An Improved Receptor-Based Pharmacophore Generation Algorithm Guided by Atomic Chemical Characteristics and Hybridization Types

Gaoqi He [1,3], Bojie Gong [1], Jianqiang Li [1], Yiping Song [1], Shiliang Li [2]* and Xingjian Lu [1,2]*

[1] Department of Computer Science and Engineering, East China University of Science and Technology, Shanghai, China, [2] Shanghai Key Laboratory of New Drug Design, East China University of Science and Technology, Shanghai, China, [3] School of Computer Science and Software Engineering, East China Normal University, Shanghai, China

Pharmacophore-based virtual screening is an important and leading compound discovery method. However, current pharmacophore generation algorithms suffer from difficulties, such as ligand-dependent computation and massive extractive chemical features. On the basis of the features extracted by the five probes in Pocket v.3, this paper presents an improved receptor-based pharmacophore generation algorithm guided by atomic chemical characteristics and hybridization types. The algorithm works under the constraint of receptor atom hybridization types and space distance. Four chemical characteristics (H-A, H-D, and positive and negative charges) were extracted using the hybridization type of receptor atoms, and the feature point sets were merged with 3 Å space constraints. Furthermore, on the basis of the original extraction of hydrophobic characteristics, extraction of aromatic ring chemical characteristics was achieved by counting the number of aromatics, searching for residual base aromatic ring, and determining the direction of aromatic rings. Accordingly, extraction of six kinds of chemical characteristics of the pharmacophore was achieved. In view of the pharmacophore characteristics, our algorithm was compared with the existing LigandScout algorithm. The results demonstrate that the pharmacophore possessing six chemical characteristics can be characterized using our algorithm, which features fewer pharmacophore characteristics and is ligand independent. The computation of many instances from the directory of useful decoy dataset show that the active molecules and decoy molecules can be effectively differentiated through the presented method in this paper.

Keywords: receptor-based pharmacophore generation, hybridization type, aromatic ring characteristic extraction, space constraint, DUD dataset testing

INTRODUCTION

The concept of pharmacophore was firstly introduced by Ehrlich in 1909(Ehrlich, 1909). More recently, IUPAC gave an official definition as "an ensemble of steric and electronic features that is necessary to ensure the optimal supramolecular interactions with a specific biological target and to trigger (or block) its biological response"(Wermuth et al., 1998).

There are two kinds of typical pharmacophore extraction methods, that is, ligand-based and structure-based (Handler and Buschmann, 2017; Machaba et al., 2017; Seidel et al., 2017). Structure-based virtual screening of pharmacophore has been developed as an important method for computer-aided drug design (Güner, 2002; Chen, 2013; Wang et al., 2017). In this case, the protein may contain ligand (holo structure) or not (apo structure). During pharmacophore extraction with holo protein structures, 4–10 different kinds of pharmacophore may be obtained due to the different strategies of conformational analysis and molecular superposition (Barnum et al., 1996). Such condition significantly reduces the efficiency of virtual screening based on pharmacophore. With the rapid progress of X-ray crystallography and nuclear magnetic resonance technologies in the determination of the protein structure of drug targets (Lam et al., 1994; Erickson et al., 2004), the pharmacophore constructed using apo protein structures becomes feasible. This method of producing a pharmacophore model directly from protein crystal structures can reveal the key elements of the receptor-ligand binding model (Böhm, 1992a).

Pocket v.2 (Chen and Lai, 2006), which was proposed by Chen and Lai at Peking University, has achieved good results by using pharmacophore extraction with apo protein structures. However, the presented pharmacophore model in Pocket v.2 only contains three kinds of chemical characteristics, which decreases the accuracy of pharmacological characteristics. In a follow-up work on Pocket v.3 (Chen et al., 2014), Chen et al. addressed its disadvantages and proposed a 5-element pharmacological characteristic set, including hydrogen-bonding acceptor (H-A), hydrogen-bond donor (H-D), positive charge (Pos), negative charge (Neg), and hydrophobic region (Hydrophobic). However, the number of chemical characteristic points extracted from Pocket v.3 is remarkably high. The points with less contribution dramatically increase the computational load of virtual screening. Thus, in this paper, we present an improved pharmacophore generation method using only the receptors based on the five kinds of probes in Pocket v.3 algorithm. Our method is guided by atomic chemical characteristics and the atomic SP hybridization (Schmitt et al., 2002; Dror et al., 2004; Shoichet et al., 2010). Atomic SP hybridization is a common way of orbital hybridization and is useful for understanding the nature of the chemical bond. This method also utilizes the constraint condition, hybrid type and space distance of receptor atoms. The major contributions of our work include the following:

(1) Four kinds of chemical characteristics (H-A, H-D, Pos, and Neg) were extracted by the receptor Atomic SP hybridization, and the characteristic point set was merged under 3 Å spatial constraints.

(2) On the basis of hydrophobic characteristics, the chemical characteristics of aromatic rings were extracted using the statistics number of aromatic atoms, the search for residual aromatic ring, and direction assessment of the aromatic ring axis.

(3) Extraction of the pharmacophore containing six kinds of chemical characteristics was realized.

METHODS

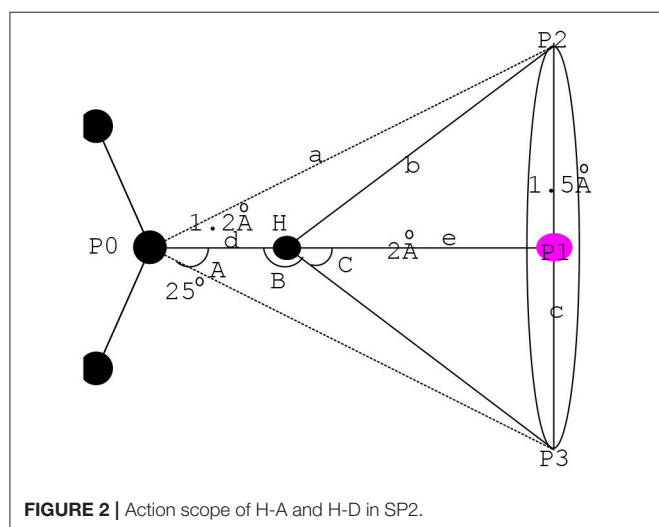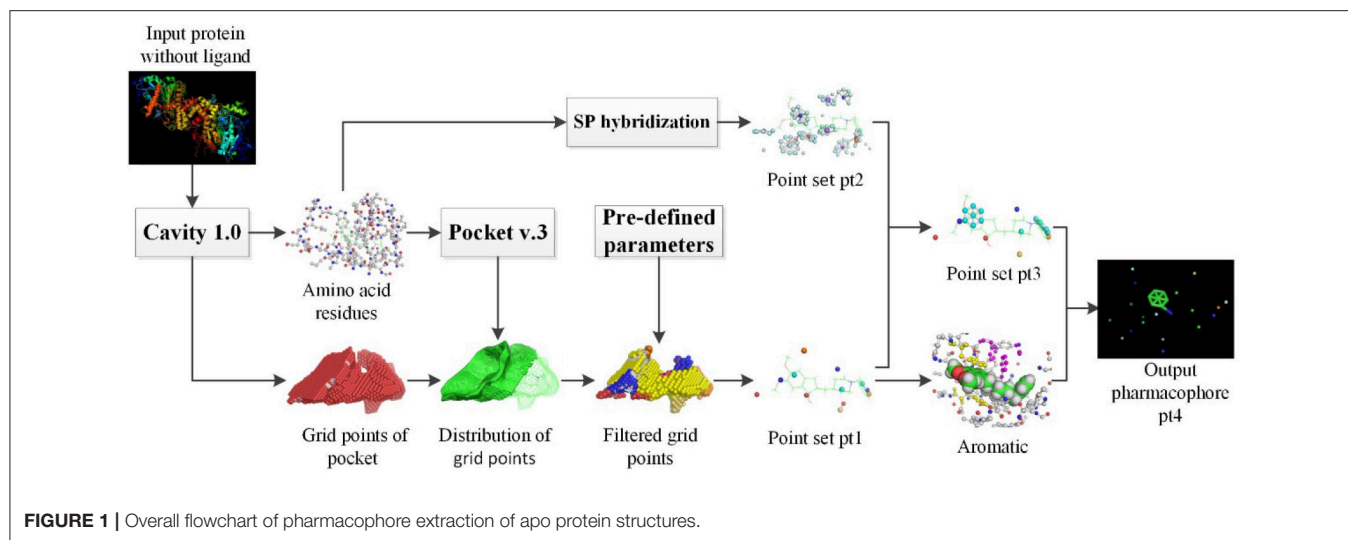### Flowchart of Virtual Screening of Apo Protein Structures

With the increasing number of apo protein structures, a solution must be developed for extracting pharmacophore from this type of proteins. Compared with the early work on Pocket v.2 and v.3, we propose a novel solution that aims at filtering out feature points with less pharmaceutical effect by adding aromatic features and SP hybridization form. Thus, the proposed method can reduce the quantity of pharmacophore features and improve the accuracy of virtual screening.

Figure 1 shows the overall flowchart of pharmacophore extraction with apo protein structures. Cavity 1.0 (Yuan et al., 2011) was exploited to carry out the extraction of receptor pockets. Amino acid residues and grid points of pocket are the two crucial output files for the subsequent processing steps. Using Pocket v.3 software and predefined configure parameters, we get the point set pt1. On the other hand, SP hybridization technology was used to simulate the ligand characteristics of residues in the pocket to extract the pharmacophore model and we get the point set pt2. After a series of filtering operation, point set pt3 is formed. Final pharmacophore pt4 is constructed based on pt3 and extracted aromatic rings. Finally, the extracted pharmacophore models were used as input for virtual screening and tested in 28 Directory of Useful Decoy (DUD) datasets (Huang et al., 2006). The results were compared with those of LigandScout algorithm (Wolber and Langer, 2005) to verify the accuracy of the extracted pharmacological features. Default parameters were used to generate the pharmacophore model by LigandScout, which was then saved as a hypoedit file for further usage of virtual screening by the in-house PharmFit algorithm.

Improved Pharmacophore Extraction Algorithm

Filtering of Pharmacophore Characteristics

Filtering of features in pharmacophore aims at removing the characteristics with less pharmaceutical effect. In our work, the energy scoring strategy used in Pocket v.3 was employed, through which the scores of the five probe atoms against the receptors were calculated accordingly. After scanning the score of each probe in each lattice point (Meng et al., 1992), we set a probe whose score is below a certain threshold to zero, because the chemical characteristics corresponding to such probe contribute less to ligand–protein binding (Carlson, 2002a,b). The threshold is set to 0.15 for HBD and HBA, 0.2 for Hydrophobic, and 0.5 for both positive and negative charges, respectively. Then, we counted the number of probes around each type of probe and determined whether it will be reserved if its statistical number is larger than a threshold value (Goodford, 1985). Here the threshold is set to 30 for HBD and HBA, 40 for Hydrophobic, and 5 for both positive and negative charges, respectively. Finally, the point with the highest score in at 2.5 Å (Chen, 2013) in a particular probe was selected as the representative chemical

FIGURE 1 | Overall flowchart of pharmacophore extraction of apo protein structures.

FIGURE 2 | Action scope of H-A and H-D in SP2.

characteristic point of this region. All the chemical characteristic points obtained in this step were denoted as the point set pt1.

Extraction of Candidate Chemical Feature Point Set

SP hybridization is a common way of orbital hybridization. SP hybridization is not only useful for understanding the nature of the chemical bond, but it may also be sculpted in the laboratory by means of coherent laser pulses. In this paper, the characteristics in the amino acid residue corresponding to the chemical features in the pockets were extracted by the SP hybrid model (Boobbyer et al., 1989) to form an accurate pharmacophore. **Figure 2** shows the related principle with an example considering the SP2 in the SP hybrid model.

As shown in **Figure 2**, the center coordinates extracted by the mirror form the center point P1 at the bottom plane (Martin et al., 1993) with the radius c. H refers to the apex of the cone and serves as the action scope of H-A or H-D (Iwata and Morokuma,

1973; Taylor and Kennard, 1982).

$$a^2 = (d + e)^2 + c^2 \tag{1}$$

$$\tan(A) = \frac{c}{(d + e)} \tag{2}$$

$$e = VDW(H) + VDW(P1) \tag{3}$$

In Equation 1, d refers to the covalent bond length of hydrogen bond. e is computed by Equation 3, and angle C is determined by the maximum feasible angle of D-H ... A and A-H ... D. Considering the superposition of the pharmacophore and the experience value, we obtained the maximum of c = 1.5 Å, and e = 2.0 Å. Hence, max (C) = 37° as deduced by tan (C) = c/e. Correspondingly, min (B) = 143°. The hydrogen bond will break when angle B approaches 90°. Hence, angle B must be >90°. The features in the residue corresponding to the chemical characteristics in the pockets can be extracted by the SP2 hybrid model and the distance and angle constraints. We recorded the candidate chemical feature point set obtained here as pt2.

Extraction of Aromatic Ring Chemical Characteristics

In our algorithm, an aromatic ring is extracted by assessing the number of aromatic atoms and the vector direction of aromatic rings. As aromatic characteristics occupy an important position in the virtual screening of pharmacophore (Rauh et al., 2004), they were added in this paper to expand the five chemical characteristics of Pocket v.3. We searched for aromatic atoms on the amino acid residue and counted the total number n in the 4.5 Å range of the hydrophobic chemical feature points (Böhm, 1992b; Greene et al., 1994). Then, the characteristic points with $n \geq 6$ were kept and considered an aromatic ring. We also considered the vector direction of the aromatic ring to determine whether it can be treated as an aromatic feature. Commonly, the normal vector of aromatic rings will either be oriented (1) toward the edge or parallel to the pocket or (2) toward the inner side of the pocket. The latter exerts a more significant influence on virtual
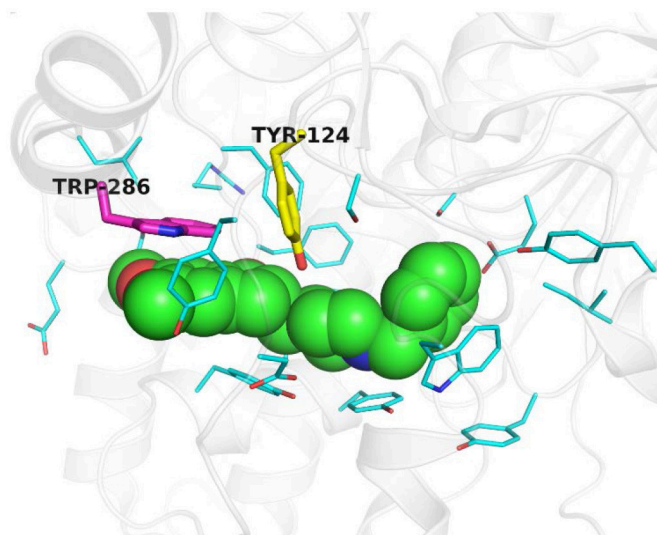
FIGURE 3 | Example of aromatic rings in the binding site of ACHE (PDB ID: 4EY7). The ligand E2020 is shown as green sphere. Key residues are displayed as thin cyan sticks while key aromatic rings are highlighted as yellow and purple thick sticks. This figure is rendered by PyMol (version 1.3).

screening. Thus, aromatic rings with normal vector toward the inner side of the pocket were selected as the final aromatic features.

Figure 3 shows an example of aromatic rings in the binding site of ACHE (PDB ID: 4EY7). In **Figure 3**, the normal vector of the yellow aromatic ring is parallel to the direction of the ligand, whereas that of the purple aromatic ring heads toward the inner side of the pocket. Thus, points were extracted as an aromatic ring chemical feature when the chemical characteristic point of the hydrophobic satisfies the following: (1) the number of aromatic atoms $\geq$ 6; (2) the normal vector of the aromatic ring is toward the inner side of the pocket at 4.5 Å.

Construction of the Pharmacophore Model
Through the above steps, we have obtained pt1 and pt2 by using the filtering mechanism and the SP hybrid models together with the aromatic chemical characteristics. As a result, the comprehensive pharmacophore model can be constructed. First, pt1 is handled by the SP hybrid model and then intersected with pt2 in the 3 Å range. The operation result was denoted as pt3, which contains four kinds of chemical characteristics (H-A, H-D, Pos, and Neg). Finally, hydrophobic feature in pt1, the extracted aromatic ring, and pt3 represent six important types of chemical characteristics. These characteristics make up the proposed pharmacophore model in our paper (Hoffren et al., 2001). **Figure 4** displays the pharmacophores of the binding site in ACHE extracted using our algorithm independent of the bound ligand. Detailed descriptions of this pharmacophore model were in the next section.

RESULTS AND DISCUSSION

A series of experiments was conducted to validate the effectiveness of the presented pharmacophore extraction method.

Comparison With Ligandscout Software
The pharmacophore model of ACHE was also generated through LigandScout software using the PDB code of 4EY7 (**Figure 5**). The comparison of the pharmacophore model generated by our method and LigandScout was illustrated in **Figure 5**.

Analysis 1: Comparison of the Distribution of Chemical Characteristic Points Corresponding to Pharmacophores
Figure 5A displayed the pharmacophore model generated by our method for ACHE with the ligand E2020 displayed only for the convenience of analysis. There were three HBA pharmacophores predicted by our method, and the one marked by black rectangle corresponds to the HBA pharmacophore generated from the hydroxyl group of E2020 by LigandScout. The proposed algorithm is independent of ligand information; such condition is necessary for the LigandScout algorithm. However, these two algorithms feature corresponding chemical characteristic of HBA, because both use the SP hybrid concept of atoms. In addition, there were five hydrophobic pharmacophores predicted by our method, and three of them are aromatic. By overlaying the bound ligand E2020 to the pharmacophore model generated by our method, it could be obviously observed that the two hydrophobic pharmacophores around the two methoxy groups and the two aromatic pharmacophores around the two benzene rings in E2020 are all accurately predicted by our method. There was another hydrophobic center including an aromatic feature around the piperidine ring predicted by our method while LigandScout predicted one hydrophobic feature at that place. Collectively, the pharmacophores predicted by our method can match most of the pharmacophores predicted by LigandScout.
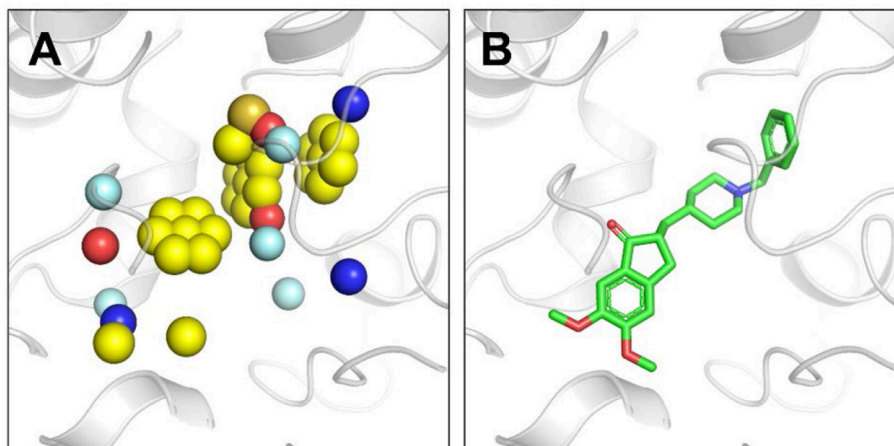
FIGURE 4 | (A) Pharmacophores extracted by our algorithm in the ACHE binding pocket (PDB ID: 4EY7). Blue sphere represents HBD pharmacophore, red sphere represents HBA pharmacophore, single yellow sphere represents hydrophobic pharmacophore, single yellow sphere that surrounded by six planar yellow spheres represents aromatic pharmacophore, sulfur sphere represents positive, orange sphere represents negative, and palecyan sphere represents the root of h-bond. All the pharmacophore models generated by our method in this paper use this color scheme. **(B)** The bound ligand E2020 in the crystal structure of ACHE. Our algorithm generates pharmacophores independent of the bound ligand.
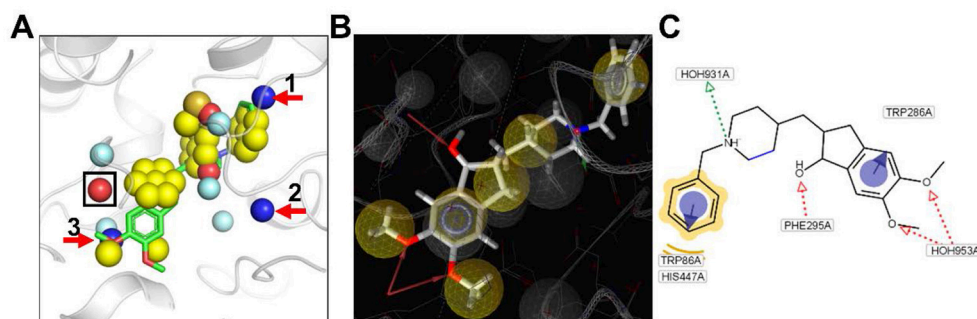

FIGURE 5 | (A) Pharmacophores extracted by our algorithm with the bound ligand overlaid (PDB ID: 4EY7). **(B)** Pharmacophores extracted by LigandScout software. **(C)** 2D diagram for molecular interactions between E2020 and ACHE. One common HBA pharmacophore generated by both methods are highlighted by a black rectangle. Three HBD pharmacophores generated only by our method are marked by red arrows.

Analysis 2: Comparison of Different Chemical Characteristic Points of Pharmacophores

Apart from the above common pharmacophores, our method predicted two more HBA, three HBD, and one Neg center compared with LigandScout. Our method failed to predict the Pos pharmacophore around the charged nitrogen atom of the piperidine ring which directly formed one hydrogen bond with HOH931 other than an amino acid residue of the protein as displayed in **Figures 5B,C**. The above differences between the pharmacophore models generated by our method and LigandScout mainly due to the reason that our method predicts pharmacophores using the apo protein structure and does not consider water molecules.

To validate whether the additional pharmacophores extracts by our method appear in other crystal structures of ACHE, we calculated pharmacophores on another two structures of ACHE with PDB codes of 4EY6 and 4EY5. For both cases

of 4EY6 and 4EY5 (**Figures 6A–F**), the five hydrophobic pharmacophores in **Figure 5A** were also extracted, indicating that those pharmacophores are important for ligand binding. In addition, two of the three HBD pharmacophores in **Figure 5A** were also extracted simultaneously for 4EY6 and 4EY5, suggesting that HBD pharmacophore 1 and 2 highlighted in **Figures 6A,C** are important pharmacophores. By overlaying the bound ligand in 4EY6 and 4EY5 (**Figures 6B,E**) with the respective pharmacophore model (**Figures 6C,F**), we found that except for the well-matched hydrophobic pharmacophore, both of the ligands have a hydrogen bond donor around the HBD pharmacophore 1: (-)-galanthamine bound in 4EY6 has a hydroxy group as HBD while (-)-huperzine A bound in 4EY5 has an amine group in the amide as HBD. Although no ligand atoms bind at the places around HBD pharmacophore 2 and 3, we still found crystal water molecules as HBD. For the place around HBD pharmacophore 2, we found
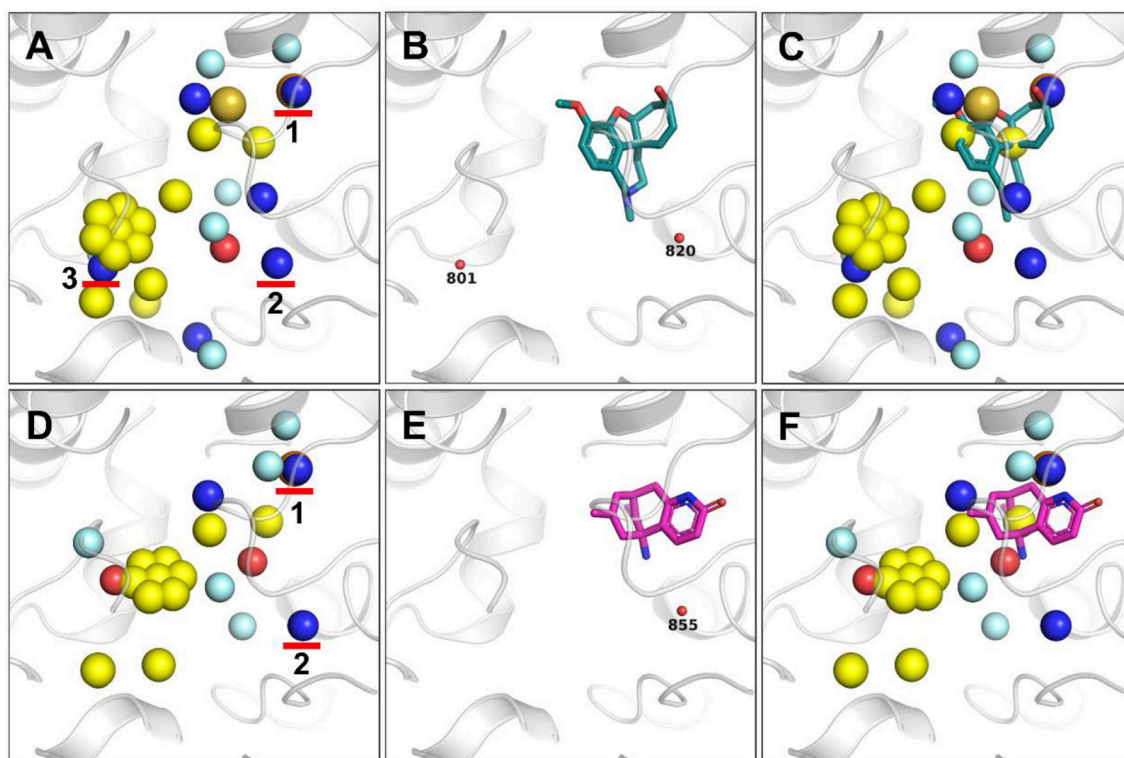
FIGURE 6 | Pharmacophores generated based on the crystal structures of ACHE with PDB codes of 4EY6 **(A–C)** and 4EY5 **(D–F)**. Three HBD pharmacophores generated only by our method are marked by red arrows **(A,D)**. The ligand (-)-galanthamine bound in 4EY6 is displayed as green sticks **(B,C)**, while ligand (-)-huperzine A bound in 4EY5 is displayed as magenta sticks **(E,F)**. Water molecules are shown as small red balls **(B,E)**.
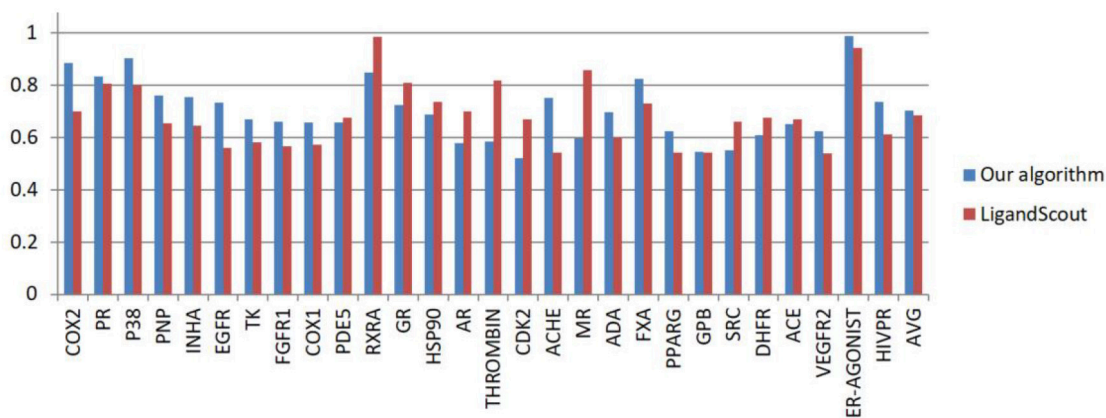

FIGURE 7 | AUC values obtained by LigandScout and our algorithm from the 28 DUD datasets.

water molecules W820 and W855 bound at that place as HBD in 4EY6 and 4EY5, respectively (**Figures 6B,E**). For the place around HBD pharmacophore 3, we found one water molecule W801 bound at that place as HBD in 4EY6 (**Figure 6B**). The results suggested that those HBD features can be considered in the drug discovery and design pipeline

against ACHE. Those results directly proved that our ligand-independent pharmacophore extraction method is reliable and superior to LigandScout in detecting pharmacophores where no ligand atoms bind. The interactions between the receptor and ligand should be evaluated when the LigandScout algorithm extracts pharmacophores. Only when interaction exists will
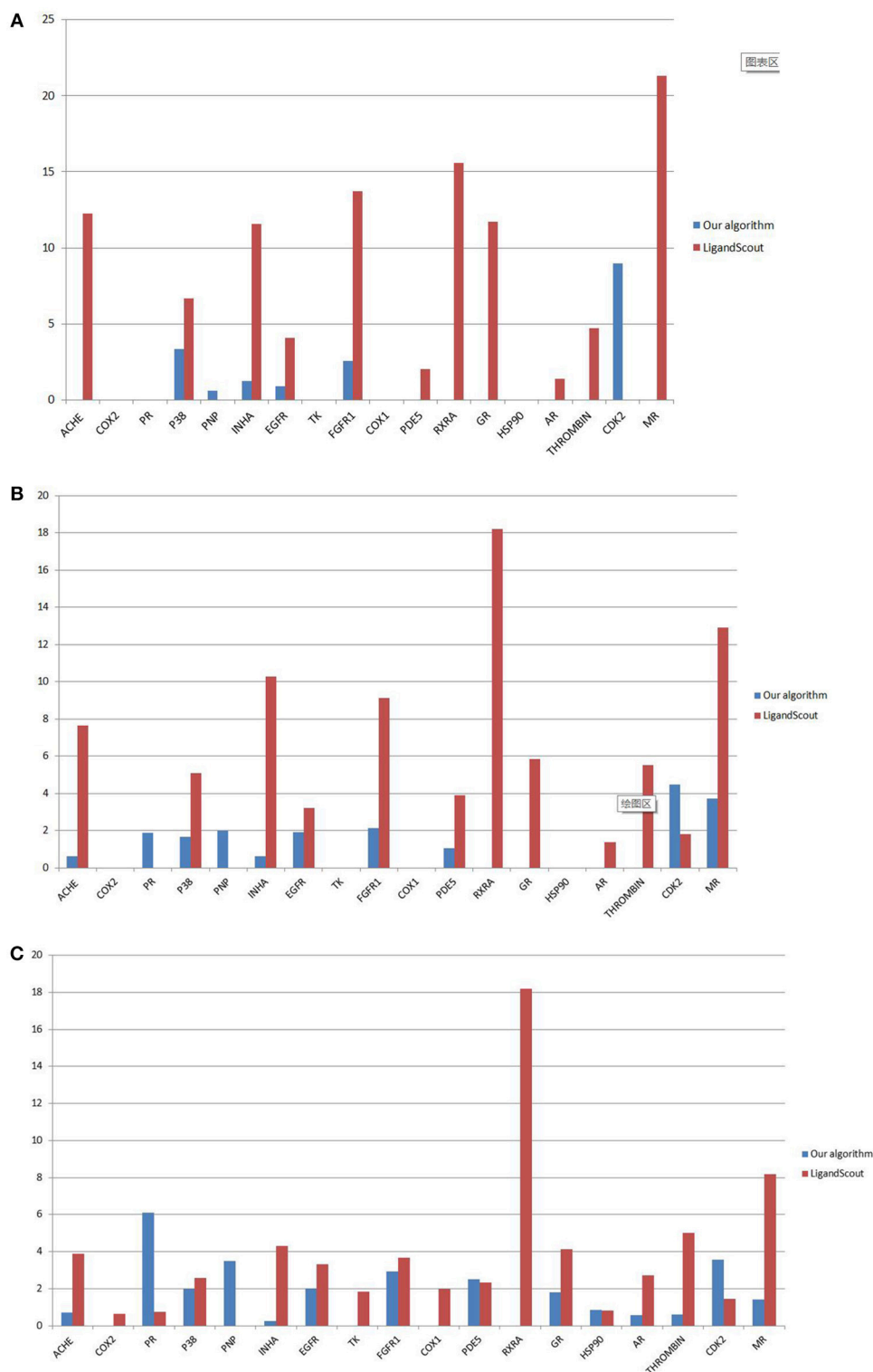
FIGURE 8 | Comparison of EF values under 1% **(A)**, 2% **(B)**, and 5% **(C)** filter libraries.

the corresponding chemical characteristics be extracted. When a ligand is determined, only one type of interaction with the receptor exists in the static model. However, the actual interaction between the receptor and the ligand is a dynamic process. Our ligand-independent pharmacophore extraction method can adapt to the conformational changes of the receptor.

Consequently, the proposed ligand-independent pharmacophore extraction method shows its advantages against LigandScout in terms of pharmacophore extraction, and can be helpful for pharmacophore-based virtual screening to find more diverse active molecules.

Verifying the Feasibility of Using DUD Datasets

In the experiment, we executed both the proposed algorithm and LigandScout algorithm to produce the corresponding pharmacophore model by using 28 DUD datasets. These datasets include ACHE, COX2, PR, P38, PNP, Inha, EGFR, TK, FGFR1, COX1, PDE5, Rxra, GR, HSP90, AR, thrombin, CDK2, MR, ADA, FXA, PPARG, GPB, SRC, DHFR, ACE, VEGFR2, ER-Agonist, and HIVPR.

The feasibility of the pharmacophore model extracted by the proposed algorithm was verified by the in-house PharmFit algorithm. PharmFit uses a triangular hashing algorithm to realize the pharmacophore superimposition with high efficiency and adopts the default parameter setting. Xu et al. used PharmFit for virtual screening in their work of CavityPlus (Xu et al., 2018). Hence, we scored the activity and the decoy molecule of each dataset by using PharmFit. Then, the area under the receiver operating characteristic curve (AUC) and the enrichment factors (EF)(Machaba et al., 2017) were calculated. EFs are used in virtual screening to distinguish the capability between active and inactive molecules and is defined by Equation 4 (Chen, 2013). Here, the EF values were computed under 1, 2, and 5% filter libraries.

$$EF = \frac{Hits_{sampled}}{N_{sampled}} \times \frac{N_{total}}{Hits_{total}} \qquad (4)$$

where $Hits_{sampled}$ represents the number of hit active molecules in the corresponding x% filter Library, $N_{sampled}$ corresponds to the total number of molecules, N_{total} denotes the total number of molecules, and $Hits_{total}$ stands for the total number of active molecules in the filter library.

Comparison of AUC Using 28 Datasets

Using the 28 DUD datasets, pharmacophores were generated using our algorithm and LigandScout algorithm. **Figure 7** shows the AUC scores calculated by the two algorithms. Detail information about DUD dataset and AUC computation for ER_AGONIST can be found from the **supplementary materials**.

As shown in **Figure 7**, the AUC value of our algorithm exceeds that of the LigandScout algorithm in 17 out of 28 test systems. The average AUC of our algorithm reaches 0.703, whereas that of LigandScout algorithm is 0.686, indicating a remarkable progress in pharmacophore generation.

Under 1, 2, and 5% filter libraries, the EF value of our algorithm is generally lower than that of LigandScout algorithm (**Figure 8**). This discovery illustrates the characteristics of our pharmacophore-extracting algorithm.

(1) The pharmacophore model is not built on the ligand interaction but only from the receptor to deduce all the possible binding models. Hence, this design features universality.

(2) The early enrichment effect is poor, but the whole active molecule can be effectively screened.

By analyzing and comparing the screening results through the AUC, we observed that the effectiveness of our pharmacophore model with apo protein structure surpasses that of related algorithms with holo protein structure on most datasets.

CONCLUSION

This study designed and developed a new algorithm for pharmacophore extraction. This algorithm exploits the concept of SP hybridization and adds the aromatic probe and the 3 Å distance constraint to the original Pocket v.3 algorithm. The experimental results show that the pharmacophore extracted by the proposed algorithm features less distribution of pharmacophore feature points and high accuracy, thus enriching the current research on pharmacophore extraction without ligand information.

Additional future works are headed for this direction. Pocket size is an essential parameter in pharmacophore generation. When the pocket size measures more than three times of the average ligand volume, the number of feature points from the pharmacophore model increases significantly. It is necessary to further explore the reason why EF index is low and find a good solution. At the same time, the performance highly depends on the output quality of Cavity 1.0.

AUTHOR CONTRIBUTIONS

GH is the group leader and he is responsible for the project management and in charge of revising this manuscript. BG, JL, and YS are in charge of the data analysis and the planning and performing the experiments. SL designed the experiments and gave professional descriptions. XL provided valuable advice about this manuscript.

ACKNOWLEDGMENTS

SUPPLEMENTARY MATERIAL

The Supplementary Material for this article can be found online at: https://www.frontiersin.org/articles/10.3389/fphar. 2018.01463/full#supplementary-material

REFERENCES

Barnum, D., Greene, J., and Smellie, A., Sprague, P. (1996). Identification of common functional configurations among molecules. *J. Chem. Inform. Comp. Sci.* 36, 563–571. doi: 10.1021/ci950273r

Böhm, H. J. (1992a). LUDI: rule-based automatic design of new substituents for enzyme inhibitor leads. *J. Comp. Aided Mol. Design* 6, 593–606. doi: 10.1007/BF00126217

Böhm, H. J. (1992b). The computer program LUDI: A new method for the de novo design of enzyme inhibitors. *J. Comp. Aided Mol. Design* 6, 61–78. doi: 10.1007/BF00124387

Boobbyer, D. N., Goodford, P. J., McWhinnie, P. M., and Wade, R. C. (1989). New hydrogen-bond potentials for use in determining energetically favorable binding sites on molecules of known structure. *J. Med. Chem.* 32, 1083–1094. doi: 10.1021/jm00125a025

Carlson, H. A. (2002a). Protein flexibility and drug design: how to hit a moving target. *Curr. Opin. Chem. Biol.* 6, 447–452. doi: 10.1016/S1367-5931(02)00341-1

Carlson, H. A. (2002b). Protein flexibility is an important component of structure-based drug discovery. *Curr. Pharm. Design* 8, 1571–1578. doi: 10.2174/1381612023394232

Chen, C. Y. (2013). A novel integrated framework and improved methodology of computer-aided drug design. *Curr. Topics Med. Chem.* 13, 965–988. doi: 10.2174/1568026611313090002

Chen, J., and Lai, L. (2006). Pocket v.2: further developments on receptor-based pharmacophore modeling. *Cheminform* 46, 2684–2691. doi: 10.1021/ci600246s

Chen, J., Ma, X., Yuan, Y., Pei, J., and Lai, L. (2014). Protein-protein interface analysis and hot spots identification for chemical ligand design. *Curr. Pharm. Design* 20, 1192–1200. doi: 10.2174/13816128113199990065

Dror, O., Shulman-Peleg, A., Nussinov, R., and Wolfson, H. J. (2004). Predicting molecular interactions *in silico*: I. A guide to pharmacophore identification and its applications to drug design. *Curr. Med. Chem.* 11, 71–90. doi: 10.2174/0929867043456287

Ehrlich, P. (1909). Über den jetzigen Stand der Chemotherapie. *Berichte Deutsch. Chem. Gesells.* 42, 17–47. doi: 10.1002/cber.19090420105

Erickson, J. A., Jalaie, M., Robertson, D. H., Lewis, R. A., and Vieth, M. (2004). Lessons in molecular recognition: the effects of ligand and protein flexibility on molecular docking accuracy. *J. Med. Chem.* 47, 45–55. doi: 10.1021/jm030209y

Goodford, P. J. (1985). A computational procedure for determining energetically favorable binding sites on biologically important macromolecules. *J. Med. Chem.* 16, 849–857. doi: 10.1021/jm00145a002

Greene, J., Kahn, S., Savoj, H., Sprague, P., and Teig, S. (1994). Chemical function queries for 3D database search. *J. Chem. Inform. Model.* 34, 1297–1308. doi: 10.1021/ci00022a012

Güner, O. F. (2002). History and evolution of the pharmacophore concept in computer-aided drug design. *Curr. Topics Med. Chem.* 2, 1321–1332. doi: 10.2174/1568026023392940

Handler, N., and Buschmann, H. (2017). "Pharmacophore Generation for Multiple Ligands," *Drug Selectivity: An Evolving Concept in Medicinal Chemistry*, 275–312.

Hoffren, A. M., Murray, C. M., and Hoffmann, R. D. (2001). Structure-based focusing using pharmacophores derived from the active site of 17beta-hydroxysteroid dehydrogenase. *Curr. Pharm. Design* 7, 547–566. doi: 10.2174/1381612013397870

Huang, N., Shoichet, B. K., and Irwin, J. J. (2006). Benchmarking sets for molecular docking. *J. Med. Chem.* 49, 6789–6801. doi: 10.1021/jm0608356

Iwata, S., and Morokuma, K. (1973). Molecular orbital studies of hydrogen bonds. IV. Hydrogen bonds in excited states of H_2CO with H_2O ?. *Chem. Phys. Lett.* 19, 94–98. doi: 10.1016/0009-2614(73)87071-X

Lam, P. Y., Jahdav, P. K., Eyermann, C. J., Hodge, C. N., Ru, Y., Bacheler, L. T., et al. (1994). Rational design of potent, bioavailable, nonpeptide cyclic ureas as HIV protease inhibitors. *Science* 263, 380–384. doi: 10.1126/science.8278812

Machaba, K. E., Mhlongo, N. N., Dokurugu, Y. M., and Soliman, M. E. (2017). Tailored-pharmacophore model to enhance virtual screening and drug discovery: a case study on the identification of potential inhibitors against drug-resistant Mycobacterium tuberculosis (3R)-hydroxyacyl-ACP dehydratases. *Fut. Med. Chem.* 9, 1055–1071. doi: 10.4155/fmc-2017-0020

Martin, Y. C., Bures, M. G., Danaher, E. A., DeLazzer, J., Lico, I., and Pavlik, P. A. (1993). A fast new approach to pharmacophore mapping and its application to dopaminergic and benzodiazepine agonists. *J. Comp. Aided Mol. Design* 7, 83–102. doi: 10.1007/BF00141577

Meng, E. C., Shoichet, B. K., and Kuntz, I. D. (1992). Automated docking with grid-based energy evaluation. *J. Comp. Chem.* 13, 505–524. doi: 10.1002/jcc.540130412

Rauh, D., Klebe, G., and Stubbs, M. T. (2004). Understanding protein–ligand interactions: the price of protein flexibility. *J. Mol. Biol.* 335, 1325–1341. doi: 10.1016/j.jmb.2003.11.041

Schmitt, S., Kuhn, D., and Klebe, G. (2002). A new method to detect related function among proteins independent of sequence and fold homology. *J. Mol. Biol.* 323, 387–406. doi: 10.1016/S0022-2836(02)00811-2

Seidel, T., Bryant, S. D., Ibis, G., and Poli, G., langer T., et al. (2017). 3D pharmacophore modeling techniques in computer aided molecular design using ligandscout. *Tutorials Chemoinform.* 281, 279–309. doi: 10.1002/9781119611110.ch20

Shoichet, B. K., Bodian, D. L., and Kuntz, I. D. (2010). Molecular docking using shape descriptors. *J. Comput. Chem.* 13, 380–397. doi: 10.1002/jcc.540130311

Taylor, R., and Kennard, O. (1982). Crystallographic evidence for the existence of C-H.O, C-H.N, and C-H.C1 hydrogen-bonds. *J. Am. Chem. Soc.* 104, 5063–5070. doi: 10.1021/ja00383a012

Wang, X., Shen, Y., Wang, S., Li, S., Zhang, W., Liu, X., et al. (2017). PharmMapper 2017 update: a web server for potential drug target identification with a comprehensive target pharmacophore database. *Nucleic Acids Res.* 45, 356–360. doi: 10.1093/nar/gkx374

Wermuth, C. G., Ganellin, C., Lindberg, P., and Mitscher, L. A. (1998). Glossary of terms used in medicinal chemistry (IUPAC Recommendations 1998). *Pure Appl. Chem.* 70, 1129–1143. doi: 10.1351/pac199870051129

Wolber, G., and Langer, T. (2005). LigandScout:3-D pharmacophores derived from protein-bound ligands and their use as virtual screening filters. *J. Chem. Inform. Model.* 45, 160–169. doi: 10.1021/ci049885e

Xu, Y., Wang, S., Hu, Q., Gao, S., Ma, X., Zhang, W., et al. (2018). CavityPlus: a web server for protein cavity detection with pharmacophore modelling, allosteric site identification and covalent ligand binding ability prediction. *Nucleic Acids Res.* 46, 374–379. doi: 10.1093/nar/gky380

Yuan, Y., Pei, J., and Lai, L. (2011). LigBuilder 2: a practical de novo drug design approach. *J. Chem. Inform. Model.* 51, 1083–1091. doi: 10.1021/ci100350u