

## QSldb: quorum sensing interference molecules

Shengbo Wu<sup>1,3,4</sup>, Chunjiang Liu<sup>1,3,4</sup>, Jie Feng<sup>2</sup>, Aidong Yang<sup>7</sup>, Fei Guo<sup>2\*</sup>, and Jianjun Qiao<sup>1,4,5,6\*</sup>

<sup>1</sup> School of Chemical Engineering and Technology, Tianjin University, Tianjin, 300072, China

<sup>2</sup> School of Computer Science and Technology, College of Intelligence and Computing, Tianjin University, Tianjin 300350, China

<sup>3</sup> State Key Laboratory of Chemical Engineering, Tianjin University, Tianjin 300072, China

<sup>4</sup> Collaborative Innovation Center of Chemical Science and Engineering (Tianjin), Tianjin, 300072, China

<sup>5</sup> Key Laboratory of Systems Bioengineering, Ministry of Education (Tianjin University), Tianjin, 300072, China

<sup>6</sup> Frontiers Science Center for Synthetic Biology (Ministry of Education), Tianjin University, Tianjin, 300072, China

<sup>7</sup> Department of Engineering Science, University of Oxford, Oxford OX1 3PJ, UK

\* Correspondence and requests for materials should be addressed to J.Q. (email: [jianjunq@tju.edu.cn](mailto:jianjunq@tju.edu.cn)) and Fei Guo (email: [fguo@tju.edu.cn](mailto:fguo@tju.edu.cn))

### Abstract

Quorum sensing interference (QSI), the disruption and manipulation of quorum sensing (QS) in the dynamic control of bacteria populations could be widely applied in synthetic biology to realize dynamic metabolic control and develop potential clinical therapies. Conventionally, limited QSI molecules (QSIMs) were developed based on molecular structures or for specific QS receptors, which are in short supply for various interferences and manipulations of QS systems. In this study, we developed QSldb (<http://qsldb.lbc.net/>), a specialized repository of 633 reported QSIMs and 73,073 expanded QSIMs including both QS agonists and antagonists. We have collected all reported QSIMs in literatures focused on the modifications of N-acyl homoserine lactones, natural QSIMs, and synthetic QS analogues. Moreover, we developed a pipeline with SMILES-based similarity assessment algorithms and docking-based validations to mine potential QSIMs from existing 138,805,608 compounds in the PubChem database. In addition, we proposed a new measure, *pocketedit*, for assessing the similarities of active

protein pockets or QSIMs crosstalk, and obtained 273 possible potential broad-spectrum QSIMs. We provided user-friendly browsing and searching facilities for easy data retrieval and comparison. QSIdb could assist the scientific community in understanding QS-related therapeutics, manipulating QS-based genetic circuits in metabolic engineering, developing potential broad-spectrum QSIMs, and expanding new ligands for other receptors.

### Biographical Note

**Shengbo Wu** is a PhD candidate at School of Chemical Engineering and Technology, Tianjin University, Tianjin, 300072, China. He has been working in the field of QS and QSI applications.

**Chunjiang Liu** is an associate professor at the State Key Laboratory of Chemical Engineering, Tianjin University, Tianjin 300072, China. His research interest is focused on system engineering of various networks, including cell-cell communications.

**Jie Feng** is a master candidate at School of Computer Science and Technology, College of Intelligence and Computing, Tianjin University, Tianjin 300350, China.

**Aidong Yang** is an associate professor at the Department of Engineering Science, University of Oxford, Oxford OX1 3PJ, UK. His research interest is focused on developing methods and tools for modelling chemical/biochemical processes and related systems, such as QS systems.

**Fei Guo** is an associate professor at the School of Computer Science and Technology, College of Intelligence and Computing, Tianjin University, Tianjin 300350, China. Her research interest is focused on comparative genomics and bioinformatics

**Jianjun Qiao** is a professor at the Key Laboratory of Systems Bioengineering, Ministry of Education (Tianjin University) and Frontiers Science Center for Synthetic Biology (Ministry of Education), Tianjin University, Tianjin, 300072, China. His research interest is focused on synthetic biology, applications of QS in metabolism regulation and potential medicine development.

**Keywords:** antibiotics, multidrug-resistant, quorum sensing analogues, quorum sensing interference, SMILES, AutoDock Vina.

## Introduction

Quorum sensing (QS), a microbial cell-to-cell communication process, dynamically regulates a variety of metabolism and physiological activities, such as biofilm formation, conjugation, competence, bacteriocin production, and pathogenesis [1]. QS signaling molecules, once produced by cells, diffuse into the culture medium freely. When the level of QS signaling molecules reaches a certain threshold, they will diffuse back into cells and be recognized by specific QS receptors to allow cells to sense their density and regulate the expression of downstream genes [2]. QS can be utilized for population-level control of bacteria and potential clinical therapies development based on autoinducers (AIs) including acyl-homoserine lactones (AHLs) for Gram-negative bacteria [3], auto-inducing peptides (AIPs) for Gram-positive bacteria [4], autoinducer 2 (AI-2) [5] and indole [6] for interspecies communication. Underpinned by functions of AIs and quorum sensing interference molecules (QSIMs), the disruptions and manipulations of QS can be widely applied in synthetic biology [7] to realize the dynamic metabolism control [8] and develop medicine for multidrug-resistant bacteria [9], such as *Staphylococcus aureus* [10] and *Pseudomonas aeruginosa* [11].

Various research studies have been conducted for developing QSIMs by screening or mimicking AIs. The most typical method for screening QSIMs is using active reporters of AIs to detect natural or synthesized compounds. Rasmussen et al. [12] constructed a collection of screening systems with *lux* and *las* QS devices, which enabled them to identify a number of novel QSIMs among natural and synthetic compound libraries, such as garlic extract and 4-nitro-pyridine-N-oxide (4-NPO). Müh et al. [13] developed an ultra-high-throughput cell-based assay (nanowell technology) to screen a library of approximately 200,000 synthetic molecules for inhibitors of LasR-dependent gene expression. Compared with traditional screening methods heavily dependent of verification experiments, computer-based virtual screening for natural QSIMs and synthesizing compounds as analogues of AIs has attracted increasing attention due to its convenience and economy. Virtual screening for QSIMs can be based on either ligands such as AHLs [14] and AI-2 [15] or the structure of receptors such as LasR [16, 17], RhIR [16], LuxR [18], TraR [19], QscR [20], PqsR [21].

Some databases relevant to QS signaling molecules of Gram-negative and Gram-positive bacteria have been constructed for better understanding of various QS mechanisms. Rajput et al. [22] have constructed SigMol, which held various QS molecules for Gram-negative bacteria. Quorum sensing peptides (QSPs), signaling molecules used by Gram-positive bacteria, have also been collected and

analyzed by Rajput et al. [23] in their QSPpred database. Many databases for proteins with ligand binding, such as TarFisDock [24], PSCDB [25], BioLiP [26], and PDID [27], have been constructed to provide important information for understanding interactions between target proteins and drug molecules. Antimicrobial peptides [28], being regarded as important templates for developing a new generation of antimicrobials, have been collected and summarized in some databases. Thakur et al. [29] collected 1,245 antiviral peptides for human viruses battle, such as HIV, to establish the AVPpred database. Wang et al [30] constructed the antimicrobial peptide database (APD), which contains 3,142 antimicrobial peptides from six kingdoms. Rajput et al. [31] have completed a summary for biofilm inhibition named “aBiofilm”, which contains 5,027 anti-biofilm agents. However, there is no database specifically for QSIMs. Due to the current dramatic increase in drug-resistant bacteria and the decline in the availability of new antibiotics, a database for QSIMs is necessary and urgent. In addition, currently reported molecules with QSI activity are relatively limited, and it is highly desirable to mine potential QSIMs from the existing small molecules based on the reported QSIMs.

Based on the reported QSIMs in literatures and some available database, such as PubChem [32], ChEMBL [33], Reaxys [34], and Zinc [35], we have collected 633 QSIMs to construct a small virtual library including natural and synthetic agonists and antagonists for various QS systems. An edit distance of the Simplified Molecular Input Line Entry System (SMILES) [36] of reported QSIMs was used to create a non-redundant set of ligands to reduce the total computational workload. We developed a pipeline with SMILES-based similarity assessment algorithms and docking-based validations to mine the potential QSIMs from existing 138,805,608 compounds in the PubChem database. To validate the QSI activity of the screened 73,073 molecules, we have docked them with nine typical QS receptors (SmcR, TraR, LasR, CviR, QscR, SdiA, PqsR, LuxP, and LsrB) by AutoDock Vina software [37] and ranked them according to their free binding energies (FBEs). In addition, we proposed a new measure, namely *pocketedit*, for the similarity calculation of active protein pockets or QSIMs crosstalk, and developed the potential broad-spectrum QSIMs for multiple QS receptors. Finally, we obtained 273 potential broad-spectrum QSIMs for nine QS receptors based on analysis of *pocketedit* and the potential QSIMs distributions. QSldb, a specialized database of the reported and potential quorum sensing interference molecules, has been constructed for browsing and searching data friendly and conveniently at webserver <http://qsldb.lbcj.net/>.

## Methods and materials

## Data

QS can be utilized to control the expression of groups of genes in a cell-density-dependent manner [38]. Generally, Gram-negative bacteria use AHLs, 2-heptyl-3-hydroxy-4-quinolone (PQS) and diffusible signal factors (DSFs) as autoinducers for intracellular communication [3]. The signals for intercellular communication are mainly AI-2 [39-41] and indole [42], which can regulate the cooperation or competition in various microbial communities. Except the AIPs and antimicrobial peptides for Gram-positive bacteria, the reported QSIMs mainly includes the modifications of the acyl-chain and lactone-ring for AHLs, natural QSIMs from various organism, synthetic analogues of CAI-1 (*cholera* autoinducer 1), DKPs (Diketopiperazines), HAQs(4-hydroxy-2-alkylquinolines), DSFs, AI-2 and indole.

## Data acquisition

We have investigated and collected QS signals and QSIMs stated above for various QS systems from many literatures and several online databases, such as PubChem [32], ChEMBL [33], Reaxys [34], and Zinc [35]. The nine typical QS receptors with existing crystal structures, i.e. SmcR (3KZ9), TraR (1L3L), LasR (2UV0), CviR (3QP1), QscR (3SZT), SdiA (4Y15), PqsR (4JVD), LuxP (1JX6), and LsrB (1TJY), are from the Protein Data Bank (PDB) [43].

## SMILES-based virtual screening for expanded QSIMs

Chemical similarity search is a fundamental technique for ligand-based drug discovery [44]. In this study, similarity-based virtual screening is firstly used to mine the potential QSIMs from existing compounds. Note that converting an existing molecule structure into chemical information applicable for potential QSIMs development requires similarity analysis from chemical descriptor. Generally, there are one-dimensional (SMILES, [45]), two-dimensional (Mold<sup>2</sup> [46]), three-dimensional (GETAWAY, [47]), and four-dimensional (GRID, [48]) chemical descriptors of molecules [49]. SMILES describes molecular structure in the form of strings, and has been applied to chemical properties prediction [50], virtual screening [51], quantitative structure-activity relationship (QSAR) modelling [52, 53], drug-target interaction prediction [54], and machine learning-based drug design [55]. Here, SMILES is applied to the similarity-based virtual screening for the development of potential QSIMs for the first time. The specific SMILES-based similarity calculations are based on the edit distance [54] and LINGO [56].

## Edit distance

Edit distance is one of the most widely used measures to conduct comparison between strings. The edit distance between strings S1 and S2 is the minimum number of edit operations, such as insertion,

deletion, and substitution for converting S1 to S2 [54]. The edit distance similarity is calculated as follows:

$$EditS(S1, S2) = 1 - \frac{edit(S1, S2)}{Max(length(S1), length(S2))} \quad (1)$$

Having observed significant redundancy (i.e. high similarities) in a subset of the 633 reported molecules and with an aim to reduce the computational burden, we carried out a de-redundancy process with the edit distance cutoff being 0.6. According to the distribution of similarities for different distance cutoffs shown in Fig. S3, the non-redundant dataset of 142 ligands with this cutoff threshold includes most of the molecules with low SMILES similarities, while maintaining a medium computational workload.

### LINGO

Chemical linguistics with maximum common substructures, such as LINGO, are considered widely in quantifying molecular similarity, and they can be used as the computerized linguistics in comparing and searching through large text collections [57]. LINGO is based on the fragmentation of SMILES strings into overlapping substrings of defined sizes, which was developed by Vidal et al. [56]. Note that the LINGO profile is only related to the  $q$  value of LINGO and its corresponding occurrence number and does not depend on the order of LINGO occurrence in a SMILES string. LINGO-based similarity between S1 and S2 is calculated as follows:

$$LINGOsim = \frac{\sum_{i=1}^m 1 - \frac{|N_{S1,i} - N_{S2,i}|}{|N_{S1,i} + N_{S2,i}|}}{m} \quad (2)$$

where  $m$  is the total number of unique LINGO created from S1 and S2.  $N_{S1,i}$  and  $N_{S2,i}$  represent the frequency of LINGO of type  $i$  in S1 and S2, respectively.

The LINGO formats in common use are LINGO3, LINGO4, and LINGO5 with  $q$  being 3, 4, and 5, respectively. This work established that virtual screening with LINGO3 for each non-redundant reported QSIMs from PubChem database is more flexible in SMILES-based structure similarity, and it can get more potential QSIMs than LINGO4 and LINGO5. Therefore, the LINGO length was fixed as  $q = 3$ . In addition, to reduce the computational workload for docking-based validations by AutoDock Vina, the final cutoff for LINGO3 was set to 0.7. Based on the SMILES of the reported 142 QSIMs in the non-redundant dataset I, we developed a pipeline with SMILES-based algorithms and docking-based validations to mine potential QSIMs with creating a dataset II containing 73,073 compounds from existing 138,805,608 compounds (Fig. 1).

(Fig. 1)

### Docking-based validation by AutoDock Vina

AutoDock Vina is a software tool for docking calculations based on a simple scoring function and fast gradient optimization conformation search [37]. A series of pretreatments are required for receptors and ligands before docking. As an example, pretreatments for 3-oxo-C12-HSL (*las* AHL) and its QS receptor (LasR) are illustrated in Fig. 2. The strain A of 2UV0 from PDB database was selected to be LasR, and subsequently water molecules and complexed ligands were removed from the strain A by Pymol [58]. Hydrogen atoms were added, and all atoms were set to be of the Assign AD4 type. KOLLMAN charges were taken into account during charge assignment with the help of AutoDockTools [59]. At the same time, 3-oxo-C12-HSL needed to be converted into a series of file formats (sdf, pdb, and pdbqt) with the help of a variety of software tools (Open babel [60], Raccon [61] or AutoDockTools [59]) to dock with LasR.

(Fig. 2)

Many species of Gram-negative bacteria communicate with kin and coordinate group behaviors through AHL QS systems. There are many reports of “promiscuous” receptors (LuxR, LasR, TraR) that respond broadly to nonself AHL signals [62]. It has been indicated that the promiscuous communications of the AHL QS systems suit well for the chemical similarity principle [44], which states that compounds and proteins with similar structures will probably have similar bioactivities. Therefore, an underlying assumption for our proposed SMILES-based virtual screening pipeline is that chemically similar QSIMs will bind to the same or similar QS receptors and vice-versa. Note that most of the reported QS receptors do not have existing crystal structures in the PDB database, except nine QS receptors (SmcR, TraR, LasR, CviR, QscR, SdiA, PqsR, LuxP, and LsrB) we collected. However, the homology modeling of QS receptors without crystal structures can be carried out based on the nine QS receptors stated above. As listed in Table 1, the binding sites for collected nine QS receptors are fixed for their AIs and QSIMs. In addition, due to the missing of crystal structure of LuxR and the high conservatism of SmcR and LuxR [63], we represented LuxR with the crystal structure of SmcR, a transcriptional regulator from *V. vulnificus* (PDB: 3KZ9). We selected nine natural QS signals (C6HSL, C8HSL, C12HSL, 3OC6HSL, 3OC8HSL, 3OC12HSL, PQS, S-THMF-borate, and R-THMF) to dock with nine QS receptors to verify the promiscuous QS communications, determine the FBEs cutoff used in separating binding and non-



binding, and test the validity of docking-based approach. We applied a ligand (PubChem ID: 101776684), which has a high similarity with S-THMF-borate, to dock with QS receptors for none-recognition of boron in the Vina software. With the help of the FBEs cutoffs, we ranked and analyzed the docking results for the expanded 73,073 ligands and nine QS receptors.

(Table 1)

### Protein pockets similarity or QSIMs crosstalk definition

While the similarity of pockets with similar amino acid binding sites, such as the pockets of LasR, TraR, CviR, QscR, SdiA (Table 1) is relatively easy to quantify, deliberations are required for pockets with different binding sites such as SmcR and PqsR (Table 2). Here, we firstly performed a similarity analysis of the amino acid sequences for nine QS receptors. The distance analysis of the three-dimensional structures for receptors was then carried out on the root-mean-squared deviation (RMSD) scoring function [64]. RMSD measures the deviation of a target set of coordinates to a reference set of coordinates, with RMSD equal 0 indicating a perfect overlap. To deal with the protein pockets with different amino acid binding sites, we proposed a new similarity measure, termed *pocketedit*, based on the distribution of the docking FBEs, which is stated as follows:

$$pocketedit_{x,y} = \frac{N_{xy}}{N_x + N_y - N_{xy}} \quad (3)$$

where  $x$  and  $y$  represent two different protein pockets, such as Als binding pockets of LasR and TraR.  $pocketedit_{x,y}$  is the similarity between two different protein pockets.  $N_x$ ,  $N_y$  are the number of ligands that bind to  $x$  and  $y$  protein pockets, respectively.  $N_{xy}$  is the number of ligands that bind to both  $x$  and  $y$  protein pockets.

The higher the similarity of pockets, the more significant the crosstalk of two receptors will be for QSIMs. Therefore, this equation can be used to calculate the similarity of pockets and also well represents the crosstalk degree of different receptors to QSIMs. As such, *pocketedit* could be instrumental for developing the drug-like QSIMs which can act on multiple QS receptors and which are termed as the broad-spectrum QSIMs in this work.

## Results

### Reported QSIMs

Based on a large number of literatures and various databases, 633 reported QSIMs have been collected



corresponding to various QS receptors, i.e. LuxR (254), LasR (275), TraR (67), CviR (19), RhIR (28), LuxN (6), PhzR (8), QscR (15), PqsR (25), LuxP (42), LsrB (50) and other receptors (45). The analysis of the 633 reported QSIMs is conducted in two parts, namely the statistics and the docking-based validation.

**Statistics of reported QSIMs**

According to the specific distribution of 633 reported QSIMs on various QS receptors (Fig. 3a), SmcR (LuxR) from *V. vulnificus* and LasR from *P. aeruginosa* are the main targets of the reported QSIMs, followed by TraR (*R. radiobacter*), LsrB (*S. typhimurium*) and others. Among the collected QSIMs, there are 79 natural compounds from various organisms, accounting for 12.48% (Fig. 3b). Due to the small concentration produced and the toxicity of the natural QSIMs, more and more synthetic QSIMs have been developed gradually by modifying the original AIs. For example, the development of synthetic QSIMs for interfering AHL QS systems is mainly focused on the modifications in the AHL side chain and the AHL ring moiety, which account for 31.91% and 35.55%, respectively. Totally among the 633 reported QSIMs, AHL-based ligands account for the majority (71.57%), followed by AI-2 analogues (13.43%), natural QSIMs and PQS analogues (2.53%).

The 633 QSIMs we have collected include QS agonists and QS antagonists. There are 267 QS antagonists and 62 QS agonists (Fig. 3c) that have specific IC<sub>50</sub> values and EC<sub>50</sub> values, respectively. Furthermore, it is shown that 21 molecules can act as both QS agonists and QS antagonists for different QS receptors (Table 2). Note that the total number of the QSIMs for the receptors in Fig. 3a is more than 633, which means that there are some QSIMs crosstalk among different QS receptors. To better understand crosstalk, we analyzed the distribution of QSIMs for LuxR (SmcR), LasR, TraR, LuxP, and LsrB (Fig. 3d), which shows a certain degree of QSIMs crosstalk among LuxR (SmcR), LasR and TraR. Additionally, QSIMs crosstalk exists between LuxP and LsrB, too.

(Fig. 3), (Table 2)

**Docking results for reported QSIMs and receptors**

The distribution of docking free binding energies (FBEs) among the nine typical QS receptors and their natural corresponding AIs is listed in the heatmap of Fig. 4a. The FBEs of six AHLs (C6HSL, C8HSL, C12HSL, 3OC6HSL, 3OC8HSL, 3OC12HSL) with different side chains and the reported AHL-binding QS receptors (SmcR, LasR, TraR, CviR, QscR, and SdiA) are with high similarities. It verified the

existing QS signals crosstalk among the promiscuous AHL-based communications. Furthermore, it can be seen that the FBEs of six AHLs and AHL-binding receptors stated above are lower than -6 kCal/mol, except the FBE of C6-HSL and SmcR which is -5.9 kCal/mol. As intercellular signal molecules, S-THMF-borate and R-THMF are often regarded as the non-binder to the AHL-binding receptors, and their FBEs to AHL-binding receptors are greater than -6 kCal/mol. Therefore, the FBEs cutoff used in separating binding and non-binding was set to -6 kCal/mol. It means that if the FBE is greater than -6 kCal/mol, there is little binding between the QS-based ligand and the receptor.

Note that the FBE of 101776684 and LuxP is -9.3 kCal/mol, which is much lower than -6 kCal/mol. As illustrated in Fig. 4b, the hydrogen bonding of protein LuxP and 101776684 includes Gln77, Ser79, Trp82, Asn159, Arg215, Thr266 and Arg310, which are included in the binding sites of LuxP and S-THMF-borate (Table 2). Therefore, it is feasible to replace the S-THMF-borate with 101776684 to dock with LuxP. The FBEs of PQS to SmcR, LasR, TraR, CviR, QscR, and SdiA are -6.9, -9.8, -9.5, -9.0, -9.6, and -8.7 kCal/mol, respectively, which are lower than that of its original receptor (PqsR, -6.8 kCal/mol). This indicates that the bacteria with any of these six QS receptors, such as *Vibrio fischeri*, *Chromobacterium violaceum*, and *Escherichia coli* can respond to the PQS, which is one of the QS signals secreted by *Pseudomonas aeruginosa*. In addition, the distribution of FBEs involving LuxP and LsrB, two QS receptors for the intercellular communications, is rather different from that of seven other QS receptors for intracellular communications, which agrees well with the situation of real bacterial QS interactions. These results offer a validation of the docking-based approach and support its application for the screening of potential QSIMs.

(Fig. 4)

### Expanded QSIMs

Similar to the analysis of the reported QSIMs, the results for expanded QSIMs are also reported in two parts: the statistics and the docking-based results.

#### Statistics of expanded QSIMs

The docking results for the potential QSIMs and nine QS receptors (SmcR, TraR, LasR, CviR, QscR, SdiA, PqsR, LuxP, and LsrB) are illustrated in Fig. 5. To categorize FBEs better, we set -8 kCal/mol as another cutoff in separating strong and weak binding. As for SmcR, there are 3,029 ligands (4.15%) which have FBEs lower than -8 kCal/mol, while the numbers of ligands for LasR, TraR, CviR, QscR, SdiA, and PqsR are 44,674 (61.14%), 30,185 (41.31%), 18,910 (25.88%), 28,867 (39.5%), 24,690

(33.79%), and 8,091 (11.07%), respectively (Fig. 5 and Fig. S2). The fractions of FBEs ( $\leq -8$  kCal/mol) for LuxP and LsrB are 0.09% and 0%, respectively, which are much lower than those of the above seven QS receptors. On the fractions of FBEs being greater than 0 kCal/mol (completely incapable of QSI), LsrB ranks first (39,729, 54.37%), followed by LuxP (16,459, 22.52%), SdiA (7,275, 9.96%), CviR (1,165, 1.59%), TraR (1,137, 1.56%), QscR (674, 0.92%), PqsR (6, 0.01%). With the help of the FBEs cutoffs being -6 kCal/mol and -8 kCal/mol, we ranked the FBEs among the expanded 73,073 ligands and nine QS receptors.

(Fig. 5)

### Docking-based results for expanded QSIMs and receptors

Based on the FBEs ranking of the expanded 73,073 ligands and nine QS receptors, we identified the ligands that bind most strongly to their corresponding QS receptors, which are listed in Fig. 6. The ligands 56324809, 101018536, 3424041, 20279703, 109057410, 118004297, 13782492, 131400990, and 45084042 (PubChem CID) have the lowest FBEs for binding to SmcR, TraR, LasR, CviR, QscR, PqsR, SdiA, LuxP, and LsrB, respectively. According to the FBEs distribution, it can be seen that there are serious QSIMs crosstalks among the nine molecules stated above and seven QS receptors (SmcR, LasR, TraR, CviR, QscR, PqsR and SdiA). The FBEs distribution of LuxP and LsrB is quite different from the other seven receptors. The specific binding sites of each QS receptor and the corresponding strongest binding ligands are shown in Fig. 6. The binding sites of LuxP and 131400990 include Ser79, Arg215, Thr266, Trp289 and Arg310, which are rather similar to those of 101776684 and LuxP (Gln 77, Ser79, Trp82, Asn159, Arg215, Thr266, Arg310) (Fig. 4b). This suggests that structurally the 131400990 has a significant potential to be one of the QSIMs for LuxP receptor.

(Fig. 6)

### Data retrieval

#### Browse

A user-friendly 'BROWSE' option allows to explore the QSIMs data by two-tiers: the reported QSIMs and the extended QSIMs (Fig. 7). For the reported QSIMs, a separate browse option is provided where the visualized data can be sought by the type of molecules, such as AHLs, acyl-chain modifications for AHLs, lactone-ring modifications for AHLs, PQS analogues, AI-2 analogues and natural QSIMs.

For the expanded QSIMs, another separate browse option is provided where the visualized data can be sought by the receptors, such as LasR and TraR. The users can choose required categories for further details. For a better browsing experience, a drop-down menu is added to filter ligands according to their FBEs range. In addition, we added the PubChem database links for the reported and the extended QSIMs to make it easier to obtain the detailed chemoinformatic description.

(Fig. 7)

## Search

QSIdb includes "SEARCH" and "EXPAND" searching facilities for the reported and expanded QSIMs, respectively. 'SEARCH' is provided to query some reported QSIMs according to different options: ALL, Category (molecular type of ligands), PubChem CID, QS receptors or SMILES. The output displays information of the reported QSIMs, fielded by PubChem CID, Binding protein for antagonists, Receptor for agonists, SMILES, and ligand type (Category). 'EXPAND' is provided to query extended QSIMs according to different fields: PubChem CID, SMILES of ligands or the QS receptors (Fig. 8). For a better searching experience, a drop-down menu is added to filter ligands according to their FBEs range. Output displays information of potential QSIMs, fielded by PubChem ID, SMILES, FBEs,  $K_i$  (inhibitor constant of the enzyme-inhibitor complex,  $\mu\text{M}$ ),  $K_d$  (dissociation constant,  $\mu\text{M}$ ), and  $\text{IC}_{50}$  ( $\text{EC}_{50}$ ) for nine QS receptors. To make it more convenient for getting the detailed chemoinformatic description, we added the PubChem database links for the reported and extended QSIMs. The users can specify any field against which one wishes to search, or else keep the default 'all' option which will search against all fields in the database. Besides the option to choose the fields, search type allows to retrieve either an exact match or the match containing query.

(Fig. 8)

## Discussion

We conducted similarity analysis on the amino acid sequences and three-dimensional structures for nine QS receptors, both reflecting the features of the entire proteins, which are not necessarily always suitable for identifying potential QSIMs. For example, the amino acid sequence and three-dimensional structure similarities of SmcR and PqsR are 0.07 and 12.41 (RMSD distance), respectively (Fig. S1a and Fig. S1b). It means that both amino acid sequence and the three-dimensional structure similarities

of SmcR and PqsR are relatively low. However, as illustrated in Fig. 5, there is a considerable similarity between the FBEs distributions of SmcR and PqsR receptors. In this work, we proposed and applied *pocketedit* (equation 3) to analyze the similarity of the receptor pockets with different amino acid binding sites. LasR leads among the nine proteins with 67,359 potential active ligands (FBEs lower than -6 kcal/mol), followed by QscR (63,345), TraR (62,184), PqsR (61,320), CviR (58,237), SdiA (53,011), SmcR (48,949), LuxP (16,785) and LsrB (2,329) (Fig. 9a). The number of ligands binding nine receptors was calculated pairwise, and the results of *pocketedit* for various pairwise are shown in Fig. 9b. Note that the amino acid binding sites of the pockets of two receptors (SmcR, PqsR) are different from those of five receptors (LasR, TraR, CviR, QscR and SdiA) (Table 1), but they have quite similar FBEs distributions (Fig. 5) and *pocketedit values* (Fig. 9b). Therefore, different receptors with low similarities in amino acid sequences and three-dimensional structures can have high similarities in binding pockets, such as SmcR and PqsR. The similarity of protein pockets can be characterized by the similarity of binding ligands distribution and even calculated by the *pocketedit*.

(Fig. 9)

*Pocketedit* can be applied to calculate and analyze the similarity of protein pockets as well as the crosstalk of QSIMs for various QS receptors. The higher the *pocketedit* is, the higher degree the crosstalk of two receptors will be of. In the QS-based applications for the dynamic control of microbial consortia, most researchers hope to eliminate QS crosstalk to obtain orthogonal QS systems. However, more attentions should be paid to those QSIMs with higher crosstalk for multiple targets (receptors) to develop broad-spectrum QSIMs. In order to obtain the potential broad-spectrum QSIMs for nine QS receptors, we analyzed the *pocketedit*s and the distributions of 73,073 ligands, and obtained 273 potential broad-spectrum QSIMs (Table S1). Ten of them are listed in Table 3, which are expected to be verified by experiments in the future.

(Table 3)

To reduce the computational workload and save time, SMILES-based similarity assessment is based on the simple similarity calculation method, such as edit distance or LINGO3, in the pipeline to screen the potential QSIMs from the PubChem database. The cutoff should be set as an appropriate value (0.7 for LINGO3 in this study) based on the analyzation of the simulation results distribution with different

**cutoffs.** Recently, several evolved SMILES-based representations have been reported, such as SMILES2Vec [50], Mol2Vec [55], FP2Vec [53], which have combined with the machine learning techniques to predict the properties of the potential molecules. These evolved SMILES-based representations could be applied in our pipeline for the screening of the potential QSIMs with low sequence but high functional similarities from a much smaller database. In addition, to improve the accuracy and diversity of the chemical similarity-based virtual screening, some other chemical descriptors (e.g. 3D [47] or 4D [48]) and fingerprints (e.g. ECFP [65]) can be applied in our pipeline to mine the potential QSIMs.

## Conclusion

Quorum sensing interference (QSI) is an effective strategy to deal with infection and disease caused by various pathogenic or even multidrug-resistant bacteria. QSI can also be applied in synthetic biology to realize the population-level control of bacteria. QSldb is a comprehensive resource of the reported and potential QSI molecules (QSIMs). There are 633 reported QSIMs in QSldb, which can be searched conveniently by protein types or ligand types, out of which 308 QSIMs have specific  $IC_{50}$  values or  $EC_{50}$  values. Based on the SMILESs of the non-redundant dataset (containing 142 reported QSIMs), we developed a pipeline with SMILES-based algorithms and docking-based validations to create a potential QSIMs dataset (73,073 compounds) from the existing 138,805,608 compounds in the PubChem database. With the help of the FBEs cutoff, we ranked and analyzed the docking results among the expanded 73,073 ligands and nine QS receptors. In addition, *pocketedit* as a new measure we proposed for receptor pockets similarity calculation or QSIMs crosstalk evaluation will do much help in developing the potential multi-target QSIMs. Thus, we obtained 273 potential broad-spectrum QSIMs for nine QS receptors based on the analysis of the *pocketedit* values and the potential QSIMs distributions. Here, most QSIMs have been collected and integrated in our platform that can accelerate the research in field of quorum quenching therapeutics, superbug therapy without antibiotics, potential broad-spectrum QSIMs development and population-level control of bacteria. Researchers can explore the ligands listed in our collection and test them by experiments to understand the complex patterns of QS networking and contribute to various applications.

## Key Points

- QSIdb, a database of 633 reported and 73,073 potential QSIMs, is carefully constructed focusing on the modifications and synthetic analogues of various autoinducers for the typical QS receptors.
- The ligands regarded as both agonists and antagonists with specific concentrations are carefully curated in QSIdb for the first time.
- A pipeline with SMILES-based algorithms and docking-based validations is developed to mine the potential QSIMs from the existing 138,805,608 compounds.
- *pocketedit* is proposed as a new measure to calculate receptor pockets similarity and evaluate QSIMs crosstalk.
- 273 possible potential broad-spectrum QSIMs are obtained for nine QS receptors.

## Data availability

QSIdb, a database of 633 reported and 73,073 potential quorum sensing interference molecules for various QS systems, which is freely available at: (<http://qsidb.lbcj.net/>).

We will continuously update the database QSIdb.

The data are available from the corresponding author upon request with email.

## Supplementary data

Supplementary Data will be available online.

## Acknowledgements

This study was supported by the National Key Research and Development Project of China (2017YFD0201400), the National Natural Science Foundation of China (31570089, 31770076, 61772362), and the Funds for Creative Research Groups of China (21621004), Dr. Jianjun Qiao was supported by The New Century Outstanding Talent Support Program, Education Ministry of China.

## Conflict of interest

The authors declare no competing financial interests.



## References

1. Wu S, Liu J, Liu C, et al. Quorum sensing for population-level control of bacteria and potential therapeutic applications. *Cell Mol Life Sci* 2020;**77**:1319-43.
2. Wang S, Payne GF, Bentley WE. Quorum sensing communication: molecularly connecting cells, their neighbors, and even devices. *Annu Rev Chem Biomol Eng* 2020;**11**:447-68.
3. Papenfort K, Bassler BL. Quorum sensing signal-response systems in Gram-negative bacteria. *Nat Rev Microbiol* 2016;**14**:576-88.
4. Monnet V, Juillard V, Gardan R. Peptide conversations in Gram-positive bacteria. *Crit Rev Microbiol* 2014;**42**:339-51.
5. Xavier KB, Bassler BL. Interference with AI-2-mediated bacterial cell-cell communication. *Nature* 2005;**437**:750-53.
6. Lee JH, Lee J. Indole as an intercellular signal in microbial communities. *FEMS Microbiol Rev* 2010;**34**:426-44.
7. Stephens K, Bentley WE. Synthetic biology for manipulating quorum sensing in microbial consortia. *Trends Microbiol* 2020. DOI: 10.1016/j.tim.2020.03.009.
8. Gupta A, Reizman IM, Reisch CR, et al. Dynamic regulation of metabolic flux in engineered bacteria using a pathway-independent quorum-sensing circuit. *Nat Biotechnol* 2017;**35**:273-79.
9. Zhao X, Yu Z, Ding T. Quorum-sensing regulation of antimicrobial resistance in bacteria. *Microorganisms* 2020;**8**:425-46.
10. Piewngam P, Zheng Y, Nguyen TH, et al. Pathogen elimination by probiotic *Bacillus* via signalling interference. *Nature* 2018;**562**:532-37.
11. O'Loughlin CT, Miller LC, Siryaporn A, et al. A quorum-sensing inhibitor blocks *Pseudomonas aeruginosa* virulence and biofilm formation. *Pnas* 2013;**110**:17981-86.
12. Rasmussen TB, Bjarnsholt T, Skindersoe ME, et al. Screening for quorum-sensing inhibitors (QSI) by use of a novel genetic system, the QSI selector. *J Bacteriol* 2005;**187**:1799-14.
13. Müh U, Schuster M, Heim R, et al. Novel *Pseudomonas aeruginosa* quorum-sensing inhibitors identified in an ultra-high-throughput screen. *Antimicrob Agents Chemother* 2006;**50**:3674-79.
14. Geske GD, Wezeman RJ, Siegel AP, et al. Small molecule inhibitors of bacterial quorum sensing and biofilm formation. *J Am Chem Soc* 2005;**127**:12762-63.
15. Zhu P, Li M. Recent progresses on AI-2 bacterial quorum sensing inhibitors. *Curr Med Chem* 2012;**19**:174-86.
16. Annapoorani A, Umamageswaran V, Parameswari R, et al. Computational discovery of putative quorum sensing inhibitors against LasR and RhIR receptor proteins of

- Pseudomonas aeruginosa. *J Comput Aided Mol Des* 2012;**26**:1067-77.
17. Kalia M, Singh PK, Yadav VK, et al. Structure based virtual screening for identification of potential quorum sensing inhibitors against LasR master regulator in Pseudomonas aeruginosa. *Microb Pathog* 2017;**107**:136-43.
18. Soulere L, Sabbah M, Fontaine F, et al. LuxR-dependent quorum sensing: computer aided discovery of new inhibitors structurally unrelated to N-acylhomoserine lactones. *Bioorg Med Chem Lett* 2010;**20**:4355-58.
19. Liu HB, Kim JS, Park S. Development and comparison of whole-cell assay systems for quorum-sensing inhibitors based on TraR, LasR, and QscR. *J Biomol Screen* 2011;**16**:986-94.
20. Hai-Bo L, Joon-Hee L, Kim JS, et al. Inhibitors of the Pseudomonas aeruginosa quorum-sensing regulator, QscR. *Biotechnol Bioeng* 2010;**106**:119-26.
21. Ilangovan A, Fletcher M, Rampioni G, et al. Structural basis for native agonist and synthetic inhibitor recognition by the Pseudomonas aeruginosa quorum sensing regulator PqsR (MvfR). *PLoS Pathog* 2013;**9**:e1003508.
22. Rajput A, Kaur K, Kumar M. SigMol: repertoire of quorum sensing signaling molecules in prokaryotes. *Nucleic Acids Res* 2016;**44**:634-39.
23. Rajput A, Gupta AK, Kumar M. Prediction and analysis of quorum sensing peptides based on sequence features. *PLoS One* 2015;**10**:e0120066.
24. Li H, Gao Z, Kang L, et al. TarFisDock: a web server for identifying drug targets with docking approach. *Nucleic Acids Res* 2006;**34**:W219-24.
25. Amemiya T, Koike R, Kidera A, et al. PSCDB: a database for protein structural change upon ligand binding. *Nucleic Acids Res* 2012;**40**:D554-58.
26. Yang J, Roy A, Zhang Y. BioLiP: a semi-manually curated database for biologically relevant ligand-protein interactions. *Nucleic Acids Res* 2013;**41**:D1096-103.
27. Wang C, Hu G, Wang K, et al. PDID: database of molecular-level putative protein-drug interactions in the structural human proteome. *Bioinformatics* 2016;**32**:579-86.
28. Fernandes FC, Rigden DJ, Franco OL. Prediction of antimicrobial peptides based on the adaptive neuro-fuzzy inference system application. *Biopolymers* 2012;**98**:280-87.
29. Thakur N, Qureshi A, Kumar M. AVPpred: collection and prediction of highly effective antiviral peptides. *Nucleic Acids Res* 2012;**40**:W199-204.
30. Wang G, Li X, Wang Z. APD3: the antimicrobial peptide database as a tool for research and education. *Nucleic Acids Res* 2016;**44**:D1087-93.
31. Rajput A, Thakur A, Sharma S, et al. aBiofilm: a resource of anti-biofilm agents and their potential implications in targeting antibiotic drug resistance. *Nucleic Acids Res* 2018;**46**:D894-D900.
32. Kim S, Thiessen PA, Bolton EE, et al. PubChem substance and compound databases. *Nucleic Acids Res* 2016;**44**:D1202-13.

33. Gaulton A, Hersey A, Nowotka M, et al. The ChEMBL database in 2017. *Nucleic Acids Res* 2017;**45**:D945-D54.
34. Goodman J. Computer software review: Reaxys. *J Chem Inf Model* 2009;**49**:2897-98.
35. Shoichet B, Irwin J. ZINC – A free database of commercially available compounds for virtual screening. *J Chem Inf Model* 2005;**45**:177-82.
36. Weininger D. SMILES, a chemical language and information system. 1. Introduction to methodology and encoding rules. *J Chem Inf Comput Sci* 1988;**28**:31-36.
37. Trott O, Olson AJ. AutoDock Vina: improving the speed and accuracy of docking with a new scoring function, efficient optimization, and multithreading. *J Comput Chem* 2010;**31**:455-61.
38. Goo E, An JH, Kang Y, et al. Control of bacterial metabolism by quorum sensing. *Trends Microbiol* 2015;**23**:567-76.
39. Bassler BL, Wright M, Showalter RE, et al. Intercellular signalling in *Vibrio harveyi*: sequence and function of genes regulating expression of luminescence. *Mol Microbiol* 1993;**9**:773-86.
40. Chen X, Schauder S, Potier N, et al. Structural identification of a bacterial quorum-sensing signal containing boron. *Nature* 2002;**415**:545-49.
41. Sedlmayer F, Hell D, Muller M, et al. Designer cells programming quorum-sensing interference with microbes. *Nat Commun* 2018;**9**:1822-35.
42. Lee JH, Wood TK, Lee J. Roles of Indole as an Interspecies and Interkingdom Signaling Molecule. *Trends Microbiol* 2015;**23**:707-18.
43. Berman HM, Westbrook J, Feng Z, et al. The Protein Data Bank. *Nucleic Acids Res* 2000;**106**:149-58.
44. Lo YC, Rensi SE, Torng W, et al. Machine learning in chemoinformatics and drug discovery. *Drug Discov Today* 2018;**23**:1538-46.
45. Ozturk H, Ozkirimli E, Ozgur A. A novel methodology on distributed representations of proteins using their interacting ligands. *Bioinformatics* 2018;**34**:i295-i303.
46. Hong H, Xie Q, Ge W, et al. Mold(2), molecular descriptors from 2D structures for chemoinformatics and toxicoinformatics. *J Chem Inf Model* 2008;**48**:1337-44.
47. Saiz-Urra L, Gonzalez MP, Fall Y, et al. Quantitative structure-activity relationship studies of HIV-1 integrase inhibition. 1. GETAWAY descriptors. *Eur J Med Chem* 2007;**42**:64-70.
48. Ash J, Fourches D. Characterizing the chemical space of ERK2 kinase inhibitors using descriptors computed from molecular dynamics trajectories. *J Chem Inf Model* 2017;**57**:1286-99.
49. Rognan D. Chemogenomic approaches to rational drug design. *Br J Pharmacol* 2010;**152**:38-52.
50. Goh G, Hodas N, Siegel C, et al. SMILES2Vec: an interpretable general-purpose deep

- neural network for predicting chemical properties. *arXiv preprint* 2017;**arXiv**:1712.02034. DOI: 10.475/123\_4.
51. Miljkovic F, Rodriguez-Perez R, Bajorath J. Machine learning models for accurate prediction of kinase inhibitors with different binding modes. *J Med Chem* 2019. DOI: 10.1021/acs.jmedchem.9b00867.
52. Toropov AA, Benfenati E. SMILES in QSPR/QSAR Modeling: results and perspectives. *Curr Drug Discov Technol* 2007;**4**:77-116.
53. Jeon W, Kim D. FP2VEC: a new molecular featurizer for learning molecular properties. *Bioinformatics* 2019;**35**:4979-85.
54. Ozturk H, Ozkirimli E, Ozgur A. A comparative study of SMILES-based compound similarity functions for drug-target interaction prediction. *BMC Bioinformatics* 2016;**17**:128-39.
55. Jaeger S, Fulle S, Turk S. Mol2vec: unsupervised machine learning approach with chemical intuition. *J Chem Inf Model* 2018;**58**:27-35.
56. Vidal D, Thormann M, Pons M. LINGO, an efficient holographic text based method to calculate biophysical properties and intermolecular similarities. *J Chem Inf Model* 2005;**45**:386-93.
57. Wozniak M, Wolos A, Modrzyk U, et al. Linguistic measures of chemical diversity and the "keywords" of molecular collections. *Sci Rep* 2018;**8**:7598.
58. Delano WL. The PyMol molecular graphics system. *Protns Structure Function Bioinformatics* 2002;**30**:442-54.
59. Morris GM, Ruth H, William L, et al. AutoDock4 and AutoDockTools4: Automated docking with selective receptor flexibility. *J Comput Chem* 2010;**30**:2785-91.
60. O'Boyle NM, Banck M, James CA, et al. Open Babel: An open chemical toolbox. *J Cheminform* 2011;**3**:33-47.
61. Forli S, Huey R, Pique ME, et al. Computational protein–ligand docking and virtual drug screening with the AutoDock suite. *Nat Protoc* 2016;**11**:905-19.
62. Wellington S, Greenberg EP. Quorum sensing signal selectivity and the potential for interspecies cross talk. *mBio* 2019;**10**:e00146-19.
63. Rajamanikandan S, Jeyakanthan J, Srinivasan P. Binding mode exploration of LuxR-thiazolidinedione analogues, e-pharmacophore-based virtual screening in the designing of LuxR inhibitors and its biological evaluation. *J Biomol Struct Dyn* 2017;**35**:897-916.
64. Roe DR, Cheatham TE, 3rd. PTRAJ and CPPTRAJ: Software for processing and analysis of molecular dynamics trajectory data. *J Chem Theory Comput* 2013;**9**:3084-95.
65. Rogers D, Hahn M. Extended-connectivity fingerprints. *J Chem Inf Model* 2010;**50**:742-54.

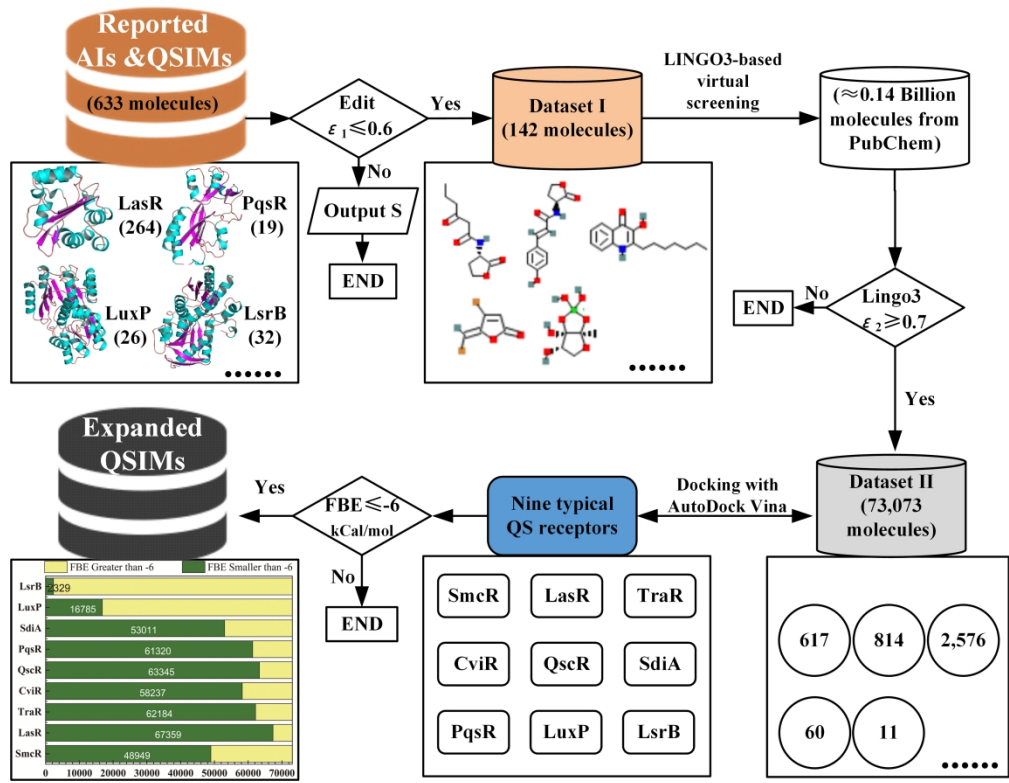


Fig. 1. Schematic diagram of the virtual screening pipeline for developing potential QSIMs

189x148mm (600 x 600 DPI)

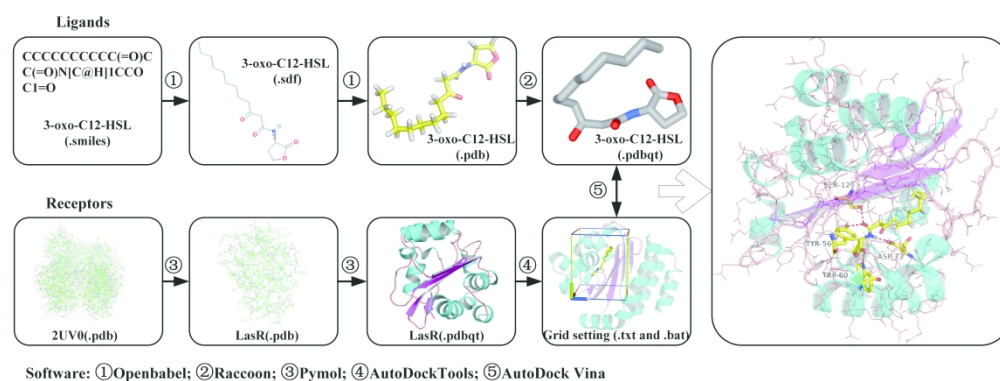


Fig 2. Flowsheet for dealing with receptors and ligands

178x67mm (1000 x 1000 DPI)

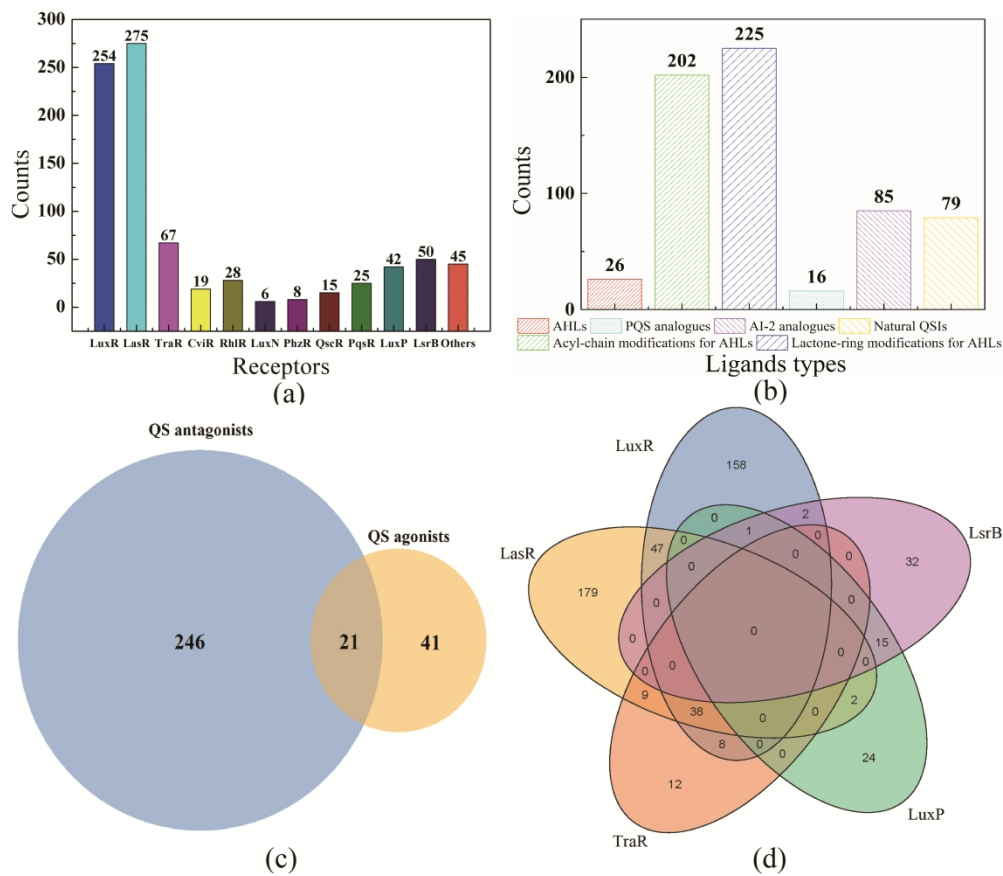


Fig. 3. Statistics of the reported QSIMs. (a) The distribution and counts of ligands on various QS receptors (LuxR, LasR, TraR, CviR, RhlR, LuxN, PhzR, QscR, PqsR, LuxP, LsrB, and other receptors); (b) The distribution and percent of ligands on their own type. (c) Distribution for QS agonists and QS antagonists. (d) QSIMs crosstalk analysis on the QSIMs distributions of five proteins (LuxR, LasR, TraR, LuxP, and LsrB).

174x152mm (600 x 600 DPI)



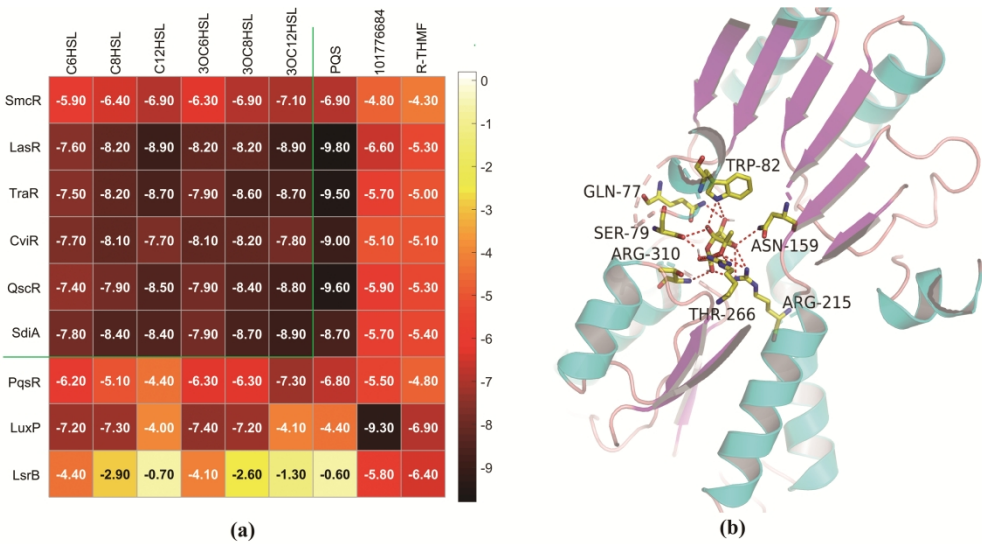


Fig. 4. FBEs results for the reported QSIMs and QS receptors. (a) FBEs heat map for the nine typical QS receptors and their corresponding ligands; (b) Hydrogen bonding of LuxP and the 101776684 molecule.

178x95mm (600 x 600 DPI)

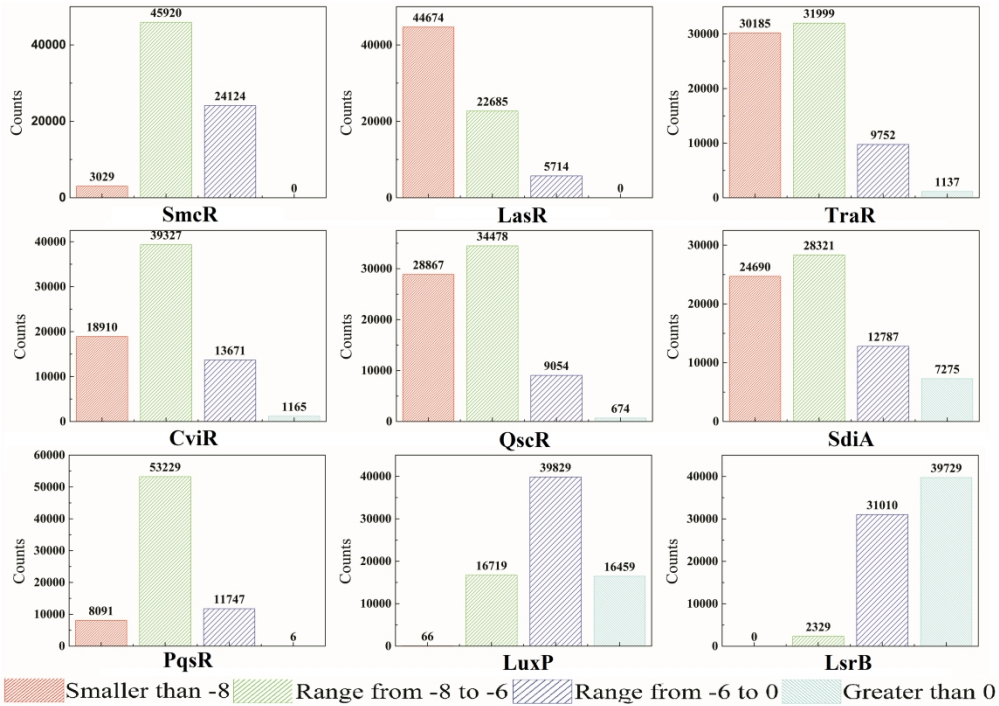


Fig. 5. FBEs distributions for the expanded ligands and nine QS receptors.

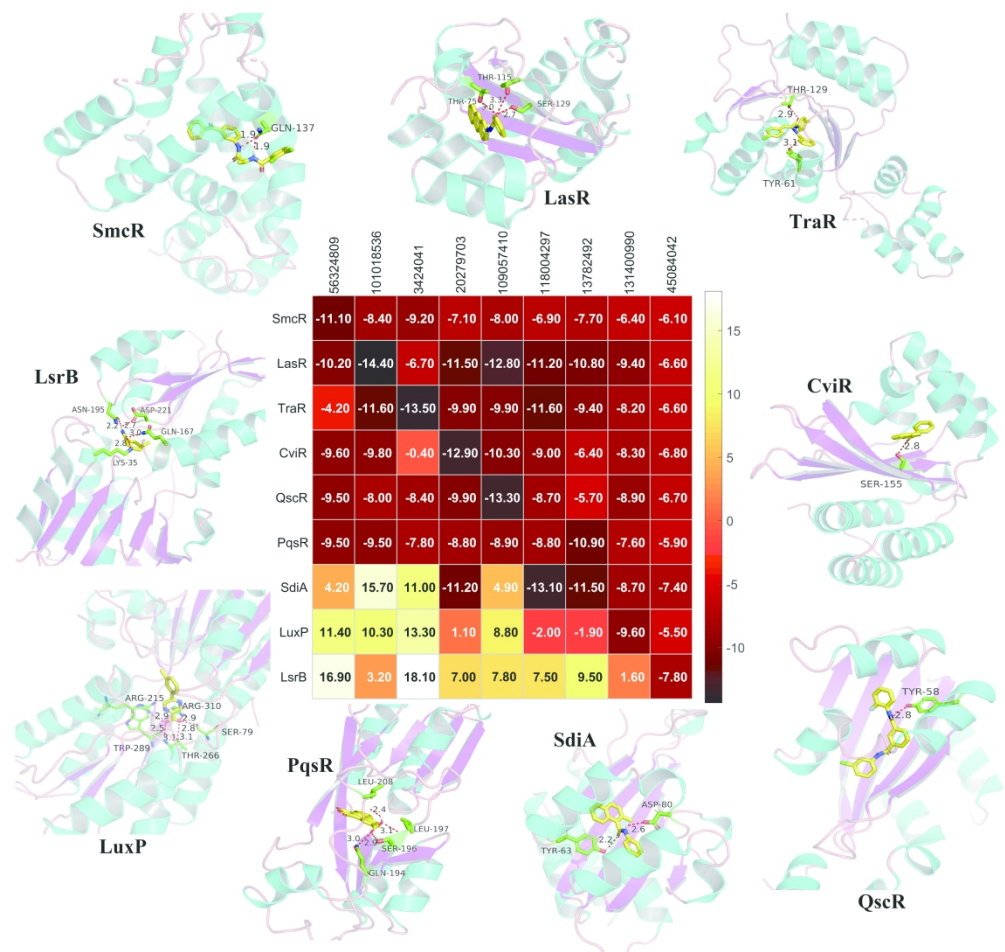


Fig 6. Typical and the strongest binding FBEs results for the screened nine ligands and receptors.

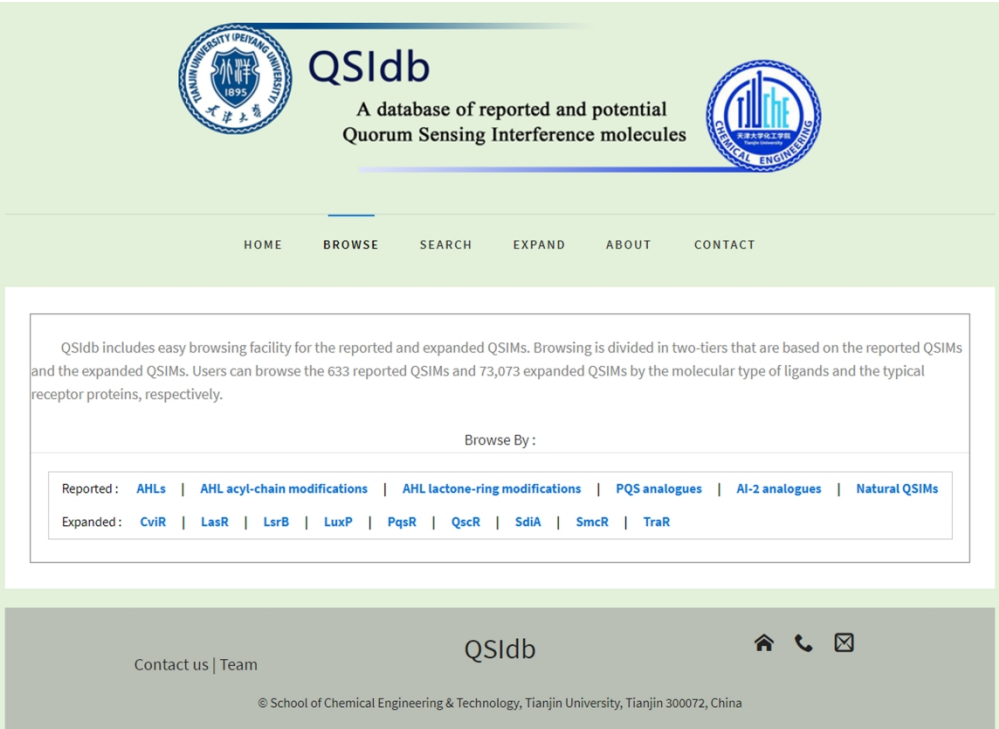


Fig 7. Browse facility for the reported and expanded QSIMs.

139x101mm (600 x 600 DPI)

1  
2  
3  
4  
5  
6  
7  
8  
9  
10  
11  
12  
13  
14  
15  
16  
17  
18  
19  
20  
21  
22  
23  
24  
25  
26  
27  
28  
29  
30  
31  
32  
33  
34  
35  
36  
37  
38  
39  
40  
41  
42  
43  
44  
45  
46  
47  
48  
49  
50  
51  
52  
53  
54  
55  
56  
57  
58  
59  
60

QSIdb  
A database of reported and potential  
Quorum Sensing Interference molecules

HOME BROWSE SEARCH EXPAND ABOUT CONTACT

QSIdb includes "SEARCH" and "EXPAND" searching facilities for reported and expanded QSIMs, respectively. In this search option, query box is provided in which user can enter the query on the basis of PubChem ID and SMILES of ligands or the typical receptor proteins for the searching of about 73,073 expanded QSIMs.

Expand By :

☒ PubChem CID [e.g, 10130163]  
☐ SMILES [e.g, C=CCSCC=C]  
☐ Receptor [e.g, CviR]

Go Reset

Contact us | Team QSIdb

© School of Chemical Engineering & Technology, Tianjin University, Tianjin 300072, China

Fig 8. "EXPAND" searching facility for the expanded QSIMs.

139x117mm (600 x 600 DPI)

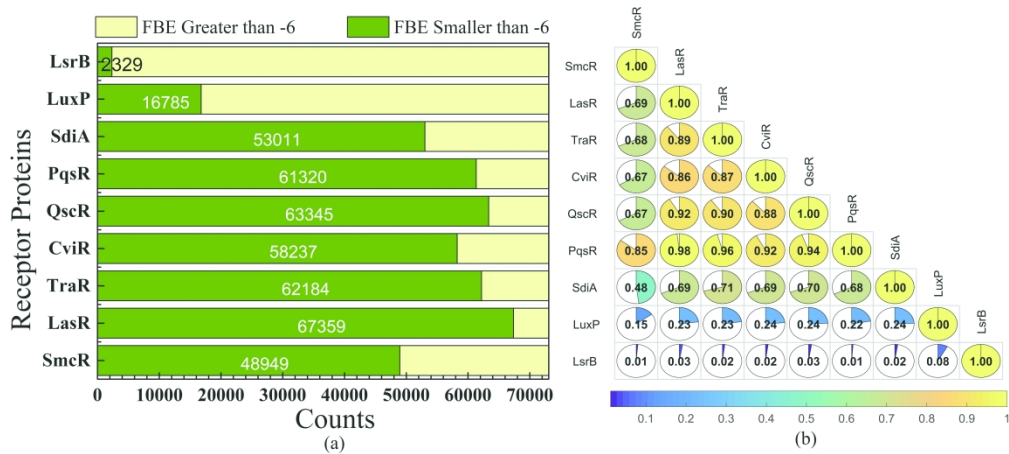


Fig 9. The analysis and discussion for the receptors binding pockets similarity. (a) Potential QSIs distribution for nine receptors; (b) Results of pocketeditx,y for various receptor pairwise.

178x79mm (600 x 600 DPI)

Table 1. Fixed binding site for the nine typical receptors

Receptors	PDB code	Hydrogen bonding	References
LasR	2UV0	Tyr56, Trp60, Asp73, Ser129	[1]
TraR	1L3L	Tyr53, Trp57, Asp70	[2]
CviR	3QP1	Tyr80, Trp84, Asp97, Ser155	[3]
QscR	3SZT	Tyr58, Trp62, Asp75, Ser38	[4]
SdiA	4Y15	Tyr63, Trp67, Asp80, Ser43	[5]
SmcR	3KZ9	Asn133, Gln137	[6]
PqsR	4JVD	Ile149, Ile168, Leu207, Ile236, Thr265	[7]
LuxP	1JX6	Gln77, Ser79, Trp82, Asn159, Arg215, Thr266, Asp267, Trp289, Arg310	[8]
LsrB	1TJY	Lys35, Asp116, Asp166, Gln167, Ala222, Pro220	[9]

#### Reference

1. Bottomley MJ, Muraglia E, Bazzo R et al. Molecular insights into quorum sensing in the human pathogen *Pseudomonas aeruginosa* from the structure of the virulence regulator LasR bound to its autoinducer, *Journal of Biological Chemistry* 2007;282:13592-13600.
2. Zhang RG, Pappas KM, Brace JL et al. Structure of a bacterial quorum-sensing transcription factor complexed with pheromone and DNA, *Nature* 2002;417:971-974.
3. Chen G, Swem LR, Swem DL et al. A strategy for antagonizing quorum sensing, *Molecular Cell* 2011;42:199-209.
4. Lintz MJ, Oinuma K, Wysoczynski CL et al. Crystal structure of QscR, a *Pseudomonas aeruginosa* quorum sensing signal receptor, *Proceedings of the National Academy of Sciences of the United States of America* 2011;108:15763-15768.
5. Nguyen Y, Nguyen NX, Rogers JL et al. Structural and mechanistic roles of novel chemical ligands on the SdiA quorum-sensing transcription regulator, *mBio* 2015;6:e02429-02414.
6. Rajamanikandan S, Jeyakanthan J, Srinivasan P. Binding mode exploration of LuxR-thiazolidinedione analogues, e-pharmacophore-based virtual screening in the designing of LuxR inhibitors and its biological evaluation, *Journal of Biomolecular Structure and Dynamics* 2017;35:897-916.
7. Ilangovan A, Fletcher M, Rampioni G et al. Structural basis for native agonist and synthetic inhibitor recognition by the *Pseudomonas aeruginosa* quorum sensing regulator PqsR (MvfR), *PLoS Pathogens* 2013;9:e1003508.
8. Chen X, Schauder S, Potier N et al. Structural identification of a bacterial quorum-sensing signal containing boron, *Nature* 2002;415:545-549.
9. Miller, Stephen T, Xavier et al. *Salmonella typhimurium* Recognizes a Chemically Distinct Form of the Bacterial Quorum-Sensing Signal AI-2, *Molecular Cell* 2004;15:677-687.



Table 2. The ligands regarded as both agonists and antagonists with specific concentrations

Pub CID	Receptors	IC <sub>50</sub> (μM)	Receptors	EC <sub>50</sub> (μM)	SMILES
127293	LasR, TraR	0.11, *	PhzR	0.011	<chem>CCCCC(=O)CC(=O)N[C@H]1CCOC1=O</chem>
10221437	LasR	0.04	PhzR,	0.005	<chem>CCCCCCCCCCCC(=O)N[C@H]1CCOC1=O</chem>
25014956	LuxR, TraR	4.7, 1.3	LasR	3.4	<chem>C1COC(=O)[C@H]1NC(=O)CC2=CC=C(C=C2)SC(F)(F)F</chem>
10131281	LuxR, TraR, LasR	0.4, 0.25, 1.05	PhzR	0.006	<chem>CCCCCCCCCCC(=O)N[C@H]1CCOC1=O</chem>
15384158	LuxR	3.4	LasR	>10	<chem>C1COC(=O)[C@H]1NC(=O)CCC2=CC3=C(C=C2)OCO3</chem>
11536248	TraR	0.83	LasR	0.54	<chem>C1COC(=O)[C@H]1NC(=O)CC(=O)CC2=CC=CC=C2</chem>
25014704	LasR	3.97	LuxR	>200	<chem>C1COC(=O)[C@H]1NC(=O)CC2=CC(=CC=C2)Cl</chem>
25014624	LasR	4.06	LuxR	>50	<chem>C1COC(=O)[C@H]1NC(=O)CC2=CC(=CC=C2)Br</chem>
25014709	LasR	0.61	LuxR	0.35	<chem>C1COC(=O)[C@H]1NC(=O)CC2=CC(=CC=C2)[N+](=O)[O-]</chem>
24179365	TraR, LasR	0.46 4.67	LasR	6.28	<chem>C1COC(=O)[C@H]1NC(=O)COC2=CC=C(C=C2)OC(F)(F)F</chem>
10059709	LuxR, TraR	0.131, 3.2	LasR	0.13	<chem>CCCC(=O)CC(=O)N[C@H]1CCSC1=O</chem>
10935915	LuxR, TraR	0.45, 1.8	LasR	0.007	<chem>CCCCCCCCCCC(=O)CC(=O)N[C@H]1CCSC1=O</chem>
53484106	LuxR	0.35	TraR	20	<chem>CCCCCC(=O)CC(=O)N[C@H]1CCSC1=O</chem>
57401661	LuxP	159.7	LuxP	19.5	<chem>C[C@H]([C@H](C(=O)C(=O)C)O)O</chem>
25023205	LsrB	5.04	LuxP	1.01	<chem>CCCCC(=O)C(=O)[C@H](CO)O</chem>
57403389	LuxP	169	LuxP	0.65	<chem>C[C@@H]([C@H](C(=O)C(=O)C)O)O</chem>
25023206	LsrB	24.9	LuxP	1.52	<chem>CCCCCCC(=O)C(=O)[C@H](CO)O</chem>
25023204	LsrB	5.3	LuxP	0.75	<chem>CCCC(=O)C(=O)[C@H](CO)O</chem>
57392938	LuxP	57.54	LuxP	6.21	<chem>C[C@@H]([C@H](C(=O)C(=O)C)O)O</chem>
25023203	LsrB	50	LuxP	0.58	<chem>CCC(=O)C(=O)[C@H](CO)O</chem>
45267703	LsrB	5	LuxP	1.35	<chem>CCCCCC(=O)C(=O)[C@H](CO)O</chem>

Table 3. The potential broad-spectrum QSIs for the nine QS receptors

CID	SMILES	SmcR	LasR	TraR	CviR	QscR	PqsR	SdiA	LuxP	LsrB
328	<chem>C1=CC(=CC=C1C(C(=O)O)O)O</chem>	-6.1	-7.9	-6.5	-7.2	-6.8	-6.1	-7.1	-7.2	-6
7532	<chem>C1=CC=C(C=C1)CS(=O)(=O)O</chem>	-6.2	-7.9	-7.1	-7.2	-6.6	-6.3	-7.4	-7.3	-6.8
10313	<chem>C1=CC=C(C=C1)C(=O)NO</chem>	-6.5	-7.8	-7.3	-7.4	-7.5	-6.4	-7.4	-6.2	-7
11955	<chem>C1=CC=C(C=C1)C(=O)NN</chem>	-6.6	-7.7	-7.4	-7.5	-7.5	-6.4	-7.3	-6.1	-6.9
65723	<chem>C1=CC=C(C=C1)S(=O)(=O)NN</chem>	-6.1	-7.5	-7	-7.2	-7.5	-6.6	-7.3	-7.3	-6.7
69033	<chem>C1=CC=C(C=C1)S(=O)(=O)NO</chem>	-6.1	-7.5	-7	-7.3	-7.5	-6.6	-7.3	-7.5	-6.7
74128	<chem>C1=CC=C2C=C(C=CC2=C1)CO</chem>	-6.6	-7.9	-8.2	-8	-7.7	-7.1	-8.6	-6	-6.4
77864	<chem>CC(C)CC1=CC(=O)C=CC1=O</chem>	-7.2	-7.8	-7.3	-7.3	-7	-6.2	-7.8	-6.6	-6.2
78391	<chem>C1=CC=C(C=C1)C(O)S(=O)(=O)O</chem>	-6.4	-7.5	-6.7	-7	-7.1	-6.2	-7.4	-7.4	-6.3
85782	<chem>C1=CC(=C(C=C1C(C(=O)O)O)O)O</chem>	-6	-7.7	-6.5	-6.5	-6.9	-6	-7.3	-6.9	-6.4

Unit of the free binding energy, kCal/mol

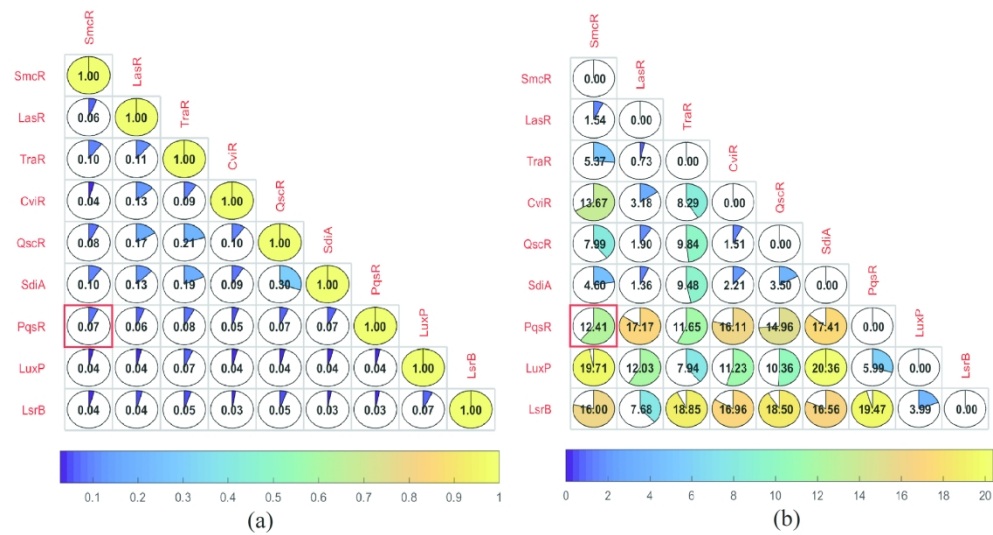


Fig S1. The analysis and discussion for QS receptors similarity of amino acids sequence and three-dimensional structure. (a) Similarity analysis of amino acids sequence of nine receptors; (b) Similarity analysis of the three-dimensional structure of nine receptors.

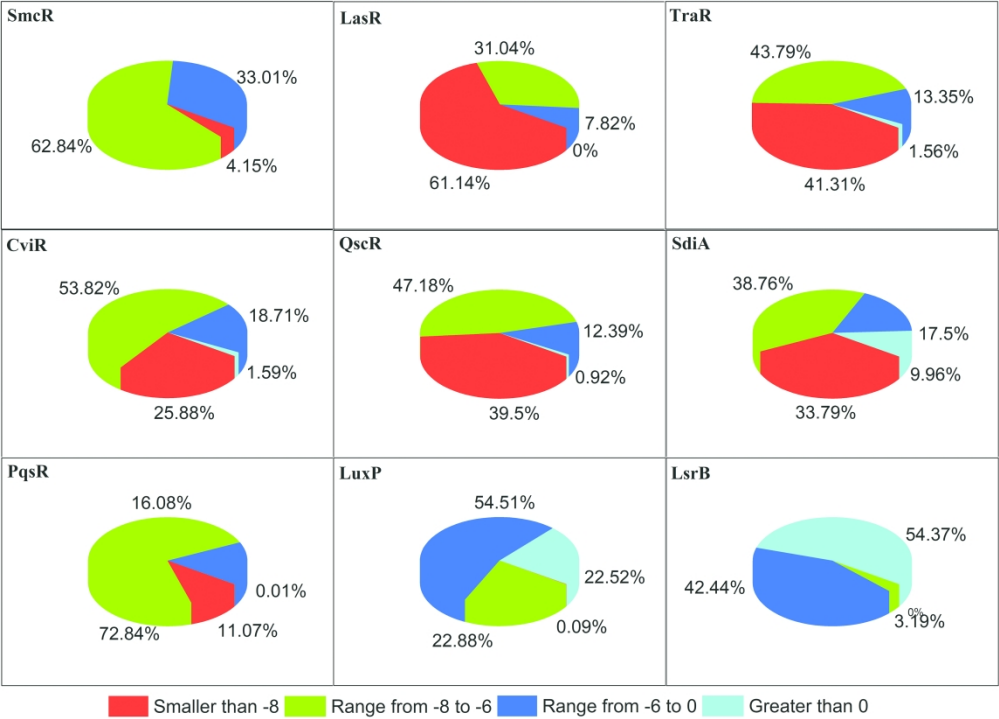


Fig. S2. Piecharts for FBEs distributions for the expanded ligands and nine QS receptors.

178x129mm (600 x 600 DPI)

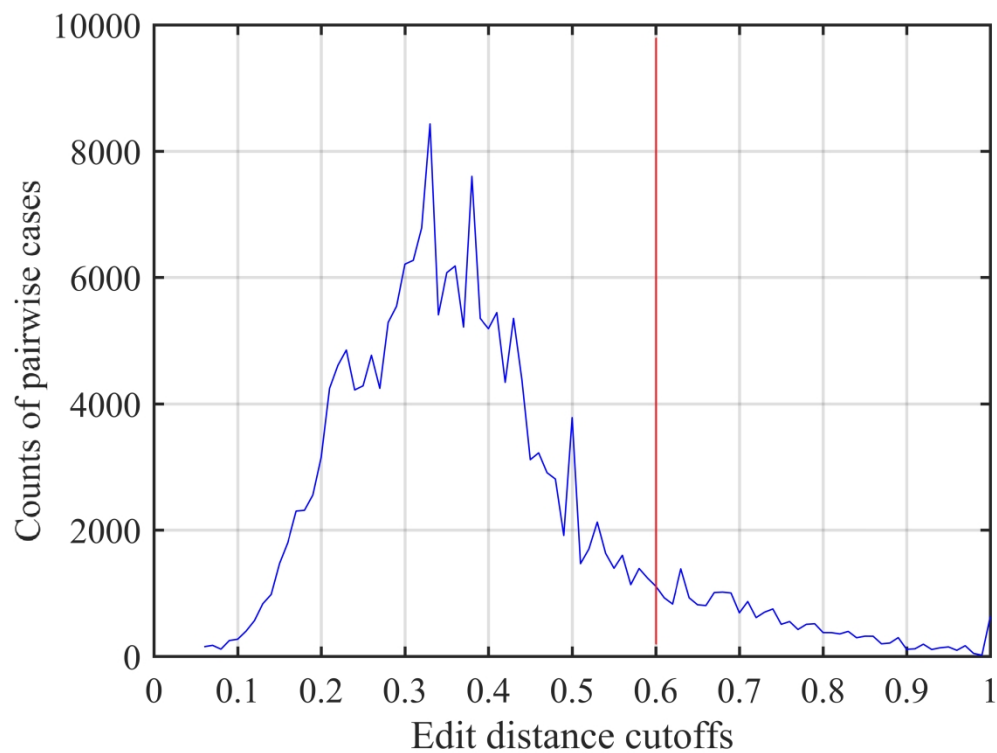


Fig. S3. The distribution of similarities for different distance cutoffs.

999x749mm (120 x 120 DPI)

Table S1. The total potential broad-spectrum QSIs screened on the nine QS receptors

PubChem CID	SMILES	SmcR	LasR	TraR	CviR	QscR	PqsR	SdiA	LuxP	LsrB
328	<chem>C1=CC(=CC=C1C(C(=O)O)O)O</chem>	-6.1	-7.9	-6.5	-7.2	-6.8	-6.1	-7.1	-7.2	-6
7532	<chem>C1=CC=C(C=C1)CS(=O)(=O)O</chem>	-6.2	-7.9	-7.1	-7.2	-6.6	-6.3	-7.4	-7.3	-6.8
10313	<chem>C1=CC=C(C=C1)C(=O)NO</chem>	-6.5	-7.8	-7.3	-7.4	-7.5	-6.4	-7.4	-6.2	-7
11955	<chem>C1=CC=C(C=C1)C(=O)NN</chem>	-6.6	-7.7	-7.4	-7.5	-7.5	-6.4	-7.3	-6.1	-6.9
65723	<chem>C1=CC=C(C=C1)S(=O)(=O)NN</chem>	-6.1	-7.5	-7	-7.2	-7.5	-6.6	-7.3	-7.3	-6.7
69033	<chem>C1=CC=C(C=C1)S(=O)(=O)NO</chem>	-6.1	-7.5	-7	-7.3	-7.5	-6.6	-7.3	-7.5	-6.7
74128	<chem>C1=CC=C2C=C(C=CC2=C1)CO</chem>	-6.6	-7.9	-8.2	-8	-7.7	-7.1	-8.6	-6	-6.4
77864	<chem>CC(C)CC1=CC(=O)C=CC1=O</chem>	-7.2	-7.8	-7.3	-7.3	-7	-6.2	-7.8	-6.6	-6.2
78391	<chem>C1=CC=C(C=C1)C(O)S(=O)(=O)O</chem>	-6.4	-7.5	-6.7	-7	-7.1	-6.2	-7.4	-7.4	-6.3
85782	<chem>C1=CC(=C(C=C1C(C(=O)O)O)O)O</chem>	-6	-7.7	-6.5	-6.5	-6.9	-6	-7.3	-6.9	-6.4
96574	<chem>CCOC(=O)CC1=CC=CC(=C1)C</chem>	-6.6	-7.3	-7.8	-7.6	-7.4	-6.1	-8.3	-7	-6
98068	<chem>C1=CC(=CC=C1C(C(=O)O)O)F</chem>	-6.4	-7.8	-7	-7.5	-7.4	-6.3	-7.7	-7.2	-6.3
137282	<chem>C1=CC=C2C=C(C=CC2=C1)CN</chem>	-6.6	-8	-8.3	-8	-7.7	-7.1	-8.6	-6	-6.2
138430	<chem>C1CC2C3=CC=CC=C3C1N2</chem>	-6.7	-7.3	-7.4	-7.7	-7.3	-6.5	-8	-6	-6.3
139951	<chem>C1C2C1C3CC2C4=CC=CC=C34</chem>	-7.2	-8.4	-8.1	-8	-7.7	-7.1	-7.9	-6	-7.1
141510	<chem>C=C1CC1C2=CC=CC=C2</chem>	-7.2	-7.3	-7.9	-7.6	-6.9	-6	-7.9	-6.2	-6.4
144215	<chem>C1CC2C1C3CC2C4=CC=CC=C34</chem>	-7.1	-9.1	-8.5	-8.2	-8.2	-7.3	-8.3	-6.1	-6.6
185912	<chem>C1=CC=C2C(C=CC2=C1)C(=O)O</chem>	-6.5	-7.6	-7.4	-7.8	-7.5	-6.6	-8.3	-6.5	-7
193409	<chem>C1=CC=C2C=C(C=CC2=C1)C(=N)N</chem>	-7.3	-8.9	-8.9	-8.6	-8.3	-7.7	-9.2	-6.9	-6
287162	<chem>C1CC2CC1C3C2C4=CC=CC=C34</chem>	-6.8	-9.1	-8.6	-8.3	-8.4	-7.2	-8.9	-6.3	-6.3
287421	<chem>CC1CC1CC2=CC=CC=C2</chem>	-7.2	-7.7	-8.1	-8.1	-7.3	-6.4	-7.8	-6.2	-6.3
302932	<chem>C1=CC(=C(C(=C1C(=O)O)F)F)F</chem>	-6.5	-7.9	-7.4	-7.4	-7.8	-6.3	-7.5	-6.4	-6.1
317708	<chem>C1CC1C2CCC3=CC=CC=C23</chem>	-6.4	-8.1	-8.2	-8.1	-7.4	-6.7	-8.4	-6.5	-6.9

1  
2  
3  
4  
5  
6  
7  
8  
9  
10  
11  
12  
13  
14  
15  
16  
17  
18  
19  
20  
21  
22  
23  
24  
25  
26  
27  
28  
29  
30  
31  
32  
33  
34  
35  
36  
37  
38  
39  
40  
41  
42  
43  
44  
45  
46

430144	<chem>C1=CC=C(C=C1)N(O)S(=O)(=O)O</chem>	-6.2	-7.7	-6.7	-7.3	-7.2	-6.4	-7.6	-7.5	-6.6
431079	<chem>C1=CC=C(C=C1)NS(=O)(=O)O</chem>	-6.1	-7.9	-7	-7	-6.9	-6.3	-7.1	-7.1	-6.1
439435	<chem>C1=CC(=C(C=C1CC(C(=O)O)O)O)O</chem>	-6.6	-8.1	-7.2	-6.7	-7.2	-6.5	-7.4	-6.8	-6.4
521908	<chem>C1=CC=C(C=C1)S(=NO)(=O)O</chem>	-6	-7.5	-7.2	-7.3	-7.5	-6.4	-7.3	-6.5	-6.6
562002	<chem>C1CC1CCC2=CC=CC=C2</chem>	-7	-7.8	-7.8	-7.7	-6.8	-6.1	-7.6	-6.2	-6
565120	<chem>C=C=CCOC1=CC=CC=C1</chem>	-6.5	-7.6	-7.4	-7.5	-7.1	-6	-7.4	-6	-6.1
588433	<chem>C1CC2CC1C=C2C3=CC=CC=C3</chem>	-8.3	-8.6	-9.1	-9.5	-8.2	-7.5	-8.9	-6.7	-6.5
602809	<chem>C1C=CC2CC1C=C2C3=CC=CC=C3</chem>	-8.2	-9.1	-9.1	-9.6	-8.4	-8	-9	-6.4	-6.2
1512675	<chem>COC(=O)CCC1=CC(=CC=C1)O</chem>	-6.5	-7.8	-7.6	-6.9	-7.3	-6.2	-7.6	-6.3	-6.5
4615423	<chem>CC1=CC=C(C=C1)OS(=O)(=O)O</chem>	-6.6	-7.4	-7.4	-7.3	-7.2	-6.7	-8	-7.4	-6
5070076	<chem>CCC1CC2=CC=CC=C2C1=O</chem>	-6.1	-8.4	-8.2	-8.2	-8.2	-6.7	-8.5	-6	-6.4
6421254	<chem>C1CC2CC3=CC=CC=C3C2C1</chem>	-6.3	-8.4	-8	-8.2	-7	-7.2	-8.3	-6.1	-7.3
7567094	<chem>C1=CC=C(C=C1)C(=O)C2C=CC=C2</chem>	-8	-9	-8.8	-9.1	-8.1	-7.1	-8.5	-6.8	-6.9
9815760	<chem>C1=CC(=C(C=C1C(C(C(=O)O)O)O)O)O</chem>	-6	-7.8	-7.2	-6.7	-7.3	-6.8	-7.4	-7.4	-6.7
10035212	<chem>CCCCC(C1=CC(=CC=C1)O)O</chem>	-6.8	-7.9	-7.7	-7.1	-7.4	-6.2	-7.7	-6.3	-6
10192593	<chem>CCCCC1=C(C(=CC=C1)O)O</chem>	-6.7	-7.5	-6.8	-6.8	-7.4	-6	-7.2	-6.3	-6.4
10855821	<chem>CC(=C)CC1=C(C(=CC=C1)O)O</chem>	-6.8	-7.3	-6.9	-6.9	-7.4	-6.2	-7.1	-6	-6.8
10931557	<chem>C1CC1=CCC2=CC=CC=C2</chem>	-7.3	-8.2	-8.3	-8.1	-7.5	-6.3	-8.2	-6.4	-6.4
10931558	<chem>C1C=CCC1C2=CC=CC=C2</chem>	-7.8	-8	-8.2	-8.3	-7.6	-6.7	-8.1	-6.5	-6.3
10931712	<chem>C=C1CC1CCC2=CC=CC=C2</chem>	-7.3	-8.1	-8.4	-8.2	-7.7	-6.6	-8.4	-6.7	-6.4
11051794	<chem>C1C2CC1C2C3=CC=CC=C3</chem>	-6.8	-7.5	-7.5	-7.8	-7.2	-6.2	-7.9	-6.2	-6
11084048	<chem>C1CC2=CC=CC=C2C1OO</chem>	-6.2	-7.7	-7.1	-7.3	-7.5	-6.5	-7.7	-6.6	-7
11332716	<chem>C1C=CCC2C1CC3=CC=CC=C23</chem>	-6.1	-9	-8	-7.6	-7.5	-7.3	-8.6	-6	-6.1
11469142	<chem>C=C1CCC2=CC=CC=C2O1</chem>	-6	-7.9	-8	-7.7	-7	-6.6	-8.2	-6	-6.3
11735538	<chem>CCOC(=O)CC1=CC(=CC=C1)F</chem>	-6.6	-7.7	-7.9	-7.4	-7.1	-6	-7.8	-6.7	-6.3
11816356	<chem>C1=CC=C(C=C1)S(=O)(=O)C(F)F</chem>	-6.5	-7.4	-7.1	-7.5	-7.3	-6	-7.7	-6.3	-6



12031685	C1C=CCC1CC2=CC=CC=C2	-7.9	-8.5	-8.5	-8.6	-7.8	-6.5	-8.3	-6.4	-6.6
12040849	COC(=O)CCC1=CC(=CC=C1)F	-6.6	-8	-8	-7.2	-7.2	-6.2	-7.9	-6.6	-6.6
12084350	C1=CC=C(C=C1)S(=O)(=O)N=N	-6.1	-7.6	-7.2	-7	-7.4	-6.6	-7.5	-7.3	-6.5
12249731	C1C=CCC1C2=C(C(=CC=C2)O)O	-7.5	-8.5	-7.9	-8.1	-8.3	-6.5	-8.2	-6.8	-6
12326507	C([C@@H](C(=O)O)NC(=O)C(F)(F)F)C(=O)O	-6.3	-6.6	-7	-6.3	-7.1	-6.5	-7	-6.9	-6.4
12383287	CC1CC2CC(=O)OC2C1	-6.2	-6.6	-6.3	-6	-6.8	-6.1	-7.2	-6.6	-7.1
12463222	C1=CC=C(C=C1)S(=O)(=O)ONN	-6.5	-7.4	-7.3	-6.8	-7.3	-6.7	-7.8	-7.3	-7.2
12506472	CC1=CC=C(C=C1)C(=O)CC(C)C	-7.4	-8.6	-8.2	-8.6	-8	-7	-8.5	-6.7	-6.2
12600454	CC1CCCC1C2=CC=CC=C2	-7.6	-7.8	-7.9	-8.1	-7.7	-6.7	-8.5	-6.6	-6.3
12662762	C1=CC=C(C=C1)C(=O)OC2C=CC=C2	-7.1	-9.5	-8.9	-9.3	-8.5	-6.8	-9	-6.8	-6.5
12662766	CC1=CC=C(C=C1)C(=O)OC2C=CC=C2	-6.4	-10.2	-9.2	-9.8	-9	-7.4	-9.5	-6.6	-6
12665209	CC1C2C1C3=CC=CC=C3C2	-7.4	-7.6	-8.1	-8.3	-7.7	-6.8	-8.7	-6.1	-6.1
12733519	CCCCC(C1C=CC(=O)O1)O	-6.7	-7.4	-7.1	-7.3	-7.1	-6.1	-7.2	-6.8	-6
12810281	C1C2CC3C1C2C4=CC=CC=C34	-7.5	-8.5	-8.2	-8.3	-7.8	-7.1	-8.1	-6.6	-6.4
12827258	C1=CC=C2C(=C1)C3=C(O2)C=CN3	-6.4	-8.5	-8	-8	-7.5	-6.7	-8.3	-6.1	-6.4
12836317	C1CC12CC2C3=CC=CC=C3	-7.4	-7.7	-7.9	-8	-7.4	-6.5	-8.1	-6	-6.1
12864396	CC1=CC2CC1C3=CC=CC=C23	-7	-8.4	-7.8	-7.5	-7.2	-7	-7.7	-6.1	-7.1
12939347	CC1=CCC2C1CC=CC2	-6.4	-7.1	-6.8	-7	-7.1	-6	-7.3	-6	-6.5
13092026	C1=CC=C(C(=C1)C2C(=O)NC(=O)O2)N	-6.7	-8.8	-7.7	-7.6	-7.4	-6.8	-8.6	-6.2	-7.6
13197606	C=CCCCOC1=CC=CC=C1O	-6.1	-7.8	-7.2	-7.4	-7.2	-6	-7.6	-6.5	-6.2
13253247	C1CC2C1CC3=CC=CC=C23	-6	-7.8	-7.6	-8	-7.1	-6.7	-8.2	-6.1	-7.3
13495438	C1C2CC2C1CC3=CC=CC=C3	-8	-8.3	-8.5	-8.4	-7.7	-6.7	-8.2	-6.6	-6.9
13660041	C1C2CC3C1C3C4=CC=CC=C24	-6.7	-8.4	-8.2	-8.5	-7.3	-7.1	-8.3	-6	-6.7
13807382	CC1=CC=CC=C1CCCC=O	-6.8	-7.1	-7	-6.9	-6.7	-6.3	-7.8	-6	-6
14087329	C1CC2=CC=CC=C2C1NO	-6.5	-7.3	-7.2	-7.7	-7.7	-6.6	-8	-6.4	-6.2
14195743	CC(=O)CC1=CC=CC=CC1=O	-6.7	-7.2	-7.3	-7.3	-7.6	-6.4	-8	-6.2	-7

14204944	CCCCC1=C(C(=CC=C1)O)O	-7	-7.8	-7.3	-7.1	-7.7	-6.3	-7.5	-6.6	-6
14384349	C1C=CCC1C2=CC=CC=C2O	-7.8	-8.7	-7.7	-8.2	-8	-6.4	-8.1	-6	-6.9
14490423	C1C2CC3C1C3C2C4=CC=CC=C4	-7.7	-8.6	-8.4	-8.2	-8.2	-7.1	-8.4	-6.1	-6.4
14527740	CC1CC=CC1C2=CC=CC=C2	-7.6	-8	-8.4	-8.9	-7.8	-7.2	-8.7	-6.5	-6.4
14948280	C=C1CC2CC1C3=CC=CC=C23	-7.1	-8.4	-7.8	-7.4	-7.2	-7	-7.7	-6	-6.8
14954003	CC1=CC=CCC1C2=CC=CC=C2	-6.1	-8.2	-7.7	-7.8	-8	-7.2	-8.8	-6	-6
15209342	CCOC(=O)CCC1=CC(=CC=C1)F	-6.9	-8.3	-8.4	-7.6	-7.4	-6.2	-8	-6.6	-6.1
15327731	CC1=CC=C(C=C1)C(=O)C2C=CC=C2	-8.6	-9.6	-9.2	-9.4	-8.8	-7.7	-9.4	-6.5	-6.2
15373214	C1=CC=C(C=C1)S(=O)(=O)NF	-6.1	-7.6	-7.1	-7.4	-7.4	-6.4	-7.3	-6.8	-6.1
15408803	C1=CC=C(C=C1)S(=NF)(=O)O	-6.1	-7.6	-7.4	-7.6	-7	-6	-7.4	-6.1	-6.5
15474146	C1C=CCC2C1C2=C3C=CC=C3	-7.6	-8.3	-8.2	-8.1	-7.4	-7.1	-8.5	-6	-7.2
15692112	CC(CC1=C(C(=CC=C1)O)O)NC	-6.3	-7.5	-6.8	-6.6	-7.5	-6.1	-7.4	-6.2	-6.3
15706998	C/C=C/CC1=CC(=O)C=CC1=O	-6.5	-7.9	-7.8	-7.4	-7.3	-6.2	-7.8	-7.1	-6.3
15759874	C=C=C1CC1C2=CC=CC=C2	-7.4	-7.8	-8.2	-7.7	-7.7	-6.6	-8	-6.4	-6.6
15824936	C1=CC=C(C=C1)C(=O)CC2C=CC=C2	-7.3	-9.3	-9.3	-9.6	-8.7	-7	-9.4	-6.4	-6.5
16685843	CCC1=CC=C2C1C=CC=C2	-7.3	-7.4	-7.3	-7.7	-7	-6.7	-7.8	-6	-6.2
16699325	C1=CC2C=C(C=C2C=C1)C(=O)O	-7.1	-8	-8.1	-8.2	-8.1	-7	-8.3	-6.1	-7.6
17969228	CCC(C1=C(C(=CC=C1)O)O)F	-6.6	-7.5	-6.9	-6.6	-7.5	-6	-7.2	-6.3	-6.6
17969677	CCC(C)C1=C(C(=CC=C1)O)O	-6.7	-7.5	-6.7	-6.8	-7.3	-6.1	-7.1	-6	-6.3
18366203	C=CC(=O)CC1=C(C=CC=C1F)F	-6.5	-7.4	-7.4	-7.6	-7.4	-6.1	-7.9	-6.3	-6.4
18972422	C1=CC=C2C(C(=CC2=C1)O)O	-6.4	-7.5	-7.4	-7.5	-7.7	-6.3	-7.8	-6.2	-6.7
19782555	C1=CC=C(C=C1)S(=O)(=O)OC(=O)O	-6.5	-7.5	-7.6	-7	-7.6	-6.3	-8	-6.9	-6
19783500	C1=CC=C2C(C=CC2=C1)S(=O)(=O)N	-6.4	-7.7	-7.4	-8	-7	-7.1	-8.1	-6.7	-6.1
19880772	C1=CC=C2C(=C1)C=CN(O2)O	-6.1	-8	-7.5	-7.5	-7.7	-6.4	-8	-6.2	-7.1
19886395	COC1=CC2=C(C=C2C=C1)C(=O)O	-6.7	-8	-7.4	-7.1	-7.6	-6.9	-7.7	-6	-6
19886408	C1=CC=C2C(=C1)C=C2CCO	-6.8	-7.7	-7.8	-7.6	-7.2	-6.2	-7.7	-6.1	-7

19910317	C/C=C/C1=C(C(=CC=C1)O)O	-6.5	-7.4	-6.9	-6.8	-7.4	-6	-7.3	-6.4	-6.5
20077952	C1=CC(=C(C=C1C(C(=O)O)ON)O)O	-6.3	-7.9	-6.5	-6.5	-7.1	-6.3	-7	-7.1	-6.1
20221896	CCCOC1=CC2=CC=CC=C21	-6.8	-8	-7.9	-8	-7.9	-6.2	-8	-6.2	-6.4
20279430	C1=CC(=CC2=C(C=C21)C(=O)O)O	-6.8	-7.7	-7.6	-7	-7.5	-6.7	-7.5	-6	-6.8
20361494	CCCCC1=C(C(=CC=C1)OO)OO	-6	-6.9	-7.3	-6.9	-7.6	-6.4	-7.6	-7.5	-6.7
20507381	C1=CC(=C(C=C1C(CC(=O)O)O)O)O	-6.1	-7.8	-7	-6.6	-7.3	-6.4	-7.4	-6.6	-7
21219659	C1=CC=C(C=C1)N(N)S(=O)(=O)O	-6.4	-7.5	-6.7	-7.1	-7	-6.2	-7.4	-7.2	-6
21225956	C=C(CC1=CC=CC=C1)C(=O)N	-7.2	-7.8	-8.1	-7.4	-7	-6.6	-8	-6.7	-6
21433320	C1=CC=C(C=C1)S(=O)(=O)N=C=N	-6.7	-7.5	-7.4	-7.2	-7.7	-6.6	-7.9	-6.9	-6.1
21452091	C1=CC=C(C=C1)C(F)S(=O)(=O)O	-6.5	-8.1	-7.1	-7.3	-7.3	-6.3	-7.7	-7	-6.9
21521486	CCCCC1=CC=CC(=O)C1=O	-6.1	-8	-7.2	-7.1	-7.3	-6.2	-7.7	-6.7	-6.3
21729606	C=C1CC1CC2=CC=CC=C2	-7.5	-8	-8	-7.9	-7.3	-6.5	-7.8	-6.2	-6.3
21896267	C1=CC=C(C=C1)C(=O)S(=O)(=O)O	-6.6	-8.5	-7.7	-7.9	-7.7	-6.8	-7.8	-7.4	-6
22044427	CCCCC1C2=CC=CC=C12	-7	-7.7	-7.9	-7.6	-7.2	-6.3	-7.7	-6.1	-6.1
22055108	CC(C)CC1=C(C(=CC=C1)O)O	-6.6	-7.6	-7	-6.9	-7.4	-6.2	-7.4	-6.1	-7.1
22171862	C1CC2=C1C3=CC=CC=C23	-7	-7.5	-8	-7.9	-7.5	-6.3	-8.6	-6.2	-8
22215574	C1CC2=C(C=C1C2)C(=O)C3=CC=CC=C3	-8.6	-8.9	-9.4	-9	-8.5	-7.8	-9.3	-7.3	-6
22364268	C1C=CC2=C1C=C(C=C2)C(=O)O	-6.9	-8.1	-7.9	-8.3	-7.7	-7	-8.3	-6.6	-6.4
22395239	C1=CC=C(C=C1)C(C=O)S(=O)(=O)O	-6.6	-6.7	-6.4	-6.8	-7.2	-6.3	-7.5	-7.5	-6.1
22401604	C/C(=C/CC(=C)C(=O)O)/C(=O)Cl	-6.6	-7.4	-6.8	-6.5	-6.6	-6	-7.1	-6.2	-6.4
22798161	C1=CC=C(C=C1)CC(=O)[C@H](CO)N	-6.9	-7.5	-7.6	-7.6	-7.2	-6.6	-7.7	-6.8	-6.4
22904532	C1=CC(=CC=C1C(C(=O)O)O)O)F	-7.3	-7.8	-7.6	-7.3	-7.6	-6.6	-7.8	-6.9	-6.6
23335488	C1CC2=CC=CC=C2C1NN	-6.6	-7.6	-7.1	-7.1	-7.6	-6.5	-7.8	-6.7	-6.6
23391769	C1C=CCC1CC2=CC=CC=C2O	-7.7	-9	-8.1	-8.3	-8.4	-6.5	-8.5	-6.9	-6.5
23392720	C1=CC=C(C=C1)S(=O)ONCl	-6.3	-7.4	-7	-7	-7	-6	-7.2	-6.6	-6.1
28297093	CC1=CC=CC=C1CCCON	-6.8	-6.9	-6.8	-6.7	-6.6	-6.3	-7.9	-6.1	-6.7

29005021	C1COC(=O)[C@H]1NCC(C(F)F)(F)F	-6.6	-7.8	-7.2	-7.3	-7	-6.2	-7.9	-6.7	-6.4
30048061	CC(C)CC(=O)N[C@@H](CO)C(=O)O	-6.3	-7.3	-6.3	-6.7	-6.3	-6	-7	-6.7	-6.1
30053902	C([C@@H](C(=O)O)NC(=O)CC(F)(F)F)O	-6.3	-7.3	-6.5	-6.7	-6.7	-6	-7	-6.7	-6.6
45099155	C1=CC=C2C(=C1)C=C3C2=CNO3	-6.2	-8.4	-8.5	-8.2	-8.5	-7.1	-8.8	-6.3	-6.5
53685382	C([C@@H](C(=O)O)NC(=O)C(F)(F)F)C(=O)N	-6.5	-6.9	-7.1	-6.4	-7.3	-6.2	-7.2	-7	-6.3
53707866	CC=CC1=CC2CCC1C2	-6.5	-6.9	-7.7	-7.1	-6.9	-6.2	-6.9	-6.2	-6.9
53739682	C1=CC(=C(C(=C1C(=O)F)F)F)F	-6.4	-8	-7.7	-7.6	-7.8	-6.3	-7.8	-6.2	-6.2
53817830	CC=CC1CCC2=CC=CC=C12	-6.8	-7.7	-8.1	-8.1	-7.6	-6.5	-8.3	-6	-6.1
53912394	CC=CCC1=CC2CCC1C2	-6.8	-7.2	-7.8	-7.2	-6.9	-6.3	-7.2	-6.1	-6.2
53925996	C=CC=CC1=CC2CCC1C2	-6.6	-7.1	-7.7	-7.6	-7.3	-6.6	-7.3	-6.4	-6.6
53983408	CCC=CC1=CC2CC1C=C2	-7	-7.3	-7.8	-7.3	-7.3	-6.4	-7.4	-6	-6.2
54103613	CC=CCC1=C(C=CC(=C1)O)O	-6.6	-7.5	-7.5	-7	-7	-6.1	-7.4	-7	-6.3
54110535	CCOC1CC2=CC=CC=C2C1	-6.4	-7.3	-7.9	-7.5	-7.6	-6.1	-8.1	-6	-6.4
54193313	C1CC2=CC=CC=C2C1N=O	-6.4	-7.1	-6.9	-7.5	-7.2	-6.1	-7.8	-6.1	-6.3
54197370	C1=CC=C(C=C1)S(=O)(=O)N(O)F	-6.5	-7.3	-7.1	-7.2	-7.5	-6.8	-7.5	-6.7	-6.2
54380049	CC=CC1=C(C(=CC=C1)O)O	-6.5	-7.4	-6.9	-6.8	-7.4	-6	-7.3	-6.4	-6.5
54541314	CC=CC1=C(C(=CC=C1)O)F	-6.3	-7.5	-7.3	-7	-7.4	-6.4	-7.5	-6.1	-6.2
54542571	CCC=CC1=CC2CCC1C2	-7	-7.1	-7.8	-7.2	-7.2	-6.1	-7.3	-6	-6
54776127	CC1=CC=CC=C1C2CC2CO	-7.1	-7.5	-7.2	-7.3	-7.1	-6.5	-7.9	-6	-7
55268119	C1=CC=C2C(=C1)C=C2CCN	-6.7	-7.7	-7.8	-7.6	-7.1	-6.3	-7.8	-6.2	-6.8
57098943	CCCC(CC1=C(C(=CC=C1)O)O)O	-6	-8	-7.2	-7.1	-8.1	-6.3	-7.7	-6.5	-6.2
57126213	C1=CC=C2C(=C1)C=CN2OO	-6	-7.7	-7.3	-7	-7.3	-6.3	-7.3	-6.1	-6.4
57219742	CCCC=C1CC2=CCC1C2	-7	-7.1	-7.5	-7.1	-7.1	-6.3	-7.4	-6.3	-6.1
57254362	C1=CC=C(C=C1)S(=O)(=O)NN=O	-6.3	-7.9	-7.3	-7.2	-7.7	-6.6	-7.8	-7	-6.3
57271792	CCC=C1CC2=CCC1C2	-6.7	-6.8	-7.2	-6.8	-6.9	-6	-7.1	-6	-6.8
57470122	C1=CC=C(C=C1)S(=O)(=O)N(F)F	-6.5	-7.7	-7	-7.3	-7.6	-6.4	-7.6	-6.7	-6.1

58083757	C1C=C1CCC2=CC=CC=C2	-7	-7.9	-7.9	-7.7	-7	-6.4	-7.8	-6.2	-6.1
59910019	C[C@@H](CC1=CC=C(C=C1)O)CC(=O)O	-7.3	-7.6	-7.4	-7.5	-7.7	-7	-7.5	-6.7	-6
59922559	C=CC1CC2=CC=CC=C2C1=O	-6.1	-8.3	-8.2	-8.2	-8.2	-6.7	-8.5	-6	-6.2
59930519	C=CC1CC2=CC=CC=C2C1	-6.8	-7.4	-8.1	-7.5	-7.5	-6.5	-7.9	-6	-6.4
61312351	COC(=O)CNC1=NC=C(C=C1)F	-6.1	-7.7	-7.2	-6.8	-7.2	-6	-7.2	-6.9	-6.1
61499672	CC(C(=O)O)NCC1=C(C=CC=C1F)F	-7.1	-8	-7.5	-7.7	-7.9	-7	-8.1	-6.2	-6.1
62366913	CC1=CC=CC=C1CCCN	-7.1	-7.1	-6.9	-6.7	-6.7	-6.5	-7.9	-6.2	-6.7
65805082	C1CNC1CCC2=CC=CC=C2	-7.1	-8.1	-8.1	-7.8	-7.5	-6.4	-7.8	-6.6	-6
65810874	C1=CC(=C(C=C1C(C(=O)O)O)F)F	-6.6	-7.7	-7.1	-7.5	-7.8	-6.7	-7.9	-6.3	-6.6
66701678	CC(=C(C1=CC=CC=C1)O)O	-6.8	-7.6	-7.4	-7.3	-7	-6	-7.1	-6.1	-6.1
67254952	CC1=CC(=C(C=C1)O)C[C@@H](C(=O)O)N	-6.1	-7.5	-7.2	-7	-7.2	-7.2	-8	-6.9	-7.3
67644608	CC=CC1=CC(=O)C=CC1=O	-6.8	-7.4	-7.2	-7.3	-7	-6	-7.4	-6.9	-6.1
67670465	C1=CC=C(C=C1)C2=CC2C(=O)O	-6.9	-8.4	-8.2	-7.8	-8	-7	-8	-6.7	-6.2
67814057	CC(=C(CC1=CC=CC=C1)O)O	-7.1	-8.2	-7.3	-7.5	-7.1	-6.3	-7.5	-6.2	-6.6
67862673	C1=CC=C(C=C1)C(=O)C2=C3C=C3C=C2	-8.2	-10	-9.3	-9.9	-9.1	-7.4	-9.5	-7.1	-6.1
68098372	C1=CC=C(C=C1)S(=O)(=O)CN=N	-6.4	-7.1	-7.2	-6.9	-7.3	-6.5	-7.7	-7	-6
68102430	C1=CC(=C(C=C1C(C(=O)O)O)O)F	-6.3	-7.5	-6.7	-7	-7.2	-6.4	-7.6	-7	-6.5
68530213	C1=CC(=C(C=C1CC(=O)O)O)F	-6	-8	-7	-7.1	-7.5	-6.3	-7.7	-7	-7
68892867	C1CC1CCC2=CC=CC=C2F	-7.3	-7.6	-8	-7.7	-7.3	-6.4	-7.9	-6.4	-6.5
68971788	C1=CC=C(C=C1)S(=O)(=O)N(O)O	-6.3	-7.4	-6.7	-6.8	-7.5	-6.6	-7.6	-7.1	-6.8
69106651	C1CC2CC1C3=C2C=C(C=C3)C(=O)N	-7.4	-9.3	-8.5	-8.7	-8.5	-7.6	-8.5	-6.9	-6.3
69115977	C1=CC=C(C=C1)C2C=CC(=O)N2	-7.5	-8.4	-8.1	-8	-7.5	-7	-8.5	-6.5	-6.7
69169673	CC=C(CC1=CC=CC=C1)C(=O)N	-7.9	-7.7	-8.7	-8	-7.9	-6.4	-8.4	-6.5	-6.4
69255777	CCCC(C)C1=C(C=CC=C1)O)O	-7	-7.8	-7.1	-6.9	-7.9	-6.6	-7.6	-6.1	-6.5
69404194	CC(=C([C@H]([C@@H]1[C@@H]([C@H]([C@H](O1)O)O)O)O)O)C	-6.7	-6.4	-6.7	-7.4	-6.8	-6.3	-6.9	-6.9	-6.4
69424111	CCCCC1CC(=O)NC1	-6.5	-6.7	-6.7	-6.5	-6.5	-6	-6.9	-6.2	-6

70063405	C1=CC=C2C3C(=CC2=C1)C=CN3	-6.6	-8.2	-8.1	-8.4	-8.5	-6.9	-8.8	-6	-6.1
70560901	CCCCC1=CC(=O)C=CC1=N	-7.1	-7.6	-7.5	-7.2	-6.9	-6.2	-7.5	-6.7	-6.5
70587063	C1=CC=C(C=C1)C(=O)C2=CC2O	-7.1	-8.5	-8	-8.1	-7.5	-6.3	-7.9	-6.6	-6.8
71756609	C1CC2=CC=CC=C2C3C1C3CN	-6.6	-7.8	-7.9	-7.8	-7.8	-7.7	-8.4	-6.3	-6
72222744	C1CC1CC2=C(C(=CC=C2)O)F	-6.4	-7.7	-7.5	-7.5	-7.4	-6.7	-7.6	-6.3	-6.1
73120763	CCC=CC(C1C=CC(=O)O1)O	-6.8	-7.1	-7	-7	-6.9	-6	-7.2	-7.1	-6.1
73746948	C1CC2=C1CC2C3=CC=CC=C3	-8.1	-8.5	-8.6	-8.6	-8.1	-6.9	-8.9	-6.8	-6
73910492	CC(C)CC(=O)N[C@@H](CC(=O)O)C(=O)O	-6.2	-7.2	-7.1	-6.4	-7	-6.6	-7.4	-7.5	-6
79367162	C1=CC(=C(C(=C1C(C(=O)O)F)F)F)F	-6.5	-8.1	-7.4	-7.9	-8	-6.8	-8.5	-6.9	-6.9
79722928	CCCCC(=O)C1=NC=C(C=C1)F	-6.4	-7.9	-7.4	-7.5	-7.5	-6	-7.5	-6.6	-6.4
80378637	CCC(=O)CC1=C(C=CC=C1F)F	-6.5	-7.3	-7.4	-7.5	-7.3	-6.1	-7.8	-6.2	-6.3
81009944	CC1=CC=CC=C1C2CCC2CN	-6.8	-7.6	-7.5	-6.9	-7.5	-7	-8.6	-6.1	-6.4
81644550	C1CC2=CC=CC=C2C3C1C3CO	-6	-7.9	-7.9	-7.7	-7.8	-7.6	-8.6	-6.3	-6.1
82396082	C1=CC=C(C(=C1)CC2C(=O)NC(=O)O2)N	-7.7	-8.4	-7.9	-7.8	-8.3	-7.2	-8.5	-6.1	-6.2
84769930	C1=CC(=C(C(=C1CC(=O)O)O)F)O	-6.1	-8.3	-6.4	-6.7	-7.4	-6.1	-7.3	-6.6	-6.1
84819575	C1CC1CCC2=CC=CC=C2O	-7	-7.8	-7.5	-7.5	-7.6	-6.3	-7.7	-6.7	-6.6
85085291	CC1C=CCC1C2=CC=CC=C2	-7.3	-7.9	-8.3	-8.8	-7.8	-6.7	-8.6	-6.8	-6.6
85540217	C=CC(=O)C1=CC=CC=CC1=O	-6.4	-7.5	-7.3	-7.5	-7.8	-6.3	-7.6	-6.1	-6.4
85549715	C=CC1CC2=CC=CC=C2C1O	-6.5	-7.4	-6.9	-7	-7.5	-6.3	-7.8	-6.1	-7
85967923	C1=CC=C(C=C1)[N+](=C[N+](=O)[O-])[O-]	-7.1	-8.6	-8	-8	-7.9	-6.4	-8	-6.8	-6.3
85991543	CC1=CC=C(C=C1)C(=O)C2CC2=C	-7	-9.1	-8.6	-8.8	-8.1	-7.3	-8.6	-6.5	-6
86023307	C1CC1CCC2=CC=CC=C2N	-6.4	-7.9	-7.6	-7.5	-7.4	-6.3	-7.9	-6.1	-6
86132247	C1CC2CCC1C2=C3C=CC=C3	-7.3	-8.5	-7.5	-7.8	-7.4	-6.6	-6.9	-6.7	-6.3
86177481	C=C1C=CCCC1C2=CC=CC=C2	-7.4	-8.6	-8.6	-8.5	-7.7	-7	-9	-6.3	-7
86306285	C([C@@H](C(C(F)(F)F)C(F)(F)F)O)C(=O)O	-6.5	-7.6	-7.4	-6.7	-7.4	-6.1	-7.4	-6.4	-7.7
86337526	C1=CC=C(C=C1)CC(=O)[C@H](CO)O	-6.8	-7.6	-7.6	-7.6	-7.3	-6.5	-7.6	-6.8	-6.4

86715115	<chem>C1CC1COC(=O)C2=NC=C(C=C2)O</chem>	-6	-8.1	-7.4	-7.5	-7.5	-6	-7.4	-6.6	-6
87842741	<chem>C1=CC=C2C(C=CC2=C1)C(=O)NO</chem>	-6.8	-8.3	-7.7	-8.1	-8.3	-7.5	-9.1	-6.7	-7.3
88445062	<chem>C1=C(C(=CC(=C1C(=O)O)O)O)CC=C=O</chem>	-6.7	-7.5	-6.9	-6.8	-7.3	-6.9	-7.5	-6.3	-6.3
88672539	<chem>C1=CC=C(C=C1)S(=O)(=O)ON=O</chem>	-6.3	-7.8	-7.4	-6.9	-7.4	-6.2	-7.7	-7	-6.7
89008217	<chem>CC1=CC=CC=C1CCC=C=O</chem>	-7.1	-7.4	-7.4	-7.2	-7.2	-6.7	-8	-6.2	-6
89028635	<chem>C=C(CCC1=CC(=CC=C1)O)O</chem>	-6.5	-7.5	-7.3	-6.9	-7.5	-6.1	-7.4	-6.2	-6.6
89088385	<chem>CC1=CC=CC=C1C2CC=NO2</chem>	-7	-7.9	-7.1	-7.5	-7.4	-6.3	-8.4	-6.3	-6.5
89402374	<chem>CC1=CC(=CC=C1)CNC(=O)CCO</chem>	-6.7	-8.1	-8.1	-7.6	-7.6	-6.8	-8.5	-6.7	-6
90255515	<chem>C1=CC(=O)OC2=CC(=C(C=C21)F)F</chem>	-6.9	-8.9	-8	-8.4	-7.9	-7.1	-8.9	-6.1	-6.3
90444348	<chem>C[C@@H](CCC1=CC=C(C=C1)O)C(=O)O</chem>	-6.8	-8.4	-7.9	-7.5	-7.3	-6.9	-7.8	-7	-6
90762075	<chem>C=CC1CC2=CC=CC=C2C1N</chem>	-6.3	-7.6	-6.7	-6.4	-7.3	-6.4	-7.6	-6	-6.6
91229800	<chem>C1CC2CC1CC2C3=CCC=C3</chem>	-7.9	-8.1	-8.5	-7.7	-7.8	-6.8	-8	-7	-6.2
91454738	<chem>CC1=CCC=CC1C2=CC=CC=C2</chem>	-6.3	-8.4	-8.2	-8.5	-8	-6.5	-9	-6	-6.1
92977874	<chem>CCCC[C@@H]1CC[C@H](N1)C(=O)N</chem>	-6.3	-7	-6.8	-6.8	-7	-6.2	-7.3	-6.9	-6
94253908	<chem>C[C@@H](CC1=CC(=CC=C1)O)CC(=O)O</chem>	-6.9	-7.7	-7.6	-6.9	-7.7	-7	-7.8	-7.2	-6
96125788	<chem>CC1=CC(=C(C=C1)O)C[C@@H](C(=O)O)O</chem>	-6.1	-7.3	-7.2	-7.1	-7.3	-7.1	-8	-6.8	-7.4
96771609	<chem>CCC1=CC(=C(C=C1)O)C[C@@H](C)C(=O)O</chem>	-6	-8.1	-7.8	-7.6	-7.7	-7.5	-8.4	-6.2	-6.4
96777923	<chem>CC[C@@H](CC(=O)O)C1=C(C=CC(=C1)F)O</chem>	-6.6	-7.6	-7.3	-7.3	-7.7	-6.8	-8.1	-6	-6.7
96930715	<chem>C([C@@H](C(=C(F)F)C(=O)O)F)C(F)(F)F</chem>	-6.6	-7.6	-7	-7.2	-7.2	-6.3	-7.8	-6.8	-7.1
101040094	<chem>C/C=C/C1=CC(=O)C=CC1=O</chem>	-6.8	-7.4	-7.2	-7.3	-7	-6	-7.4	-6.9	-6.1
101452601	<chem>CCC(=O)OC1=CC(=C(C=C1)O)F</chem>	-6.5	-7.8	-7.4	-7.4	-7	-6.1	-7.4	-6.9	-6
101889682	<chem>C=C1CCCC1C2=CC=CO2</chem>	-6.4	-7	-7	-7	-6.9	-6.2	-7.8	-6.2	-6.2
102208315	<chem>CC(C)C1=CC(C(=CC=C1)O)O</chem>	-6.1	-7.3	-7.2	-7.1	-7.3	-6.5	-7.8	-6.1	-6.1
102289708	<chem>C1C2CC3=CC=CC=C3C1C2=O</chem>	-6.9	-8.1	-7.7	-7.9	-7.6	-7.4	-8.1	-6.2	-6.1
102392185	<chem>C1=CC(C=C1)C(=O)OC2=CN=CC=C2</chem>	-7	-9.2	-7.7	-8.3	-7.6	-6.5	-8	-7.5	-6.3
103309153	<chem>C([C@@H](C(=O)O)NC(=O)C(C(F)(F)F)C(F)(F)F)O</chem>	-6.7	-8.4	-7.4	-7.6	-7.6	-7.1	-7.7	-6.5	-7



1												
2												
3												
4												
5	105432192	CC1=CC=CC=C1C2CCON2	-7.3	-7.9	-7.2	-7.5	-7.4	-6.6	-8.4	-6.4	-6.1	
6	105432198	CC1=CC=CC=C1C2CCNO2	-7.3	-7.3	-7.6	-7.5	-7.4	-6.4	-8.2	-6.2	-6.9	
7	105436620	CC1=CC=CC=C1C2CCC2CO	-7	-7.6	-7.2	-6.8	-7.4	-6.9	-8.6	-6	-6.5	
8	108052857	C([C@@H](C(=O)O)NC(=O)C(C(F)F)(F)F)O	-6.8	-7.6	-6.7	-6.7	-6.9	-6.3	-7.1	-6.4	-6.5	
9	108053147	C([C@@H](C(=O)O)NC(=O)C(C(F)F)(F)F)C(=O)O	-6.8	-7.5	-7.5	-7.1	-7.4	-6.8	-7.6	-7	-6.1	
10	108053300	C([C@@H](C(=O)O)O)NC(=O)C(C(F)F)(F)F	-6.4	-7.7	-7.3	-6.8	-7.1	-6.4	-7.3	-6.6	-6.5	
11	108093548	C([C@@H](C(=O)O)O)NC(=O)C(C(F)F)(F)F	-6.7	-8	-7.7	-7.3	-7.4	-6.6	-7.4	-6.7	-6.9	
12	114978538	C#CC(=O)CC1=C(C=CC=C1F)F	-6.5	-7.3	-7.6	-7.7	-7.5	-6.2	-7.7	-6.3	-6.7	
13	116605120	CCCCC(=O)C1=NC=C(C=C1)N	-6	-8	-7	-7.3	-7	-6	-7.1	-6.8	-6.1	
14	117287560	C1=C(C(=CC(=C1C(=O)O)O)O)CCO	-6	-7	-6.2	-6.1	-6.7	-6.4	-6.9	-6	-6.1	
15	117292284	C1=CC(=C(C(=C1C(=O)O)O)F)O)F	-6.1	-7.4	-6.6	-7.1	-7.4	-6.4	-7.9	-6.2	-6	
16	117299852	C1=C(C(=CC(=C1C(=O)O)O)O)CCC=O	-6.3	-7.1	-6.5	-6.4	-6.7	-6.7	-7.5	-6.3	-6.5	
17	117302638	C1=C(C(=CC(=C1C(=O)O)O)O)CCCO	-6	-7.2	-6.6	-6.6	-6.8	-6.6	-7.3	-6.2	-6.6	
18	117324768	C1=C(C(=CC(=C1C(=O)O)O)O)CCC(=O)O	-6	-8	-7	-6.9	-7.5	-7	-7.6	-6.5	-6.1	
19	118453489	C=C(C1=C2C=CC=C2C=C1)C(=O)N	-7.6	-8.6	-8.5	-8.1	-8.2	-7.4	-9.2	-6.4	-6.3	
20	121007170	C1C=CC1CCC2=CC=CC=C2	-7.4	-8.4	-8.4	-8.3	-7.6	-6.6	-8.4	-6.5	-6.3	
21	121220525	CCCC1C2=CC=CC=C2C1=O	-6.6	-8.2	-7.7	-8	-7.8	-6.5	-8.1	-6	-6.2	
22	121282030	CC1CCCC1C2=CC=CC=C2OO	-6	-8	-7.2	-7.6	-7.5	-6.7	-8.6	-6.1	-6.3	
23	121394326	CCCC(=O)C1=NC=C(C=C1)C	-6.5	-8.1	-7.3	-7.5	-7	-6	-7.4	-6.5	-6	
24	121454529	C1CC1C2CC3=CC=CC=C3O2	-7.1	-7.9	-7.8	-7.9	-7.8	-6.6	-8.4	-6.3	-6	
25	123264162	C1CC2C1C3=CC=CC=C3O2	-6	-7.6	-7.4	-7.7	-7.3	-6.1	-7.7	-6.1	-7.6	
26	126986684	C#CCCC(=O)N[C@@H]1CC[C@H]1O	-6.2	-7.4	-6.7	-6.9	-6.6	-6.1	-6.9	-6.1	-6.1	
27	129653236	CC(=C=C(C1=CC=CC=C1)O)O	-7.8	-8.3	-8	-7.9	-7.5	-6.5	-8.1	-6.7	-6.3	
28	129738993	C1=CC2=C3C(=C1)OC3=NC2=O	-6.2	-7.9	-6.3	-6.9	-7.2	-6.5	-7.5	-7	-7.3	
29	129911228	C1=C(OC(=C1)S(=O)(=O)N)[N+](=O)[O-]	-6.1	-7.5	-6.3	-6	-6.7	-6.1	-6.8	-6.9	-6.1	
30	130131264	CC(C)C1=CC=CC(C(=C1)O)O	-6.6	-7.1	-7.5	-7.5	-7.6	-6.7	-7.9	-6.1	-6	
31												
32												
33												
34												
35												
36												
37												
38												
39												
40												
41												
42												
43												
44												
45												
46												

130142273	C1C2CC3=CC=CC=C3C2C1O	-6.3	-7.8	-7.7	-7.6	-7	-6.5	-8.5	-6.5	-6.8
130142771	CC/C=C/C(C1C=CC(=O)O1)O	-6.8	-7.1	-7	-7	-6.9	-6	-7.2	-7.1	-6.1
130427856	C[C@@H](CC1=CC(=C(C=C1)O)O)OC	-6.2	-7.5	-7.2	-6.6	-7.3	-6.3	-7.2	-6.6	-6.3
130669129	CCC(C)C(=O)N[C@@H]1CC[C@H]1O	-6.5	-7.5	-7	-7	-6.8	-6	-7.1	-6.1	-6
130671082	C1=CC=C(C=C1)C(=O)[C@H](CO)O	-6.4	-7.8	-7.5	-7.5	-7	-6.4	-7.2	-6.4	-6.6
130717007	CC1CC1CC(=O)N[C@@H]2CC[C@H]2O	-6.5	-7.8	-7.6	-7.4	-7.2	-6.7	-8	-6.9	-6.6
130818495	CC1=CC(=C(C=C1)O)C(=O)C[C@@H](C)O	-6.9	-8.6	-7.9	-7.4	-7.2	-6.6	-8.2	-6.3	-6
130899233	C1CCC(C1)C(=O)N[C@@H]2CC[C@H]2O	-7	-8.1	-7.5	-7.8	-7.3	-6.4	-7.9	-6.6	-6.2
131071844	C1CC2=CC=CC=C2CCC1NN	-6	-7.8	-7.6	-7.9	-7	-7.5	-9.3	-6.2	-6.5
131093170	CCC[C@@H]1CC[C@H](N1)C(=O)O	-6	-6.8	-6.2	-6.2	-6.4	-6.2	-6.7	-6.8	-6.5
131178805	CCC(=O)N[C@@H]1CCCC[C@H]1F	-6	-7	-6.6	-6.4	-7.1	-6.3	-7	-6.1	-6
131385444	CC=CC1=C(C(=CC(=C1)O)O)O	-6.6	-7.3	-6.9	-6.8	-7.3	-6.3	-7.1	-7.1	-6
131408804	C1CC1CC2=C(C(=CC=C2)O)O	-6.9	-7.6	-7.1	-7.1	-7.6	-6.2	-7.4	-6.4	-6.8
131468705	CC(C(CC1=CC(=CC=C1)O)O)O	-6.5	-7.5	-7.5	-6.8	-7.6	-6.2	-7.6	-6.2	-6.4
132564184	CC(CCC1=C(C(=CC=C1)O)O)O	-6.2	-7.8	-7	-6.8	-7.8	-6	-7.3	-6.5	-6.2
134902841	C=CC1CC2CC1C3=CCC=C23	-6.9	-8	-7.7	-7.6	-7.7	-6.3	-7.5	-6.2	-6.3

Unit of the free binding energy, kCal/mol