Published in final edited form as:

AAPS J. 2017 July ; 19(4): 1235–1248. doi:10.1208/s12248-017-0093-5.

Integrated *In Silico* Fragment-Based Drug Design: Case Study with Allosteric Modulators on Metabotropic Glutamate Receptor 5

Yuemin Bian^{1,2,3}, Zhiwei Feng^{1,2,3}, Peng Yang^{1,2,3}, and Xiang-Qun Xie^{1,2,3,4,5,6}

¹Department of Pharmaceutical Sciences and Computational Chemical Genomics Screening Center, School of Pharmacy, University of Pittsburgh, Pittsburgh, Pennsylvania 15261, USA

²NIDA National Center of Excellence for Computational Drug Abuse Research, University of Pittsburgh, Pittsburgh, Pennsylvania 15261, USA

³Drug Discovery Institute, University of Pittsburgh, Pittsburgh, Pennsylvania 15261, USA

⁴Department of Computational Biology, School of Medicine, University of Pittsburgh, Pittsburgh, Pennsylvania 15261, USA

⁵Department of Structural Biology, School of Medicine, University of Pittsburgh, Pittsburgh, Pennsylvania 15261, USA

Abstract

GPCR allosteric modulators target at the allosteric binding pockets of G protein-coupled receptors (GPCRs) with indirect influence on the effects of an orthosteric ligand. Such modulators exhibit significant advantages compared to the corresponding orthosteric ligands, including better chemical tractability or physicochemical properties, improved selectivity, and reduced risk of oversensitization towards their receptors. Metabotropic glutamate receptor 5 (mGlu₅), a member of class C GPCRs, is a promising therapeutic target for treating many central nervous system diseases. The crystal structure of mGlu₅ in the complex with the negative allosteric modulator mavoglurant was recently reported, providing a fundamental model for designing new allosteric modulators. Computational fragment-based drug discovery represents a powerful scaffold-hopping and lead structure-optimization tool for drug design. In the present work, a set of integrated computational methodologies was first used, such as fragment library generation and retrosynthetic combinatorial analysis procedure (RECAP) for novel compound generation. Then, the compounds generated were assessed by benchmark dataset verification, docking studies, and QSAR model simulation. Subsequently, structurally diverse compounds, with reported or unreported scaffolds, can be observed from top 20 *in silico* synthesized compounds, which were predicted to be potential mGlu₅ modulators. *In silico* compounds with reported scaffolds may fill SAR holes in known, patented series of mGlu₅ modulators. And the generation of compounds

⁶To whom correspondence should be addressed. (xix15@pitt.edu).
Yuemin Bian and Zhiwei Feng contributed equally to this work.

Electronic supplementary material The online version of this article (doi:10.1208/s12248-017-0093-5) contains supplementary material, which is available to authorized users.

COMPLIANCE WITH ETHICAL STANDARDS

Conflict of Interest The authors declare that they have no competing interests.

without reported tests on mGluR indicates that our approach is doable for exploring and designing novel compounds. Our case study of designing allosteric modulators on mGlu₅ demonstrated that the established computational fragment-based approach is a useful methodology for facilitating new compound design in the future.

Keywords

allosteric modulator; computational fragment-based drug discovery; GPCRs; metabotropic glutamate receptor 5

INTRODUCTION

The human G protein-coupled receptors (GPCRs), which can also be described as seven-transmembrane (TM) receptors account for more than 1% of human genome. The GPCR superfamily, comprised more than 800 receptors, can be further categorized into four different classes, classes A, B, C, and F (Frizzled), according to their sequence homology (1,2). As essential receptors associated with a variety of physiological processes, including neurotransmission, immune defense, and cell growth, over 30% of currently marketed drugs are using GPCRs as their targets (3,4).

Each GPCR possesses an orthosteric binding pocket for its respective endogenous ligands. Compounds derived from natural sources or chemical synthesis binding to this pocket are termed orthosteric ligands (5). Early drug development was focused on the orthosteric modulation and almost all of the FDA-approved compounds for therapeutic use target at the orthosteric binding sites of the receptors (5,6). However, the ligands coming out from this strategy have drawbacks including limited or poor selectivity, a lack of efficacy, and resistance or decreased efficacy upon chronic administration (7–9). GPCRs can have allosteric binding sites, which have topological and functional distinctions from corresponding orthosteric binding sites. The existence of allosteric binding pockets allows additional interactions between ligands and receptors (5,10,11). Across a receptor family, the allosteric pockets usually stand with less evolutionary pressure for conservation, and the corresponding allosteric ligands usually have better selectivity. Meanwhile, allosteric ligands have a ceiling effect, which means their effects are saturable (12,13). Based on the modes of pharmacology, allosteric ligands can be categorized into positive allosteric modulators (PAMs), which enhance agonist-mediated receptor response, negative allosteric modulators (NAMs), which noncompetitively attenuate orthosteric activities, and silent allosteric modulators (SAMs), which have no effect on responses triggered by orthosteric ligands but block the effects caused by PAMs and NAMs (5). The clinical success of the benzodiazepines attracts the attention of using allosteric regulation as a promising therapeutic strategy (14).

Metabotropic glutamate receptor 5 (mGlu₅) is a member of class C GPCRs, which mainly responds to glutamate, one of the major neurotransmitters (15). Class C GPCRs include metabotropic glutamate receptors, taste receptors, calcium-sensing receptors, GABA_B, and others. mGlu₅ is a promising drug target for the treatment of diseases ranging from fragile X syndrome to depression and movement disorder (16). mGlu₅-negative allosteric modulators,

which could attenuate orthosteric ligands mediated mGlu₅ activation, are under clinical evaluation for the treatment of multiple diseases (17,18). Dr. Fiona H. Marshall and her group have reported the crystal structure of the transmembrane domain of human mGlu₅ receptor, in the complex with mavoglurant, the negative allosteric modulator. The structure provides the precise location of the allosteric binding site within the transmembrane domain (16). Multiple research groups have focused on the allosteric regulation of mGlu₅, for example, Andreas Ritzén's group and P. Jefferey Conn's group designed and synthesized a series of allosteric ligands for mGlu₅ (19). Although a classical structure-based medicinal chemistry approach could be useful for the generation of corresponding allosteric modulators, state-of-the-art computational methods could also be powerful tools in generating new ideas in allosteric modulator design in this field (20,21).

Fragment-based approaches for designing and generating lead compounds have proven quite fruitful in drug discovery. There are several ways that fragment-based approach could be applied to drug discovery: (1) binding sites and pharmacophores identification for receptor binding could be achieved with fragment screening techniques; (2) HTS libraries could be biased and the optimization of lead compounds could be guided with fragments screening hits; (3) leads with the potential to be optimized into drug-like compounds can be identified with fragment screening (21–25). Fragment binding can reveal hot spots on proteins, which could result in high-affinity receptor-ligand interactions. The fragment-based approach is especially suitable for detecting the interaction spots among the protein-ligand binding, as fragments are capable of forming relatively strong interaction with a certain region of the target protein. Once the identification of fragments inside the binding pocket is done, these fragments could be grown, linked, or merged to develop the potential ligands (21). A number of fragment-based screening campaigns have successfully delivered clinical candidates (26). Steven Howard and his group successfully applied fragment-based screening and fragments linking to discover novel thrombin inhibitors (27). Philip D. Edwards and his group reported the novel cyclic amidine β -secretase with the fragment-based approach to generate lead compounds (28). However, there are still several undesired drawbacks for this approach. The defects of the experimental fragment-based method include time consuming, costly, and low affinity of hits (21,29,30). Another complementary way to approach fragment-based drug discovery is the application of computational methods, which can offer several advantages: (1) computational methodologies can construct high-quality and diversified fragment libraries with both time and cost efficiencies(31,32); (2) computational approaches can easily explore larger fragment databases; (3) optimization strategies for improving the drug-likeness of the hits with computational tools have efficiency and flexibility (33). Recently, a drug design effort, which involved virtual screening of a fragments library, binding confirmation by NMR, X-ray crystallography, and structure-based optimization of fragments, by a research group in AstraZeneca, led to the discovery of highly potent FXIa inhibitors (34). Another example is the development of inhibitors of cyclophilin A. Li's group reported that amide fragment, which functions as the key linker, is one of the critical pharmacophores for CypA inhibitors (35). Through computational FBDD approaches, acylurea was designed by fusing amide and urea to function as a new linker and contributed to the discovery of novel inhibitors.

In our *in silico* research, we used a computational fragment-based approach to propose structures as potential novel allosteric modulators of mGlu₅. We generated a fragment library according to the reported GPCRs' allosteric modulators on Allosteric Database (ASD v2.0) (36). Retrosynthetic combinatorial analysis procedure (RECAP) analysis and synthesis were used to generate the novel compounds. Molecular docking was applied to screen the hits for mGlu₅ by docking the *in silico* synthesis compounds back to the pocket with predicted binding energy and docking scores. Multiple computational methodologies, including benchmark dataset verification, docking studies, QuaSAR model, etc. were utilized for further validation of our hits. Twenty *in silico* synthesized compounds are predicted to be potential mGlu₅ allosteric modulators with preferable binding energies and docking scores. Structure diversity among the *in silico* design could be observed. Series of compounds with reported allosteric activities on mGluR could be recurred. Our case study on designing allosteric modulators on mGlu₅ suggested that this computational fragment-based approach is a useful methodology for facilitating the future compound design processes.

MATERIALS AND METHODS

X-ray Structures and GPCR Allosteric Modulators

Two crystal structures of the transmembrane domain of the human mGlu₅ were used in this work. The first model (PDB entry: 4OO9; resolution, 2.6 Å; method, X-ray diffraction) (16) was in complex with the negative allosteric modulator, mavoglurant. The second model (PDB entry: 5CGD; resolution, 2.6 Å; method, X-ray diffraction) (37) was in complex with the negative allosteric modulator, HTL14242. The structures of mGlu₅ were downloaded from the Protein Data Bank (PDB). Then, the SYBYL-X 1.3 (38) was used for the preparation of the crystal structure, including energy minimization and residue repair (39).

Ligands were downloaded from the Allosteric Database (ASD v2.0) (36). SYBYL-X 1.3 and PyMol (<http://www.pymol.org>) were used for molecular visualization, structural superimposition, and data analysis (40).

Fragment Generation and *In Silico* Synthesis

RECAP Analysis and RECAP Synthesis in ChemAxon's Fragmenter (<https://www.chemaxon.com>) were applied to generate our fragment library. The purpose of RECAP analysis is to fragment compounds according to simple retrosynthetic analysis rules and gather statistics towards fragment products. The RECAP synthesis could be applied to combine fragment products from RECAP analysis randomly, in order to produce the novel chemical structures, which should be synthetically reasonable (41,42). RECAP analysis was used for the establishment of the fragment library from allosteric modulators. And RECAP synthesis was used for the generation of *in silico* synthesized novel structures from processed, analyzed, and categorized fragments. The molecule filter was set to be leadlike (43). The heavy atoms were set to be 25 in average with 11 in standard deviation.

Molecular Docking for the Studies of mGlu₅-Ligand Interactions

We performed the molecular docking between crystal structure of mGlu₅ and fragments in the fragment library using SYBYL-X 1.3. Surflex-Dock GeomX, the docking algorithm that implemented in SYBYL, was applied to predict detailed receptor-ligand interaction. The total score was expressed as $-\log_{10}(K_d)$ (44). In the docking simulations, the allosteric binding site was first defined to cover all residues within 4 Å of the NAM in the initial mGlu₅-mavoglurant complex. The Kollman all-atom approach was used to calculate atomic charges for the protein (45) and the Gasteiger-Hückel approach for the ligand (46). The movement of hydrogen atoms of the protein was allowed. Additional starting conformations per molecule were set to 10, and the angstroms to expand search grid was set to 6. Three independent runs were performed for our fragment library.

Molecular docking between the *in silico* synthesized compounds and mGlu₅ was done with SYBYL-X 1.3 and AutoDock 4.0 (47). The parameters stated above of SYBYL-X 1.3 were used in these docking experiments. For AutoDock, Lamarckian genetic algorithm (LGA), in which the individual's genotype was reverse transcribed based on the adaptation of individual's phenotype according to environmental factors (48), was used in this study. The grid box contained the entire allosteric binding site of the mGlu₅ and permits translation and rotation of ligands. Numbers of points in grid box for three dimensions were 24, 24, and 22; spacing (angstrom) was 1.000. AutoGrid has been used for calculating the energy map of each atom in the ligands. The receptor was set to be rigid. Genetic algorithm with default parameters was chosen for the search. A binding energy, which is constituted by intermolecular energy, internal energy, torsional energy, and unbound extended energy, was reported for each run.

Generation of Benchmarking Dataset

A key metric for docking studies of molecules remains the enrichment of top hits (49). Once a docking screening could distinguish active compounds as top hits against a large number of decoys in the database, this docking screening is considered a success (50). In order to distinguish the *in silico* synthesized compounds from decoys in the docking processes, a benchmarking dataset is introduced. In the spirit of the DUD (49) and DUD-E (51) reference data sets for validation studies, 50 decoys, which have similar physical properties but distinctive topological properties with the reported mGlu₅ NAM, were generated through <http://decoys.docking.org>, using the reported protocols and parameters.

Quantitative Structure-Activity Relationship Model

The correlation between the structures of molecules and their corresponding chemical or physical properties could be summarized and predicted with quantitative structure-activity relationships (QSAR) techniques (52). A QSAR model could be used to evaluate specific parameters that affect some properties of the molecules or to estimate same properties for other molecules in the same series (53). A QSAR model, dealing with the correlation between the Log*K_i* value and theoretical descriptors for molecules, was developed for a set of 1,2-diphenylethyne mGlu₅ NAMs. Sixty-six analogs of 1,2-diphenylethyne with existing *K_i* values were acquired from ASD. *K_i* values underwent logarithmic transformation to get Log*K_i* values. Kennard-Stones algorithm was used for determining the training and test sets.

One hundred eighty-six descriptors were added towards each compound. Four criteria, contingency coefficient, Cramer's V, entropic uncertainty, and linear correlation, were used to evaluate the contribution of each descriptor after the contingency analysis (54–56). Partial least squares (PLS) regression method was used for building the model with selected descriptors. The model has been verified with the test set and used for the prediction of *in silico* synthesized compounds.

RESULTS

The Establishment of Fragment Library

Twenty-seven thousand two hundred sixty-two GPCR allosteric modulators were downloaded from ASD. RECAP analysis in ChemAxon's Fragmenter (<https://www.chemaxon.com>) was used to generate our fragment library. "Rule of three" (57,58) rather than "rule of five" (59) was used here for guiding the selection of ideal fragments according to the recent reports. The contents for "rule of three" included molecular weight <300, hydrogen bond donors ≤ 3, hydrogen bond acceptors ≤ 3, and cLogP ≤ 3. After deleting duplicated items, a library with 863 fragments was generated. Among them, 47 fragments had their molecular weights over 300, 3 fragments had over 3 hydrogen bond donors, 56 fragments had over 3 hydrogen bond acceptors, and 72 fragments had their LogP(o/w) values over 3. Table I demonstrates top ten most frequently appeared fragments with corresponding properties as an example. A comparison between the fragment library we generated and the Maybridge fragment library was conducted. MACCS structural keys were used for calculating the fingerprint of each fragment. Tanimoto coefficient was set to be 80%. Seventy-eight percent of fragments in the library we generated could be matched with items in Maybridge library, while the remaining 22% of fragments may contribute to the particularity of GPCR allosteric modulators.

Docking Studies of Fragments into mGlu₅

Before the docking studies of the fragments library we generated, we first docked the reported NAM, mavoglurant, back to the defined allosteric binding site to function as a control for the validation of our docking process. Residues from TM2, TM3, TM5, TM6, and TM7 formed the pocket. Gly624, Ile625, Gly628, Pro655, Ser805, Val806, and Ser809 were directly involved in the pocket (Fig. 1a). Comparing our docking results with the crystal structure of human mGlu₅ (PDB entry: 4OO9; resolution, 2.6 Å; method, X-ray diffraction) (16) in complex with the negative allosteric modulator, mavoglurant bound well inside the pocket with the alkyne linker traversing a narrow channel between Tyr659, Ser809, Val806, and Pro655. Three hydrogen bonds could be formed between mavoglurant and Asn747, Ser805, and Ser809 (Fig. 1b). Our docking studies between mavoglurant and the defined allosteric binding site provided congruent results with the crystal structure. Identical residues were involved in the formation of hydrogen bond interaction (Fig. 1c). The docking result overlapped well with the crystallized complex, with the RMSD of ~0.3 Å (Fig. 1d), indicating that our docking process is relatively reliable.

Then, the generated fragments in our fragment library were docked into the mGlu₅ allosteric binding site. We conducted three independent docking studies for all fragments and found

that fragments invariably aggregated into two distinct regions of the binding pocket of mGlu₅, an upper region and a bottom region (Fig. 2a, b). The number of fragments for each region is summarized in Fig. 2c. Our docking results were very consistent across our three independent runs, with ~416 (85%) of fragments in the upper region and ~320 (87%) of fragments in the bottom region were repeated in the same region throughout the three docking runs.

***In Silico* Synthesis of Novel Compounds from Categorized Fragments**

A single fragment may bind with low affinity to a protein because of the lower molecular weight, which results in a limited amount of ligand-receptor interaction (60). The combination of the fragments from upper and bottom regions is a promising strategy for increasing the affinity of binding. RECAP synthesis in ChemAxon Fragmenter was used to combine the fragments in upper region with the fragments in the bottom region. To find out whether the docking scores can be used as guidance for RECAP Synthesis, a trial was first conducted. A small scale of RECAP synthesis was performed with two sets of fragments: (1) ~600 compounds were generated by 30 fragments in upper region and 20 in bottom region with the highest docking scores; (2) ~600 compounds were generated by 30 fragments in upper region and 20 in bottom region with the lowest docking scores, respectively. The distribution of their docking scores (Fig. 3a) revealed that the combination of fragments with higher docking scores tended to result in compounds with better binding affinities. The Mann-Whitney test showed that asymptotic significance (two-tailed) is less than 0.001, which meant that these two distributions were significantly different. Then, 9600 novel compounds were generated through the large-scale RECAP synthesis with 120 highest scored fragments in the upper region (Supplementary Table 1) and 80 highest scored fragments in the bottom region (Supplementary Table 2). Our results showed that 124 *in silico* synthesized compounds had the docking score higher than seven (Supplementary Table 3). Since the docking score was expressed as $-\log_{10}(Kd)$, as described in “Materials and Methods”, compound with a docking score higher than 7 means that its corresponding predicted *Kd* is less than 10^{-7} Mol, which is a positive sign for binding affinity. Interestingly, six *in silico* synthesized compounds have higher docking scores than mavoglurant. The structures of top 20 *in silico* synthesized compounds, as well as the mavoglurant, are shown in Table II.

Moreover, we introduced binding energy of molecular docking as a supplementary evidence for receptor-ligand interaction by using AutoDock 4.0. The results of top 20 compounds can be found in Table II. A correlation plot was drafted (Fig. 4) to show the relationship between the docking score and the binding energy for each compound. All of these 20 *in silico* synthesized compounds had a predicted negative binding energy, which had a high correlation with the docking scores (61,62), providing the evidence that the docking studies with SYBYL-X 1.3 and AutoDock 4.0 were consistent with each other.

The binding poses of the highest ranked *in silico* synthesized compound was compared with the reported NAM, mavoglurant. The *in silico* synthesized compound occupied the identical position as mavoglurant (Fig. 1b), forming the hydrogen bond with Ser805 (Fig. 3b). Two compounds overlapped very well with the RMSD of 1 Å, as shown in Fig. 3c. Interestingly,

our docking studies suggest that compound 1, our highest scoring compound proposed by our *in silico* methods, is predicted to form hydrogen bonds with Thr735 and Tyr659, which are not reported for mavoglurant. These two additional receptor-ligand interactions might explain the higher docking score acquired by the highest ranked *in silico* synthesized compound than mavoglurant. Whether these would form if compound 1 were synthesized and crystallization attempted would be an interesting experiment to conduct.

Evaluation of Docking Process and *In Silico* Synthesized Compounds with a Benchmarking Dataset

Dr. John. J. Irwin and his group generated the Directory of Useful Decoys, Enhanced (DUD-E), which is available at <http://dude.docking.org>. The physical properties used are molecular weight, calculated LogP, H-bond donors and acceptors, the number of rotatable bonds and net molecular charge in order to build physically similar decoys (51). A tool (<http://decoys.docking.org>), which is automated and available online, could be used for generating the matched decoys for user-supplied ligands. The reported mGlu₅ NAM, mavoglurant, was used as the ligand for decoys generation. seventy-one compounds, including 50 decoys (Supplementary Table 4), mavoglurant, as well as top 20 *in silico* synthesized compounds were docked into mGlu₅ (Fig. 5). Our results showed a clear enrichment of NAM and *in silico* synthesized compounds against 50 decoys. These results indicated that 20 *in silico* synthesized compounds could be enriched with NAM and were different from challenging decoys.

Prediction of Competition Binding with a Quantitative Structure-Activity Relationship Model

The selectivity of a ligand towards receptor subtypes could be measured with the competitive binding assay, which also reveals the percentage and density of each subtype among a certain tissue (63). Competition binding is widely used in testing potential allosteric modulators for GPCRs. Sixty-six analogs of 1,2-diphenylethyne with existing *K_i* values were selected from ASD (Supplementary Table 5). Following the method described in the “Materials and Methods,” a QSAR model with 12 descriptors was built:

$$\begin{aligned} \text{Log}K_i = & -6.81 + 3.11 * \text{BCUT_PEOE_3} - 2.46 * \text{BCUT_SLOGP_3} \\ & + 0.46 * \text{b_1rotN} - 12.76 * \text{b_1rotR} - 9.57 * \text{GCUT_PEOE_0} \\ & - 4.37 * \text{GCUT_SLOGP_0} + 4.49 * \text{GCUT_SMR_0} - 0.17 * \text{Kier3} \\ & + 0.06 * \text{PEOE_VSA4} + 5 - 0.03 * \text{SlogP_VSA4} - 0.02 * \text{SMR_VSA0} \\ & - 0.02 * \text{SMR_VSA1} \end{aligned}$$

Table III shows the calculated descriptors for each molecule, the descriptors coefficients, contingency coefficient, Cramer’s V, entropic uncertainty, and linear correlation. A correlation plot between the predicted Log*K_i* values and experimental Log*K_i* values was shown in Fig. 6. With a correlation coefficient of 0.82 and a cross-validated correlation coefficient of 0.74, the established QSAR model was acceptable (64). This model was used to predict the *K_i* values for 14 compounds among the top 20 *in silico* synthesized compounds, as listed in Table II. Compounds 10, 13–16, and 18 were excluded from the

prediction with this QSAR model, because they have distinctive structural features towards 1,2-diphenylethyne analogs. The predicted values for 14 compounds are ranging from 1.3 to 1162.3 nM. These data are congruent with their docking scores and binding energies, indicating that these compounds are potential mGlu₅ allosteric modulators.

Validation with PAINS-Remover and Toxtree

PAINS-Remover is designed and constructed to remove the Pan Assay Interference Compounds (PAINS) from screening libraries (65). The top 20 *in silico* synthesized compounds listed in Table II were tested by “Pan Assay Interference Compounds” (PAINS, www.cblligands.org/PAINS/), and all compounds passed the filter. Toxtree (<http://toxtree.sourceforge.net/index.html>) is a software to estimate the toxic hazard through decision tree approach (66). All these top 20 *in silico* synthesized compounds were tested with Toxtree, and all of them had the same level of estimated toxic hazard with the reported NAM, mavoglurant.

The Effects on Ligand and Fragment Binding Caused by the Rotation of Trp785

The Trp785 on TM6 is highly conserved, which is similar to the central toggle switch in class A GPCRs. The different conformations of Trp785 could be observed between different crystal models due to the cocrystallization with different chemotypes. John A. Christopher and his group cocrystallized the mGlu₅ with HTL14242 and another molecule in the series in 2015. Their crystal model of mGlu₅ (PDB entry: 5CGD) has a different Trp785 conformation (Fig. 7a, b). The evaluation of ligand and fragment binding between these two models were performed to illustrate the effects caused by the rotation of a critical residue Trp785. Trp785 rotates out of the allosteric binding pocket on 4OO9, which results in a relatively large pocket, especially for the upper region (Fig. 7a). While for 5CGD, the Trp785 rotates into the allosteric binding pocket, which narrows down the space inside. One hydrogen bond could be formed between Trp785 and Ser809 in 5CGD, which further limits the size of the substructures for ligands to the top (Fig. 7b). The docking study between the mavoglurant and 5CGD showed that the saturated bicyclic ring system was no longer favored in the upper region due to the severe collision with Trp785. Instead, the aromatic ring on mavoglurant was placed in the upper region with potential hydrophobic interactions with rings system on Trp785, and the saturated bicyclic ring system was placed in the bottom region (Fig. 7c, d). This pose of mavoglurant was not favored by the allosteric binding pocket with a docking score of 4.5. The docking studies between 5CGD and 80 highest scored fragments in the bottom region, as well as 120 highest scored fragments in upper region were followed. The aggregation of fragments in the bottom region remained relatively well. Figure 7e shows that the rotation of Trp785 has limited influence on the fragment aggregation in the bottom region. Although some big fragments may invade into the upper region, the same situations could be observed with 4OO9, because of the size of these big fragments. Figure 7f shows that the aggregation of fragments in the upper region can no longer been maintained, especially for big fragments. The fragments that used to be categorized in the upper region tend to distribute among both upper and bottom regions in 5CGD. The distribution of fragments (Fig. 7g) showed that the bottom region is favored by the majority of the fragments. Although the overall properties of the upper region remain hydrophobic, the rotation of Trp785 significantly affects the shape of the upper region. The

dramatic conformational changes on critical residues would have unignorable effects towards the effectiveness of the computational fragment linking described in this paper. Fragment growth and fragment merge would be alternative methodologies for novel compound generation using the model 5CGD. Based on the identified fragment hits, linkers and functional groups can be introduced to improve the chemical physical properties and drug likability. These works may be beyond the scope of this paper, and 4OO9 was focused to illustrate the study approach.

DISCUSSION

According to the primary principle of Medicinal Chemistry, similar compounds can have similar bioactivities. In order to further validate our *in silico* synthesized compounds, we tried to find out whether there are existing mGlu₅ allosteric modulators that are structurally similar to our *in silico* synthesized compounds.

A similarity search was performed on SciFinder (scifinder.cas.org) using compound 1, which is the most potent one according to our prediction, as listed on Table II. The search was started with chemical structures and the search type was similarity. The hits with Tanimoto coefficients over 85% were selected for further analysis. Interestingly, we identified one compound (CAS number 1197356-89-0), which is 90% similar towards compound 1, to be almost identical towards compound 6 (97%) listed on Table II. The only difference between 1197356-89-0 and compound 6 is the substitution on the benzene ring (Fig. 8a), with compound 6 featuring chlorine substitution instead of a fluorine. Compound 1197356-89-0 was first published on a US patent, *Piperazine Metabotropic Glutamate Receptor Negative Allosteric Modulators for Anxiety/Depression*, US 2009/0325964 A1. A competitive binding assay, which assesses compounds' ability to displace MPEP from Hek-293 cell membrane expressing rat mGlu₅ receptors, was used to identify the compounds' activity in that patent (67). Median *K_i* value for 1197356-89-0 is 20 nM (Log*K_i* = 1.30), while our prediction for compound 6 is 324.2 nM (Log*K_i* = 2.51). Since logarithmic transformation was used in our QSAR model simulation, the difference in absolute *K_i* value is reasonable and acceptable. The promising biological activity of 1197356-89-0 demonstrates that our method for *in silico* design can identify promising chemotypes for targets of interest when an existing crystal structure is available to build a model around. We further assessed whether the difference between the chlorine and fluorine would be the key component for the activity. It is known that fluorine could function as a hydrogen bond acceptor to intermediate receptor-ligand interactions. But, for chlorine, there is almost no reported role in hydrogen bond interactions. Another compound in that patent, 1197358-18-1, has chlorine connected towards the benzene ring (Fig. 8a). The major difference between 1197358-18-1 and 1197356-89-0 is the substitution on the piperazine ring. For 1197358-18-1, a pyridine is attached, while for 1197356-89-0, a pyrimidine is attached. Median *K_i* value for 1197358-18-1 is 3 nM (Log*K_i* = 0.48), which shows that the change from fluorine to chlorine does not decrease the allosteric regulation potency. Among our *in silico* synthesized compounds, with an additional methyl group on pyridine, compound 7 listed on Table II shares 96% similarity towards 1197358-18-1 (Fig. 8a). The predicted *K_i* value for compound 7 is 545.0 nM (Log*K_i* = 2.73). So, the potency of 1197358-18-1 turns out to be another evidence to support our *in silico* design. And the

consistency between predicted K_i value and experimental data shows our evaluation is reliable to some extent. Through searching the original allosteric modulator database, which used for our fragments generation, there are no identical structures to these two compounds in patent. Although similar compounds do exist, they share a low Tanimoto coefficient.

Meanwhile, since compounds 10, 13–16, and 18 have distinctive structures to 1,2-diphenylethyne analogs and piperazine scaffold described in the abovementioned patent, similarity searches for them on Scifinder were continued. Compound 1312414-34-8, which shares 93% similarity with compound 10, is a compound that published on a US patent, *Bicyclic Thiazoles as Allosteric Modulators of Mglur5 Receptors*, US 2012/0258955 A1 (68). Compound 1312414-34-8 was tested with functional assay, and the pEC_{50} value is 5.29. Compound 10 has one additional methoxy group connected with the aromatic ring, when compared with 1312414-34-8 (Fig. 8b). Compound 879873-30-0 is a compound that is included on a US patent, *Novel Thieno-Pyridine and Thieno-Pyrimidine Derivatives and Their Use as Positive Allosteric Modulators of Mglur2-Receptors*, US 20070275984 A1 (69). Compound 879873-30-0 shares 97% similarity towards compound 14. Compound 14 has one additional methyl group connected with alkane chain (Fig. 8c). [35 S]GTP γ S binding assay was performed to assess the allosteric activities. Compound 879873-30-0 left-ward shifts the agonist mGlu $_2$ concentration-response curve by 2–3.5-fold. Compounds 15/16 share same O-benzyl nicotinamide scaffold. The difference is the aromatic or alkane substitution connected to the peptide bond. Compound 1276013-84-3, which shares 92% similarity with compound 15, is published on a US patent, *O-benzyl Nicotinamide Analogs as Mglur5 Positive Allosteric Modulators*, US 2011/0183980 A1 (Fig. 8d) (70). The EC_{50} (nM) for 1276013-84-3 recorded on the patent is 120. Compound 19 shares same 3-cyano-pyridone scaffold with 950199-70-9 with an 83% similarity (Fig. 8e). Compound 950199-70-9 is listed on a US patent, *1, 4-Disubstituted 3-Cyano-Pyridone Derivatives and Their Use As Positive Allosteric Modulators of MGLUR2-Receptors*, US 2014/0315903 A1 (71). The pEC_{50} of 950199-70-9 for the GTP γ S-PAM is 6.7. For compound 13, it shares 83% similarity with 36295-38-2 (Fig. 8f). Dr. Rosaria Gitto's group reported that *N*-substituted isoquinoline derivatives could be functioned as potential AChE inhibitors. But, there are no reported activities on metabotropic glutamate receptors for *N*-substituted isoquinoline derivatives, as 36295-38-2.

The recall of compounds with existing scaffolds and reported effects on mGluR showed that the approach we generated is feasible. And the discovery of compounds with unreported activities on mGluR revealed that our approach is doable for exploring and designing novel compounds.

CONCLUSION

In this study, we demonstrated a case study of designing allosteric modulators on metabotropic glutamate receptor 5 using *in silico* fragment-based novel compound design. A GPCR allosteric modulator specific fragment library was generated, which could be used for future studies. Various computational methodologies, including benchmarking dataset validation, docking studies, QuaSAR model simulation, etc., were used to construct the *in silico* compounds and validate the effectiveness of this lead generation strategy. The effects

associated with the rotation of toggle switch, Trp785, were considered and evaluated. The dramatic conformational changes on critical residues would have unignorable effects towards the effectiveness of the computational methodology described in this paper. The determination of using one model or combining two or more models should be specified at the beginning of the study design. Among the top 20 *in silico* synthesized compounds, as listed in Table II, diversified structures could be observed. Through the compound similarity search on the Scifinder, multiple patents were identified to contain compounds that have similar structural features with our *in silico* design. Series of compounds with reported mGluR allosteric activities were recurred in this case study for designing mGlu₅ allosteric modulators. The *in silico* designed compounds with reported scaffolds may fill SAR holes in the known, patented series of mGlu₅ modulators. And the recall of compounds with existing scaffolds and reported effects on mGluR showed that the approach we generated is feasible. Meanwhile, the generation of compounds without reported activities on mGluR indicates that our approach is doable for exploring and designing novel compounds. The medicinal chemistry synthesis for these *in silico* compounds is in progress in our laboratory, and we will perform the experiments to confirm our predictions in the future to find out whether they (1) are mGlu₅ allosteric modulators, (2) retain the ranking order the same way as our prediction, and (3) have better activities than known modulators or not in a head-to-head comparison. Since the focus of this manuscript is more the methodology instead of the final products, the experimental validation is beyond the scope and will not be included. Our case study on designing allosteric modulators on mGlu₅ suggested that this computational fragment-based approach is a reliable and powerful methodology for facilitating the future compound design processes.

Supplementary Material

Refer to Web version on PubMed Central for supplementary material.

Acknowledgments

Authors would like to acknowledge the funding supports to the Xie laboratory from the NIH NIDA (P30 DA035778A1), NIH (R01 DA025612), and DOD (W81XWH-16-1-0490).

References

1. Wu H, Wang C, Gregory KJ, Han GW, Cho HP, Xia Y, et al. Structure of a class C GPCR metabotropic glutamate receptor 1 bound to an allosteric modulator. *Science*. 2014; 344(6179):58–64. [PubMed: 24603153]
2. Venter JC, Adams MD, Myers EW, Li PW, Mural RJ, Sutton GG, et al. The sequence of the human genome. *Science*. 2001; 291(5507):1304–51. [PubMed: 11181995]
3. Overington JP, Al-Lazikani B, Hopkins AL. How many drug targets are there? *Nat Rev Drug Discov*. 2006; 5(12):993–6. [PubMed: 17139284]
4. Feng Z, Hu G, Ma S, Xie X-Q. Computational advances for the development of allosteric modulators and bitopic ligands in G protein-coupled receptors. *AAPS J*. 2015; 17(5):1080–95. [PubMed: 25940084]
5. Wenthur CJ, Gentry PR, Mathews TP, Lindsley CW. Drugs for allosteric sites on receptors. *Annu Rev Pharmacol Toxicol*. 2014; 54:165. [PubMed: 24111540]
6. Wellendorph P, Bräuner-Osborne H. Molecular basis for amino acid sensing by family CG-protein-coupled receptors. *Br J Pharmacol*. 2009; 156(6):869–84. [PubMed: 19298394]

7. Lane JR, Abdul-Ridha A, Canals M. Regulation of G protein-coupled receptors by allosteric ligands. *ACS Chem Neurosci*. 2013; 4(4):527–34. [PubMed: 23398684]
8. Melancon BJ, Hopkins CR, Wood MR, Emmitte KA, Niswender CM, Christopoulos A, et al. Allosteric modulation of seven transmembrane spanning receptors: theory, practice, and opportunities for central nervous system drug discovery. *J Med Chem*. 2012; 55(4):1445–64. [PubMed: 22148748]
9. Christopoulos A. Allosteric binding sites on cell-surface receptors: novel targets for drug discovery. *Nat Rev Drug Discov*. 2002; 1(3):198–210. [PubMed: 12120504]
10. Conn PJ, Christopoulos A, Lindsley CW. Allosteric modulators of GPCRs: a novel approach for the treatment of CNS disorders. *Nat Rev Drug Discov*. 2009; 8(1):41–54. [PubMed: 19116626]
11. Bridges TM, Lindsley CW. G-protein-coupled receptors: from classical modes of modulation to allosteric mechanisms. *ACS Chem Biol*. 2008; 3(9):530–41. [PubMed: 18652471]
12. Digby GJ, Conn PJ, Lindsley CW. Orthosteric-and allosteric-induced ligand-directed trafficking at GPCRs. *Curr Opin Drug Discov Devel*. 2010; 13(5):587.
13. Fang Z, Grütter C, Rauh D. Strategies for the selective regulation of kinases with allosteric modulators: exploiting exclusive structural features. *ACS Chem Biol*. 2012; 8(1):58–70. [PubMed: 23249378]
14. Möhler H, Fritschy J, Rudolph U. A new benzodiazepine pharmacology. *J Pharmacol Exp Ther*. 2002; 300(1):2–8. [PubMed: 11752090]
15. Pin J-P, Galvez T, Prézeau L. Evolution, structure, and activation mechanism of family 3/C G-protein-coupled receptors. *Pharmacol Ther*. 2003; 98(3):325–54. [PubMed: 12782243]
16. Doré AS, Okrasa K, Patel JC, Serrano-Vega M, Bennett K, Cooke RM, et al. Structure of class C GPCR metabotropic glutamate receptor 5 transmembrane domain. *Nature*. 2014; 511(7511):557–62. [PubMed: 25042998]
17. Li G, Jørgensen M, Campbell BM. Metabotropic glutamate receptor 5-negative allosteric modulators for the treatment of psychiatric and neurological disorders (2009–July 2013). *Pharmaceutical patent analyst*. 2013; 2(6):767–802. [PubMed: 24237242]
18. Levenga J, Hayashi S, de Vrij FM, Koekkoek SK, van der Linde HC, Nieuwenhuizen I, et al. AFQ056, a new mGluR5 antagonist for treatment of fragile X syndrome. *Neurobiol Dis*. 2011; 42(3):311–7. [PubMed: 21316452]
19. Ritzén A, Sindet R, Hentzer M, Svendsen N, Brodbeck RM, Bundgaard C. Discovery of a potent and brain penetrant mGluR5 positive allosteric modulator. *Bioorg Med Chem Lett*. 2009; 19(12):3275–8. [PubMed: 19443216]
20. Hajduk PJ, Greer J. A decade of fragment-based drug design: strategic advances and lessons learned. *Nat Rev Drug Discov*. 2007; 6(3):211–9. [PubMed: 17290284]
21. Joseph-McCarthy D, Campbell AJ, Kern G, Moustakas D. Fragment-based lead discovery and design. *J Chem Inf Model*. 2014; 54(3):693–704. [PubMed: 24490951]
22. Allen KN, Bellamacina CR, Ding X, Jeffery CJ, Mattos C, Petsko GA, et al. An experimental approach to mapping the binding surfaces of crystalline proteins. *J Phys Chem*. 1996; 100(7):2605–11.
23. Chessari G, Woodhead AJ. From fragment to clinical candidate—a historical perspective. *Drug Discov Today*. 2009; 14(13):668–75. [PubMed: 19427404]
24. de Kloe GE, Bailey D, Leurs R, de Esch IJ. Transforming fragments into candidates: small becomes big in medicinal chemistry. *Drug Discov Today*. 2009; 14(13):630–46. [PubMed: 19443265]
25. Jencks WP. On the attribution and additivity of binding energies. *Proc Natl Acad Sci*. 1981; 78(7):4046–50. [PubMed: 16593049]
26. Schneider G, Fechner U. Computer-based de novo design of drug-like molecules. *Nat Rev Drug Discov*. 2005; 4(8):649–63. [PubMed: 16056391]
27. Howard N, Abell C, Blakemore W, Chessari G, Congreve M, Howard S, et al. Application of fragment screening and fragment linking to the discovery of novel thrombin inhibitors. *J Med Chem*. 2006; 49(4):1346–55. [PubMed: 16480269]
28. Edwards PD, Albert JS, Sylvester M, Aharony D, Andisik D, Callaghan O, et al. Application of fragment-based lead generation to the discovery of novel, cyclic Amidine β -secretase inhibitors

- with nanomolar potency, cellular activity, and high ligand efficiency §. *J Med Chem.* 2007; 50(24): 5912–25. [PubMed: 17985862]
29. Sun C, Petros AM, Hajduk PJ. Fragment-based lead discovery: challenges and opportunities. *J Comput Aided Mol Des.* 2011; 25(7):607–10. [PubMed: 21732249]
30. Loving K, Alberts I, Sherman W. Computational approaches for fragment-based and de novo design. *Curr Top Med Chem.* 2010; 10(1):14–32. [PubMed: 19929832]
31. Schuffenhauer A, Ruedisser S, Marzinzik A, Jahnke W, Selzer P, Jacoby E. Library design for fragment based screening. *Curr Top Med Chem.* 2005; 5(8):751–62. [PubMed: 16101415]
32. Siegal G, Eiso A, Schultz J. Integration of fragment screening and library design. *Drug Discov Today.* 2007; 12(23):1032–9. [PubMed: 18061882]
33. Lepre C. Fragment-based drug discovery using the SHAPES method. *Expert Opin Drug Discovery.* 2007; 2(12):1555–66.
34. Fjellström O, Akkaya S, Beisel H-G, Eriksson P-O, Erixon K, Gustafsson D, et al. Creating novel activated factor xi inhibitors through fragment based lead generation and structure aided drug design. *PLoS One.* 2015; 10(1):e0113705. [PubMed: 25629509]
35. Ni S, Yuan Y, Huang J, Mao X, Lv M, Zhu J, et al. Discovering potent small molecule inhibitors of cyclophilin A using de novo drug design approach. *J Med Chem.* 2009; 52(17):5295–8. [PubMed: 19691347]
36. Huang Z, Mou L, Shen Q, Lu S, Li C, Liu X, et al. ASD v2. 0: updated content and novel features focusing on allosteric regulation. *Nucleic Acids Res.* 2014; 42(D1):D510–D6. [PubMed: 24293647]
37. Christopher JA, Aves SJ, Bennett KA, Doré AS, Errey JC, Jazayeri A, et al. Fragment and structure-based drug discovery for a class C GPCR: discovery of the mGlu5 negative allosteric modulator HTL14242 (3-chloro-5-[6-(5-fluoropyridin-2-yl) pyrimidin-4-yl] benzonitrile). *J Med Chem.* 2015; 58(16):6653–64. [PubMed: 26225459]
38. SYBYL-X 1.3 TI. Vol. 2010. 1699 South Hanley Rd., St. Louis, MO, 63144, USA: 2010.
39. Feng Z, Ma S, Hu G, Xie X-Q. Allosteric binding site and activation mechanism of class C G-protein coupled receptors: metabotropic glutamate receptor family. *AAPS J.* 2015; 17(3):737–53. [PubMed: 25762450]
40. The PyMOL Molecular graphics system, version 1.5.0.1. Schrodinger, LLC;
41. Lewell XQ, Judd DB, Watson SP, Hann MM. RECAP Retrosynthetic Combinatorial Analysis Procedure: a powerful new technique for identifying privileged molecular fragments with useful applications in combinatorial chemistry. *J Chem Inf Comput Sci.* 1998; 38(3):511–22. [PubMed: 9611787]
42. Fechner U, Schneider G. Flux (1): a virtual synthesis scheme for fragment-based de novo design. *J Chem Inf Model.* 2006; 46(2):699–707. [PubMed: 16563000]
43. Teague SJ, Davis AM, Leeson PD, Oprea T. The design of leadlike combinatorial libraries. *Angew Chem Int Ed.* 1999; 38(24):3743–8.
44. Jain AN. Scoring noncovalent protein-ligand interactions: a continuous differentiable function tuned to compute binding affinities. *J Comput Aided Mol Des.* 1996; 10(5):427–40. [PubMed: 8951652]
45. Weiner SJ, Kollman PA, Nguyen DT, Case DA. An all atom force field for simulations of proteins and nucleic acids. *J Comput Chem.* 1986; 7(2):230–52. [PubMed: 29160584]
46. Meng EC, Shoichet BK, Kuntz ID. Automated docking with grid-based energy evaluation. *J Comput Chem.* 1992; 13(4):505–24.
47. Sanner MF. Python: a programming language for software integration and development. *J Mol Graph Model.* 1999; 17(1):57–61. [PubMed: 10660911]
48. Morris GM, Goodsell DS, Halliday RS, Huey R, Hart WE, Belew RK, et al. Automated docking using a Lamarckian genetic algorithm and an empirical binding free energy function. *J Comput Chem.* 1998; 19(14):1639–62.
49. Huang N, Shoichet BK, Irwin JJ. Benchmarking sets for molecular docking. *J Med Chem.* 2006; 49(23):6789–801. [PubMed: 17154509]

50. Weiss DR, Bortolato A, Tehan B, Mason JS. GPCR-bench: a benchmarking set and practitioners' guide for G protein-coupled receptor docking. *J Chem Inf Model*. 2016; 56(4):642–51. [PubMed: 26958710]
51. Mysinger MM, Carchia M, Irwin JJ, Shoichet BK. Directory of useful decoys, enhanced (DUD-E): better ligands and decoys for better benchmarking. *J Med Chem*. 2012; 55(14):6582–94. [PubMed: 22716043]
52. Ioele G, De Luca M, Oliverio F, Ragno G. Prediction of photosensitivity of 1, 4-dihydropyridine antihypertensives by quantitative structure-property relationship. *Talanta*. 2009; 79(5):1418–24. [PubMed: 19635379]
53. Raska I, Toropov A. Comparison of QSPR models of octanol/water partition coefficient for vitamins and non vitamins. *Eur J Med Chem*. 2006; 41(11):1271–8. [PubMed: 16920228]
54. Hogg, R., Tanis, E. Probability and statistical inference. 4. New York: MacMillan Publishing Co; 1993. Nonparametric methods; p. 589-646.
55. Hogg, RV., Tanis, EA. Solutions Manual, Probability and Statistical Interference: Solutions. Macmillan; 1988.
56. Ertl P. Cheminformatics analysis of organic substituents: identification of the most common substituents, calculation of substituent properties, and automatic identification of drug-like bioisosteric groups. *J Chem Inf Comput Sci*. 2003; 43(2):374–80. [PubMed: 12653499]
57. Congreve M, Carr R, Murray C, Jhoti H. A 'rule of three' for fragment-based lead discovery? *Drug Discov Today*. 2003; 8(19):876–7.
58. Rees DC, Congreve M, Murray CW, Carr R. Fragment-based lead discovery. *Nat Rev Drug Discov*. 2004; 3(8):660–72. [PubMed: 15286733]
59. Lipinski CA, Lombardo F, Dominy BW, Feeney PJ. Experimental and computational approaches to estimate solubility and permeability in drug discovery and development settings. *Adv Drug Deliv Rev*. 2012; 64:4–17.
60. Bogan AA, Thorn KS. Anatomy of hot spots in protein interfaces. *J Mol Biol*. 1998; 280(1):1–9. [PubMed: 9653027]
61. Leffler JE. The enthalpy-entropy relationship and its implications for organic chemistry. *The Journal of Organic Chemistry*. 1955; 20(9):1202–31.
62. Eftink MR, Anusiem A, Biltonen RL. Enthalpy-entropy compensation and heat capacity changes for protein-ligand interactions: general thermodynamic models and data for the binding of nucleotides to ribonuclease A. *Biochemistry*. 1983; 22(16):3884–96. [PubMed: 6615806]
63. Davenport, AP., Russell, FD. Current Directions in Radiopharmaceutical Research and Development. Springer; 1996. Radioligand binding assays: theory and practice; p. 169-79.
64. Cui W, Yan X. Adaptive weighted least square support vector machine regression integrated with outlier detection and its application in QSAR. *Chemom Intell Lab Syst*. 2009; 98(2):130–5.
65. Baell JB, Holloway GA. New substructure filters for removal of pan assay interference compounds (PAINS) from screening libraries and for their exclusion in bioassays. *J Med Chem*. 2010; 53(7): 2719–40. [PubMed: 20131845]
66. Patlewicz G, Jeliazkova N, Safford R, Worth A, Aleksiev B. An evaluation of the implementation of the Cramer classification scheme in the Toxtree software. *SAR QSAR Environ Res*. 2008; 19(5–6):495–524. [PubMed: 18853299]
67. Bursavich, MG., Gilbert, AM., Stock, JR. Piperazine metabotropic glutamate receptor 5 (MGLUR5) negative allosteric modulators for anxiety/depression. US Patent. 2009.
68. Macdonald, G., Tresadern, G., Trabanco-Suarez, A., Pastor-Fernandez, J. Preparation of bicyclic thiazoles as allosteric modulators of mGluR5 receptors. WO Patent. 2011.
69. Imogai, H., Mutel, V., Duvey, GA., Cid-Nunez, J., Le Poul, E., Lutjens, R. Novel thieno-pyridine and thieno-pyrimidine derivatives and their use as positive allosteric modulators of mGluR2-receptors. US Patent. 2005.
70. Conn, J., Lindsley, C., Weaver, C., Stauffer, S., Williams, R., McDonald, G., et al. Preparation of O-benzyl nicotinamide analogs as mGluR5 positive allosteric modulators. WO Patent. 2011.
71. Imogai, HJ., Cid-Nunez, JM., Andres-Gil, JI., Trabanco-Suarez, AA., Santamarina, JO., Dautzenberg, FM., et al. 1, 4-Disubstituted 3-cyano-pyridone derivatives and their use as positive allosteric modulators of MGLUR2-receptors. US Patent. 2014.

Fig. 1. Allosteric binding site of mGlu₅. **a** The allosteric binding site (highlighted in *salmon pink*), which is formed by surrounding residues, marked in *cyan*. **b** The interaction between mavoglurant and surrounding residues based on the spatial information from the reported crystal complex. **c** Docking pose and ligand-residue interaction predicted through the docking study. **d** The overlap of mavoglurant between crystallized complex and predicted pose

Fig. 2. Fragment docking studies and fragments sorting. **a** Upper region (highlighted in *pink*) and bottom region (highlighted in *yellow*), which are connected through a narrow channel and recognized through fragments docking. **b** The aggregation of fragments inside the allosteric binding site. **c** The distribution of fragments towards the upper and bottom region of 4009. *Blue bars* represent the number of fragments in the upper region. *Yellow bars* represent the number of fragments in the bottom region

Fig. 3.

In silico synthesis of potential mGlu₅ allosteric modulators. **a** Distribution of the docking scores for small-scale RECAP synthesis. *Yellow bars* represent compounds generated from fragments with highest docking scores. *Blue bars* represent compounds generated from fragments with lowest docking scores. **b** The binding pose of top one *in silico* synthesized compound. **c** Comparison of binding poses between top one *in silico* synthesized compound (*pink*) and mavoglurant (*yellow*)

Fig. 4. Correlation plot for docking score and binding energy. *Red spot* represents the docking score and binding energy for mavoglurant. *Yellow spots* represent the docking scores and binding energies for top 20 *in silico* synthesized compounds

Fig. 5. The enrichment test with decoys. The *red bar* represents the reported negative allosteric modulator, mavoglurant, for mGlu₅. *Yellow bars* represent the top 20 *in silico* synthesized compounds. *Blue bars* represent the decoys generated under DUD-E

Fig. 6. Correlation plot for quantitative structure-activity relationship model. *Blue spots* represent for items in training set. *Yellow spots* represent for items in test set. The root mean square error is 0.49. The correlation coefficient is 0.82. The cross-validated RMSE is 0.60. The cross-validated R^2 is 0.74

Fig. 7. Assessment on the rotation of Trp785. **a** Extracellular view on the mGlu₅ allosteric binding pocket from 4O09. **b** Extracellular view on the mGlu₅ allosteric binding pocket from 5CGD. **c, d** Extracellular view and parallel to the membrane view on the binding pose of mavoglurant inside 5CGD according to the molecular docking. **e, f** Distribution of categorized fragments in the allosteric binding pocket of 5CGD. **g** The distribution of fragments towards the upper and bottom region of 5CGD. *Blue bars* represent the number of fragments in the upper region. *Orange bars* represent the number of fragments in the bottom region

Fig. 8. Structural comparison between *in silico* synthesized compounds and existing compounds with known mGluR allosteric activities. Comparisons were *circled with dashed lines in red, blue, and pink colors*. The major difference inside each group was *circled with the same color* as their group. The CAS number for patented compounds and the corresponding number for *in silico* synthesized compounds are listed under each structure

Table I

Structures and Properties for Top Ten Most Frequently Appeared Fragments

Structure	Frequency	LogP(o/w)	M.W.	Hydrogen bond donor	Hydrogen bond acceptor
407	1.60	94.11	1	1
357	0.56	107.11	0	2
324	0.91	203.20	0	1
197	-0.06	46.07	1	1
188	1.99	124.11	0	1

Structure	Frequency	LogP(o/w)	M.W.	Hydrogen bond donor	Hydrogen bond acceptor
188	-0.42	153.19	1	3
173	0.99	70.01	0	0
167	-0.10	45.08	1	1
163	2.06	96.10	0	0
162	0.67	79.10	0	1

Table II

Docking Scores, Binding Energies, and Predicted *K_i* Values for Top 20 *In Silico* Synthesized Compounds

No.	Structure	Docking score	Binding energy	Predicted <i>K_i</i>
1		10.4	−7.7 kcal/mol	1.3 nM
2		8.9	−7.2 kcal/mol	10.9 nM
3		8.3	−5.5 kcal/mol	67.3 nM
4		8.1	−6.1 kcal/mol	132.6 nM
5		8.1	−5.8 kcal/mol	88.1 nM
6		8.0	−5.4 kcal/mol	324.2 nM
7		7.7	−4.7 kcal/mol	545.0 nM

No.	Structure	Docking score	Binding energy	Predicted Ki
8		7.6	-5.1 kcal/mol	225.5 nM
9		7.6	-4.8 kcal/mol	199.4 nM
10		7.6	-5.4 kcal/mol	--
11		7.6	-5.3 kcal/mol	207.0 nM
12		7.5	-4.5 kcal/mol	130.9 nM
13		7.5	-4.3 kcal/mol	--
14		7.4	-4.4 kcal/mol	--
15		7.4	-5.5 kcal/mol	--
16		7.4	-5.1 kcal/mol	--

No.	Structure	Docking score	Binding energy	Predicted Ki
17		7.3	-4.2 kcal/mol	1162.3 nM
18		7.3	-5 kcal/mol	357.6 nM
19		7.3	-4.1 kcal/mol	--
20		7.3	-4.4 kcal/mol	792.2 nM
NAM		7.9	-5.9 kcal/mol	107.0 nM (Experimental Ki value: 66 nM)

Table III

Descriptors of Linear Regression Analysis

Notation	Descriptors	Coefficient	Contingency coefficient	Cramer's V	Entropic uncertainty	Linear correlation
Intercept	—	−6.81	—	—	—	—
BCUT_PEOE_3	PEOE Charge BCUT (3/3)	3.11	0.82	0.41	0.43	0.16
BCUT_SLOGP_3	LogP BCUT (3/3)	−2.46	0.80	0.39	0.40	0.17
b_IrotN	Number of rotatable single bonds	0.46	0.69	0.27	0.22	0.15
b_IrotR	Fraction of rotatable single bonds	−12.76	0.79	0.37	0.36	0.20
GCUT_PEOE_0	PEOE Charge GCUT (0/3)	−9.57	0.81	0.40	0.44	0.29
GCUT_SLOGP_0	LogP GCUT (0/3)	−4.37	0.77	0.35	0.39	0.17
GCUT_SMR_0	Molar Refractivity GCUT (0/3)	4.49	0.82	0.41	0.41	0.20
Kier3	Third Kappa shape index	−0.17	0.76	0.33	0.30	0.19
PEOE_VSA+5	Total positive 5 vdw surface area	0.06	0.66	0.25	0.30	0.19
SlogP_VSA4	Bin 4 SlogP (0.10, 0.15)	−0.03	0.78	0.36	0.36	0.20
SMR_VSA0	Bin 0 SMR (0.000, 0.110)	−0.02	0.72	0.30	0.32	0.16
SMR_VSA1	Bin 1 SMR (0.110, 0.260)	−0.02	0.73	0.31	0.28	0.25