Fragment- and Negative Image-Based Screening of Phosphodiesterase 10A Inhibitors

Elmeri M. Jokinen[1], Pekka A. Postila[2], Mira Ahinko[2], Sanna Niinivehmas[1], Olli T. Pentikäinen[1,2,3*]

[1] Institute of Biomedicine, Kiinamyllynkatu 10 C6, FI-20014 University of Turku, Finland.

[2] University of Jyvaskyla, Department of Biological and Environmental Science, P.O. Box 35, FI-40014 University of Jyvaskyla, Finland.

[3] Aurlide ltd., Turku, Finland.

[*]Corresponding author: olli.pentikainen@utu.fi

ABSTRACT

A novel virtual screening methodology called fragment- and negative image-based (F-NiB) screening is introduced and tested experimentally using phosphodiesterase 10A (PDE10A) as a case study. Potent PDE10A-specific small-molecule inhibitors are actively sought after for their antipsychotic and neuroprotective effects. The F-NiB combines features from both fragment-based drug discovery and negative image-based (NIB) screening methodologies to facilitate rational drug discovery. The selected structural parts of protein-bound ligand(s) are seamlessly combined with the negative image of the target's ligand-binding cavity. This cavity- and fragment-based hybrid model, namely its shape and electrostatics, is used directly in the rigid docking of *ab initio* generated ligand 3D conformers. In total, 14 compounds were acquired using the F-NiB methodology, 3D quantitative structure-activity relationship modeling and pharmacophore modeling. Three of the small-molecules inhibited PDE10A at ~27 μM to ~67 μM range in a radiometric assay. In a larger context, the study shows that the F-NiB provides a flexible way to incorporate small-molecule fragments into the drug discovery.

Key words: Phosphodiesterase 10A (PDE10A), schizophrenia, Parkinson's disease, Huntington's disease, negative image-based (NIB), fragment- and negative image-based (F-NiB) screening, virtual screening, structure-based virtual screening, radiometric activity assay, fragment-based drug discovery.

INTRODUCTION

In theory, virtual screening assays, which sieve through thousands to billions of compounds within hours or days, produce sufficient level of enrichment to limit the costly experimental testing to the very best drug candidates. In reality, these *in silico* assays frequently come short of this goal and produce lackluster results. Regardless, the computations can still bring forth solid innovations such as drug scaffolds to be developed further using systematic organic synthesis programmes (Frydenvang et al., 2009; Lash et al., 2008; Niinivehmas, Manivannan, Rauhamäki, Huuskonen, & Pentikäinen, 2016; Postila, Swanson, & Pentikäinen, 2010).

The virtual screening approaches are divided roughly in two categories: the ligand-based and structure-based methodologies. The ligand-based methods such as pharmacophore (PHA) and three-dimensional quantitative structure-activity relationship (3D-QSAR) modelling, which rely on the advance knowledge of active compounds, are computationally inexpensive. However, despite being commonly used in the drug discovery and lead optimization, the PHA and 3D-QSAR model generation does not necessarily utilize the bioactive ligand conformers but those that generate the most explanatory model(s) (Cramer, Patterson, & Bunce, 1988; Kinase, Zhang, Li, Zhang, & Ai, 2010; Lowe, Ferrebee, Rodriguez, Conn, & Meiler, 2010; Niinivehmas et al., 2016; Patel, Noolvi, & Sharma, 2014; Shubina, Niinivehmas, & Pentikäinen, 2015; Tian et al., 2011; Yadav et al., 2010). In contrast, the structure-based methods such as flexible molecular docking, which attempt to predict the ligand's bioactive binding pose and estimate its binding energy, rely solely on the target protein's 3D structure and require a lot of computational resources (Niinivehmas et al., 2016; Nurminen et al., 2010; Shubina et al., 2015).

In this regard, the negative image-based (NIB) screening (Figs. 1 and S1) (Ahinko, Kurkinen, Niinivehmas, Pentikäinen, & Postila, 2019; Lätti, Niinivehmas, & Pentikäinen, 2016; Niinivehmas, Salokas, Lätti, Raunio, & Pentikäinen, 2015; Niinivehmas, Virtanen, Lehtonen, Postila, & Pentikäinen, 2011; Virtanen & Pentikäinen, 2010) is set to get the best from both worlds as it can be used to screen compounds with only the protein 3D structure and the screening process itself lends its speed from the traditional ligand-based screening methodology.

[Figure_1]

In the NIB screening, the shape of the ligand-binding cavity is utilized directly in the cavity-based rigid docking (Figs. 1 and S1). The technique is reminiscent of the traditional ligand-based approach in which the similarity comparison is performed against a known template ligand, whose active binding pose has been solved experimentally using for example X-ray crystallography. The NIB method relies on building a drug-like pseudo-ligand or negative image of the protein's ligand-binding cavity, where the electrostatics, solvation and alternative protonation can be incorporated as well (Figs. 1 and S1). The NIB models are composed of neutral filler atoms and negatively and positively charged cavity points reflecting the hydrogen bonding (H-bonding) features of the cavity lining residues (Fig. 1). The models need to be generated using specifically tailored cavity detection software PANTHER (Niinivehmas et al., 2015), but the *ab initio* generated ligand 3D conformers and the similarity comparison or rigid docking is performed with established ligand-based screening tools (Vainio & Johnson, 2007; Vainio, Puranen, & Johnson, 2009).

Because the shape complementarity between the protein's ligand-binding cavity and the ligand is a key factor in assuring strong binding, the NIB can outperform traditional virtual screening approaches such as flexible molecular docking on a case by case basis (Niinivehmas et al., 2015, 2011; Virtanen & Pentikäinen, 2010). In addition to the benchmark testing, the NIB methodology has been used successfully in tool compound discovery projects for metabolizing enzymes (Juvonen, Ahinko, Huuskonen, Raunio, & Pentikäinen, 2018). The NIB benefits especially from the Molecular Mechanics Generalized Born Surface Area (MM/GBSA) post-processing, which, in general, has been shown to be an efficient way to recognize correct binding poses of small-molecule ligands (Ahinko, Niinivehmas, Jokinen, & Pentikäinen, 2019; Niinivehmas et al., 2011). The NIB has also been paired with the PHA filtering, where specific PHA points are generated based on well-known ligand-receptor hot spots or interaction sites (Rauhamäki et al., 2018). Moreover, the negative images can be used to rescore explicit docking poses to improve the flexible docking enrichment – a set-up that works excellently with a multitude of targets based on the benchmark testing (Kurkinen et al., 2018). What is common to these prior efforts is that the similarity comparison has relied solely on the cavity-based models composed of the charged cavity points and neutral filler atoms (Fig. 1).

In the fragment-based drug discovery (FBDD), the virtual screening is used to find low-molecular mass compounds or functional fragments that bind into the target protein rather than attempting to discover full-size drug candidates directly (Scott, Coyne, Hudson, & Abell, 2012). Frequently, the fatal challenge of the FBDD is the stage in which the fragments are supposed to be fused together to produce a potent drug – a process that typically involves organic synthesis (Hao et al., 2012; Xiong et al., 2017, 2016). Here, the idea of the fragment- and negative image-based (F-NiB) screening (Figs. 1 and S1) is to simply incorporate the ligand fragments directly into the cavity-based negative image (Figs. 1 and S1) to improve the model fitness in the rigid docking. Thus, the F-NiB relies on generating hybrid models that, at least in theory, fuse seamlessly together the best parts from both the protein-bound fragment(s) and the cavity-based model.

The viability of the F-NiB approach for drug discovery is demonstrated in practice by screening inhibitors for phosphodiesterase 10A (PDE10A; Fig. 2A). The hydrolysis of intracellular second messenger cyclic nucleotide monophosphate is catalyzed by the PDE10A (Fig. 2A-B) especially at the striatum (Fujishige, Kotera, & Omori, 1999). The enzyme was chosen as the test case for the F-NiB methodology, because, if discovered, potent PDE10A-specific inhibitors (Fig. 2C-F) could be used in the treatment of schizophrenia and neurodegenerative disorders such as Parkinson's and Huntington's diseases (Garcia, Redondo, Martinez, & Gil, 2014). Moreover, the second messengers maintain neuronal functioning such as cellular metabolism and neurotransmitter synthesis and, thus, preventing their breakdown by the PDEs improves indirectly the synaptic neurotransmission (Duinen et al., 2015; Nestler & Duman, 1999). In addition, a vast amount of protein 3D structure and ligand activity data is available for PDE10A. This made it possible to apply various virtual screening methods such as the 3D-QSAR and flexible molecular docking in combination with the F-NiB in the search of the novel PDE10A inhibitors.

[Figure_2]

MATERIALS AND METHODS

Computational methods

Ligand preparation. Small-molecule ligands with experimentally measured activities were obtained from the ChEMBL (Bento et al., 2014) (n=804, mw: 215-600 g/mol, retrieved 22nd of July 2016) in the SMILES (Simplified Molecular-Input Line-Entry System) format. The 3D structures were generated for the ligands by using LIGPREP in MAESTRO 10.5.013 (Schrödinger Inc, Portland, OR, USA). In LIGPREP, the OPLS3 (Harder et al., 2016) force field was used, ionization was done at pH 7.4 with EPIK (Shelley et al., 2007), tautomers were created and ligands were desalted, specific chiralities were retained, at most 32 tautomers and chiralities per ligand were generated, keeping one low energy ring conformation per ligand. Multiple conformations were generated for the 3D converted ligands with CONFGEN (Watts et al., 2010) in MAESTRO, using the intermediate search strategy. The SPECS molecular database (www.specs.net) was used as a source of molecules in virtual screening with all the models described below. The SPECS molecules (Specs 10mg drug-like Apr 2014 collection, max 8 rotbonds, mw: 250-600 g/mol) were provided as multiple conformers, produced in similar way as the ligands retrieved from the ChEMBL (Malamas, Ni, & Erdei, 2011).

Pharmacophore modeling. In the pharmacophore (PHA) modelling: a training set including preferably both active and inactive molecules is acquired; low-energy ligand 3D conformers are generated; the molecules are superimposed to produce the best fits for the functional groups; the overlaid/matching information is abstracted to simple PHA points/spheres such as H-bond donor or H-bond acceptor elements (PHA hypotheses); and, finally, a scoring function is used to rank the hypotheses according to their ability to represent the common pharmacophoric features of the active molecules (Leach, Gillet, Lewis, & Taylor, 2010). Here, a PHA model (Fig. S3A) was generated with PHASE (Dixon et al., 2006) in MAESTRO using default settings and a set of 74 PDE10A active ChEMBL ligands (Malamas et al., 2011). ChEMBL ligands, which were used in the PHA model building, are shown in Fig. S3B bound at the enzyme's ligand-binding site (docked using PLANTS; see details below). Ligands that had experimental $pIC_{50} > 8.0$ were defined active. Hypotheses that utilized minimum of five pharmacophoric sites that matched at least 35 of the 42 active compounds were considered. The model AAHHR.155 that contained two H-bond acceptor sites, two hydrophobic sites and one aromatic ring site was selected because of having the best PHASE survival score (3.904).

Field-based QSAR. Ligand alignment performed by, for example, PHA model can be used to establish quantitative structure-activity relationship (QSAR) on the aligned molecules (Leach et al., 2010). The field-based QSAR (FQSAR) reminds the PHA modelling in that it generates specific 3D fields (not necessarily spheres) designating the areas housing groups responsible of H-bonding or hydrophobic interactions based on the superimposition and abstraction. However, the FQSAR also incorporates QSAR data directly into the fields and, thus, allows the building of more representative and predictive models to provide more information of the contributions of different molecular features to the activity. An FQSAR model was created based on the active ligand alignment by AAHHR.155. The FQSAR model was generated by using Gaussian steric, electrostatic, hydrophobic, H-bond donor and acceptor 3D fields (Klebe & Abraham, 1999; Klebe, Abraham, & Mietzner, 1994). Linear correlation between independent (molecular properties

represented by the 3D fields) and dependent variables (ligand activities) was constructed by using the Partial Least Squares (PLS) regression methodology. The model was internally validated by the leave-one-out method. 20 % of the active PDE10A ligands were assigned to the test set by random assignment and 80 % of the molecules were kept in the training set. The external validation of the QSAR model was performed by predicting the activities of the test set molecules. A series of models with an increasing number of PLS factors was examined and a model with the most statistical robustness was selected. The generated FQSAR model was used to predict activities of the SPECS molecules ranked best by the PHA model.

Building fragment- and cavity negative image-based hybrid models. The hybrid models used in the F-NiB screening are built in a straightforward manner (Figs. 1 and S1): (1) a negative image of the binding pocket is generated using PANTHER; (2) one or more ligand fragments making the key interactions at the ligand-binding pocket are selected; (3) the coordinates are merged and the overlapping cavity sections are removed. Rest of the screening process involving protein preparation (*e.g.* solvation, protonation), ligand 3D conformer generation and the similarity screening do not differ from the established NIB methodology (Figs. 1 and S1) (Niinivehmas et al., 2015; Vainio et al., 2009).

In total, four hybrid models (I-IV; Figs. 3 and S4) were generated utilizing the ligand-binding pockets of two human PDE10A X-ray crystal structures (PDB: 4HEU; 3SN7) (Malamas et al., 2011; Rzasa et al., 2012). Before building the cavity-based negative images or NIB (negative image-based) models with PANTHER (Niinivehmas et al., 2015), protons were added to the protein structures using REDUCE 3.24.130724 (Word, Lovell, Richardson, & Richardson, 1999). The selected fragments of bound compounds, explained in more detail below, were incorporated in the models in the BODIL Molecular Modeling Environment (Lehtonen et al., 2004). Instead of trying to build one sweeping model that ticks all the boxes, Models I-IV (Figs. 3 and S4) were intended to act as realistic templates for finding moderately sized inhibitors that occupy a few specific high-potency regions of the spacious PDE10A pocket (Fig. 2G).

[Figure_3]

Model I: For the first model (Fig. 3D-E), PDB-entry 4HEU (Rzasa et al., 2012) was used as the input template structure in building the cavity's negative image. For PANTHER (Niinivehmas et al., 2015), box radius was set to 12 Å and the model dimensions were kept close to the protein bound biaryl ether inhibitor (residue name: **15J***)* with ligand distance limit of 1.0 Å. The first fragment, a hydroxyl group extracted from the inhibitor (Fig. 2C-F), was incorporated directly into the model in BODIL (Fig. 3D). The hydroxyl group's atoms were given charges matching serine residue in the charge library of PANTHER. The radius of the carbon atoms of the target protein was set to 2.25 Å in the PANTHER radius library to retain the size of the model in a moderate level.

Model II: The second model (Fig. S4) was also built based on PDB-entry 4HEU with the same settings as was done with Model I and, thus, the hydroxyl of **15J** was again directly incorporated into the hybrid model. However, in the middle section of the model was incorporated an additional fragment, three-ring system, extracted from compound **12** (Fig. S2), which was discovered from the SPECS database using the PHA model (Fig. S3A). Although compound **12** was eventually

found inactive in the experimental testing, in its optimized binding pose the compound is forming three bonds with an optimal geometry: the amino group of Gln716 side chain donated H-bonds both to the keto group and the nitrogen in the four-ring system; and the hydroxyl group of Tyr683 side chain (or $Tyr683^{OH}$) donated an H-bond to the same keto group (Fig. S5). Those cavity points or filler atoms generated by PANTHER in the area occupied by the ring system were replaced by the ring system atoms in BODIL (Lehtonen et al., 2004). In the model, these ring atoms were given the same charges as designated in the SPECS molecular database. The polar groups in the fragment were placed in the same way as those obtained for the cavity model using PANTHER, but the introduction of the ring-system provided a more planar shape to the model at this site.

Model III: Two changes were implemented for the third hybrid model (Fig. 3G-H) in comparison to the Model II (Fig. S4). First, the three-ring system fragment of compound **12** (Fig. S2) was replaced with a 8-fluoro-6methoxy-3,4-dimethyl-imdazo[1,5-a]quinoxaline fragment extracted from a bound inhibitor ligand present in the PDB-entry 3SN7 (residue name: **540**). The new fragment was given charges that are the opposite to the surrounding polar residues or the main chain nitrogen and oxygen of Gln716 and Tyr514, respectively (Fig. 2B-C). Second, the hydroxyl atoms acquired from the biaryl ether inhibitor **15J** (Fig. 2C) for the Models I and II (Figs. 3D-E and S4) were not included in the Model III.

Model IV: The fourth model (Fig. 3A-B) was generated using the PDB-entry 3SN7 (Malamas et al., 2011). The NIB model dimensions were restricted by applying the ligand distance limit of 1.0 Å in PANTHER for the protein bound inhibitor **540**. The inhibitor in question does not place any parts into the PDE10A "*selectivity pocket*" region (Verhoest et al., 2009) (cyan in Fig. 2G) and, accordingly the resulting model does not extend to this section of the pocket either. The same three-ring structure as in the Model II was incorporated to produce the hybrid model.

Screening with fragment- and negative image-based models. The small-molecules were geometry optimized and aligned on top of the F-NiB Models I-IV using ShaEP (Vainio et al., 2009). The similarity search algorithm compares the shape and electrostatics of the template F-NiB models against the ligand conformers, superimposes them, and ranks the results based on both the shape and electrostatics of the match. The hybrid models Model I and IV were directly screened against the inhibitor molecules obtained from the ChEMBL (Bento et al., 2014) to validate their use in the screening with the molecules of the SPECS database. The Models I and IV were further processed using the FQSAR (see below).

Field-based QSAR rescoring of F-NiB screening results. Activities of the top 5,000 SPECS-molecules originating from the F-NiB screenings using Models I and IV were predicted by utilizing FQSAR. In other words, a set of 53 active PDE10A inhibitors (Rzasa et al., 2014, 2012) was geometry optimized and aligned against the Model I with ShaEP (Vainio et al., 2009) and a set of 78 active PDE10A inhibitors (Malamas et al., 2011) was similarly superposed against Model IV. The same settings for the model generation and validation were used as with the PHA-based FQSAR.

Flexible molecular docking. A set of 78 active ChEMBL ligands (Malamas et al., 2011) was docked into the PDE10A binding site (see Fig. S3B) using PLANTS (Korb, Stützle, & Exner, 2009) with the ChemPLP scoring function. The centroid for docking was taken directly from the

inhibitor molecule **540** (PDB: 3SN7; chain A) (Malamas et al., 2011). Speed setting 1 and radius of 15 Å were used. The flexible docking did not produce a significant correlation with the available PDE10A activity data (data not shown) and the docking scores were not used as a basis for selecting any ligands for the activity testing. However, PLANTS did produce the binding pose of **540** with an RMSD of 0.56 Å when compared to the pose seen in the X-ray crystal structure. Due to this, the best-ranked SPECS small-molecules suggested by the F-NiB Model III and PHA screening were docked flexibly with PLANTS for the follow-up MM/GBSA calculations (see below. With the Model III, PLANTS was expected to generate slightly better poses at the protein's cavity than the original poses originating from the F-NiB screening. The optimized docking pose of compound **3** that differs from the original F-NiB screening pose is discussed in detail.

Molecular Mechanics/Generalized Born Surface Area (MM/GBSA) calculations or binding free energy (ΔG) calculations were performed using PRIME in MAESTRO (Jacobson, Friesner, Xiang, & Honig, 2002) on ligands suggested by the F-NiB, F-NiB-FQSAR and PHA screening. The ligand-protein complexes suggested by either ShaEP (Vainio et al., 2009) (Models I, II and IV) or PLANTS (Model III and PHA) were used as the starting structures for the subsequent MM/GBSA calculations. The VSGB solvation model (Li et al., 2011) and OPLS3 force field were used, defining amino acids within 4.0 Å from the ligand as flexible in the MM/GBSA calculations. The energy minimization was performed on the flexible region with the "Minimize" - sampling method. The binding free energy values obtained by MM/GBSA calculations were compared to values calculated for the inhibitors present in the PDB-entries 4HEU (-87.518 kcal/mol) and 3SN7 (-46.883 kcal/mol) to guide the final compound selections for the experimental testing.

Brain/blood permeability prediction, pan-assay interference compounds filtering. The brain/blood permeability (QPlogBB) of the compounds was estimated using QIKPROP 5.1 in MAESTRO. The methodology has been demonstrated to work robustly previously (Jhala, Chettiar, & Singh, 2012). The predicted QPlogBB values range from -0.064 to 1.030, which is well within the reference range of -3.0 to 1.2 needed for passing the blood brain barrier. Similarly, the active PDE10A compounds included in the ChEMBL were also predicted to pass the barrier without problems (data not shown). In addition, structural filtering was performed for the selected compounds in CANVAS (Duan, Dixon, Lowrie, & Sherman, 2010; Sastry, Lowrie, Dixon, & Sherman, 2010) implementation of MAESTRO to detect possible pan-assay interference compounds (PAINS). The selected compounds passed the PAINS1-3 filters.

Structural novelty. Structural similarity between the selected compounds and ChEMBL-molecules was evaluated by 2D fingerprint similarity comparison in CANVAS (Duan et al., 2010; Sastry et al., 2010). Hashed 32-bit linear 2D-fingerprints with Daylight atom typing and maximum path length 7 (14 for ring closures) were generated. Similarity analysis was performed by calculating Tanimoto coefficients to describe fingerprint overlap between those compounds selected *via* virtual screening and the known active (ChEMBL) molecules for the PDE10A. The aim of the similarity analysis was to avoid selecting molecules with high structural similarity when compared to the known PDE10A inhibitors. As completely similar and different 2D fingerprints give Tanimoto coefficients 1 and 0, respectively, molecules with low values (0-0.2) were considered structurally novel PDE10A inhibitor candidates.

Figure preparation. Fig. 2 and C, D, F, G and I in Fig. 3 were prepared using VMD1.9.2 (Humphrey, Dalke, & Schulten, 1996). Fig. 1, E and H in Fig. 3 were done with BODIL (Lehtonen et al., 2004), MOLSCRIPT 2.1.2 (Kraulis, 1991) and RASTER 3D 3.0.2 (Merritt & Murphy, 1994).

Experimental materials and methods

The purity of the active ligands is 95 % or more, based on proton nuclear magnetic resonance (H-NMR; **2** and **3**) and liquid chromatography-mass spectrometry (LC-MS; **1**-**3**) as provided by the SPECS molecular database (see Text S1). The radiometric activity assay was performed commercially by SB Drug Discovery (Glasgow, UK) using recombinant human PDE10A enzyme expressed in a baculoviral system. The similarity of the expressed system has been validated by comparing it against PDE10A enzyme taken directly from human tissue using known inhibitor standards where available. The radiometric assay method is a modification of the two-step method of by Thompson & Appleman (Thompson & Appleman, 1971), which has been adapted for 96-well plate format. Dose response curves for compounds **1**-**13** and the control papaverine against human PDE10A1 are shown in Fig. S6.

RESULTS

Virtual screening assays. The fragment- and negative image-based (F-NiB) screening (Figs. 1 and S1) was performed using four different hybrid NIB models combining fragment and cavity information (Figs. 3 and S4; N=13; Table 1). Both pharmacophore (PHA) modelling (Fig. S3; N=1; Table 1) and flexible molecular docking (data not shown) were applied to discover novel PDE10A inhibitor compounds from the SPECS database. In addition, a field-based QSAR or FQSAR-based rescoring scheme was applied on the small-molecules of the SPECS database ranked best by the screens with the PHA or the F-NiB with the Models I or IV.

[Table 1]

The molecular alignment is inarguably the most crucial part of the FQSAR model building. Both the PHA and the F-NiB models worked well in aligning the active ligands to produce predictive FQSAR-models. The PHA-based FQSAR model used 3 PLS factors and had R^2 (training set) = 0.87, Q^2 (test set) = 0.81 and R^2 Scramble = 0.41. The FQSAR model based on the Model I used a single PLS factor as utilizing a higher number of PLS factors resulted in over-fit models (R^2 Scramble > 0.5). However, the PLS = 1 model having R^2 = 0.89, Q^2 = 0.89 and R^2 Scramble = 0.38 indicated adequate predictive power for usage in the virtual screening. The FQSAR model based on rigid docking using the Model IV utilized 3 PLS factors and had R^2 = 0.74 and Q^2 = 0.72 with R^2 Scramble = 0.46. Full FQSAR model statistics, field contributions to the ligand activity and field contour maps are displayed in Tables S1 and S2 and Figure S7, respectively.

The FQSAR methodology was used to predict pIC_{50}-values of the best-ranked SPECS molecules: 75 molecules from the PHA screen were predicted to have pIC_{50} > 10.0, 47 molecules from the F-

NiB with the Model I screening had predicted pIC_{50} > 7.5 and 100 molecules put forth by the F-NiB Model IV screening had predicted pIC_{50} > 8.7. The protein-ligand complexes of these molecules were refined and their binding energies calculated using the MM/GBSA routine (Jacobson et al., 2002) in MAESTRO. In addition, 100 top-scored molecules from the F-NiB screens with the Models II and III were processed with the MM/GBSA routine (Fig. S5; Table 1). The optimized ligand-protein complexes were subjected to the final visual evaluation before the final compound selection.

Compound selection. Regardless of the used screening/scoring method (Table 1), the chances of each compound to bind and inhibit PDE10A activity were estimated on case-by-case basis by inspecting and visualizing each of the bound ligands in complex with the protein. In visual examination of the binding poses, molecules forming favorable interactions especially at the "*Gln interaction"* and "*hydrophobic clamp*" sites were sought (Table S3). Ultimately, only those compounds with sufficient level of structural novelty or perceived "uniqueness" in comparison to the prior PDE10A inhibitors (listed in the ChEMBL database; www.ebi.ac.uk/chembl/) were chosen (Fig. S2). The structural novelty was evaluated by analyzing the structural similarity between the selected molecules and the ChEMBL-molecules by 2D fingerprint comparison. Tanimoto coefficients of all selected molecules were <0.163, indicating structural novelty when compared to previously known PDE10A inhibitors (Table 1).

In practice, the manual selection process was limited to, at most, top 100 compounds put forth by each applied screening method and the scoring from different sources was cross-referenced (Table S4). In total, 14 compounds (**1**-**14** in Fig. S2) were purchased for the radiometric PDE10A activity testing. Although all of the compounds were predicted beforehand to be water-soluble (see LogS values in Table 1), compound **14** (Fig. S2) was later found out to be insoluble (Table 1) when preparing the dilutions from the DMSO stocks. In addition, the selected compounds were deemed likely to pass the blood brain barrier (see QPlogBB values in Table 1) using QIKPROP in MAESTRO.

Experimental results. Majority of the compounds (**4**-**12** in Fig. S2) produced >100 μM PDE10A inhibition and only compounds **13** and **14** (Fig. S2) were completely inactive (Table 1). Importantly, compounds **1**, **2** and **3** (Fig. S2), selected based on the F-NiB screening (Table 1), produced IC_{50} values of 27 μM, 49 μM, and 67 μM, respectively. The IC_{50} value of inhibitor papaverine (Fig. S2), which was used as a positive control in the testing, was determined to be 147 nM (Table 1). Accordingly, **1**-**3** function as PDE10A inhibitors at the micromolar range (Table 1) and, on broader terms, they are mid-range PDE10A inhibitors.

The predicted binding modes of the new inhibitors. The binding modes of compounds **1**-**3**, (Fig. S2) acquired *via* the F-NiB screening with the Models IV, I and III, respectively, are shown in Fig. 3 (Table 1). Their binding was determined also using flexible docking (Model III) and MM/GBSA calculations (Fig. S5); however, it is noteworthy that the binding energy predictions for the active compounds did not significantly differ from the inactive ones (**1**-**3** *vs.* **4**-**13** in Table 1). In general, the different ring systems of the compounds (Fig. 3C, F, I) aligned between or close to the side chains of Phe686 and Phe719 forming the "*hydrophobic clamp*" region (yellow in Fig. 2G) at the

PDE10A active site (Chappie, Helal, & Hou, 2012). The binding modes of compounds **1-3** are described in more detail in Text S2.

DISCUSSION

Difficulties of fragment-based drug discovery. The major difficulty in implementing the fragment-based drug discovery (FBDD) in large scale is not necessarily the lack of promising fragments or even drug scaffolds but the limited ways in which these diverse parts can be patched up together. If starting the compound search using a specific fragment, however simple, one usually quickly discovers that there are no compounds available housing the other required parts in the existing molecular databases. This problem can sometimes be overcome by quick and simple organic synthesis efforts guided by expert insight and/or automated algorithms, but, more often than not, the issue is that the envisioned compound is difficult, costly or even impossible to produce at the chemist's workbench. The problem is even more pronounced, when there is available only a single, possibly tiny, fragment to begin the drug design with, because it means that the researcher has to come up with rest of the molecule on the spot for the organic synthesis.

Fragment- and negative image-based screening. In the novel virtual screening methodology, referred as fragment- and negative image-based (F-NiB; Figs. 1 and S1) screening, the user-selected ligand fragments are fused together with the cavity-based negative image to produce a hybrid F-NIB model, which, in turn, is used directly in the rigid docking of *ab initio* generated ligand 3D conformers. Ideally the chosen fragments, originating for example from X-ray crystallography, improve the cavity's shape/charge characteristics and, thus, facilitate the discovery of potent compounds from the vast virtual screening libraries during the screening phase. On the one hand, the hybrid methodology bypasses some of these practical run-of-the-mill chemistry issues linked to the FBDD, on the other hand, the F-NiB shares some of the limitations common to all virtual screening techniques.

Firstly, because the similarity search is performed using the hybrid NIB model containing both the fragment(s) and the cavity-based negative image, there is no need to come up with explicit solutions for merging the diverse fragments into a single full-sized compound. Secondly, the similarity searches focus on the validated small-molecule libraries containing drug-like compounds and, thus, neither the low-molecular mass fragments nor organic synthesis steps need to be considered. Thirdly, when working with just a single fragment, the F-NiB approach truly excels as even a tiny fragment can be expanded to facilitate full-size compound search. This is achieved by increasing the size of the template fragment by introducing information directly from the cavity itself. Fourthly, the screening is not limited by the original chemical composition of the ligand fragment(s); *i.e.* different bioisosteres are interchangeable and readily swapped during the similarity searches without undue bias given to the original fragment(s).

Due to the sheer speed and the reduced fidelity towards the chosen fragment(s), the F-NiB method differs markedly from the prior, mostly flexible docking-based, *in silico* FBDD protocols (Scott et al., 2012). Although the docking algorithms usually produce correct binding poses for the ligands or fragments, the scoring functions are not always able to rank these poses high enough (Verdonk

et al., 2011). In the F-NiB, the fragments can originate from molecular docking simulations (Model IV; Fig. 2A), experiments such as X-ray crystallography (Models I and III; Fig. 3D and G) or from both sources (Model II; Fig. S4) and they can be taken from protein-bound full-size ligands or low-molecular mass fragments. Thus, if implemented properly, the F-NiB approach could be used to discover truly novel compounds for experimental testing using minimal amount of computing resources. This tendency for scaffold hopping is usually a desired property, but one could limit the ligand set to include only those compounds containing certain desired fragment part(s), if need be.

Furthermore, as is demonstrated in this study, the activity data of previously known ligands can be considered in the virtual screening protocol to utilize rescoring of the top F-NiB screening results by 3D-QSAR modeling (Fig. 2). The implementation of FQSAR in unison with the F-NiB makes it possible to inspect for example specific H-bonding interactions with the target protein *via* the representative fields (Fig S7). Consequently, the fields describing the activity contributions of different physicochemical properties of small-molecules could be used as guidance for generating novel F-NiB models containing fragments that exhibit these properties. However, as is shown by discovery of the active compound **3** (Figs. 2F and 3G; Table 1), the F-NiB approach performs well even if not enough active ligand data is available for generating a reliable QSAR model. Thus, in comparison to for example e-pharmacophore technique (Loving, Salam, & Sherman, 2009; Salam, Nuti, & Sherman, 2009), which relies on a vast amount of docking experiments with known active ligands, the F-NiB can be performed successfully with limited amount of prior ligand data.

The downside of streamlining the fragment-based search with the F-NiB (Figs. 1 and S1) is shared by all structure-based virtual screening methodologies; *i.e.* the readily available molecular databases do not necessarily contain active compounds; the used protein 3D structure could be in a conformation that does not facilitate drug binding; and/or the scoring functions fail to recognize the active compounds. Despite these common concerns, the F-NiB clearly provides a tangible way to accomplish the task of using potent but low-molecular mass fragments in the drug discovery and virtual screening in an innovative manner. Thus, while the F-NiB requires further optimization and automation, the technique shows great promise due to its speed and cost-effective features in comparison to the FBDD relying on flexible docking and extensive organic synthesis programmes.

CONCLUSIONS

A novel virtual screening protocol called fragment- and negative image-based (F-NiB; Fig. 1) screening is introduced and tested experimentally using the phosphodiesterase 10A (PDE10A) as a case study. In the negative image-based (NIB; Fig. 1) screening small-molecules are rigidly docked by focusing on the shape/electrostatics complementarity with the protein's ligand-binding pocket (Niinivehmas et al., 2015, 2011; Virtanen & Pentikäinen, 2010). The F-NiB adds another dimension to the methodology: hybrid F-NiB models that incorporate both the fragments of bound ligands and the protein's cavity information are generated and used in the similarity comparison (Fig. 1). In theory, the hybrid models encompass key features from both the bound ligand(s) and the cavity itself. If activity data is available, it can be included into the F-NiB methodology by

utilizing 3D-QSAR as a rescoring scheme. As a proof of concept, three micromolar-range PDE10A inhibitors were discovered from the SPECS database using both the F-NiB and the combined F-NiB-QSAR screening. In short, the results indicate that the target protein's cavity shape/electrostatics can be used in unison with protein-bound ligand fragments to discover active compounds.

ACKNOWLEDGMENTS

The project was funded by the Jenny and Antti Wihuri Foundation (MA), the Emil Aaltonen Foundation (MA), the Finnish Cultural Foundation, Varsinais-Suomi Regional fund (EMJ; 85182232), and the Academy of Finland (SN; 315492). CSC, The Finnish IT Centre for Science, is acknowledged for generous computational grants (O.T.P. projects jyy2516, jyy2585, and jyy2586).

ABBREVIATIONS USED

Phosphodiesterase 10A (PDE10A), negative image-based (NIB), fragment- and negative image-based (F-NiB), Molecular Mechanics/Generalized Born Surface Area (MM/GBSA), pharmacophore (PHA), Fragment-based drug discovery (FBDD), Quantitative structure-activity relationship (QSAR), Field-based QSAR (FQSAR), Partial Least Squares (PLS), pan-assay interference compounds (PAINS).

DATA AVAILABILITY STATEMENT

The data that support the findings of this study are available from the corresponding author upon reasonable request.

CONFLICT OF INTEREST

The authors declare that they have no conflict of interest.

REFERENCES

Ahinko, M., Kurkinen, S. T., Niinivehmas, S. P., Pentikäinen, O. T., & Postila, P. A. (2019). A Practical Perspective: The Effect of Ligand Conformers on the Negative Image-Based

Screening. *International Journal of Molecular Sciences*, *20*. https://doi.org/10.3390/ijms20112779

Ahinko, M., Niinivehmas, S., Jokinen, E., & Pentikäinen, O. T. (2019). Suitability of MMGBSA for the selection of correct ligand binding modes from docking results. *Chemical Biology & Drug Design*, *93*, 522–538. https://doi.org/10.1111/cbdd.13446

Bento, A. P., Gaulton, A., Hersey, A., Bellis, L. J., Chambers, J., Davies, M., … Overington, J. P. (2014). The ChEMBL bioactivity database: An update. *Nucleic Acids Research*, *42*, 1083–1090. https://doi.org/10.1093/nar/gkt1031

Chappie, T. A., Helal, C. J., & Hou, X. (2012). Current landscape of phosphodiesterase 10A (PDE10A) inhibition. *Journal of Medicinal Chemistry*, *55*, 7299–7331. https://doi.org/10.1021/jm3004976

Cramer, R. D., Patterson, D. E., & Bunce, J. D. (1988). Comparative molecular field analysis (CoMFA). 1. Effect of shape on binding of steroids to carrier proteins. *Journal of the American Chemical Society*, *110*, 5959–5967. https://doi.org/10.1021/ja00226a005

Dixon, S. L., Smondyrev, A. M., Knoll, E. H., Rao, S. N., Shaw, D. E., & Friesner, R. A. (2006). PHASE: A new engine for pharmacophore perception, 3D QSAR model development, and 3D database screening: 1. Methodology and preliminary results. *Journal of Computer-Aided Molecular Design*, *20*, 647–671. https://doi.org/10.1007/s10822-006-9087-6

Duan, J., Dixon, S. L., Lowrie, J. F., & Sherman, W. (2010). Analysis and comparison of 2D fingerprints: Insights into database screening performance using eight fingerprint methods. *Journal of Molecular Graphics and Modelling*, *29*, 157–170. https://doi.org/10.1016/j.jmgm.2010.05.008

Duinen, M., Reneerkens, O., Lambrecht, L., Sambeth, A., Rutten, B., Os, J., … Prickaerts, J. (2015). Treatment of Cognitive Impairment in Schizophrenia: Potential Value of Phosphodiesterase Inhibitors in Prefrontal Dysfunction. *Current Pharmaceutical Design*, *21*, 3813–3828. https://doi.org/10.2174/1381612821666150605110941

Frydenvang, K., Lash, L. L., Naur, P., Postila, P. A., Pickering, D. S., Smith, C. M., … Kastrup, J. S. (2009). Full domain closure of the ligand-binding core of the ionotropic glutamate receptor iGluR5 induced by the high affinity agonist dysiherbaine and the functional antagonist 8,9-dideoxyneodysiherbaine. *Journal of Biological Chemistry*, *284*, 14219–14229. https://doi.org/10.1074/jbc.M808547200

Fujishige, K., Kotera, J., & Omori, K. (1999). Striatum- and testis-specific phosphodiesterase PDE10A isolation and characterization of a rat PDE10A. *European Journal of Biochemistry*, *266*, 1118–1127.

Garcia, A. M., Redondo, M., Martinez, A., & Gil, C. (2014). Phosphodiesterase 10 Inhibitors: New Disease Modifying Drugs for Parkinson's Disease? *Current Medicinal Chemistry*, *21*, 1171–1187. https://doi.org/10.2174/09298673216661312282217 49

Hao, G.-F., Wang, F., Li, H., Zhu, X.-L., Yang, W.-C., Huang, L.-S., … Yang, G.-F. (2012). Computational Discovery of Picomolar Qo Site Inhibitors of Cytochrome bc1 Complex. *Journal of the American Chemical Society*, *134*, 11168–11176. https://doi.org/10.1021/ja3001908

Harder, E., Damm, W., Maple, J., Wu, C., Reboul, M., Xiang, J. Y., … Friesner, R. A. (2016). OPLS3: A Force Field Providing Broad Coverage of Drug-like Small Molecules and Proteins. *Journal of Chemical Theory and Computation*, *12*, 281–296. https://doi.org/10.1021/acs.jctc.5b00864

Humphrey, W., Dalke, A., & Schulten, K. (1996). VMD: visual molecular dynamics. *Journal of*

Molecular Graphics, *14*, 33–38, 27–28. Retrieved from http://www.ncbi.nlm.nih.gov/pubmed/8744570

Jacobson, M. P., Friesner, R. A., Xiang, Z., & Honig, B. (2002). On the role of the crystal environment in determining protein side-chain conformations. *Journal of Molecular Biology*, *320*, 597–608. https://doi.org/10.1016/S0022-2836(02)00470-9

Jhala, D. D., Chettiar, S. S., & Singh, J. K. (2012). Optimization and Validation of an In Vitro Blood Brain Barrier Permeability Assay Using Artificial Lipid Membrane. *Journal of Bioequivalence & Bioavailability*, 10–14. https://doi.org/10.4172/jbb.S14-009

Juvonen, R. O., Ahinko, M., Huuskonen, J., Raunio, H., & Pentikäinen, O. T. (2018). Development of new Coumarin-based profluorescent substrates for human cytochrome P450 enzymes. *Xenobiotica; the Fate of Foreign Compounds in Biological Systems*, 1–10. https://doi.org/10.1080/00498254.2018.1530399

Kinase, A. B., Zhang, B., Li, Y., Zhang, H., & Ai, C. (2010). *3D-QSAR and Molecular Docking Studies on Derivatives of MK-0457 , GSK1070916 and SNS-314 as Inhibitors against*. 4326–4347. https://doi.org/10.3390/ijms11114326

Klebe, G., & Abraham, U. (1999). Comparative Molecular Similarity Index Analysis (CoMSIA) to study hydrogen-bonding properties and to score combinatorial libraries. *Journal of Computer-Aided Molecular Design*, *13*, 1–10. https://doi.org/10.1023/A:1008047919606

Klebe, G., Abraham, U., & Mietzner, T. (1994). Molecular Similarity Indices in a Comparative Analysis (CoMSIA) of Drug Molecules To Correlate and Predict Their Biological Activity. *Journal of Medicinal Chemistry*, *37*, 4130–4146. https://doi.org/10.1021/jm00050a010

Korb, O., Stützle, T., & Exner, T. E. (2009). Empirical scoring functions for advanced Protein-Ligand docking with PLANTS. *Journal of Chemical Information and Modeling*, *49*, 84–96. https://doi.org/10.1021/ci800298z

Kraulis, P. J. (1991). MOLSCRIPT: a program to produce both detailed and schematic plots of protein structures. *Journal of Applied Crystallography*, *24*, 946–950. https://doi.org/10.1107/S0021889891004399

Kurkinen, S. T., Niinivehmas, S., Ahinko, M., Lätti, S., Pentikäinen, O. T., & Postila, P. A. (2018). Improving docking performance using negative image-based rescoring. *Frontiers in Pharmacology*, *9*. https://doi.org/10.3389/fphar.2018.00260

Lash, L. L., Sanders, J. M., Akiyama, N., Shoji, M., Postila, P., Pentikäinen, O. T., … Swanson, G. T. (2008). Novel analogs and stereoisomers of the marine toxin neodysiherbaine with specificity for kainate receptors. *The Journal of Pharmacology and Experimental Therapeutics*, *324*, 484–496. https://doi.org/10.1124/jpet.107.129890

Lätti, S., Niinivehmas, S., & Pentikäinen, O. T. (2016). Rocker: Open source, easy-to-use tool for AUC and enrichment calculations and ROC visualization. *Journal of Cheminformatics*, *8*, 1–5. https://doi.org/10.1186/s13321-016-0158-y

Leach, A. R., Gillet, V. J., Lewis, R. A., & Taylor, R. (2010). Three-dimensional pharmacophore methods in drug discovery. *Journal of Medicinal Chemistry*, *53*, 539–558. https://doi.org/10.1021/jm900817u

Lehtonen, J. V, Still, D.-J., Rantanen, V.-V., Ekholm, J., Björklund, D., Iftikhar, Z., … Johnson, M. S. (2004). BODIL: a molecular modeling environment for structure-function analysis and drug design. *Journal of Computer-Aided Molecular Design*, *18*, 401–419. Retrieved from http://www.ncbi.nlm.nih.gov/pubmed/15663001

Li, J., Abel, R., Zhu, K., Cao, Y., Zhao, S., & Friesner, R. A. (2011). The VSGB 2.0 model: A next generation energy model for high resolution protein structure modeling. *Proteins:*

Structure, Function and Bioinformatics, *79*, 2794–2812. https://doi.org/10.1002/prot.23106
Loving, K., Salam, N. K., & Sherman, W. (2009). Energetic analysis of fragment docking and application to structure-based pharmacophore hypothesis generation. *Journal of Computer-Aided Molecular Design*, *23*, 541–554. https://doi.org/10.1007/s10822-009-9268-1
Lowe, E. W., Ferrebee, A., Rodriguez, A. L., Conn, P. J., & Meiler, J. (2010). 3D-QSAR CoMFA study of benzoxazepine derivatives as mGluR5 positive allosteric modulators. *Bioorganic & Medicinal Chemistry Letters*, *20*, 5922–5924. https://doi.org/10.1016/j.bmcl.2010.07.061
Malamas, M., Ni, Y., & Erdei, J. (2011). Highly potent, selective, and orally active phosphodiesterase 10A inhibitors. *Journal of Medicinal Chemistry*, *54*, 7621–7638.
Merritt, E. A., & Murphy, M. E. P. (1994). Raster3D Version 2.0. A program for photorealistic molecular graphics. *Acta Crystallographica Section D Biological Crystallography*, *50*, 869–873. https://doi.org/10.1107/S0907444994006396
Nestler, R. S., & Duman, E. J. (1999). Functional Roles for cAMP and cGMP. In M. D. U. George J. Siegel, Bernard W. Agranoff, R. Wayne Albers, Stephen K. Fisher (Ed.), *Basic Neurochemistry: Molecular, Cellular and Medical Aspects.* (6th ed.). Retrieved from https://www.ncbi.nlm.nih.gov/books/NBK27915/
Niinivehmas, S. P., Manivannan, E., Rauhamäki, S., Huuskonen, J., & Pentikäinen, O. T. (2016). Identification of estrogen receptor a ligands with virtual screening techniques. *Journal of Molecular Graphics and Modelling*, *64*, 30–39. https://doi.org/10.1016/j.jmgm.2015.12.006
Niinivehmas, S. P., Salokas, K., Lätti, S., Raunio, H., & Pentikäinen, O. T. (2015). Ultrafast protein structure-based virtual screening with Panther. *Journal of Computer-Aided Molecular Design*, *29*, 989–1006. https://doi.org/10.1007/s10822-015-9870-3
Niinivehmas, S. P., Virtanen, S. I., Lehtonen, J. V., Postila, P. A., & Pentikäinen, O. T. (2011). Comparison of virtual high-throughput screening methods for the identification of phosphodiesterase-5 inhibitors. *Journal of Chemical Information and Modeling*, *51*, 1353–1363. https://doi.org/10.1021/ci1004527
Nurminen, E. M., Pihlavisto, M., Lázár, L., Szakonyi, Z., Pentikäinen, U., Fülöp, F., & Pentikäinen, O. T. (2010). Synthesis, in vitro activity, and three-dimensional quantitative structure-activity relationship of novel hydrazine inhibitors of human vascular adhesion protein-1. *J Med Chem*, *53*, 6301–6315. Retrieved from http://view.ncbi.nlm.nih.gov/pubmed/20690686
Patel, H. M., Noolvi, M. N., & Sharma, P. (2014). *CHEMISTRY Quantitative structure – activity relationship (QSAR) studies as strategic approach in drug discovery*. 4991–5007. https://doi.org/10.1007/s00044-014-1072-3
Postila, P. A., Swanson, G. T., & Pentikäinen, O. T. (2010). Exploring kainate receptor pharmacology using molecular dynamics simulations. *Neuropharmacology*, *58*, 515–527. https://doi.org/10.1016/j.neuropharm.2009.08.019
Rauhamäki, S., Postila, P. A., Lätti, S., Niinivehmas, S., Multamäki, E., Liedl, K. R., & Pentikäinen, O. T. (2018). Discovery of Retinoic Acid-Related Orphan Receptor γt Inverse Agonists via Docking and Negative Image-Based Screening. *ACS Omega*, *3*, 6259–6266. https://doi.org/10.1021/acsomega.8b00603
Rzasa, R. M., Frohn, M. J., Andrews, K. L., Chmait, S., Chen, N., Clarine, J. G., … Allen, J. R. (2014). Synthesis and preliminary biological evaluation of potent and selective 2-(3-alkoxy-1-azetidinyl) quinolines as novel PDE10A inhibitors with improved solubility. *Bioorganic and Medicinal Chemistry*, *22*, 6570–6585. https://doi.org/10.1016/j.bmc.2014.10.013

Rzasa, R. M., Hu, E., Rumfelt, S., Chen, N., Andrews, K. L., Chmait, S., … Allen, J. R. (2012). Discovery of selective biaryl ethers as PDE10A inhibitors: Improvement in potency and mitigation of Pgp-mediated efflux. *Bioorganic and Medicinal Chemistry Letters*, *22*, 7371–7375. https://doi.org/10.1016/j.bmcl.2012.10.078

Salam, N. K., Nuti, R., & Sherman, W. (2009). Novel method for generating structure-based pharmacophores using energetic analysis. *Journal of Chemical Information and Modeling*, *49*, 2356–2368. https://doi.org/10.1021/ci900212v

Sastry, M., Lowrie, J. F., Dixon, S. L., & Sherman, W. (2010). Large-scale systematic analysis of 2D fingerprint methods and parameters to improve virtual screening enrichments. *Journal of Chemical Information and Modeling*, *50*, 771–784. https://doi.org/10.1021/ci100062n

Scott, D. E., Coyne, A. G., Hudson, S. A., & Abell, C. (2012). Fragment based approaches in drug discovery and chemical biology. *Biochemistry*, *51*, 4990–5003. https://doi.org/10.1021/bi3005126

Shelley, J. C., Cholleti, A., Frye, L. L., Greenwood, J. R., Timlin, M. R., & Uchimaya, M. (2007). Epik: a software program for pK(a) prediction and protonation state generation for drug-like molecules. *Journal of Computer-Aided Molecular Design*, *21*, 681–691. https://doi.org/10.1007/s10822-007-9133-z

Shubina, V., Niinivehmas, S., & Pentikäinen, O. T. (2015). Reliability of virtual screening methods in prediction of PDE4B-inhibitor activity. *Current Drug Discovery Technologies*, *12*, 117–126. Retrieved from http://www.scopus.com/inward/record.url?eid=2-s2.0-84939789964&partnerID=tZOtx3y1

Thompson, W. J., & Appleman, M. M. (1971). Multiple cyclic nucleotide phosphodiesterase activities from rat brain. *Biochemistry*, *10*, 311–316. https://doi.org/10.1021/bi00778a018

Tian, Y., Xu, J., Li, Z., Zhu, Z., Zhang, J., & Wu, S. (2011). *Combined 3D-QSAR and Docking Modelling Study on Indolocarbazole Series Compounds as Tie-2 Inhibitors*. 5080–5097. https://doi.org/10.3390/ijms12085080

Vainio, M. J., & Johnson, M. S. (2007). Generating conformer ensembles using a multiobjective genetic algorithm. *Journal of Chemical Information and Modeling*, *47*, 2462–2474. https://doi.org/10.1021/ci6005646

Vainio, M. J., Puranen, J. S., & Johnson, M. S. (2009). ShaEP: Molecular overlay based on shape and electrostatic potential. *Journal of Chemical Information and Modeling*, *49*, 492–502. https://doi.org/10.1021/ci800315d

Verdonk, M. L., Giangreco, I., Hall, R. J., Korb, O., Mortenson, P. N., & Murray, C. W. (2011). Docking Performance of Fragments and Drug like Compounds. *Journal of Medicinal Chemistry*, *54*, 5422–5431. https://doi.org/Doi 10.1021/Jm200558u

Verhoest, P. R., Chapin, D. S., Corman, M., Fonseca, K., Harms, J. F., Hou, X., … Liras, S. (2009). Discovery of a novel class of phosphodiesterase 10A inhibitors and identification of clinical candidate 2-[4-(1-methyl-4-pyridin-4-yl-1H-pyrazol-3- yl)-phenoxymethyl]-quinoline (PF-2545920) for the treatment of schizophrenia. *Journal of Medicinal Chemistry*, *52*, 5188–5196. https://doi.org/10.1021/jm900521k

Virtanen, S. I., & Pentikäinen, O. T. (2010). Efficient virtual screening using multiple protein conformations described as negative images of the ligand-binding site. *Journal of Chemical Information and Modeling*, *50*, 1005–1011. https://doi.org/10.1021/ci100121c

Wang, H., Liu, Y., Hou, J., Zheng, M., Robinson, H., & Ke, H. (2007). Structural insight into substrate specificity of phosphodiesterase 10. *Proceedings of the National Academy of*

Sciences of the United States of America, *104*, 5782–5787. https://doi.org/10.1073/pnas.0700279104

Watts, K. S., Dalal, P., Murphy, R. B., Sherman, W., Friesner, R. A., & Shelley, J. C. (2010). ConfGen: A conformational search method for efficient generation of bioactive conformers. *Journal of Chemical Information and Modeling*, *50*, 534–546. https://doi.org/10.1021/ci100015j

Word, J. M., Lovell, S. C., Richardson, J. S., & Richardson, D. C. (1999). Asparagine and Glutamine: Using Hydrogen Atom Contacts in the Choice of Side-chain Amide Orientation. *J . Mol . Biol*, *285*, 1735–1747. https://doi.org/10.1006/jmbi.1998.2401

Xiong, L., Li, H., Jiang, L.-N., Ge, J.-M., Yang, W.-C., Zhu, X. L., & Yang, G.-F. (2017). Structure-Based Discovery of Potential Fungicides as Succinate Ubiquinone Oxidoreductase Inhibitors. *Journal of Agricultural and Food Chemistry*, *65*, 1021–1029. https://doi.org/10.1021/acs.jafc.6b05134

Xiong, L., Zhu, X.-L., Gao, H.-W., Fu, Y., Hu, S.-Q., Jiang, L.-N., … Yang, G.-F. (2016). Discovery of Potent Succinate-Ubiquinone Oxidoreductase Inhibitors via Pharmacophore-linked Fragment Virtual Screening Approach. *Journal of Agricultural and Food Chemistry*, *64*, 4830–4837. https://doi.org/10.1021/acs.jafc.6b00325

Yadav, D. K., Meena, A., Srivastava, A., Chanda, D., Khan, F., & Chattopadhyay, S. K. (2010). Development of QSAR model for immunomodulatory activity of natural coumarinolignoids. *Drug Design, Development and Therapy*, *4*, 173–186. Retrieved from http://www.ncbi.nlm.nih.gov/pubmed/20856844

Table 1. Virtual screening and experimental activity testing results.

Compound	Compound (SPECS name)	ShaEP	MM/GBSA (Kcal/mol)	IC_{50} (μM)	pIC_{50}	pIC_{50} Pred[1]	QPlogBB[2]	LogS[3]	Tc_{Max}[6]	Model[7]	Method[8]
1	AT-057-43485961	0.522	-62.948	26.8	4.572	9.219	0.527	-5.82	0.099	IV	F-NiB/FQSAR
2	AM-879-42012742	0.553	-72.828	48.8	4.312	9.382	0.052	-9.65	0.063	I	F-NiB/FQSAR
3	AG-690-09287047	0.585	-58.432	67.0	4.174	-	1.03	-6.77	0.102	III	F-NiB
4	AG-205-09068036	0.582	-47.282	>100	<4	-	-0.014	-2.78	0.146	III	F-NiB
5	AG-690-36108027	0.594	-52.288	>100	<4	-	0.559	-6.73	0.163	III	F-NiB
6	AK-968-15360495	0.584	-46.702	>100	<4	-	0.669	-4.3	0.139	III	F-NiB
7	AN-329-41642008	0.650	-83.451	>100	<4	-	0.916	-6.24	0.099	II	F-NiB
8	AN-329-43449158	0.590	-50.464	>100	<4	-	0.139	-3.5	0.064	III	F-NiB
9	AN-465-14013021	0.594	-49.440	>100	<4	-	-0.064	-8.86	0.11	III	F-NiB
10	AP-853-42879195	0.607	-53.995	>100	<4	-	0.333	-4.06	0.085	III	F-NiB
11	AP-970-41728638	0.586	-48.305	>100	<4	-	0.687	-5.67	0.08	III	F-NiB
12	AO-022-43390442	N/A	-74.222	NI	NI	10.112	0.478	-6.72	0.086	PHA	PHA/FQSAR
13	AP-853-42160322	0.600	-49.780	NI	NI	-	0.397	-4.95	0.158	III	F-NiB
14	AO-022-43453889	0.550	-85.774	N/A[4]	N/A[4]	7.583	0.197	-6.87	0.092	I	F-NiB/FQSAR
Papaverine	- (control)	N/A	N/A	147 nM	6.833	-	-	-4.4[5]	N/A	N/A	-

NI = no inhibition; N/A = not available, PHA = pharmacophore

[1] pIC_{50} value predicted by an FQSAR model.

[2] QPlogBB (the brain/blood permeability) from QIKPROP module calculations in MAESTRO.

[3] LogS (the aqueous solubility) from the SPECS database (www.specs.net).

[4] Not soluble.

[5] LogS from ALOGPS (http://www.vcclab.org/lab/alogps/)

[6] Maximum value of Tanimoto coefficient when compared to all molecules obtained from the ChEMBL database.

[7] The hybrid F-NiB models I-IV, the cavity-based negative images with incorporated ligand fragment(s), are shown in Figs. 3 and S4. The PHA model is shown in Fig. S3A.

[8] F-NiB: fragment- and negative image-based (F-NiB) screening or rigid docking; F-NiB/FQSAR: F-NiB docking in combination with the field-based QSAR; PHA/FQSAR: pharmacophore-based ligand alignment combined with the FQSAR.

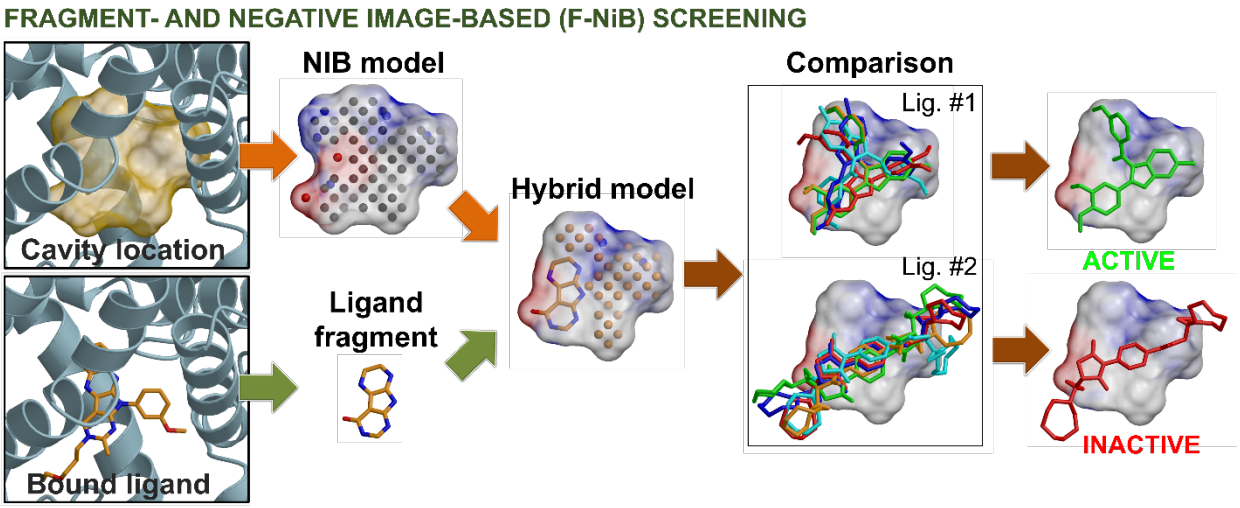

Figure 1. Utilizing ligand fragments and cavity information in the virtual screening. The fragment- and negative image-based (F-NiB) screening protocol follows four successive steps: (1) the negative image or negative image-based (NIB) model generation. The PANTHER-generated (Niinivehmas et al., 2015) NIB model (transparent surface) is composed of neutral filler atoms (black sphere) and charged cavity points (blue/red spheres). The positively (blue sphere) or negatively (red sphere) charged cavity points mirror the H-bond donors and acceptors lining the ligand-binding cavity; (2) the selection of protein-bound ligand fragments from validated 3D structures (stick model with yellow backbone) or *in silico* predictions; (3) the merging of the fragments and cavity points and removal of the overlapping points to generate a hybrid F-NiB model; and (4), finally, the similarity comparison screening or rigid docking with ShaEP (Vainio et al., 2009) (brown arrows). The shape/charge comparison against the F-NiB model is shown against multiple 3D conformers of two ligands (Lig #1 and Lig #2; stick models of different colors). Those conformers (and ligands) matching best the shape/electrostatic properties of the template F-NiB model in the rigid docking are predicted active (green stick model) and, in contrast, the weaker matching ligands are assumed inactive (red stick model). Finally, the docking poses of the top-ranked compounds are visually inspected with the protein and compounds are selected for *in vitro* testing.

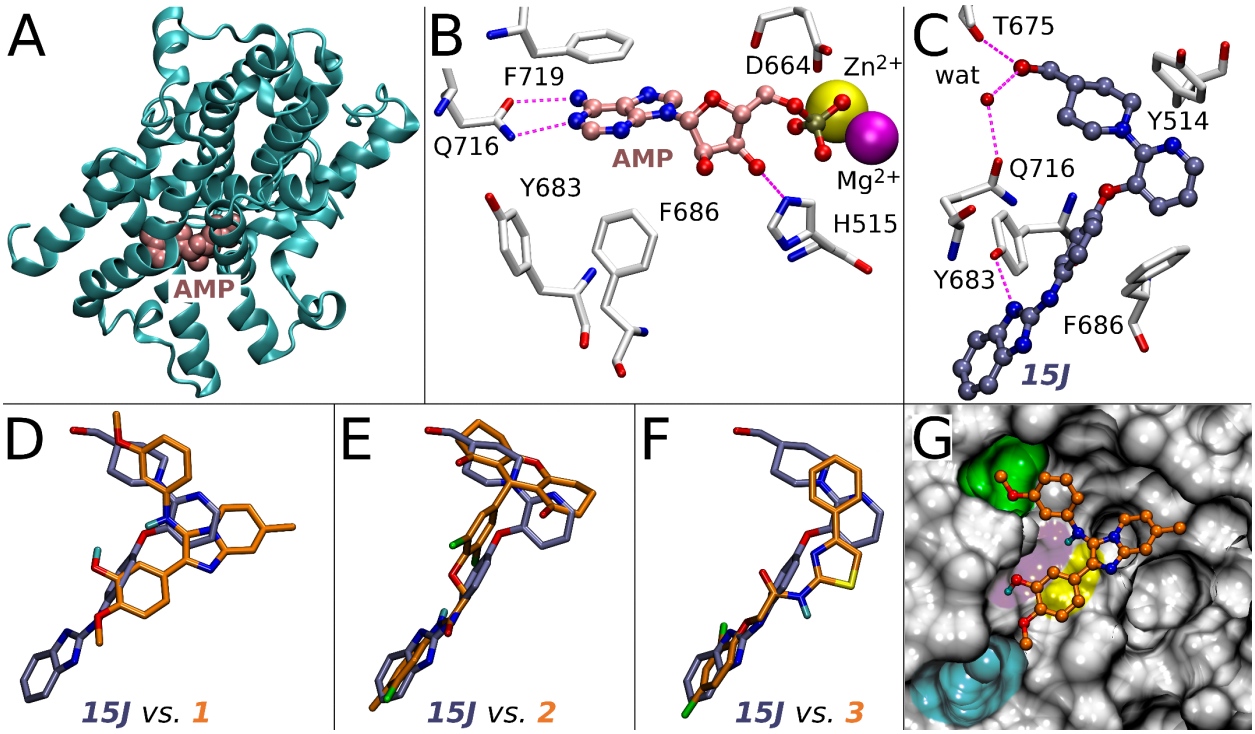

Figure 2. Phosphodiesterase 10A and its ligands. (A) The 3D structure of the PDE10A (cyan cartoon; PDB: 2OUN; chain A) (Wang et al., 2007) shown with adenosine monophosphate (AMP; pink CPK model). (B) The AMP (ball-and-stick model with pink backbone) binding is shown with the key residues (stick models with white backbone) such as Gln716 and His515 forming H-bonds (magenta dotted lines; PDB: 4HEU; A chain) (Rzasa et al., 2012) and participating with the divalent ions in the hydrolysis of cyclic AMP (or cAMP) or cyclic GMP (cGMP; nucleoside 3',5'-cyclic phosphate + H_2O = nucleoside 5'-phosphate.). (C) The binding of inhibitors such as the biaryl ether **15J** (ball-and-stick model with magenta backbone) does not involve the divalent ions. When comparing the validated binding of an inhibitor to the predicted poses for (D) **1**, (E) **2** and (F) **3**, the new molecules occupy roughly the same 3D space in the cavity. The binding poses of **2** and **3** are alike with **15J**, whereas **1** differs somewhat. (G) The new inhibitor binding centers on the four regions shown in the cross section of the cavity surface: "*hydrophobic clamp*" (yellow); "*Gln interaction*" (pink); "*buried waters*" (green); and "*selectivity pocket*" (cyan) (Chappie et al., 2012). Notably, **1** (ball-and stick model with orange backbone) does not extend to the "*selectivity pocket*" region. Chlorine, oxygen, and nitrogen atoms in the compounds are shown with green, red and blue color, respectively.

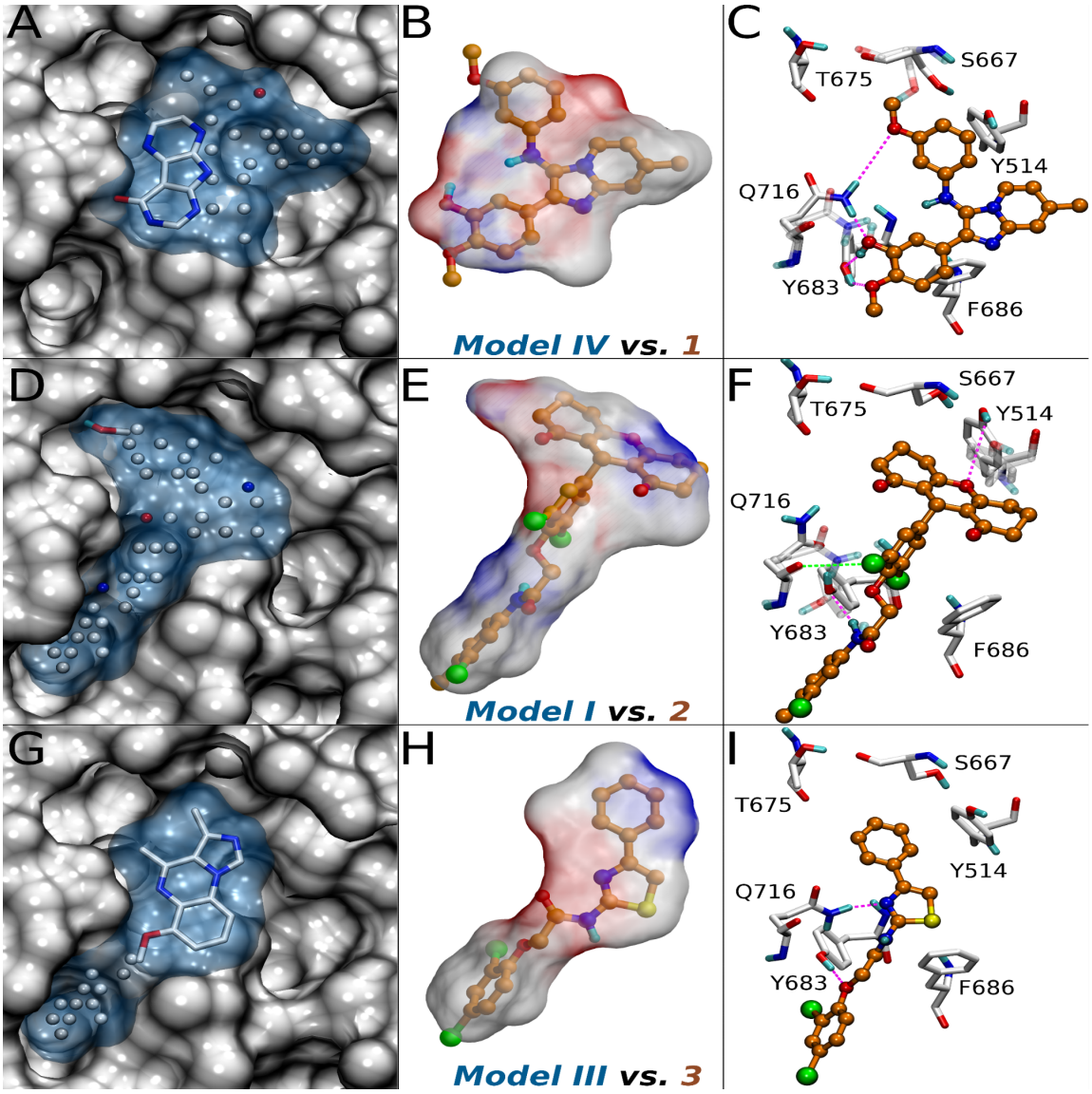

Figure 3. Fragment- and negative image-based hybrid models and their representative hits. Models (A) IV, (D) I, and (G) III (blue surface), composed of ligand fragments (sticks) and cavity points or filler atoms (balls) generated with PANTHER (Niinivehmas et al., 2015), occupy key regions of the PDE10A cavity (grey surface; see Fig. 2G). Compounds (B) **1**, (E) **2** and (H) **3** (ball-and-stick models with orange backbone) produced high shape/charge similarity scores (Table 1), when the hybrid models (transparent surface with charge potential) were compared against the SPECS compounds using ShaEP (Vainio et al., 2009). The binding modes of compounds (C) **1**, (F) **2**, and (I) **3** are shown with the key residues (stick models with white backbone) focusing on the hydrogen (magenta dotted lines) and halogen (green dotted lines) bonding. Note that the direct H-bond between Tyr514 and **2** could also be substituted by a water bridge (not shown). Those residues that were notably adjusted by the MM/GBSA routine are shown with both the original (transparent sticks; PDB: 3SN7 in A, C and 4HEU in D, F, G and I) (Malamas et al., 2011; Rzasa et al., 2012) and optimized poses; otherwise only original side chain conformations are shown (panel I).