JCIM JOURNAL OF CHEMICAL INFORMATION AND MODELING

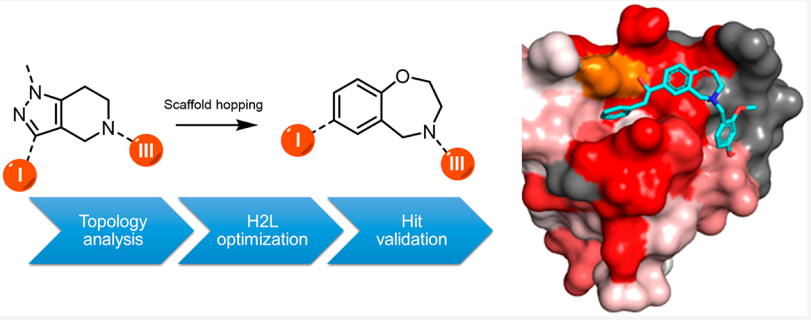
pubs.acs.org/jcim Article

Computer-Aided Design and Synthesis of a New Class of PEX14 Inhibitors: Substituted 2,3,4,5-Tetrahydrobenzo[F][1,4]oxazepines as Potential New Trypanocidal Agents

Roberto Fino,[○] Dominik Lenhart,[○] Vishal C. Kalel, Charlotte A. Softley, Valeria Napolitano, Ryan Byrne, Wolfgang Schliebs, Maciej Dawidowski, Ralf Erdmann, Michael Sattler, Gisbert Schneider, Oliver Plettenburg, and Grzegorz M. Popowicz*

Cite This: *J. Chem. Inf. Model.* 2021, 61, 5256−5268 Read Online

ACCESS | Metrics & More | Article Recommendations | Supporting Information



ABSTRACT: African and American trypanosomiases are estimated to affect several million people across the world, with effective treatments distinctly lacking. New, ideally oral, treatments with higher efficacy against these diseases are desperately needed. Peroxisomal import matrix (PEX) proteins represent a very interesting target for structure- and ligand-based drug design. The PEX5−PEX14 protein−protein interface in particular has been highlighted as a target, with inhibitors shown to disrupt essential cell processes in trypanosomes, leading to cell death. In this work, we present a drug development campaign that utilizes the synergy between structural biology, computer-aided drug design, and medicinal chemistry in the quest to discover and develop new potential compounds to treat trypanosomiasis by targeting the PEX14−PEX5 interaction. Using the structure of the known lead compounds discovered by Dawidowski et al. as the template for a chemically advanced template search (CATS) algorithm, we performed scaffold-hopping to obtain a new class of compounds with trypanocidal activity, based on 2,3,4,5-tetrahydrobenzo[f][1,4]oxazepines chemistry. The initial compounds obtained were taken forward to a first round of hit-to-lead optimization by synthesis of derivatives, which show activities in the range of low- to high-digit micromolar IC_{50} in the *in vitro* tests. The NMR measurements confirm binding to PEX14 in solution, while immunofluorescent microscopy indicates disruption of protein import into the glycosomes, indicating that the PEX14−PEX5 protein−protein interface was successfully disrupted. These studies result in development of a novel scaffold for future lead optimization, while ADME testing gives an indication of further areas of improvement in the path from lead molecules toward a new drug active against trypanosomes.

INTRODUCTION

As a consequence of the progressive raising of the temperatures on a global scale, and increased rate of migration from developing countries to Europe, increasingly more cases of African and American trypanosomiases (among other tropical diseases) are being reported in Central and Northern Europe.[1,2] Due to the difficulty of administration, the toxicity, and the possibility of resistance emerging to the available treatments (Figure 1),[3−5] new, more effective, and less harmful antiprotozoal drugs are therefore urgently needed.

For this purpose, the recently demonstrated druggability of the peroxin proteins (PEX), in particular, the PEX5−PEX14 protein−protein interface marks an excellent starting point for early stage drug development.[6] In eukaryotic cells, PEX proteins are actively involved in peroxisomal protein import, allowing cells to translocate enzymes synthesized in the cytosol into the peroxisomes. There are two principal pathways in which the enzymes are transported into the organelles—both require that the cargo enzymes contain a peroxisomal targeting

Received: April 27, 2021
Published: October 1, 2021

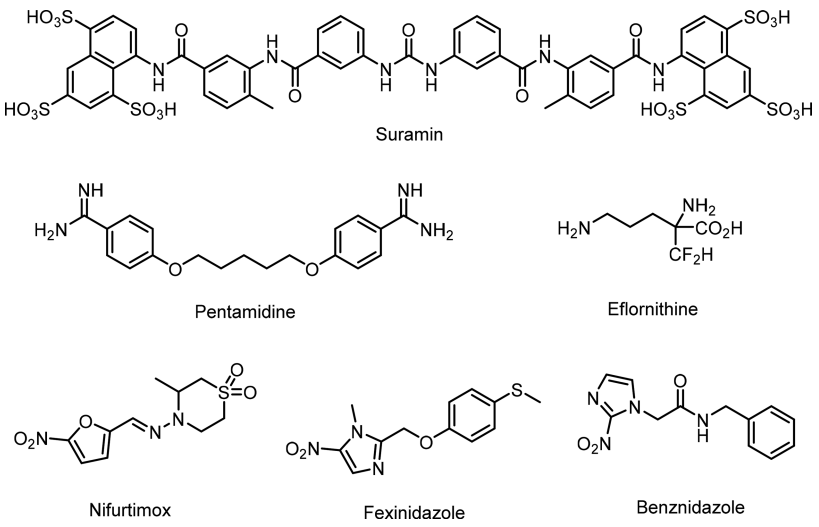


Figure 1. A summary of the available molecules used to treat trypanosomiases. *Trypanosoma brucei* and *T. cruzi* are the causative agents of African and American trypanosomiasis, respectively. Suramin is the most common treatment for the first stages of *T. brucei* infections and is administered intravenously. Common side effects are nausea, diarrhea, and abdominal pain. Pentamidine is administered to treat *T. b. gambiense* infections; Nifurtimox is often administered with Eflornithine to treat African trypanosomiasis. Benznidazole is effective against *T. cruzi* acute infections and is a mainline treatment for Chagas disease (American trypanosomiasis). There is no effective treatment for the chronic Chagas stage available. Fexinidazole is the most recent discovery in the development of treatments and one of the most effective drugs available on the market to treat African trypanosomiasis.

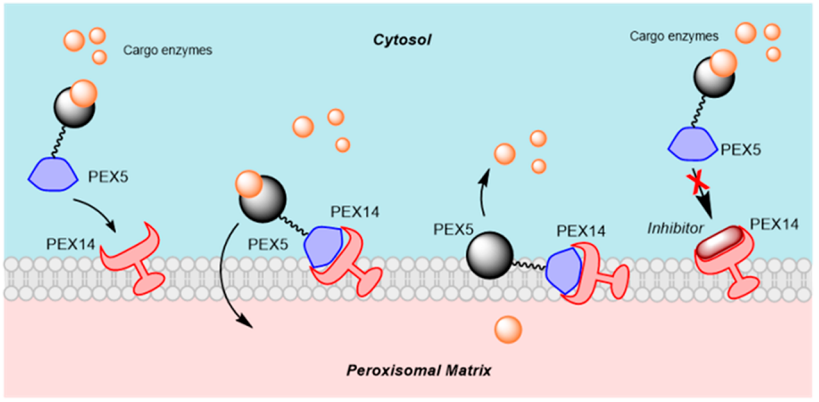


Figure 2. A simplified overview of the importomer system based on the PTS1−PEX5 interaction pathway. The cargo enzymes are recognized by the TPR domain of PEX5 and then translocated into the peroxisomal matrix via the interaction between PEX5 and PEX14, which forms a transient import pore. Inhibition of PEX5−PEX14 binding prevents import pore formation, and thus the cargo enzymes are mislocalized.

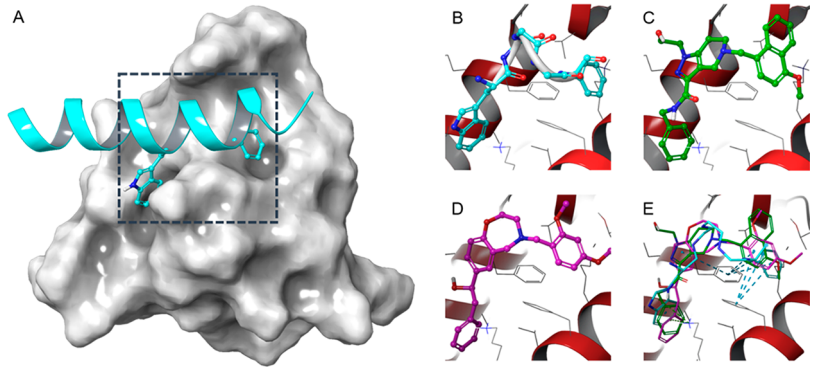


Figure 3. Overview of the PEX5−PEX14 protein−protein interface. (A) NMR structure (PDB: 2W84) of PEX5 (cyan) bound to the N-terminal domain of PEX14 (gray). Tryptophan and a phenylalanine side chain act as the main anchors, and they are separated by a distance of about 5 Å in the diaromatic pentapeptide WxxxF motif in PEX5 (panel B) allowing a T-shaped $\pi-\pi$ stacking between these and the two bridging phenylalanines (Phe17 and Phe34) in the binding site in PEX14. The main interactions maintained by the currently known pyrazolo[4,3-*c*]pyridine-based inhibitors (PDB: 5L87) are depicted in panel C. Panel D shows the predicted docking pose of **7a**, the best CATS hit. The overlay of the three structures (E) shows how the $\pi-\pi$ stacking is maintained both for **7a** and for the pyrazolo[4,3-*c*]pyridine derivative.

https://doi.org/10.1021/acs.jcim.1c00472
J. Chem. Inf. Model. 2021, 61, 5256−5268

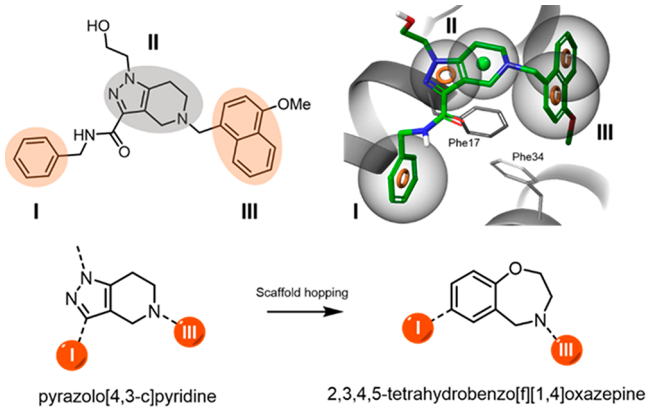


Figure 4. Schematic of a typical pharmacophore model, using a pyrazolo[4,3-*c*]pyridine derivative as the starting compound. The two hydrophobic moieties (I, III) are connected by a 4,5,6,7-tetrahydro-pyrazolo[4,3-*c*]pyridine central core (II). The π-stacking interactions between the hydrophobic cores of the inhibitor and the two bridging phenylalanines (Phe17 and Phe34) are the main contribution to ligand binding. In the CATS compound scaffold, the pharmacophore model requirements are satisfied by connecting the two hydrophobic cores via a 2,3,4,5-tetrahydrobenzo[f][1,4]oxazepine scaffold.

Table 1. Aldehydes 8a−8d Synthesized from Esters 6a−6d

entry	Ar	yield 4 [%][a]	yield 5 [%]	yield 6 [%]
a	(2,4-OMe)Ph	90	74	89
b	(4-OCF_2H)Ph	73	67	66
c	(2-F-5-OMe)Ph	62	52	65
d	(4-OMe)-1-Naphtyl	45	70	58

[a]Yield over two steps.

Table 2. Grignard Addition to Aldehydes 6a−6d Yielding CATS Compounds

entry	Ar	R	yield [%]
A	(2,4-OMe)Ph	Ph	63
B	(2,4-OMe)Ph	3-Cl-Ph	40
C	(2,4-OMe)Ph	4-F-Ph	61
D	4-(OCF_2H)Ph	Ph	39
E	(2-F,5-OMe)Ph	Ph	55
F	(4-OMe)Naphthyl	Ph	83

signal (PTS) sequence, either of type 1 or 2. These are recognized by the specialized peroxisomal import receptors that bind the cargo proteins in the cytosol and direct them to a docking complex at the peroxisomal membrane. The C-terminal PTS1 signal of the cargo enzyme is recognized by the tetratricopeptide repeat (TPR) domain of the import receptor PEX5, which binds to the membrane-associated protein PEX14, which acts as the primary docking protein (Figure 2). The cargo enzyme is then imported through a transient import pore, which is formed by PEX5 and PEX14 (Figure 2)[7−9]. In this scenario, PEX5 interacts with the N-terminal helical domain of PEX14. Blocking the PEX5−PEX14 protein−protein interaction (PPI) prevents cargo protein import. Glycosomes are peroxisome-related organelles unique to the trypanosomatid parasites and are essential for their survival. While sequences are similar, significant structural differences between *Trypanosoma* and human PEX proteins allow development of selective inhibitors that specifically block glycosomal protein import in parasites.

The PEX5−PEX14 binding interface features two main hydrophobic anchors (a tryptophan and a phenylalanine) that penetrate into two hydrophobic pockets on the PEX14. The two anchors, separated by a distance of 5 Å, are connected via three residues in the backbone of the PEX5 helix according to the sequence pattern Wxxx(F/Y).[10] The first identified scaffold of the *Trypanosoma* PEX5−PEX14 interface inhibitors[6] connects two hydrophobic moieties via a pyrazolo[4,3c]-pyridine, a 5 + 6 term condensed heterocycle, as shown in Figure 3.

Availability of the structural information from a previous work on the development of a first line of PEX14−PEX5 PPI inhibitors, computer-aided drug design (CADD) and ligand- and structure-based drug discovery (SBDD) appear to be valuable assets for the task of designing novel molecules as trypanocidal drug candidates.

Despite the considerable advances in high-throughput virtual screening (HTVS) techniques,[11] the identification of novel chemotypes with specific efficacy against a target of interest remains a challenging task. Analyses of compound libraries from major pharmaceutical companies[12] have supported the principle that the relationship between chemical and biochemical spaces remains poorly understood, with diverse chemical compounds producing an effect on a given target, and, equally, that a given compound can have an effect on targets with diverse functions and structures.

It is widely accepted that compounds interacting with a selected biological target through a specific binding site must have some fundamental similarity. The definition of a metric for parametrizing protein−ligand binding dynamics, however, is context-specific.[13] One approach for identification of novel bioactive compounds against a defined protein target is to use pharmacophoric similarity. This method requires first the definition of the submolecular pharmacophoric features, which are then used to describe a given set of molecules. We conducted such comparisons using the *Chemically-Advanced Template Search*[14] (CATS) software, which has proven successful in some scaffold-hopping efforts.[15] This approach constructs histograms of the distributions of pairwise topological (through-bond) distances between pairs of pharmacophoric features, which can then be compared to Euclidean distance measurements between vectors.

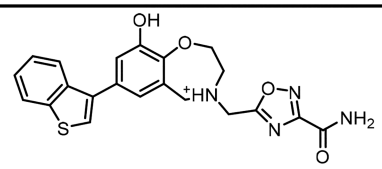
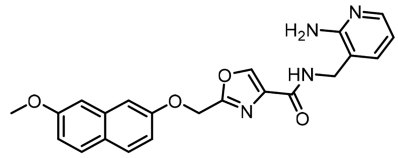
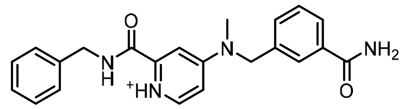
We used the CATS algorithm, starting from the pyrazolo-[4,3-*c*]pyridine derivatives as the template of the pharmacophore model, to find new active compounds for further testing.

This led to identification of compound series named CATS. This new set of ligands fulfills the spatial constraints of the PEX14 binding site, with two hydrophobic anchors connected by a 6 + 7 term heterocycle maintaining the same distribution of the pharmacophoric features as for pyrazolo[4,3-*c*]pyridine derivatives (Figure 4).

After selection of the best candidates using similarity scores,[16] we performed docking and short molecular dynamics simulations to identify the compounds that fit the binding pocket. Then, we selected the screened compounds taking into account the contacts with the key residues in the binding site of PEX14, the conformational constraints of the generated poses, and the estimated binding energy. This led us to a library of 17 compounds that we next synthesized or purchased from commercial vendors and then tested *in vitro*. Finally, we selected the most active hits from the *in vitro* assays and performed the first round of medicinal chemistry optimization (yields for the aldehydes **8a−8d** synthesized from esters **6a−**

https://doi.org/10.1021/acs.jcim.1c00472
J. Chem. Inf. Model. 2021, 61, 5256−5268

Table 3. Top CATS Scaffolds Selected after Virtual Screening

Compound	ID	Vendor	Structure
7t	27888553	ChemBridge	
13	35061195	ChemBridge	
14	Z1185855320	Enamine	

6d are reported in Table 1; in Table 2 are reported the yields for the syntheses of compounds **7a**–**7f**).

EXPERIMENTAL SECTION

CATS Algorithm-Based HTVS. The CATS algorithm assigns atom types to five pharmacophoric categories, a hydrogen-bond donor (D) or acceptor (A), positively (P) or negatively charged (N), or lipophilic (L). Taking the categories pairwise, the distances between all atoms with the chosen characteristics and a topological distance of less than or equal to 10 bonds are noted for the query molecule, as well as all molecules in the library. This results in a 150-dimensional vector per compound: 15 combinations of two categories, each considering pairs within the distance limit. This is then normalized by the number of heavy atoms in the molecule, to give the final, scaled vector for each molecule. Comparison between vectors, to find molecules with a similar pharmacophoric profile, is then accomplished simply by means of Euclidean distance comparison, as per eq 1.

CATS Euclidean distance calculation: The distance D between two compounds, represented by their CATS vectors v^A and v^B, is defined as the Euclidean distance between them.

$$D(\mathrm{A},\mathrm{B}) = \sqrt{\sum_{i=1}^{150} (v_i^{\mathrm{A}} + v_i^{\mathrm{B}})^2} \quad (1)$$

We trained the CATS algorithm with the structure of the most active compound in respect to its trypanocidal activity of the SAR library of PEX14 inhibitors.[17] The output consisted of a library of roughly 10 000 molecules structurally similar to the template structure, selected by the algorithm from an *in silico* library of approximately 14.5 million purchasable molecules.

In order to shortlist the generated library to the best candidates, a KNIME[18] workflow was deployed to filter out compounds according to Tanimoto similarity scores (the higher the score, the more similar the compounds according to the measured vectors).[19]

The structure of the pyrazolo[4,3-*c*]pyridine compound (compound **12**) was used as the reference to bias from using the CATS 2D fingerprint, previously converted to a plain SMILES string converted using Schrödinger Maestro 2017.1.[20]

In this context, to maintain a reasonable topological similarity to the reference compound (**12**) and, at the same time, extend as much as possible the chemical space covered by core-hopping, the compounds were shortlisted to 37 molecules according to their pattern fingerprint similarity (Tanimoto above 0.6) and the highest Euclidean distance as possible from the topological features defined with the CATS descriptor.

Compounds were docked, and the poses in the output were visually inspected to select three different scaffolds that could fulfill the spatial constraints required by the geometry of the binding site, reducing the selected hits to the three main cores (Table 3). Compound **7t** (ChemBridge ID 27888553) was included in this selection.

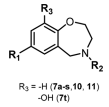
In order to access a wider range of derivatives, the Murcko scaffold[21] of the selected 2,3,4,5-tetrahydrobenzo[f][1,4]-oxazepane was then edited to increase the topological similarity to compound **12**, and a substructure search was conducted against ChemBridge Screening compound libraries. This resulted in a further 60 unique molecules that were subsequently docked and scored (the complete list of the 60 hits is available in Table S2 of the Supporting Information). After docking and after having checked the availability of the compounds, 17 molecules (**7a**, **7d**, **7e**, **7g**–**7t**) were selected for further testing. The QSAR table was subsequently extended with manually designed compounds **7b**, **7c**, **7f**, **10**, and **11**.

All the compounds in this library were then ranked and scored by selected descriptors using the hybrid docking/MD protocol.

Docking and Scoring. Docking experiments were performed using VINA,[22] and the setup was carried out with the Yasara molecular modeling program. We have chosen VINA as in our experience it performs best for mostly lipophilic, protein–protein interfaces. Docking was performed against the available X-ray structure of TbPEX14 (PDB accession code: 5L87). The protein was prepared for docking using Yasara Structure,[23] adding missing side chains and missing hydrogens, and generating protonation states using Yasara's built-in "Clean" command. The docking box was generated around the center of mass of the existing ligand, keeping a distance of 5 Å from the center of mass of the ligand atoms. The cell generated had a total volume of 9.625 Å^3 (x-axis = 23.0 Å, y-axis = 27.0 Å, z-axis = 15.5 Å; $\alpha = \beta = \gamma = 90°$),

https://doi.org/10.1021/acs.jcim.1c00472
J. Chem. Inf. Model. 2021, 61, 5256–5268

Table 4. Docking Scores Calculated for the CATS Library Compounds Using Hybrid Docking/Molecular Dynamics Protocol[a]



Compound	R_1	R_2	Enantiomer	VINA	IV-MD	WMD	AVG	LE
7a			S	-8.70	-7.47	-9.65	-8.61	0.14
7b			R	-8.58	-7.80	-10.26	-8.88	0.14
7c			S	-9.19	-8.12	-10.74	-9.35	0.15
7d			S	-8.79	-8.64	-10.39	-9.27	0.16
7e			R	-8.56	-7.45	-9.97	-8.66	0.15
7f			S	-9.69	-8.53	-10.19	-9.47	0.16
7g			S	-7.81	-6.53	-8.74	-7.69	0.13
7h			S	-8.57	-7.94	-9.38	-8.63	0.14
7i			S	-8.47	-7.80	-9.99	-8.75	0.14
7j			R	-8.05	-7.31	-9.27	-8.21	0.16
7k			R	-8.26	-7.60	-9.31	-8.39	0.15
7l			S	-8.68	-7.27	-8.85	-8.27	0.14
7m			R	-8.43	-7.19	-9.49	-8.37	0.14
7n			S	-9.83	-8.02	-10.17	-9.34	0.17
7o			S	-9.46	-8.15	-10.45	-9.35	0.17
7p			S	-8.61	-7.76	-8.99	-8.45	0.14
7q			S	-9.14	-8.79	-10.98	-9.64	0.16
7r			R	-9.22	-9.27	-10.89	-9.79	0.18
7s			S	-9.19	-7.13	-10.42	-8.91	0.17
7t			NA	-8.38	-8.19	-11.34	-9.30	0.17
10			NA	-7.82	-7.77	-9.77	-8.45	0.13
11			NA	-8.64	-7.19	-10.08	-8.63	0.15
12			NA	-9.20	-12.32	-10.48	-9.92	0.14
13			NA	-8.45	-7.41	-9.74	-8.53	0.17
14			NA	-7.10	-6.43	-7.53	-7.02	0.18

[a]Scores reported are the VINA docking score (VINA), in vacuo molecular dynamics docking score (IV-MD), solvated model molecular dynamics (WMD), and average score (AVG). The ligand efficiency (LE) reported here is calculated by dividing the VINA docking score by the number of non-hydrogen atoms in the molecule and inverting the sign. Scores are reported in kcal/mol. Scores refer to the best performing enantiomer (absolute chirality assigned according to the Cahn−Ingold−Prelog rules[30]) where chiral centers are present. For reference, docking results for pyrazolo[4,3-*c*]pyridine derivative 12 are included (Dawidowski et al.[17]).

Figure 5. Workflow representation for 2D virtual screening protocol. Starting from the structure of compound **12**, the CATS 2D screening based on the Euclidean distance led to a library of more than 10 000 compounds. Compound filtering by Tanimoto scores higher than 0.6 led to a library consisting of 37 compounds. After docking and scoring, these molecules were further shortlisted to three scaffolds with three cores as reported in Table 3; among these, the compound **7t** core based on 2,3,4,5-tetrahydrobenzo[f][1,4]oxazepine was selected for further processing. In order to match the constraints required by the pharmacophore model, the Murcko scaffold of **7t** was modified to overlap to that of **12**. The Murcko scaffold was used to run a substructure search against the ChemBridge SC library in order to find more analogs, leading to 60 more molecules. These were again docked and scored, and after manual inspection, 17 compounds were selected for purchasing. To this selection were added five more molecules designed in the first round of hit optimization. The total QSAR library consisting of 22 molecules was docked and rescored, and estimated binding energies are reported in Table 4.

https://doi.org/10.1021/acs.jcim.1c00472
J. Chem. Inf. Model. 2021, 61, 5256−5268

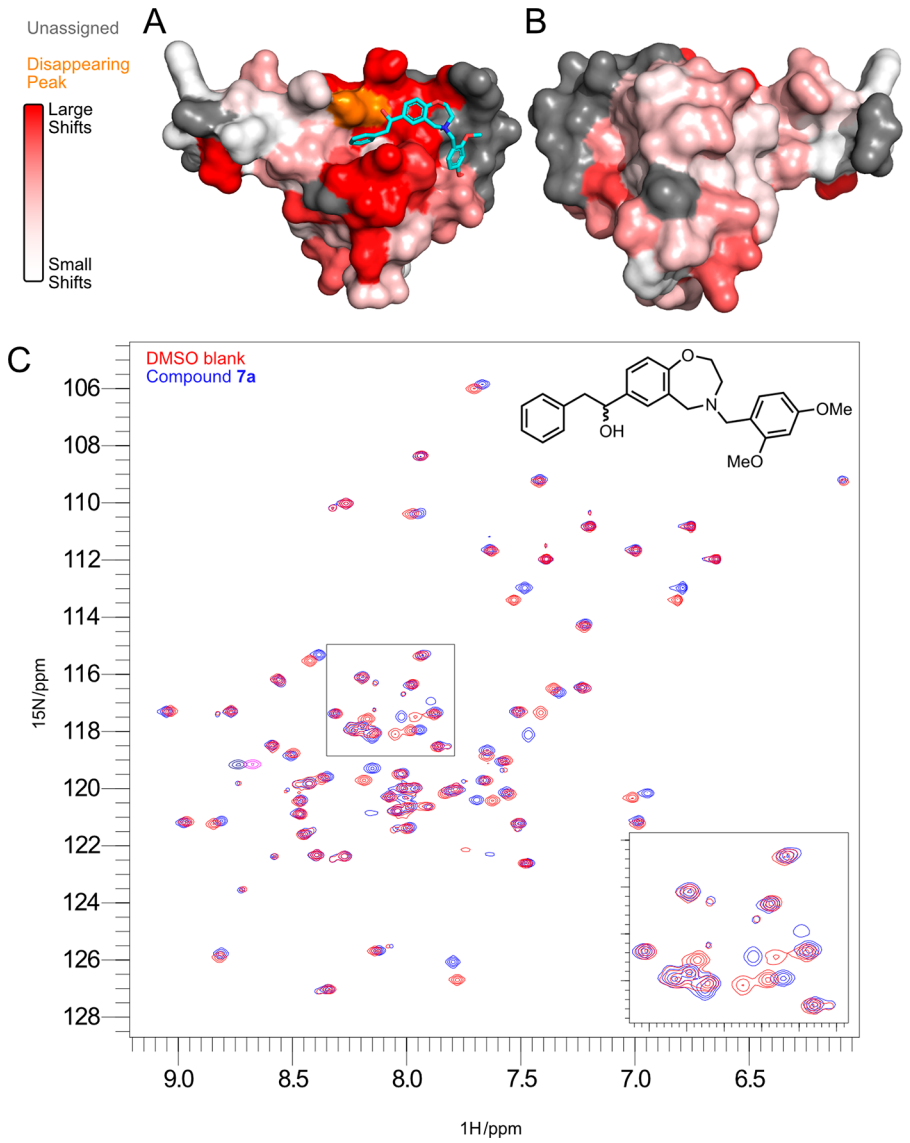


Figure 6. (A and B) Front and back views, respectively, of the CSP analysis of the effect of binding of **7a** on TcPEX14. Red represents large shifts and white minimal shifts, with a gradient between. Gray indicates residues where the assignment could not be transferred, and orange highlights residues where the peak disappeared, indicating a substantial change that cannot be represented on the white−red gradient. A superposition of the *in silico* modeling of the pose of **7a** is shown as turquoise sticks. (C) Superposition of 2D ^{1}H, ^{15}N NMR correlation spectra of *T. cruzi* PEX14 (200 μM) with **7a** at a 3:1 ratio (blue) and with DMSO as a blank (red). Negative contours are depicted in navy blue and magenta, respectively. Inset spectrum highlights significant shift changes seen. The structure of compound **7a** is inset for reference.

in order to accommodate larger ligands than the one in the original structure and to avoid artifacts in pose generation due to steric constraints.

The docking/MD protocol used in order to assess the docking scores of the compounds in four different stages: first, the best docking pose after a given number of docking runs (eight for the preliminary CATS library screening, 24 for rescoring purposes) is selected by the lowest binding energy reported as VINA scores (Table 4). Second, the protein−ligand complex is minimized in vacuo using the NOVA force field[24] adding constraints to backbone atoms in the defined simulation box; after this stage, the compound pose is scored in local scoring mode (VINALS) giving the IV-MD score. Third, the protein ligand complex is minimized using implicit solvent, and the ligand is rescored using the VINA local score to give the IV-MD. Finally, the pose is solvated with AMBER14 force field,[25] and the compound is left to simulate for 500 ps. After this step, a simulated annealing is performed, and the compound is again scored using VINALS to give the WMD score (the code used to perform the screening is available in the "Docking/MD screening protocol" of the SI).

In this protocol, *in vacuo*, solvated-complex short molecular dynamics simulations (500 ps) and simulated annealing were performed in order to assess both the ligand pose stability (eliminate most obvious steric clashes) and the quality of fit of the pharmacophore model. This approach was chosen in order to discard any potential artifact structures which could arise from docking, and the short simulation time was chosen to achieve the best trade-off between screening time required and the chemical reasonability of the poses generated.

With the purpose of maintaining the highest possible similarity to our query molecule, and to avoid biasing too far from the required scaffolding features, after screening the ETH manually curated database against the reference compound,

https://doi.org/10.1021/acs.jcim.1c00472
J. Chem. Inf. Model. 2021, 61, 5256−5268

Table 5. Assay Results: Estimated AlphaScreen (AS) EC_{50} and Trypanocidal Assay IC_{50} (TA)[a]

Compound	AS EC_{50}	Fitting error	TA IC_{50}	Range
7a	33	±4	4	2−7
7b	>1000	±137	7	6−8
7c	509	±70	9	8−11
7d	>1000	-	5	4−5
7e	168	±28	13	12−14
7f	260	±35	10	9−11
7g	268	±44	40	37−45
7h	206	±16	15	13−17
7i	156	±3	26	24−27
7j	207	±20	28	26−37
7k	407	±52	44	41−47
7l	153	±15	6	5−7
7m	168	±14	13	11−15
7n	226	±22	5	4−6
7o	>1000	-	14	11−19
7p	310	±7	13	12−15
7q	438	±17	11	11−12
7r	>1000	-	8	7−10
7s	153	±10	29	24−34
7t	177	±15	5	4−6
10	>1000	-	9	8−10
11	>1000	-	11	9−13
12[17]	17	±3	0.2	-

[a]All values are reported in micromolarity. AS data are estimated from the sigmoidal fitting of the curves (SI). When the compound solubility limit is reached for higher concentration points, the value reported is estimated to the curve inflection point. The reference compound is compound **12**, originally described in Dawidowski et al[17].

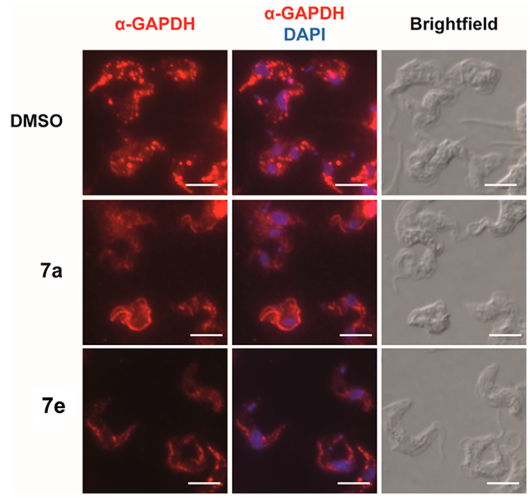


Figure 7. Inhibition of glycosomal protein import in trypanosomes by compounds **7a** and **7e**. Immunofluorescence microscopic analysis of *T. b. brucei* parasites treated with DMSO alone (negative control) or with 5 μM of compound **7a** and **7e** was performed using antibodies against glycosomal GAPDH. In DMSO treated cells, typical bright punctate staining of glycosomes is seen (red channel). Upon treatment with compounds, diffuse cytosolic labeling of GAPDH is seen which demonstrates that the glycosomal protein import is disrupted by the compounds. Nuclear and kinetoplast DNA stained with DAPI (blue channel). Brightfield images are used to clarify cell shape. Scale bar, 5 μm.

the resulting molecules were ranked by Tanimoto similarity (calculated using RDK pattern fingerprints implementation). A library of the 37 best compounds was selected; compounds in this library were obtained as SMILES strings, converted to 3D SDF, and titrated at pH 7.4 using OpenBabel v2.4.1.[26]

A preliminary docking campaign was conducted on this library using the reported docking/MD protocol, and generating eight poses for each representative structure for a total of 296 poses. The poses were then exported to Schrodinger Maestro as SDF files,[20] overlaid with the available X-ray structure of TbPEX14-bound to the pyrazolo[4,3-*c*]pyridine compound (PDB: 5L8A), and manually checked for the key interactions with the binding site residues.

After the identification of **7t**, the scaffold was edited and extended to the Murcko scaffold represented in Figure 5, and a substructure match search was conducted against the ChemBridge Screening compound database. A new library based on substituted 2,3,4,5-tetrahydrobenzo[f][1,4]oxazepines consisting of 60 analogs was obtained in this way. Compounds were downloaded as a SMILES list from the ChemBridge Web site, converted to 3D SDF using OpenBabel v2.4.1.[26] Hydrogens were added and compounds were titrated at pH 7.4.

This library was screened against tbPex14 with the same parameters as described before, generating a total number of 480 poses. The poses generated were filtered keeping only compounds with average binding affinity lower than −6.5 kcal/mol, which were subsequently overlaid to the tbPex14 complex with **12** and inspected for the goodness-of-fit of the binding hypothesis.

This led to a final selection of 17 molecules that were selected to create a QSAR library of the best candidates to prioritize for synthesis. This library was docked and rescored along with the five designed congeneric compounds, this time generating 24 poses for each compound.

At this stage of the docking screening, the chirality of the carbon atom bearing the hydroxyl moiety was also taken into consideration. In order to assess the effect of chirality on the binding efficiency of the hits, both enantiomers were generated where chiral centers were present. This led to the generation of a total of 1056 poses that were filtered in the same way as described before.

The selected compounds were then assayed using 2D $^{1}H,^{15}N$ NMR correlation experiments (Figure 6), an AlphaScreen[27] (AS) binding assay (Table 5), and trypanocidal activity (Figure 7), showing interaction with the protein target and cell toxicity toward cultured *Trypanosoma brucei* parasites.

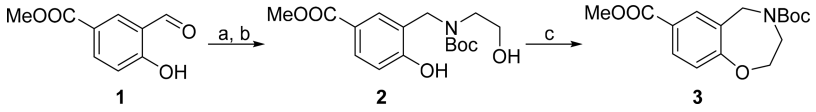
Hit Optimization. The CATS compounds were synthesized starting from readily available methyl 3-formyl-4-hydroxybenzoate (**1**).[28] In a two-step procedure reported in Scheme 1, aldehyde **1** was converted to N-Boc-protected amino alcohol **2** with a total yield of 64%. Synthesis of N-Boc-protected oxazepine **3** was accomplished with a Mitsunobu protocol.[29]

Starting from oxazepine **3**, the synthesis was continued by converting it to aldehydes **6a−6d** (Scheme 2). In a two-step procedure, the Boc group is removed under acidic conditions, and the crude product is directly converted to tertiary amines **4a−4d** by means of reductive amination. The ester group is reduced to the respective alcohols **5a−5d** by $LiAlH_4$, which in turn are converted to aldehydes **6a−6d** via MnO_2-mediated oxidation reactions.

The final step of the synthesis of CATS compounds was a Grignard addition to aldehydes **6a−6d** to yield racemic

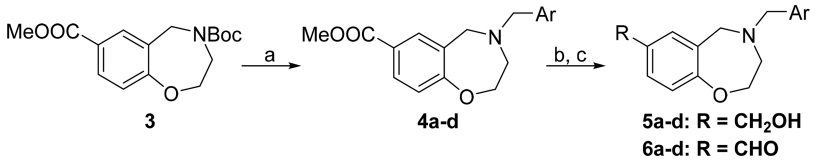
https://doi.org/10.1021/acs.jcim.1c00472
J. Chem. Inf. Model. 2021, 61, 5256−5268

Scheme 1. Synthesis of Boc-Protected Oxazepine 3[a]



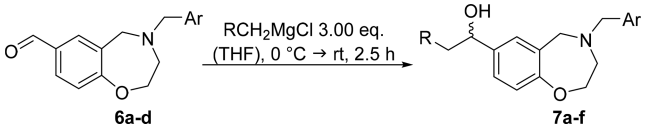
[a]Reagents and conditions: (a) 2-hydroxyethyl amine, THF/MeOH 9:1, rt, 1 h then $NaBH_4$, 45 min, 65% over 2 steps. (b) Boc_2O, $NaHCO_3$, EtOAc/H_2O 2:1, rt, 16 h, 72%. (c) DIAD, PPh_3, THF, 0 °C → rt, 3 h, 91%. THF = tetrahydrofuran, Boc = *tert*-butyloxycarbonyl, DIAD = diisopropyl azodicarboxylate.

Scheme 2. Synthesis of Precursors 6a−6d for Grignard Additions[a]



[a]Reagents and conditions: (a) HCl in 1,4-dioxane, rt, 4 h then NEt_3 in 1,2-DCE, rt, 30 min, then ArCHO, AcOH, $NaBH(OAc)_3$, rt, 12 h. (b) $LiAlH_4$, THF, 0 °C, 1.5−3 h. (c) MnO_2, CH_2Cl_2, rt, 1 h. 1,2-DCE = 1,2-dichloroethane.

Scheme 3. Synthesis of CATS Compounds 7a−7f



mixtures of the CATS compounds **7a−7f** in good yields (Scheme 3).

In addition to CATS compounds bearing secondary alcohols, compound **10** with an amide group adjacent to the oxazepine core was synthesized (Scheme 4). Boc-protected oxazoline **3** was converted to carboxylic acid **8** and further reacted with benzylamine in the presence of EDC and HOBt to yield amide **9**. The synthesis of **10** was completed by a two-step procedure converting Boc-protected amine **9** to a tertiary amine **10** with a 34% yield.

To get more insight about the importance of the secondary alcohol moiety, **7a** was converted to ketone **11** by means of MnO_2-mediated benzylic oxidation (Scheme 5).

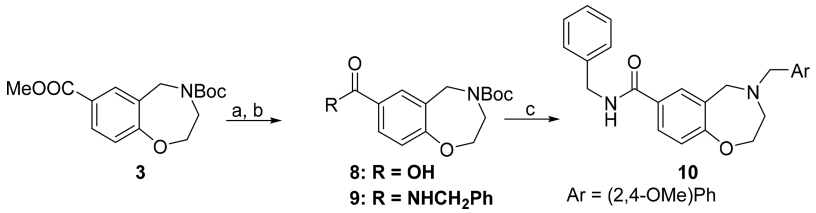
Data and Software Availability. All the information on the structures described in the main text is available in the Supporting Information. The code used for docking-MD rescoring is reported in the respective section of the Supporting Information. All the other data are available upon request. All the software used in this work is available from the respective vendors. The Yasara Structure 19 license was purchased from Yasara Bioscience. Schrödinger Maestro 2017.1 was obtained from Schrödinger (Free Maestro version available for academic use). Knime 2.5 and respective nodes were obtained free-of-charge from Knime AG, including RDkit nodes. PathwayMap is available free-of-charge on Play-Molecule.org. FAME3 is available on a web server at https://nerdd.zbh.uni-hamburg.de/fame3/. PyMol (academic version), Yasara, and Maestro were used to inspect and edit chemical structures and to process chemical files. OpenBabel is available free-of-charge. Commercial catalogs of suppliers used to build the database are available for download from the respective web sites listed in the Supporting Information. All other files (including docking poses) used for this work are available upon request.

RESULTS AND DISCUSSION

Virtual Screening. Using the virtual screening protocol described in the Experimental Section, compounds were selected according to the estimated binding energy and the fulfillment of the principal binding hypothesis. The CATS algorithm was deployed on an in-house database of commercially available compounds prepared using the merging screening libraries available from commercial vendors (a comprehensive list of the library used for the database building is available in Table S1 of the Supporting Information). The database consisted of approximately 14.5 million structures stored as SMILES strings.

After running the docking/MD protocol and a visual inspection, three different scaffolds were selected which optimally matched the pharmacophoric features of compound **12** (Table 3), while being structurally completely unrelated.

Scheme 4. Synthesis of Amide 10[a]



[a]Reagents and conditions: (a) NaOH, MeOH/H_2O 4:1, rt, 12 h, 95%; (b) benzylamine, DIPEA, HOBt, EDC·HCl, CH_2Cl_2, rt, 15 h, 83%; (c) HCl in 1,4-dioxane, rt, 4 h then NEt_3 in 1,2-DCE, rt, 30 min, then 2,4-dimethoxybenzaldehyde, AcOH, $NaBH(OAc)_3$, rt, 12 h, 34%. DIPEA = *N*,*N*-diisopropylethylamine, HOBt = hydroxybenzotriazole, EDC = 1-ethyl-3-(3-dimethylaminopropyl)carbodiimide.

https://doi.org/10.1021/acs.jcim.1c00472
J. Chem. Inf. Model. 2021, 61, 5256−5268

Scheme 5. Synthesis of Ketone 11 via MnO_2-Mediated Oxidation

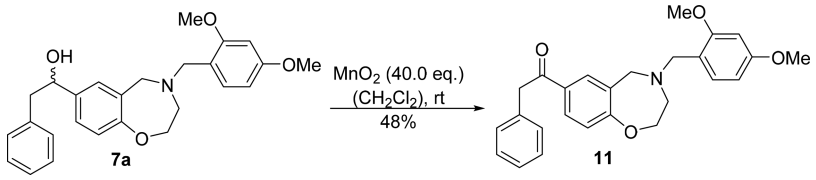


Table 6. ADME Properties for Compounds 7a and 7e[a]

					average hERG binding (%)		
compound	LogD	PPB (%)	$PST_{1/2}$(min)	$MST_{1/2}$(min)	1 μM	5 μM	25 μM
7a	3.6	96	>240	<5	92	100	100
7e	4.2	94	>240	<5	15	34	88

[a]Properties studied were mouse plasma proteins binding (PPB), metabolic stability in mouse plasma (PS), metabolic stability in mouse microsomes (MS), and hERG (human *ether-à-go-go*-related gene, potassium-dependent ion channel) channel binding. Inhibition increases risk of a cardiac failure.

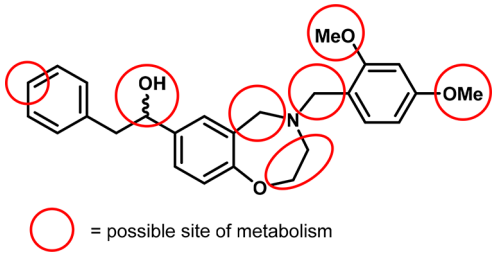


Figure 8. Depiction of areas of plausible metabolic soft spots for compound **7a**.

This screening approach led to the identification of compound **7t**. **7t** contained an interesting annulated oxazepine core scaffold, combined with a low number of rotatable bonds, which suggested a favorable space fit and certain rigidity for required structural preorientation for good binding in the binding pocket; furthermore, it lacks the presence of amides in the core scaffold, which may pose issues in terms of metabolic stability. In Table 3 are reported the two other top ranked core structures.

With the purpose of extending the QSAR Table 4 for screening, the Murcko scaffold of compound **7t** was manually modified in order to ensure better representation of topological features of compound **12**, following the scheme shown in Figure 5.

The edited Murcko scaffold was then used for a substructure search against the ChemBridge Screening compounds database. This led to a further 60 molecules once duplicates were removed (Table S2 in the Supporting Information), which were then prepared for docking[25] and scored.

In order to obtain a reasonably diverse chemical set suitable for *in vitro* testing, a final library consisting of 17 compounds was chosen from the ChemBridge Screening compounds catalog (https://www.chembridge.com/screening_libraries/) after manual inspection of the molecules and poses; five more compounds designed for the first round of hit optimization were added to this library. The docking scores in Table 4 are reported for the best performing enantiomer for each of the final selections of CATS compounds based on the 2,3,4,5-tetrahydrobenzo[f][1,4]oxazepine core. The compounds were obtained as racemic mixtures and then tested using various biophysical assays.

NMR Validation. As an initial *in vitro* screen, compounds **7a** and **7e** were tested by NMR, using $^{1}H,^{15}N$ 2D correlation spectra.[31] The molecules interacted with the N-terminal domain of *T. cruzi* PEX14 in medium to fast exchange on the NMR chemical shift time scale. Figure 6 shows the spectrum of a protein interacting with **7a** (blue) overlaid with the reference spectrum of the free protein with the same volume of d_6-DMSO added (red).

The NMR spectra clearly show that both compounds are interacting with TcPEX14, with a large number of peaks shifting up to 1 ppm. A chemical shift perturbation (CSP) analysis was carried out (Figure 5), showing that many of the affected resonances are attributable to assigned residues near the binding site for the PEX5 WxxxF motif (NMR assignment deposited under BMRB ID: 50345). This indicates that investigated 2,3,4,5-tetrahydrobenzo[f][1,4]oxazepine derivatives bind at or near the binding site expected on the basis of the docking poses generated.

AlphaScreen and Trypanocidal Assays. The compounds were assayed using AlphaScreen (AS) competition assays using TbPEX14 and a biotinylated PEX5-derived peptide containing the WxxxF motif, using methods described previously.[6] The results are reported in Table 5 (see Figure S2 in Supporting Information).

Oxidation of the benzyl alcohol or conversion into an amide led to a loss of activity. However, for 16/22 of the predicted compounds, IC_{50} values for disruption of the PEX protein interaction could be detected by AlphaScreen, and all of the proposed compounds displayed activity in the trypanocidal activity assay.

All compounds were then further tested for trypanocidal activity against the bloodstream form (mammalian stage) of the *T. b. brucei* parasite, showing moderate to high toxicity against the parasite.

To determine if the compounds impair glycosomal protein import in trypanosomes, immunofluorescence analysis was performed using antibodies against glycosomal enzyme GAPDH as shown in Figure 7.

In *trypanosoma* cells treated with DMSO alone as the negative control, a distinct punctate staining of glycosomal GAPDH is seen. However, upon treatment of trypanosomes with compounds **7a** and **7e**, a remarkable mislocalization of the glycosomal enzyme to the cytosol is observed. This indicates significant glycosomal import disruption.

https://doi.org/10.1021/acs.jcim.1c00472
J. Chem. Inf. Model. 2021, 61, 5256−5268

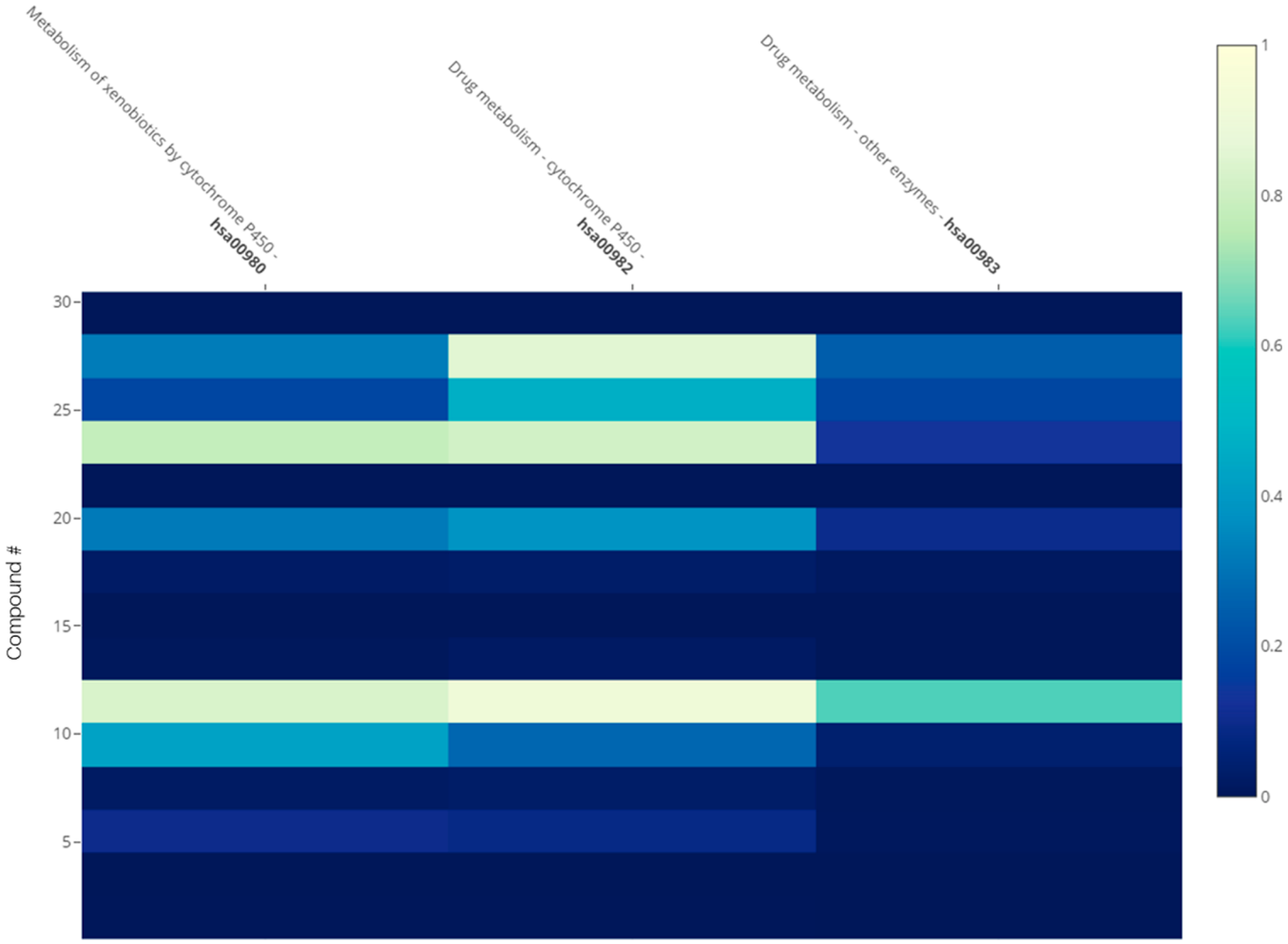


Figure 9. Screening results of the KEGG database using PathwayMap. The high recall rate in the heatmap for CYP450-related metabolic pathways suggests that the compounds may be susceptible to degradation by cytochromes, thus reducing the bioavailability of the compounds.

While our AlphaScreen data do deviate from the trypanocidal assays, it should be noted that AS EC_{50} values are usually 5 to 10 times higher than the IC_{50} values in the TA assay. The AlphaScreen system is an effective but rather artificial test system, in which it is difficult for the small molecule ligand to effectively outcompete the high-affinity peptide. Another reason for a better trypanocidal activity can be found in the mechanism of action of the compound, which, as previously observed by Dawidowski,[6] induces an avalanche effect due to the unregulated activities of the mislocalized enzymes shuttled by the PEX5−PEX14 importomer system, thus leading to a metabolic catastrophe that amplifies the potency of the compound in trypanocidal assays. Off-target effects could also occur, but given the scale of the metabolic catastrophe that is envisaged by blocking the PEX5−PEX14 PPI, they are not expected to play a major role.

ADME Data. Two compounds, **7a** and **7e**, were chosen for testing for the administration, distribution, metabolism, and excretion (ADME) properties based on both *in vitro* and *in silico* screening results. A summary of the ADME properties measured is reported in Table 6.

The compounds tested showed good stability in mouse plasma but also very poor microsomal stability. This may be explained by considering the metabolic liability of various functional groups. We identified eight sites at which metabolic reactions could take place (Figure 8). Both methoxy groups at the eastern part of the molecule pose as critical targets for *O*-demethylation reactions; furthermore, also, benzylic alcohols, electron-rich tertiary amine groups, are known to be susceptible to oxidation reactions. There has been significant progress in the development of structure- or pharmacophore-based prediction algorithms for determining susceptibility and the position of metabolism[32] of small molecule inhibitors. Using FAME,[33,34] demethylation of the para methoxy group of the eastern ring, as well as hydroxylation of the 4-position of the western phenyl ring, have been suggested as preferred sites of metabolism. This information can be important to guide further optimization of compounds **7a** and **7e**.

The potential susceptibility of these moieties to cytochrome drug metabolism was also confirmed using PathwayMap[35] available from PlayMolecule.org. The highest recall for biological pathways screening the Kyoto Encyclopedia of Genes and Genomes (KEGG)[36] database was marked for the CYP450 degradation (Figure 9).

Another challenge to overcome in future development of investigated classes of compounds is the high hERG binding, which could lead to severe cardiac side effects due to the interference with the hERG potassium channels. Even though LogD values are in the range allowed by the rule of five,[37] further optimization is needed to improve the solubility of the compounds in aqueous environments. It is interesting to note that the hERG liability, which is not unexpected for highly

https://doi.org/10.1021/acs.jcim.1c00472
J. Chem. Inf. Model. 2021, 61, 5256−5268

lipophilic tertiary amines, can be efficiently tackled by fine-tuning of the ligand decoration.[38]

CONCLUSIONS

In summary, we provide further evidence that PEX14 is a suitable drug target for the development of new lead compounds against trypanosomiasis. Our results show that this can be approached by exploring the chemical space available using different scaffolds with well-known, reliable, and simple chemistry.

We have demonstrated that scaffold hopping of pyrazolo-[4,3-*c*]pyridine derivatives using the CATS algorithm for the HTVS campaign was successful in highlighting new strategies in the early stage drug development of new ligands for the PEX5−PEX14 PPI. Established molecular modeling workflows such as docking and molecular dynamics guided the selection of the best poses that were then selected for the next steps in compound optimization. We further characterized the best hits using a combination of biophysical assays (NMR, AlphaScreen and trypanocidal assays) and ran the first campaign of chemical synthesis to produce optimized compounds for the SAR analysis, leading to a trypanosomal cellular IC_{50} of 4 μM against the bloodstream form of *T. brucei* for the best compound in the library (**7a**).

Our substituted 2,3,4,5-tetrahydrobenzo[f][1,4]oxazepine compounds are a promising starting point for further lead optimization. Furthermore, the CATS algorithm can serve as a valuable approach for scaffold diversification and derisking of optimization programs.

ASSOCIATED CONTENT

Supporting Information

The Supporting Information is available free of charge at https://pubs.acs.org/doi/10.1021/acs.jcim.1c00472.

List of top-lead scaffold-hopping inhibitors for tbPEX14 obtained via molecular docking as SMILES strings; docking-MD rescoring protocol provided as Yasara Structure-compatible Yanaconda macro; docking poses for compounds **7a**−**11** obtained using AutoDock VINA in Yasara Structure; synthesis protocols for compounds **1**−**11** discussed; AlphaScreen, trypanocidal, and immunofluorescence assays, protocols, and dose−response curves for compounds **1**−**11**; hERG binding evaluation protocols and results for compounds **7a** and **7e**; LogD distribution coefficients for compounds **7a** and **7e**; mouse plasma protein binding assays for compounds **7a** and **7e**; plasma and liver microsome stability assays for compounds **7a** and **7e**; NMR assignment of tcPEX14 and CSPs analysis; full characterization of compounds **1**−**11** via ^{13}C and ^{1}H NMR; and high resolution HPLC (PDF)

AUTHOR INFORMATION

Corresponding Author

Grzegorz M. Popowicz − *Institute of Structural Biology, Helmholtz Zentrum München, 85764 Neuherberg, Germany; Biomolecular NMR, Bayerisches NMR Zentrum and Center for Integrated Protein Science Munich at Chemistry Department, Technical University of Munich, 85747 Garching, Germany*; Email: grzegorz.popowicz@helmholtz-muenchen.de

Authors

Roberto Fino − *Institute of Structural Biology, Helmholtz Zentrum München, 85764 Neuherberg, Germany; Biomolecular NMR, Bayerisches NMR Zentrum and Center for Integrated Protein Science Munich at Chemistry Department, Technical University of Munich, 85747 Garching, Germany*; orcid.org/0000-0001-5323-4562

Dominik Lenhart − *Institute of Structural Biology and Institute of Medicinal Chemistry, Helmholtz Zentrum München, 85764 Neuherberg, Germany; Biomolecular NMR, Bayerisches NMR Zentrum and Center for Integrated Protein Science Munich at Chemistry Department, Technical University of Munich, 85747 Garching, Germany; Institute of Organic Chemistry, Center of Biomolecular Drug Research (BMWZ), Leibniz Universität Hannover, 30167 Hannover, Germany*

Vishal C. Kalel − *Institute of Biochemistry and Pathobiochemistry, Department of Systems Biochemistry, Faculty of Medicine, Ruhr-University Bochum, 44780 Bochum, Germany*

Charlotte A. Softley − *Institute of Structural Biology, Helmholtz Zentrum München, 85764 Neuherberg, Germany; Biomolecular NMR, Bayerisches NMR Zentrum and Center for Integrated Protein Science Munich at Chemistry Department, Technical University of Munich, 85747 Garching, Germany*; orcid.org/0000-0001-9652-8835

Valeria Napolitano − *Institute of Structural Biology, Helmholtz Zentrum München, 85764 Neuherberg, Germany; Biomolecular NMR, Bayerisches NMR Zentrum and Center for Integrated Protein Science Munich at Chemistry Department, Technical University of Munich, 85747 Garching, Germany*

Ryan Byrne − *Department of Chemistry and Applied Biosciences, Institute of Pharmaceutical Sciences, Swiss Federal Institute of Technology (ETH), 8093 Zürich, Switzerland*

Wolfgang Schliebs − *Institute of Biochemistry and Pathobiochemistry, Department of Systems Biochemistry, Faculty of Medicine, Ruhr-University Bochum, 44780 Bochum, Germany*

Maciej Dawidowski − *Department of Drug Technology and Pharmaceutical Biotechnology, Medical University of Warsaw, 02-097 Warsaw, Poland*

Ralf Erdmann − *Institute of Biochemistry and Pathobiochemistry, Department of Systems Biochemistry, Faculty of Medicine, Ruhr-University Bochum, 44780 Bochum, Germany*

Michael Sattler − *Institute of Structural Biology, Helmholtz Zentrum München, 85764 Neuherberg, Germany; Biomolecular NMR, Bayerisches NMR Zentrum and Center for Integrated Protein Science Munich at Chemistry Department, Technical University of Munich, 85747 Garching, Germany*; orcid.org/0000-0002-1594-0527

Gisbert Schneider − *Department of Chemistry and Applied Biosciences, Institute of Pharmaceutical Sciences, Swiss Federal Institute of Technology (ETH), 8093 Zürich, Switzerland*; orcid.org/0000-0001-6706-1084

Oliver Plettenburg − *Institute of Medicinal Chemistry, Helmholtz Zentrum München, 85764 Neuherberg, Germany; Institute of Organic Chemistry, Center of Biomolecular Drug Research (BMWZ), Leibniz Universität Hannover, 30167 Hannover, Germany*; orcid.org/0000-0001-9671-278X

https://doi.org/10.1021/acs.jcim.1c00472
J. Chem. Inf. Model. 2021, 61, 5256−5268

Complete contact information is available at:
https://pubs.acs.org/10.1021/acs.jcim.1c00472

Author Contributions

○These authors contributed equally to this work.

Funding

R.F., C.A.S., V.N., R.B., M.S., G.S., and G.P. would like to acknowledge the European Union's Framework Programme for Research and Innovation Horizon 2020 (2014−2020) under the Marie Skłodowska-Curie Grant Agreement No. 675555 and Accelerated Early staGe drug dIScovery (AEGIS) for the economic support. R.E., W.S., and V.C.K. were supported by the Deutsche Forschungsgemeinschaft grant FOR1905, FoRUM grants (F883−2016 and F913−2017) of the Ruhr-University Bochum, and Bundesministerium für Bildung and Forschung grant PEXMED. M.D. is grateful to the National Science Center Poland that provided funding through grant number UMO-2018/31/B/NZ7/02089. This work is also supported by the Helmholtz Association Initiative and Networking Funds under project number ZT-I-0003.

Notes

The authors declare no competing financial interest.

■ ACKNOWLEDGMENTS

We are grateful to all our respective groups for useful discussion and feedback. We also thank the Protein Expression and Purification Facility at Helmholtz Zentrum München for their support.

■ ABBREVIATIONS

PEX, peroxisomal proteins; PTS, peroxisomal targeting signal; TPR, tetratricopeptide repeat; PPI, protein−protein interaction; HTVS, high-throughput virtual screening; MD, molecular dynamics; AS, AlphaScreen; TA, trypanocidal assay; GAPDH, glyceraldehyde 3-phosphate dehydrogenase; PPB, plasma proteins binding; MS, microsomal stability; PS, plasma stability; hERG, human *ether-à-go-go*-related gene; CSP, chemical shift perturbations; NMR, nuclear magnetic resonance; TLC, thin layer chromatography; CATS, chemically advanced template search

■ REFERENCES

(1) Perez-Molina, J. A.; Molina, I. Chagas Disease. *Lancet* **2018**, *391*, 82−94.
(2) Requena-Mendez, A.; Aldasoro, E.; de Lazzari, E.; Sicuri, E.; Brown, M.; Moore, D. A.; Gascon, J.; Munoz, J. Prevalence of Chagas Disease in Latin-American Migrants Living in Europe: A Systematic Review and Meta-Analysis. *PLoS Neglected Trop. Dis.* **2015**, *9*, No. e0003540.
(3) Docampo, R.; Moreno, S. N. Current Chemotherapy of Human African Trypanosomiasis. *Parasitol. Res.* **2003**, *90*, S10−S13.
(4) Mesu, V. K. B. K.; Kalonji, W. M.; Bardonneau, C.; Mordt, O. V.; Blesson, S.; Simon, F.; Delhomme, S.; Bernhard, S.; Kuziena, W.; Lubaki, J.-P. F.; Vuvu, S. L.; Ngima, P. N.; Mbembo, H. M.; Ilunga, M.; Bonama, A. K.; Heradi, J. A.; Solomo, J. L. L.; Mandula, G.; Badibabi, L. K.; Dama, F. R.; Lukula, P. K.; Tete, D. N.; Lumbala, C.; Scherrer, B.; Strub-Wourgaft, N.; Tarral, A. Oral Fexinidazole for Late-Stage African Trypanosoma Brucei Gambiense Trypanosomiasis: A Pivotal Multicentre, Randomised, Non-Inferiority Trial. *Lancet* **2018**, *391*, 144−154.
(5) Wilkinson, S. R.; Kelly, J. M. Trypanocidal Drugs: Mechanisms, Resistance and New Targets. *Expert Rev. Mol. Med.* **2009**, *11*, No. e31.
(6) Dawidowski, M.; Emmanouilidis, L.; Kalel, V. C.; Tripsianes, K.; Schorpp, K.; Hadian, K.; Kaiser, M.; Maser, P.; Kolonko, M.; Tanghe, S.; Rodriguez, A.; Schliebs, W.; Erdmann, R.; Sattler, M.; Popowicz, G. M. Inhibitors of Pex14 Disrupt Protein Import into Glycosomes and Kill Trypanosoma Parasites. *Science* **2017**, *355*, 1416−1420.
(7) Emmanouilidis, L.; Gopalswamy, M.; Passon, D. M.; Wilmanns, M.; Sattler, M. Structural Biology of the Import Pathways of Peroxisomal Matrix Proteins. *Biochim. Biophys. Acta, Mol. Cell Res.* **2016**, *1863*, 804−813.
(8) Erdmann, R.; Schliebs, W. Peroxisomal Matrix Protein Import: The Transient Pore Model. *Nat. Rev. Mol. Cell Biol.* **2005**, *6*, 738−742.
(9) Montilla-Martinez, M.; Beck, S.; Klümper, J.; Meinecke, M.; Schliebs, W.; Wagner, R.; Erdmann, R. Distinct Pores for Peroxisomal Import of Pts1 and Pts2 Proteins. *Cell Rep.* **2015**, *13*, 2126−2134.
(10) Neuhaus, A.; Kooshapur, H.; Wolf, J.; Meyer, N. H.; Madl, T.; Saidowsky, J.; Hambruch, E.; Lazam, A.; Jung, M.; Sattler, M.; Schliebs, W.; Erdmann, R. A Novel Pex14 Protein-Interacting Site of Human Pex5 Is Critical for Matrix Protein Import into Peroxisomes. *J. Biol. Chem.* **2014**, *289*, 437−448.
(11) Gimeno, A.; Ojeda-Montes, M. J.; Tomás-Hernández, S.; Cereto-Massagué, A.; Beltrán-Debón, R.; Mulero, M.; Pujadas, G.; Garcia-Vallvé, S. The Light and Dark Sides of Virtual Screening: What Is There to Know? *Int. J. Mol. Sci.* **2019**, *20*, 1375.
(12) Besnard, J.; Jones, P. S.; Hopkins, A. L.; Pannifer, A. D. The Joint European Compound Library: Boosting Precompetitive Research. *Drug Discovery Today* **2015**, *20*, 181−6.
(13) Bleicher, K. H.; Bohm, H. J.; Muller, K.; Alanine, A. I. Hit and Lead Generation: Beyond High-Throughput Screening. *Nat. Rev. Drug Discovery* **2003**, *2*, 369−78.
(14) Schneider, G.; Neidhart, W.; Giller, T.; Schmid, G. "Scaffold-Hopping" by Topological Pharmacophore Search: A Contribution to Virtual Screening. *Angew. Chem., Int. Ed.* **1999**, *38*, 2894−2896.
(15) Reutlinger, M.; Koch, C. P.; Reker, D.; Todoroff, N.; Schneider, P.; Rodrigues, T.; Schneider, G. Chemically Advanced Template Search (Cats) for Scaffold-Hopping and Prospective Target Prediction for 'Orphan' Molecules. *Mol. Inf.* **2013**, *32*, 133−138.
(16) Willett, P. Similarity-Based Virtual Screening Using 2d Fingerprints. *Drug Discovery Today* **2006**, *11*, 1046−1053.
(17) Dawidowski, M.; Kalel, V. C.; Napolitano, V.; Fino, R.; Schorpp, K.; Emmanouilidis, L.; Lenhart, D.; Ostertag, M.; Kaiser, M.; Kolonko, M.; Tippler, B.; Schliebs, W.; Dubin, G.; Maser, P.; Tetko, I. V.; Hadian, K.; Plettenburg, O.; Erdmann, R.; Sattler, M.; Popowicz, G. M. Structure-Activity Relationship in Pyrazolo [4, 3-C] Pyridines, First Inhibitors of Pex14-Pex5 Protein-Protein Interaction (Ppi) with Trypanocidal Activity. *J. Med. Chem.* **2020**, *63*, 847.
(18) Berthold, M. R.; Cebron, N.; Dill, F.; Gabriel, T. R.; Kötter, T.; Meinl, T.; Ohl, P.; Thiel, K.; Wiswedel, B. Knime-the Konstanz Information Miner: Version 2.0 and Beyond. *AcM SIGKDD explorations Newsletter* **2009**, *11*, 26−31.
(19) Bajusz, D.; Rácz, A.; Héberger, K. Why Is Tanimoto Index an Appropriate Choice for Fingerprint-Based Similarity Calculations? *J. Cheminf.* **2015**, *7*, 20.
(20) *Maestro*, release S1; Schrödinger, LLC: New York, 2018.
(21) Walters, W. P.; Stahl, M. T.; Murcko, M. A. Virtual Screening—an Overview. *Drug Discovery Today* **1998**, *3*, 160−178.
(22) Trott, O.; Olson, A. J. Autodock Vina: Improving the Speed and Accuracy of Docking with a New Scoring Function, Efficient Optimization, and Multithreading. *J. Comput. Chem.* **2009**, *31*, 455−61.
(23) Krieger, E.; Vriend, G. Yasara View - Molecular Graphics for All Devices - from Smartphones to Workstations. *Bioinformatics* **2014**, *30*, 2981−2.
(24) Krieger, E.; Koraimann, G.; Vriend, G. Increasing the Precision of Comparative Models with Yasara Nova–a Self-Parameterizing Force Field. *Proteins: Struct., Funct., Genet.* **2002**, *47*, 393−402.
(25) Maier, J. A.; Martinez, C.; Kasavajhala, K.; Wickstrom, L.; Hauser, K. E.; Simmerling, C. Ff14sb: Improving the Accuracy of Protein Side Chain and Backbone Parameters from Ff99sb. *J. Chem. Theory Comput.* **2015**, *11*, 3696−713.

(26) O'Boyle, N. M.; Banck, M.; James, C. A.; Morley, C.; Vandermeersch, T.; Hutchison, G. R. Open Babel: An Open Chemical Toolbox. *J. Cheminform.* **2011**, *3*, 33.
(27) Eglen, R. M.; Reisine, T.; Roby, P.; Rouleau, N.; Illy, C.; Bossé, R.; Bielefeld, M. The Use of Alphascreen Technology in Hts: Current Status. *Curr. Chem. Genomics* **2008**, *1*, 2.
(28) Hansen, T. V.; Skattebøl, L. Discussion Addendum For: Ortho-Formylations of Phenols; Preparation of 3-Bromosalicylaldehyde. *Org. Synth.* **2012**, *89*, 220−229.
(29) Takeuchi, C. S.; Kim, B. G.; Blazey, C. M.; Ma, S.; Johnson, H. W. B.; Anand, N. K.; Arcalas, A.; Baik, T. G.; Buhr, C. A.; Cannoy, J.; Epshteyn, S.; Joshi, A.; Lara, K.; Lee, M. S.; Wang, L.; Leahy, J. W.; Nuss, J. M.; Aay, N.; Aoyama, R.; Foster, P.; Lee, J.; Lehoux, I.; Munagala, N.; Plonowski, A.; Rajan, S.; Woolfrey, J.; Yamaguchi, K.; Lamb, P.; Miller, N. Discovery of a Novel Class of Highly Potent, Selective, Atp-Competitive, and Orally Bioavailable Inhibitors of the Mammalian Target of Rapamycin (Mtor). *J. Med. Chem.* **2013**, *56*, 2218−2234.
(30) Cahn, R.; Ingold, C.; Prelog, V. Spezifikation Der Molekularen Chiralität. *Angew. Chem.* **1966**, *78*, 413−447.
(31) Bodenhausen, G.; Ruben, D. J. Natural Abundance Nitrogen-15 Nmr by Enhanced Heteronuclear Spectroscopy. *Chem. Phys. Lett.* **1980**, *69*, 185−189.
(32) He, S.; Zhang, C.; Zhou, P.; Zhang, X.; Ye, T.; Wang, R.; Sun, G.; Sun, X. Herb-Induced Liver Injury: Phylogenetic Relationship, Structure-Toxicity Relationship, and Herb-Ingredient Network Analysis. *Int. J. Mol. Sci.* **2019**, *20*, 3633.
(33) Šícho, M.; Stork, C.; Mazzolari, A.; de Bruyn Kops, C.; Pedretti, A.; Testa, B.; Vistoli, G.; Svozil, D.; Kirchmair, J. Fame 3: Predicting the Sites of Metabolism in Synthetic Compounds and Natural Products for Phase 1 and Phase 2 Metabolic Enzymes. *J. Chem. Inf. Model.* **2019**, *59*, 3400−3412.
(34) Stork, C.; Embruch, G.; Šícho, M.; de Bruyn Kops, C.; Chen, Y.; Svozil, D.; Kirchmair, J. NERDD: A Web Portal Providing Access to in Silico Tools for Drug Discovery. *Bioinformatics* **2019**, *36*, 1291−1292.
(35) Jimenez, J.; Sabbadin, D.; Cuzzolin, A.; Martinez-Rosell, G.; Gora, J.; Manchester, J.; Duca, J.; De Fabritiis, G. Pathwaymap: Molecular Pathway Association with Self-Normalizing Neural Networks. *J. Chem. Inf. Model.* **2019**, *59*, 1172−1181.
(36) Kanehisa, M.; Araki, M.; Goto, S.; Hattori, M.; Hirakawa, M.; Itoh, M.; Katayama, T.; Kawashima, S.; Okuda, S.; Tokimatsu, T.; Yamanishi, Y. KEGG for Linking Genomes to Life and the Environment. *Nucleic Acids Res.* **2007**, *36*, D480−D484.
(37) Lipinski, C. A. Lead-and Drug-Like Compounds: The Rule-of-Five Revolution. *Drug Discovery Today: Technol.* **2004**, *1*, 337−341.
(38) Kratz, J. M.; Schuster, D.; Edtbauer, M.; Saxena, P.; Mair, C. E.; Kirchebner, J.; Matuszczak, B.; Baburin, I.; Hering, S.; Rollinger, J. M. Experimentally Validated Herg Pharmacophore Models as Cardiotoxicity Prediction Tools. *J. Chem. Inf. Model.* **2014**, *54*, 2887−2901.

https://doi.org/10.1021/acs.jcim.1c00472
J. Chem. Inf. Model. 2021, 61, 5256−5268