. Low apparent aldose reductase-activity produced by monosaccharide autoxidation. *Biochem. J.* **1985**, *226*, 625−630.
(75) Collier, A.; Wilson, R.; Bradley, H.; Thomson, J. A.; Small, M. Free-radical activity in type-2 diabetes. *Diabetic Med.* **1990**, *7*, 27−30.
(76) Sato, Y.; Hotta, N.; Sakamoto, N.; Matsuoka, S.; Ohishi, N.; Yagi, K. Lipid peroxide level in plasma of diabetic-patients. *Biochem. Med.* **1979**, *21*, 104−107.
(77) Kannel, W. B.; McGee, D. L. Diabetes and Cardiovascular Risk-Factors - Framingham Study. *Circulation* **1979**, *59*, 8−13.
(78) Greene, D. A. Acute and chronic complications of diabetes-mellitus in older patients. *Am. J. Med.* **1986**, *80*, 39−53.
(79) Mohora, M.; Virgolici, B.; Paveliu, F.; Lixandru, D.; Muscurel, C.; Greabu, M. Free radical activity in obese patients with type 2 diabetes mellitus. *Rom. J. Intern. Med.* **2006**, *44*, 69−78.
(80) Nishimura, C.; Kuriyama, K. Alteration of lipid peroxide and endogenous antioxidant contents in retina of streptozotocin-induced diabetic rats - effect of vitamin-a administration. *Jpn. J. Pharmacol.* **1985**, *37*, 365−372.
(81) Paller, M. S.; Hoidal, J. R.; Ferris, T. F. Oxygen free-radicals in ischemic acute-renal-failure in the rat. *J. Clin. Invest.* **1984**, *74*, 1156−1164.
(82) Funke, I.; Melzig, M. F. Effect of different phenolic compounds on alpha-amylase activity: screening by microplate-reader based kinetic assay. *Pharmazie* **2005**, *60*, 796−797.
(83) He, Z. D.; Lau, K. M.; But, P. P. H.; Jiang, R. W.; Dong, H.; Ma, S. C.; Fung, K. P.; Ye, W. C.; Sun, H. D. Antioxidative glycosides from the leaves of Ligustrum robustum. *J. Nat. Prod.* **2003**, *66*, 851−854.
(84) Jang, D. S.; Lee, Y. M.; Jeong, I. H.; Kim, J. S. Constituents of the Flowers of Platycodon grandiflorum with Inhibitory Activity on Advanced Glycation End Products and Rat Lens Aldose Reductase In Vitro. *Arch. Pharmacal Res.* **2010**, *33*, 875−880.
(85) Wang, N.; Yang, X.-W. Two new flavonoid glycosides from the whole herbs of Hyssopus officinalis. *J. Asian Nat. Prod. Res.* **2010**, *12*, 1044−1050.
(86) Xie, H. H.; Wang, T.; Matsuda, H.; Morikawa, T.; Yoshikawa, M.; Tani, T. Bioactive constituents from Chinese natural medicines. XV.(1)) - Inhibitory effect on aldose reductase and structures of saussureosides A and B from Saussurea medusa. *Chem. Pharm. Bull.* **2005**, *53*, 1416−1422.
(87) Fang, H.; Tong, W. D.; Shi, L. M.; Blair, R.; Perkins, R.; Branham, W.; Hass, B. S.; Xie, Q.; Dial, S. L.; Moland, C. L.; Sheehan, D. M. Structure-activity relationships for a large diverse set of natural, synthetic, and environmental estrogens. *Chem. Res. Toxicol.* **2001**, *14*, 280−294.
(88) Dong, Y.; Shi, H.; Yang, H.; Peng, Y.; Wang, M.; Li, X. Antioxidant Phenolic Compounds from the Stems of Entada phaseoloides. *Chem. Biodiv.* **2012**, *9*, 68−79.
(89) Khairullina, V. R.; Gerchikov, A. Y.; Denisova, S. B. Comparative study of the antioxidant properties of selected flavonols and flavanones. *Kinet. Catal.* **2010**, *51*, 219−224.
(90) Kong, D.; Zhang, Y.; Yamori, T.; Duan, H.; Jin, M. Inhibitory Activity of Flavonoids against Class I Phosphatidylinositol 3-Kinase Isoforms. *Molecules* **2011**, *16*, 5159−5167.
(91) Mueller, M.; Lukas, B.; Novak, J.; Simoncini, T.; Genazzani, A. R.; Jungbauert, A. Oregano: A Source for Peroxisome Proliferator-Activated Receptor gamma Antagonists. *J. Agric. Food Chem.* **2008**, *56*, 11621−11630.

(92) Mei, R.-Q.; Lu, Q.; Hu, Y.-F.; Liu, H.-Y.; Bao, F.-K.; Zhang, Y.; Cheng, Y.-X. Three new polyyne (= polyacetylene) glucosides from the edible roots of Codonopsis cordifolioidea. *Helv. Chim. Acta* **2008**, *91*, 90−96.
(93) Ilango, K.; Chitra, V.; Kanimozhi, P.; Balaji, G. Antidiabetic, antioxidant and antibacterial activities of leaf extracts of Adhatoda zeylanica. Medic (Acanthaceae). *J. Pharm. Sci. Res.* **2009**, *1*, 67−73.