



# *In Silico* Workflow for the Discovery of Natural Products Activating the G Protein-Coupled Bile Acid Receptor 1

Benjamin Kirchweger<sup>1</sup>, Jadel M. Kratz<sup>1</sup>, Angela Ladurner<sup>1</sup>, Ulrike Grienke<sup>1</sup>, Thierry Langer<sup>2</sup>, Verena M. Dirsch<sup>1</sup> and Judith M. Rollinger<sup>1\*</sup>

<sup>1</sup> Department of Pharmacognosy, University of Vienna, Vienna, Austria, <sup>2</sup> Department of Pharmaceutical Chemistry, University of Vienna, Vienna, Austria

The G protein-coupled bile acid receptor (GPBAR1) has been recognized as a promising new target for the treatment of diverse diseases, including obesity, type 2 diabetes, fatty liver disease and atherosclerosis. The identification of novel and potent GPBAR1 agonists is highly relevant, as these diseases are on the rise and pharmacological unmet therapeutic needs are pervasive. Therefore, the aim of this study was to develop a proficient workflow for the *in silico* prediction of GPBAR1 activating compounds, primarily from natural sources. A protocol was set up, starting with a comprehensive collection of structural information of known ligands. This information was used to generate ligand-based pharmacophore models in LigandScout 4.08 Advanced. After theoretical validation, the two most promising models, namely BAMS22 and TTM8, were employed as queries for the virtual screening of natural product and synthetic small molecule databases. Virtual hits were progressed to shape matching experiments and physicochemical clustering. Out of 33 diverse virtual hits subjected to experimental testing using a reporter gene-based assay, two natural products, farnesiferol B (**27**) and microbidene (**28**), were confirmed as GPBAR1 activators reaching more than 50% receptor activation at 20  $\mu$ M with EC<sub>50</sub>s of 13.53  $\mu$ M and 13.88  $\mu$ M, respectively. This activity is comparable to that of the endogenous ligand lithocholic acid (**1**). Seven further virtual hits showed activity reaching at least 15% receptor activation either at 5 or 20  $\mu$ M, including new scaffolds from natural and synthetic origin.

**Keywords:** GPBAR1, TGR5, pharmacophore, virtual screening, natural product, triterpene

## INTRODUCTION

The G protein-coupled bile acid receptor 1 (GPBAR1), also commonly named M-BAR or Takeda G-protein-coupled receptor 5 (TGR5), is a rhodopsin-like G protein-coupled receptor (GPCR) expressed in various tissues. It is primarily present in the bile duct, digestive system, spleen, and placenta. It is a cell-surface receptor comprising an extracellular N-terminus, an intracellular C-terminus and seven trans-membrane helices connected by intra- and extracellular loops. Its endogenous ligands are bile acids and neurosteroids. The binding pocket is predicted to be located between the trans-membrane helices. Next to the transcription factor farnesoid X receptor (FXR), GPBAR1 was the second receptor discovered to be responsive to bile acids (Maruyama et al., 2002; Kawamata et al., 2003; Keitel et al., 2010; Gertzen et al., 2015).

## OPEN ACCESS

### Edited by:

Honglin Li,  
East China University of Science and  
Technology, China

### Reviewed by:

Christian W. Gruber,  
Medizinische Universität Wien, Austria  
Shoude Zhang,  
Qinghai University, China

### \*Correspondence:

Judith M. Rollinger  
judith.rollinger@univie.ac.at

### Specialty section:

This article was submitted to  
Medicinal and Pharmaceutical  
Chemistry,  
a section of the journal  
Frontiers in Chemistry

Received: 09 February 2018

Accepted: 06 June 2018

Published: 02 July 2018

### Citation:

Kirchweger B, Kratz JM, Ladurner A,  
Grienke U, Langer T, Dirsch VM and  
Rollinger JM (2018) *In Silico* Workflow  
for the Discovery of Natural Products  
Activating the G Protein-Coupled Bile  
Acid Receptor 1. *Front. Chem.* 6:242.  
doi: 10.3389/fchem.2018.00242

In the past decade, this receptor has attracted attention as a potential drug target for a variety of pathologic conditions (Hodge and Nunez, 2016), predominantly because GPBAR1 is a key receptor in the adjustment of energy expenditure and glucose metabolism with possible implications for the treatment of obesity and type 2 diabetes. Its activation in enteroendocrine L-cells leads to the release of the incretins peptide tyrosine tyrosine (PYY) and glucagon like peptide 1 (GLP1), which promote insulin secretion in the pancreas and are important in the suppression of appetite (Woods and D'Alessio, 2008; Bala et al., 2014). GPBAR1 activation in pancreatic cells leads to an enhanced insulin secretion and a recovery of  $\beta$ -cell mass and function switching from glucagon to GLP1 (Kumar et al., 2012, 2016). In striated myocytes and brown adipocytes, GPBAR1 activation leads to thyroid hormone activation. In white adipocytes it mediates remodeling into beige cells and improves mitochondrial dynamics and cellular respiration rate (Watanabe et al., 2006; Velazquez-Villegas et al., 2018). Moreover, endothelium- and liver-protecting, as well as immunosuppressing effects offer perspectives for new therapies for diseases like atherosclerosis and inflammatory liver diseases (Keitel et al., 2007, 2008; Keitel and Haussinger, 2011; Pols et al., 2011; Asgharpour et al., 2015). Unusual for GPCRs, the GPBAR1 seems to only transfer signaling via G proteins and not via  $\beta$ -arrestins (Jensen et al., 2013).

In animal trials, GPBAR1 agonists showed promising results, however, difficulties have also been encountered since GPBAR1 agonists may induce itching and gallbladder extension (Vassileva et al., 2006; Alemi et al., 2013). Interestingly, gallbladder extension upon GPBAR1 activation is mainly caused by smooth muscle relaxation via induction of the cAMP-PKA pathway independent of the agonist scaffold (Lavoie et al., 2010; Li et al., 2011). The plethora of GPBAR1-mediated biological functions appears to be an obvious opportunity, but a major drawback at the same time (Vassileva et al., 2006; Alemi et al., 2013; Hodge and Nunez, 2016). Novel agonistic scaffolds may incorporate a different and possibly more favorable side effect profile in terms of receptor and functional selectivity as well as pharmacokinetic properties. In this sense there is a demand for new GPBAR1 ligands as they may help to cope with pharmacologically unmet therapeutic needs against metabolic diseases.

Beside bioassay-guided fractionation of plant extracts (Sato et al., 2007), bioisosteric replacement (Park et al., 2014), and exploitation/lead optimization of bile acid scaffolds (Pellicciari et al., 2009), previous efforts in the discovery of GPBAR1 modulators have focused on high throughput screening (HTS) (Evans et al., 2009; Herbert et al., 2010; Londregan et al., 2013; Martin et al., 2013) leading to a broad range of agonists of which some are depicted in **Figure 1**.

Chenodeoxycholic acid (CDCA, **3**) is a primary bile acid, which activates both GPBAR1 and the nuclear receptor FXR. By bacterial dehydroxylation, CDCA is transformed into the more potent lithocholic acid (LCA, **1**). The bile acid's potency on GPBAR1 can be further increased by conjugation with glycine or taurine, whereas taurine-conjugated lithocholic acid (TLC) is the most potent endogenous ligand (Sato et al., 2008). Lead

optimization of the bile acid scaffold led to INT-747 (**4**), an approved drug for the treatment of primary biliary cholangitis and dual agonist of GPBAR1 and FXR as well as INT-777 (**2**), a more selective GPBAR1 agonist (Pellicciari et al., 2009; Fiorucci et al., 2014; Floreani and Mangini, 2018). Beside bile acids, several secondary plant metabolites activate this receptor (**5-8**). The antidiabetic effect of e.g., *Olea europaea* L. leaves may be linked to GPBAR1 activation by its major constituent oleanolic acid (**6**) (Sato et al., 2007). Moreover, HTS and extensive SAR efforts gave access to a large range of synthetic compounds activating this receptor in the nanomolar and low micromolar range (**9-12**).

Absence of a crystal structure of GPBAR1 forced researchers to rely either on homology models or ligand-based approaches for *in silico* studies, as this is the case for most GPCRs (Peeters et al., 2011; Vaidehi et al., 2014). Several GPBAR1 homology models have been described up to now. They all represent fundamentally different bile acid binding poses, but none of them is able to cover all results from mutagenesis studies (Macchiarulo et al., 2013; D'Amore et al., 2014; Gertzen et al., 2015; Yu et al., 2015). This prompted us to develop a ligand-based approach using pharmacophore modeling for the identification of new GPBAR1 agonists.

Pharmacophore modeling and subsequent virtual screening (VS) is a well-established method in the early drug discovery process showing some important benefits: (1) pharmacophore screening can retrieve ligands with structurally diverse scaffolds and allows for so called "scaffold-hopping"; (2) it can automatically and rapidly filter large compound libraries; (3) ligand-based pharmacophore VS has been able to retrieve satisfactory results, also without structural information on the target (Evers et al., 2005; Ha et al., 2015; Akram et al., 2017). Here, we report on the construction of two ligand-based 3D pharmacophore models, their *in silico* and *in vitro* validation, and the directed discovery of sesquiterpene coumarins as a new class of potent GPBAR1 agonists.

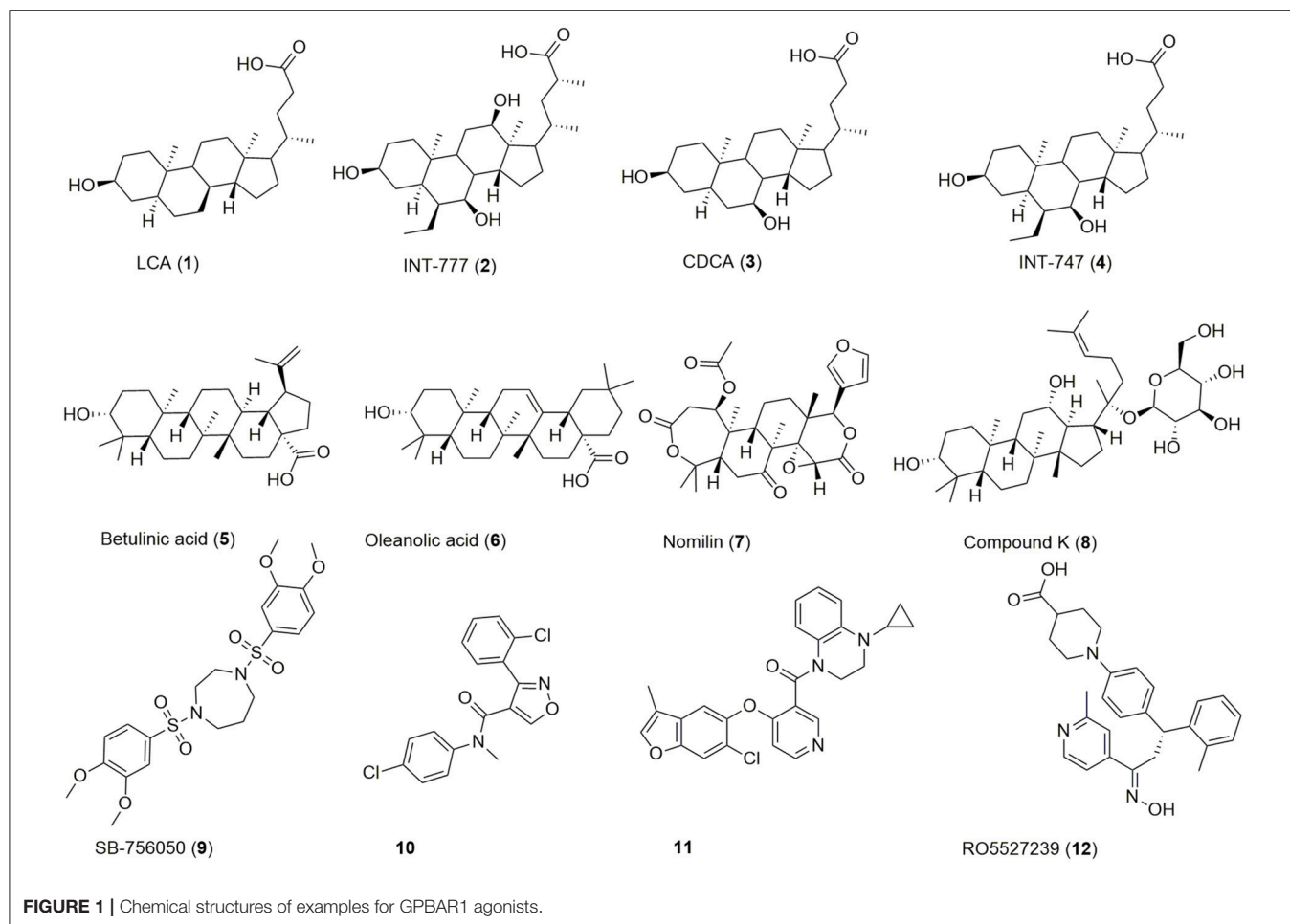
## MATERIALS AND METHODS

### Software

The generation of pharmacophore models, their subsequent refinement and VS was performed with LigandScout 4.08 Advanced, available by Inte:Ligand GmbH (Wolber and Langer, 2005). The conformational libraries for both pharmacophore modeling and the VS process were created with i:Con, LigandScout's implemented conformer generator (Friedrich et al., 2017). Shape comparison was performed with OpenEye's ROCS 3.2.1.4 (Hawkins et al., 2007; OpenEye, 2016). 2D structures were drawn with ChemDraw Professional 15.0.

### Data Sources

For model generation in LigandScout, structural data of GPBAR1 ligands with bioactivity annotations were collected. The data for GPBAR1 available in the ChEMBL database was extracted on March 15th 2016. It consisted of 24 different publications with 623 reported EC<sub>50</sub> values (Bento et al., 2014). The reliability of the



content was checked with the original literature. This molecule set was extended by extracting data from another 18 publications, 24 patents and previous in house projects, resulting in a total of 1025 activity annotations.

## Decoy Set

For the experimental validation of pharmacophores, next to a set of active molecules, also a set of inactive molecules and/or a set of decoys is necessary (Schuster et al., 2006). In contrast to true inactives, which are molecules reported in the literature not to be active at the target, decoys are hypothetical structures, which are unlikely to show activity at the target, but have not yet been tested experimentally. Due to the shortage of published negative data and therefore the presence of only a small set of reliably tested inactive compounds, a set of decoys was generated using the Dude decoys database (<http://dude.docking.org/>): 338 molecules from the “High Actives” dataset was submitted to the DUDE decoy online generator (Mysinger et al., 2012) to obtain decoys with similar 1D physicochemical properties but dissimilar 2D topology in comparison to the active compounds. Using this strategy, a “Decoy” set comprising 18 043 substances was created.

## Conformational Sampling, Ligand Set Clustering

The “High Actives,” “Decoys,” and “True Inactives” sets were transferred into multi-conformational databases via iCon with the default “BEST” settings [Timeout (s): 600, RMS threshold: 0.8, energy window: 20.0, max. pool size: 4,000, max fragment build time: 30, max number of conformers: 200]. The 338 compounds of the “High Actives” set were clustered in LigandScout 4.08 using the implemented pharmacophore clustering tool. The tool clusters molecules with similar pharmacophore characteristics in the dataset: It generates pharmacophores for each molecule in the dataset for a desired number of conformations. The similarity of these pharmacophores is measured with the cosine similarity (value between 0 and 1) of their radial distribution function score (RDF) vectors. Options, distance 0.9 and cluster distance calculation method “maximum” with three conformations for each molecule, were used.

## Pharmacophore Generation

With LigandScout pharmacophores can be generated as shared or merged feature pharmacophores. A shared feature pharmacophore only appoints common features observed

during 3D alignment of the validation set molecules. A merged feature pharmacophore merges several shared pharmacophores. Features, which are not shared by the whole validation set, are appointed as optional. The initial pharmacophore model for cluster 7 was built as a shared feature pharmacophore with six molecules as templates using “pharmacophore fit and atom overlap” as the scoring function. The initial model for cluster 12 was built as a merged feature pharmacophore with four molecules as templates using “pharmacophore fit and atom overlap” as the scoring function. The models were refined and theoretically validated until favored theoretical performance was achieved.

## Theoretical Validation

For theoretical validation, the scoring function was set to “pharmacophore-fit,” the screening mode to “match all query features” with maximum number of omitted features zero. To assess the performance of the individual models, the resulting hit list were used to calculate common enrichment metrics, as comprehensively outlined in a review by Seidel and coworkers (Seidel et al., 2010).

## Virtual Screening

Several freely available molecular structure databases were deployed for VS, having a strong focus on NP. The conformational libraries were generated with i:Con (Friedrich et al., 2017). Depending on the size of the database the recommended “BEST” or “FAST” settings were used (Table 1). For VS, the same settings were used as in the theoretical validation, although the retrieval method “get best matching conformation” was used.

## Hit List Prioritization

A principal component analysis using the chemGPS online tool (Larsson et al., 2007) was determined and a hierarchical cluster analysis with SIMCA facilitated the assignment of the compounds into 9 groups, each inhabiting a different chemical space. For clustering, the default Ward’s minimum variance agglomerative clustering algorithm for the quantitative first three chemGPS principal components, PC1, PC2, and PC3, were used. A molecule’s size, polarizability and shape are characterized by PC1, while PC2 describes its aromatic and conjugation-related properties and PC3 corresponds to its lipophilicity, polarity, and hydrogen bond (HB) capacity. Shape-focused VS was performed with Open Eye ROCS to retrieve a TC score, which combines a shape matching with a chemistry alignment Tanimoto score. This scoring function assesses the goodness of the alignment between the query and the candidate molecules. ComboScore puts exactly equal weights on both of its components, a shape-based scoring function and a function considering pharmacophore-like chemical pattern matching. Theoretically, the TC score can lie between 0 and 2 (Hawkins et al., 2007; OpenEye, 2016). The best fitting conformation of **13** derived from the alignment with model BAMS22 was used as query with the ROCS default options. The PAINS filters of the FAF-Drugs 4 online tool were applied to identify potential promiscuous hitters.

## In Vitro GPBAR1 Activity and Statistical Analysis

*In vitro* evaluation of the selected hit list was performed with a reporter gene-based luciferase assay in HEK 293T cells (obtained from ATCC, USA), which was described previously (Ladurner et al., 2017). Cells were grown and maintained in

**TABLE 1** | Screened databases with content type (origin of molecules) and size (number of molecules), their source and the used standard settings for conformer generation with i:Con.

Database	Number of molecules	Type	Source	Settings
Analyticon-MEGx	4,355	NPs	Analyticon <a href="http://www.ac-discovery.com/">http://www.ac-discovery.com/</a>	Best settings
Analyticon-Triterpenes	409	NPs	Analyticon <a href="http://www.ac-discovery.com/">http://www.ac-discovery.com/</a>	Fast settings
Drugbank	6,863	Approved/ trialed drugs	DrugBank 4.5 <a href="http://www.drugbank.ca/">http://www.drugbank.ca/</a>	Fast settings
In house database Department of Pharmacognosy; University of Vienna	1,152	NPs	In house (Update 2016)	Fast settings and best settings
NPDB	115,275	NPs	(Rollinger et al., 2004)	Fast settings
NuBBE	1,628	NPs	(Valli et al., 2013) <a href="http://nubbe.iq.unesp.br/portal/nubbedb.html">http://nubbe.iq.unesp.br/portal/nubbedb.html</a>	Best settings
SPECS NP	871	NPs	SPECS <a href="http://www.specs.net/">http://www.specs.net/</a>	Fast settings
SPECS SC	212,446	Synthetic compounds	SPECS <a href="http://www.specs.net/">http://www.specs.net/</a>	Fast settings
TCM-Taiwan	35,993	NPs	(Chen, 2011) <a href="http://tcm.cmu.edu.tw/">http://tcm.cmu.edu.tw/</a>	Fast settings

Dulbecco's modified eagle medium (DMEM) without phenol red with 10% heat-inactivated fetal bovine serum (FBS), 4.5 g/L glucose, 2 mM glutamine, 100 U/mL benzylpenicillin, and 100 µg/ml streptomycin. During the experiments, charcoal-stripped medium with 5% FBS was used.  $6 \times 10^6$  cells were grown in 15 cm dishes for 19 h and then transiently transfected using the calcium phosphate method with 5 µg of a GPBAR1 expression plasmid and 5 µg of a CRE-Luc plasmid. For later normalization, 3 µg of an EGFP expression plasmid was co-transfected. Control experiments were performed with cells transfected only with 3 µg EGFP and 5 µg CRE-Luc plasmids. After 6 h, transfected cells were reseeded to 96 well plates ( $5 \times 10^4$  cells/well) and incubated with 5 µM and 20 µM compound dilutions, respectively, for 18 h. 0.1% DMSO served as vehicle control and 10 µM LCA as positive control. After incubation the medium was removed, and the plates were immediately frozen at  $-80^\circ\text{C}$ . Plates were kept frozen for at least 1 h to facilitate lysis and measurements were performed in the following 10 days. For the measurement, cells were thawed, lysed and transferred to black 96-well plates. After addition of ATP and luciferin, emitted luminescence and fluorescence was measured with a Tecan Infinite 200 PRO plate reader (Tecan, Austria). GPBAR1 activity was expressed as fold activation compared with the solvent control (0.1% DMSO) or as % activation compared to the positive control 10 µM LCA (arbitrary 100% activation). The measured relative luciferase units (nRLU) were normalized to the transfected cell mass expressed as EGFP-derived relative fluorescence units (RFU) from at least three independent experiments (mean values  $\pm$  standard error mean) performed in quadruplicate. Quantified EGFP-derived fluorescence was used as an indicator for transfected cell mass and thus used to assess the compounds' cytotoxicity. Compounds, which resulted in significantly lower RFU values than the control, were considered as cytotoxic. For statistical analysis GraphPad Prism 4.03 was used. Statistical significance was assessed by One Way ANOVA and Bonferroni post-test ( $***p < 0.001$ ,  $**p < 0.01$ ,  $*p < 0.05$ , ns not significant). Non-linear regression was used to calculate  $\text{EC}_{50}$  values with the sigmoidal dose response (variable slope) settings.

## Compounds and Chemicals

Hederagenin (CAS#465-99-6) and bayogenin (CAS#80368) were ordered from Phytolab (Germany). 2,3-O-isopropylidene-euscaphic acid (CAS#220880-90-0) was ordered from Proactive Molecular Research (USA). Phytolaccoside B (CAS#60820-94-2) and euscaphic acid (CAS#53155-25-2) was purchased from Cambridge Chemicals (USA). Phytolaccagenic acid (CAS#54928-05-1) and 16-dehydropregnenolone (CAS#1162-53-4) were obtained from Carbosynth (UK). Spirolactone (CAS#52-01-7) and methylhoxychololate (CAS#2868-48-6) were purchased from TCI Deutschland GmbH (Germany). The screening compounds (CAS#1019061-83-6, CAS#303139-94-8, CAS#330636-58-3, CAS#353253-76-6, CAS#353779-79-0, CAS#432530-00-2, CAS#444931-63-9, CAS#496937-29-2, CAS#791840-52-3, CAS#902244-06-8, CAS#915930-57-3, CAS#932954-51-3, CAS#352644-32-7, CAS#500218-51-9, CAS#314757-83-0, CAS#380633-89-6, CAS#26179-09-9, CAS#664993-86-6, CAS#525577-20-2) were obtained from

SPECS (Netherlands). Microlobidene (CAS#89783-66-4) and farnesiferol B (CAS#54990-68-0) were available from a previous project (Rollinger et al., 2008). Nordihydroguarectic acid (CAS#500-38-9) was purchased from Fluka (Switzerland). The positive controls LCA (CAS#434-13-9) and CDCA (CAS#474-25-9) were obtained from Sigma Aldrich (Austria). Alplitolic acid (CAS#19533-92-7), was obtained by hydrolysis from a previously isolated saponin (Mair et al., 2018). The purity was checked using UPLC-PDA-MS and determined as  $\geq 98\%$  for compounds **20**, **21**, **23**, **24**, **28**, **30-35**, **37**, **40**, **41**, **44-46**, **48**, **49**, and **52**. For all other compounds it was between 90 and 98%. MS and NMR data of all in house compounds (**27**, **28**, **52**) are provided in the literature (Rollinger et al., 2008; Mair et al., 2018) and the Supplementary Information (Supplementary Figures 4–10).

## Cell Culture Reagents and Plasmids

DMEM, L-glutamine, benzylpenicillin and streptomycin were purchased from Lonza, (Switzerland), FBS, and trypsin were obtained from Gibco via Invitrogen (Austria). The GPBAR1 transcript variant 3 (NM170699) plasmid was obtained from Origene via Biomedica (Vienna, Austria). The CRE-Luc plasmid, (pGL4.29[luc2P/CRE/Hygro]), luciferase assay system and used lysis buffer were ordered from Promega (Germany) and the EGFP (pEGFP-N1) plasmid was purchased from Clontech (USA).

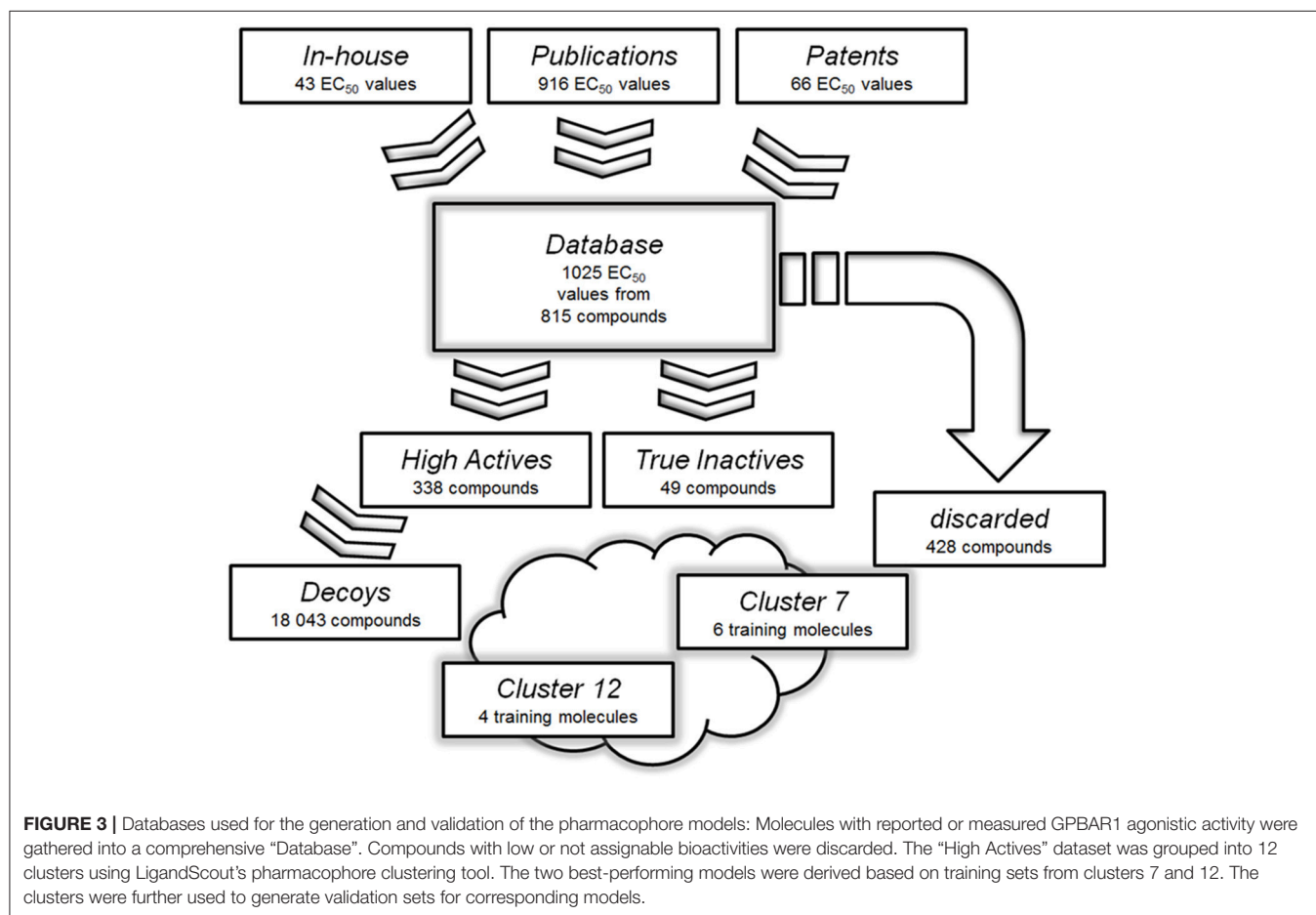
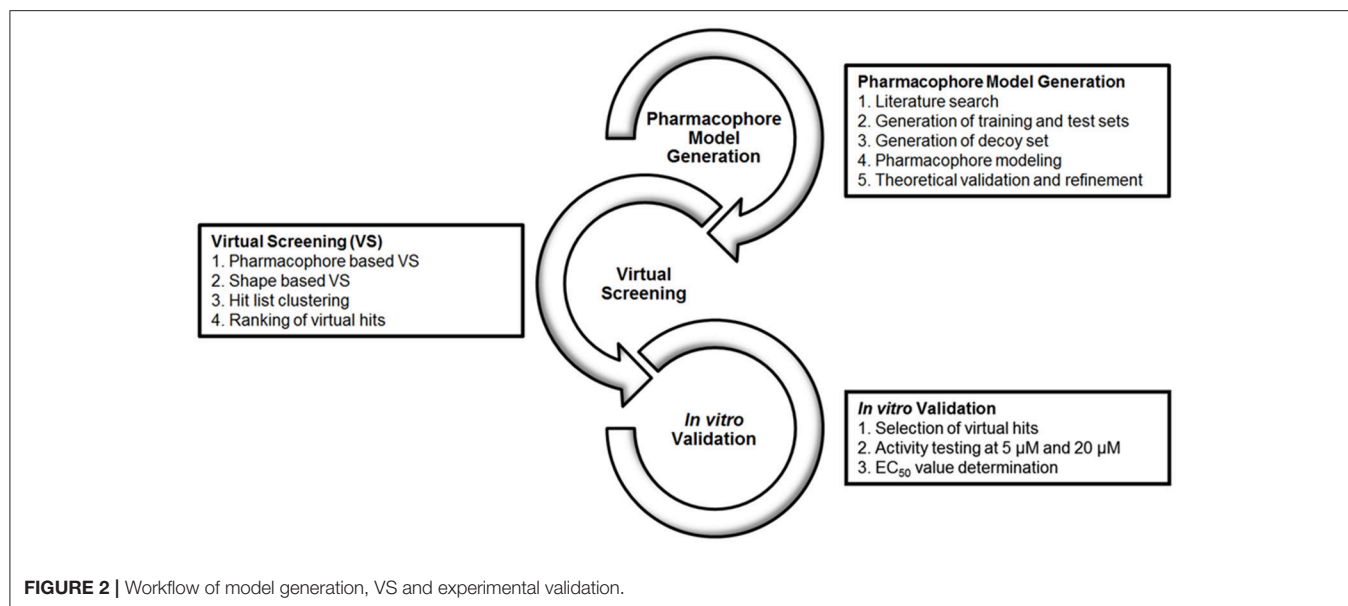
## RESULTS AND DISCUSSION

### Workflow

The workflow of this study is divided into 3 levels, as depicted in **Figure 2**: (1) Literature search for the compilation of a database of known GPBAR1 actives and inactives to be split and used as pharmacophore training set and a validation set for theoretical validation; The generation of a pharmacophore model collection with LigandScout and the subsequent theoretical validation. (2) VS of multi-conformational databases consisting of structures of natural and synthetic compounds (**Table 1**) using the two most promising models as queries; Evaluation of the hit list applying shape-based screening and physicochemical space clustering of virtual hits. (3) Selection of 33 virtual hits and their experimental validation in a HEK 293T cell based luciferase assay.

### Pharmacophore Modeling

A pharmacophore model is the abstract three dimensional representation of the molecular interactions between a target and a ligand structure, which is reduced to a collection of steric and chemical features that are necessary to trigger a desired pharmacological effect. The quality of a ligand-based pharmacophore model strongly relies on the selection of training set molecules. Therefore, it is mandatory to strictly select only highly potent activators for the training and validation sets (Seidel et al., 2010). In the case of GPBAR1, available bioactivity data were not only obtained by different working groups, but also with different cellular assays. This raised concerns about direct data comparison among the used assays. Only ligands



tested clinically or with a reported activity, which was proven to be both potent and directly comparable to respective positive controls, were therefore used in this study. Subsequently, 428

of 815 compounds had to be discarded. The remaining 338 compounds formed the “High Actives” dataset. 49 compounds were categorized as “True Inactives.” The data handling used

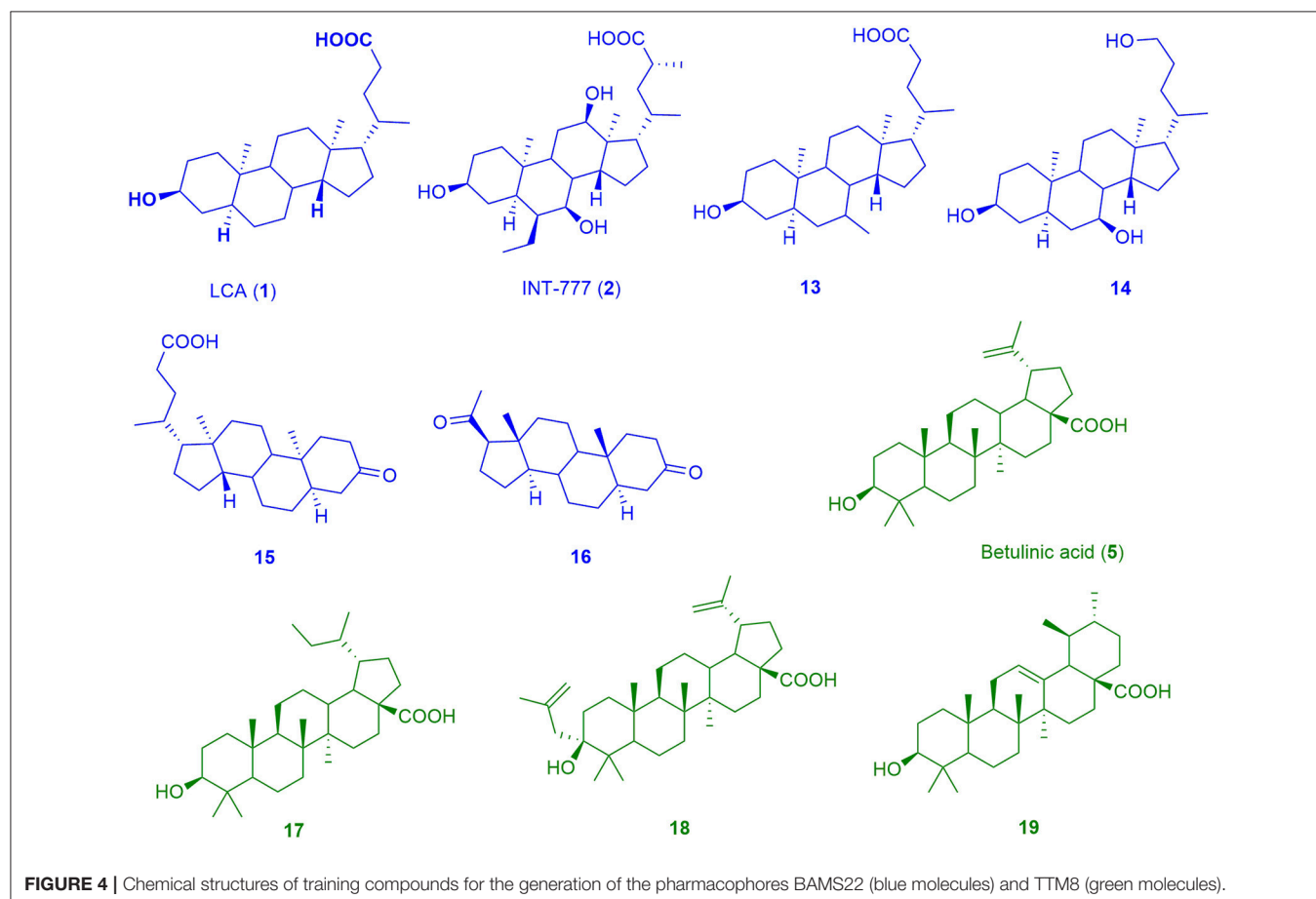
as basis for the generation of the ligand-based pharmacophore models is illustrated in **Figure 3**.

A pharmacophore model built of several query compounds binding at different ligand binding sites to the target protein would clearly distort the quality of such a model and devastate its predictive power. Therefore, the “High Actives” dataset compounds were divided into 12 clusters using the pharmacophore clustering tool implemented in LigandScout. Cluster 1 was discarded as it only consisted of one compound. The remaining 11 clusters contained between 11 and 80 molecules and were separated each into test and validation set.

Out of the retrieved 11 cluster sets, 12 pharmacophore models were generated. Altogether, in parallel screening, these models were able to predict 275 of 338 compounds (81%) in the “High Actives” database as true positives. However, a high number of false positives were retrieved, when the models were screened against the “Decoys” and “True Inactives” databases. This resulted in poor metrics of this entire model collection’s enrichment factor ( $EF = 11.22$ ). Two models, which were based on the pharmacophores of natural products, showed promising metrics and were selected for the prospective VS and experimental validation. The first model, BAMS22, was based on a training set of 6 molecules (depicted in **Figure 4**) resulting

from cluster 7. They had been selected for covering nearly the whole physicochemical space and for incorporating most of the structure-activity information contained in cluster 7. BAMS22 was used for VS of the “Decoys”/“True Inactives” ( $n = 18,112$ ) databases and the cluster 7 validation set ( $n = 20$ ), which resulted in a specificity of 1 (0.998785) and a sensitivity of 1, achieving an EF of 823.3. Along with the molecules from cluster 7, the potent ligand TLC from cluster 12 was retrieved as highly ranked virtual hit.

The BAMS22 model consists of two mandatory hydrophobic features, two mandatory HB acceptor features, an optional HB donor, an optional hydrophobic, and an optional negatively ionizable feature, as well as a rigid exclusion volume coat. In agreement to the TLC binding predictions and experiments of Gertzen and coworkers (Gertzen et al., 2015), our pharmacophore model, although not based on the homology model’s input information, depicts a very similar interaction pattern (**Figure 5**). Gertzen stated that the 3-hydroxyl moiety of TLC forms a HB to E169 and Y240, the sulfonic acid group forms a salt-bridge to R79 and hydrophobic interactions appear with L244. All of these statements were underlined with alanine-scanning experiments and are in accordance with our model. The model also suggests a second important HB interaction with the C-24 carboxamide group of TLC (**Figure 5C**), as well



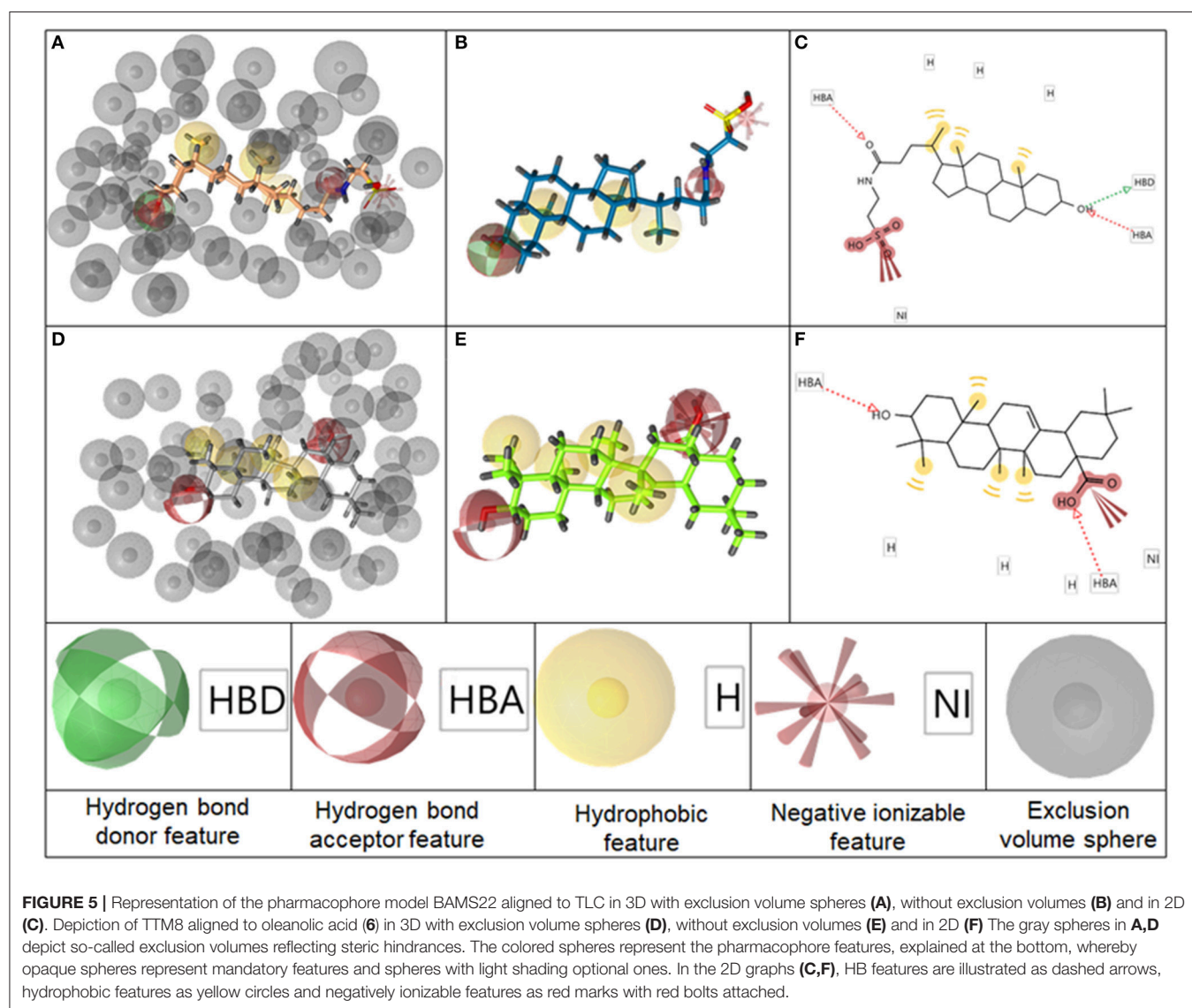
as with the C-24 hydroxyl group of **14**, or in the case of **16** with the C-20 keto group. The hydrophobic interactions were placed where hydrophobic alignment was possible. Although the model showed a very high specificity, it only consisted of four mandatory pharmacophore features, two HB accepting and two hydrophobic features, a widespread pattern of pharmacophore features. Therefore, not only steroid-like structures can putatively be retrieved in the prospective VS.

It is questionable whether triterpenes and bile acids share the same binding mode, as it was not possible to generate a restrictive pharmacophore model incorporating both scaffolds, although the binding modes appear to be very similar. It is likely that they have a different binding mode within the same binding position. Therefore, it was preferred to explain the steroidal structures with two highly specific local models and not with a single global model. It has previously been acknowledged for the identification of cyclooxygenase inhibitors that a set of highly specific local models leads to lower false positive hit

rates, compared to one pleiotropic global model (Schuster et al., 2010).

The second model, TTM8 is based on a training set of 4 molecules (Figure 4) from cluster 12. The model consists of 4 mandatory hydrophobic, two mandatory HB acceptor features, and a mandatory negatively ionizable feature (Figure 5). TTM8 was theoretically validated against the set of “Decoys”/“True Inactives” datasets ( $n = 18,112$ ) and the cluster 12 validation set ( $n = 16$ ), and showed a specificity of 1 and sensitivity of 0.81, achieving an EF of 919.5.

Genet and co-workers (Genet et al., 2010) were the first evaluating the SAR of triterpenes on the GPBAR1 receptor. They concluded that essential features for agonistic activity are a  $3\alpha$ -hydroxyl group, a carboxyl group in position  $17\alpha$ , and a rigid pentacyclic scaffold, in the best case a lupane backbone with high lipophilicity. Further publications regarding triterpenes are scarce, although some have shown higher selectivity over FXR and higher potency on the GPBAR1 than bile acids. Therefore,





**TABLE 2** | Results of the experimental validation of the virtual hits tested at 5  $\mu$ M and 20  $\mu$ M, the pharmacophore fit score (PF), the TC score (calculated in OpenEye ROCS with **13** as a query), the model with which they were predicted and their underlying database.

Compound	Database	predicted by BAMS22	predicted by TTM8	PF score	TC score	Activity (at 5 $\mu$ M)	Activity (at 20 $\mu$ M)
<b>20</b>	SPECs	<b>X</b>		45.31	0.873	2.34% ( $\pm$ 1.40)	1.74% ( $\pm$ 1.36)
<b>21</b>	SPECs	<b>X</b>		56.55	0.913	Not evaluable neither at 5 $\mu$ M nor 20 $\mu$ M, due to cytotoxic properties	
<b>22</b>	SPECs	<b>X</b>		46.23	0.703	1.42% ( $\pm$ 1.39)	-0.06% ( $\pm$ 0.39)
<b>23</b>	SPECs	<b>X</b>		56.31	0.720	1.69% ( $\pm$ 1.01)	11.00% ( $\pm$ 4.15)
<b>24</b>	SPECs	<b>X</b>		46.23	0.703	4.75% ( $\pm$ 1.51)	6.12% ( $\pm$ 2.01)
<b>25</b>	SPECs	<b>X</b>		45.24	0.627	1.16% ( $\pm$ 0.35)	0.70% ( $\pm$ 0.54)
<b>26</b>	SPECs	<b>X</b>		45.08	0.701	0.47% ( $\pm$ 1.82)	0.58% ( $\pm$ 0.89)
<b>27</b>	In house	<b>X</b>		46.39	0.737	4.70% ( $\pm$ 1.99)	60.85% ( $\pm$ 20.05)
<b>28</b>	In house	<b>X</b>		56.72	0.726	5.07% ( $\pm$ 1.36)	83.81% ( $\pm$ 12.00)
<b>29</b>	SPECs	<b>X</b>		46.23	0.783	1.54% ( $\pm$ 0.91)	1.65% ( $\pm$ 0.69)
<b>30</b>	SPECs	<b>X</b>		47.83	0.860	-0.60% ( $\pm$ 0.36)	0.92% ( $\pm$ 0.67)
<b>31</b>	SPECs	<b>X</b>		46.32	0.806	-0.31% ( $\pm$ 0.38)	0.88% ( $\pm$ 1.69)
<b>32</b>	SPECs	<b>X</b>		57.09	0.877	3.00% ( $\pm$ 0.58)	19.45% ( $\pm$ 6.00)
<b>33</b>	SPECs	<b>X</b>		56.50	0.801	0.98% ( $\pm$ 1.13)	1.00% ( $\pm$ 0.18)
<b>34</b>	SPECs	<b>X</b>		65.59	0.773	-0.02% ( $\pm$ 0.14)	9.27% ( $\pm$ 2.62)
<b>35</b>	SPECs	<b>X</b>		46.2	0.777	-0.06% ( $\pm$ 0.44)	3.10% ( $\pm$ 0.46)
<b>36</b>	SPECs	<b>X</b>		46.15	0.821	1.07% ( $\pm$ 0.10)	0.34% ( $\pm$ 0.18)
<b>37</b>	SPECs	<b>X</b>		46.28	0.838	0.33% ( $\pm$ 0.63)	3.87% ( $\pm$ 1.89)
<b>38</b>	SPECs	<b>X</b>		45,25	0.810	1.67% ( $\pm$ 2.54)	Not evaluable at 20 $\mu$ M, due to cytotoxic properties
<b>39</b>	SPECs	<b>X</b>		46.14	0.774	2.63% ( $\pm$ 0.69)	2.45% ( $\pm$ 1.42)
<b>40</b>	In house	<b>X</b>		46.16	0.783	-0.13 ( $\pm$ 0.31)	1.72% ( $\pm$ 0.90)
<b>41</b>	NPDB	<b>X</b>		65.87	0.664	3.73% ( $\pm$ 2.99)	4.35% ( $\pm$ 1.28)
<b>42</b>	TCM-DB		<b>X</b>	73.71	0.662	3.76% ( $\pm$ 1.67)	8.78% ( $\pm$ 3.87)
<b>43</b>	NPDB		<b>X</b>	73.16	0.815	3.23% ( $\pm$ 0.86)	26.28% ( $\pm$ 2.18)
<b>44</b>	NPDB		<b>X</b>	73.41	0.779	3.45% ( $\pm$ 4.66)	22.19% ( $\pm$ 15)
<b>45</b>	NPDB	<b>X</b>		55.93	1.222	5.20% ( $\pm$ 2.42)	22.58% ( $\pm$ 0.74)
<b>46</b>	drugbank	<b>X</b>		46.71	0.938	12.80% ( $\pm$ 6.60)	Not evaluable at 20 $\mu$ M, due to cytotoxic properties; at 15 $\mu$ M: 17.34% ( $\pm$ 0.91)
<b>47</b>	NPDB	<b>X</b>		56.19	0.954	Not evaluable neither at 5 $\mu$ M nor 20 $\mu$ M, due to cytotoxic properties	
<b>48</b>	NPDB	<b>X</b>		68.54	1.611	1.13% ( $\pm$ 0.77)	23.22% ( $\pm$ 5.29)
<b>49</b>	In house		<b>X</b>	72.75	0.896	1.99% ( $\pm$ 1.07)	12.08% ( $\pm$ 6.51)
<b>50</b>	In house		<b>X</b>	73.53	0.897	12.69% ( $\pm$ 5.33)	35.95% ( $\pm$ 2.37)
<b>51</b>	In house		<b>X</b>	72.75	0.864	-1.39% ( $\pm$ 1.17)	Not evaluable at 20 $\mu$ M, due to cytotoxic properties
<b>52</b>	In house		<b>X</b>	73.88	0.822	29.79% ( $\pm$ 10.79)	Not evaluable at 20 $\mu$ M, due to cytotoxic properties
Vehicle control (0.1% DMSO)						0%	
LCA ( <b>1</b> ) (10 $\mu$ M)						100%	
CDCA ( <b>3</b> ) (50 $\mu$ M)						66.69% ( $\pm$ 8.72)	

a cherry-picking pharmacophore model, highly sensitive to these pentacyclic triterpene acids, was created. It can be considered as a highly suitable filtering tool with a high applicability in *in silico* assisted NP research as previously reported e.g., for pharmacological profiling of secondary metabolites or target identification of NPs (Schuster, 2010; Waltenberger et al., 2011; Grienke et al., 2015; Kratz et al., 2016).

## Prospective Virtual Screening and Hit Selection

A prospective VS was performed with the two pharmacophores against over 350,000 molecules from nine different databases (Table 1). After removing duplicates, 1,069 virtual hits were obtained and clustered according to physicochemical diversity into 9 groups (Figure 6). As obvious from Figure 6, groups 1 and 2 differ from the other groups on a very early hierarchical level. The main structural difference of these two groups compared to the others is that they comprise synthetic compounds and NPs with aromatic rings, more conjugated double bonds and heteroatoms, while groups 3–9 consist of steroidal structures, reaching from cardenolides, pregnanes, bile acids to steroids and triterpenes. The most interesting molecules, in terms of structural diversity, are found in groups 1 and 2, as they comprise scaffolds dissimilar to the query molecules of the underlying pharmacophore models.

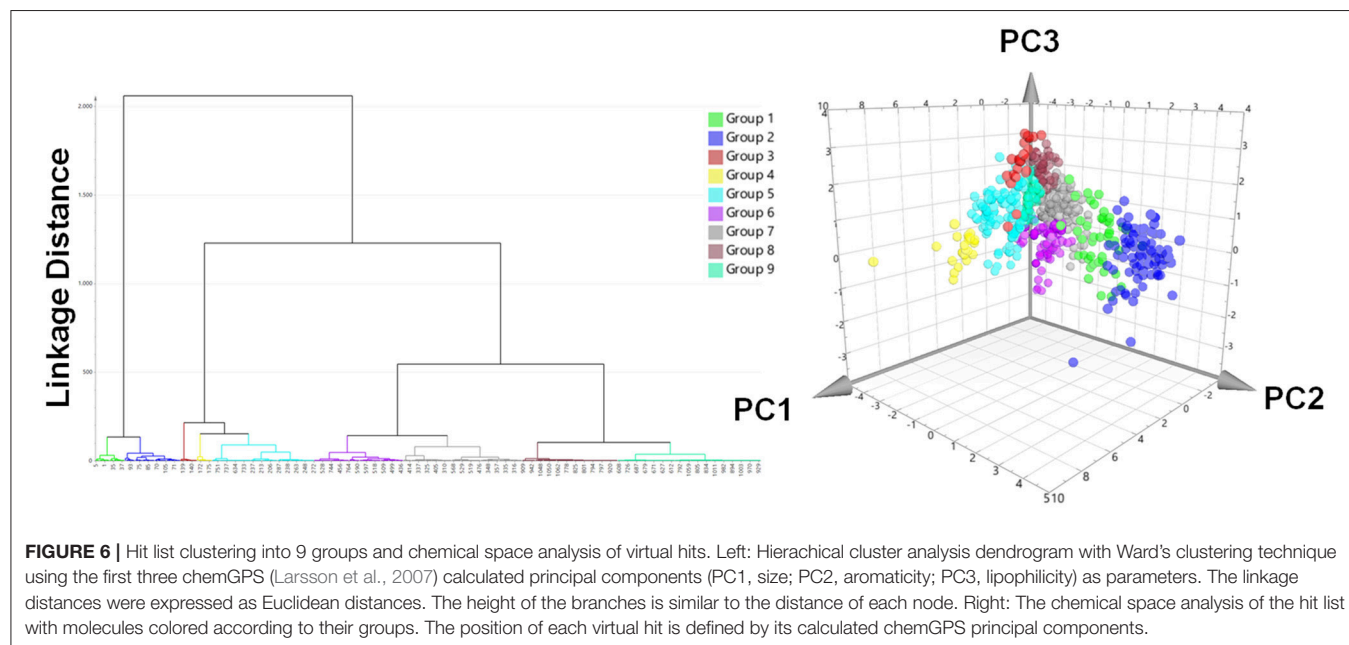
For prioritization of virtual hits to be experimentally tested a ranking was performed using shape-focused VS employing the ROCS Tanimoto Combo (TC) score (Hawkins et al., 2007; OpenEye, 2016). For this purpose best matching conformations derived from the pharmacophore-based VS were aligned with query molecule 13. Hit selection considered a high TC score, but also compound availability in sufficient purity, and structural variance. Finally, 33 compounds were subjected to experimental validation (Table 2): 11 compounds had been clustered in

groups 4–9 (Supplementary Figure 3), 9 compounds in group 1 (Supplementary Figure 1), and 13 compounds in group 2 (Supplementary Figure 2).

## Biological Evaluation

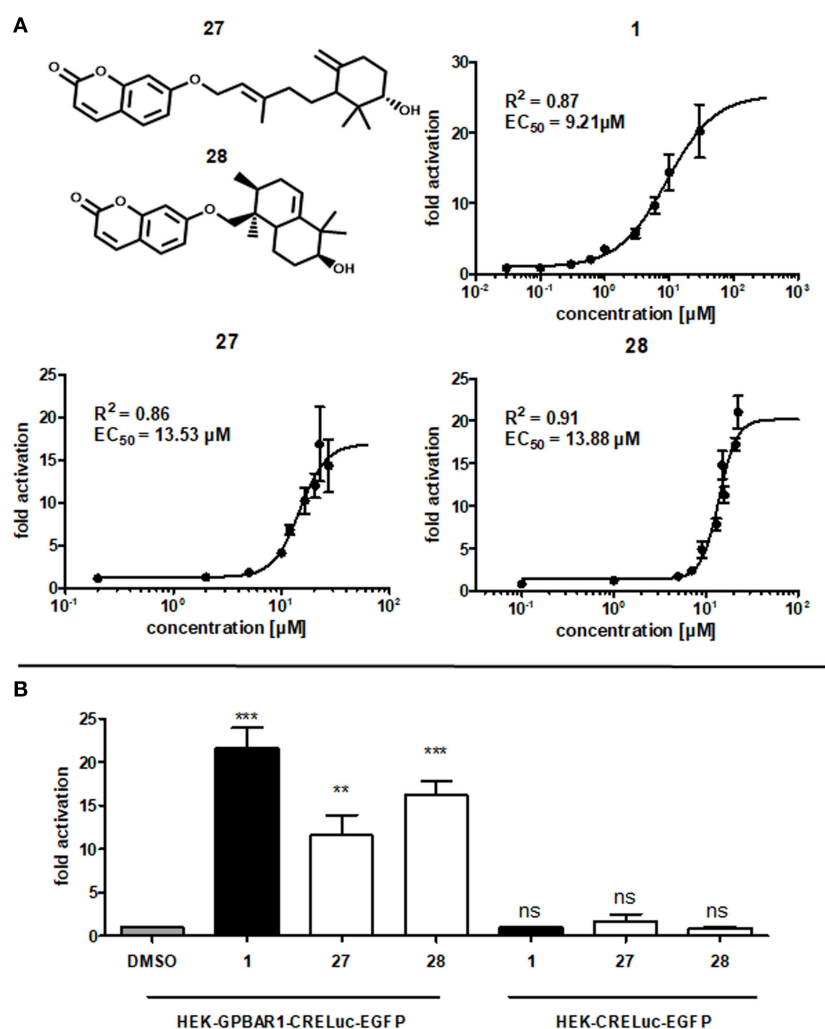
GPBAR1 activity of selected hits (Supplementary Figures 1–3) was determined in a reporter gene-based luciferase assay performed in HEK 293T cells. This assay assesses the upregulation of the cAMP-PKA-CREB pathway upon GPBAR1 activation and all conclusions are therefore limited to this receptor pathway. Compounds were considered active when they achieved at least 50% receptor activation. Compounds reaching at least 15% receptor activation were counted as weak activators. The response to 10  $\mu$ M LCA was set to 100% receptor activation. Vehicle control with a final dimethylsulfoxide (DMSO) concentration of 0.1% was set to 0% activation. Initially, compounds were tested at 2 concentrations, i.e., 5  $\mu$ M and 20  $\mu$ M. From the 33 compounds, only two (47 and 50) were cytotoxic in both concentrations tested. From the remaining 31 compounds, two showed significant activity with more than 50% receptor activation at 20  $\mu$ M and six further compounds achieved more than 15% receptor activation either at 5  $\mu$ M or 20  $\mu$ M (Table 2). Compounds 22, 24, 30, 34–36, and 41 were identified as potential pan-assay interference substance (PAINS) but none of them showed activity in the experimental validation. At 5  $\mu$ M, only one compound (52) achieved the arbitrary threshold of 15% receptor activation. Spironolactone (46) an approved drug for the treatment of heart failure, showed 17.3 % receptor activation at 15  $\mu$ M. Table 2 gives an overview of the experimental results.

As a result of this screening, the sesquiterpene coumarins 27 and 28 were discovered to be potent activators of the GPBAR1 receptor, which corroborated the scaffold-hopping competence of BAMS22. The two compounds are present in the gum resin



of *Ferula assa-foetida* L., used in central Asia as spice and medicine. The concentration response curves for **27** and **28** were determined and are shown in **Figure 7**. Compounds **27** and **28** were cytotoxic in transfected HEK cells at concentrations higher than 27.5 and 22.5  $\mu\text{M}$ , respectively. Due to this limitation the determination of  $E_{\text{max}}$  values could not be accurately determined in this assay. Accordingly, the analyzed fold activations and extrapolated  $EC_{50}$  values of **27** and **28** are limited to the non-cytotoxic concentration-response range and may not be completely accurate. Although limited by these constraints, farnesiferol B (**27**) showed  $10.54 \pm 2.25$  fold activation at 20  $\mu\text{M}$

( $60.85\% \pm 20.05$ ) and an  $EC_{50}$  of 13.53  $\mu\text{M}$ . Microlobidene (**28**) achieved  $16.21 \pm 1.64$  fold activation at 20  $\mu\text{M}$  ( $83.81\% \pm 12.00$ ) and an  $EC_{50}$  of 13.88  $\mu\text{M}$ . Whether the differences in sigmoidal slopes of LCA (**1**) and the newly identified ligands **27** and **28** are due to different interaction modes warrants further investigations. In the same assay the endogenous ligand CDCA only reached a fold activation of  $11 \pm 1.05$  at 50  $\mu\text{M}$ . The positive control LCA (**1**) reached  $18.59 \pm 0.97$  fold activation at 10  $\mu\text{M}$  and  $20.19 \pm 3.77$  at 30  $\mu\text{M}$ . The activity of compounds **27** and **28** can therefore be regarded as in the range of endogenous bile acids.



**FIGURE 7 | (A)** Concentration-dependent increase of GPBAR1 activity in response to LCA (**1**), farnesiferol B (**27**) and microlobidene (**28**). HEK-293 cells were transfected and stimulated as described in the Methods section. Luciferase activity was normalized to EGFP-derived fluorescence. Results are expressed as fold induction compared with the solvent control (DMSO, 0.1%) as the mean with SEM of at least three independent experiments. The two highest concentrations of **27** are the mean of two independent experiments. GraphPad Prism's non-linear regression with the sigmoidal dose response settings (variable slope) was used to calculate curves. **(B)** Fold activation of compounds **1** (10  $\mu\text{M}$ ), **27** and **28** (20  $\mu\text{M}$ ) in comparison to vehicle control 0.1% DMSO in (left) GPBAR1 transfected cells and (right) GPBAR1 untransfected cells. HEK 293T cells were transfected with GPBAR1, EGFP and CRE-Luc expression plasmids (left), or with EGFP and CRE-Luc expression plasmids only (right). Cells were treated for 18 h with 20  $\mu\text{M}$  of **27** and **28** as well as 10  $\mu\text{M}$  LCA (**1**) as positive control and 0.1% DMSO as vehicle control. Luciferase activity was normalized to EGFP-derived fluorescence. Results are expressed as fold induction compared with the solvent control (0.1% DMSO). All given values are the mean of at least 3 independent experiments and the variance is given as SEM. Significance was evaluated with one-way ANOVA-Bonferroni post-test (\*\* $p < 0.001$ ; \*\* $p < 0.01$ ; ns, not significant vs. vehicle control).

Many NPs are well-known PAINS or frequent hitters (Baell, 2016). In order to prevent such unintentional false-positive results, the experiments with the two GPBAR1-activating NPs have been repeated without transfecting GPBAR1. EGFP and CRE-Luc plasmids have been transfected as usual with the same concentrations. In these control experiments, none of the compounds showed a significant increase in luminescence values. In contrast to that, the increase in luminescence in GPBAR1 transfected cells in response to the positive control LCA (**1**), as well as to compounds **27** and **28**, was significant, confirming a direct interaction with GPBAR1 (Figure 7).

## CONCLUSION

The two presented 3D pharmacophore models have proven their quality as VS queries, both theoretically and experimentally. The combined computational and experimental efforts led to the successful identification of novel GPBAR1 agonists with unreported scaffolds derived both from nature (**27** and **28**) and from synthetic origin (**32**). They not only enlarge the chemical diversity of receptor activators, but can also be promising starting points for SAR and further optimization. It is also the first study reporting the activity of spironolactone (**46**) on GPBAR1, highlighting the possibility that already approved drugs may interact with GPBAR1. The elucidation of the mechanism underlying the GPBAR1 activation by these compounds may be an interesting starting point for further research. The physicochemical clustering process enabled a scaffold rich hit selection and a solid predictive power, with 6.5% correctly predicted strong activators and 18.8% weak activators, recommending the presented workflow for future works. The

## REFERENCES

- Akram, M., Waratchareeyakul, W., Haupenthal, J., Hartmann, R. W., and Schuster, D. (2017). Pharmacophore modeling and *in Silico/in Vitro* screening for human cytochrome P450 11B1 and cytochrome P450 11B2 inhibitors. *Front. Chem.* 5:104. doi: 10.3389/fchem.2017.00104
- Alemi, F., Kwon, E., Poole, D. P., Lieu, T., Lyo, V., Cattaruzza, F., et al. (2013). The TGR5 receptor mediates bile acid-induced itch and analgesia. *J. Clin. Invest.* 123, 1513–1530. doi: 10.1172/JCI64551
- Asgharpour, A., Kumar, D., and Sanyal, A. (2015). Bile acids: emerging role in management of liver diseases. *Hepatol. Int.* 9, 527–533. doi: 10.1007/s12072-015-9656-7
- Baell, J. B. (2016). Feeling nature's PAINS: natural products, natural product drugs, and pan assay interference compounds (PAINS). *J. Nat. Prod.* 79, 616–628. doi: 10.1021/acs.jnatprod.5b00947
- Bala, V., Rajagopal, S., Kumar, D. P., Nalli, A. D., Mahavadi, S., Sanyal, A. J., et al. (2014). Release of GLP-1 and PYY in response to the activation of G protein-coupled bile acid receptor TGR5 is mediated by Epac/PLC-epsilon pathway and modulated by endogenous H<sub>2</sub>S. *Front. Physiol.* 5:420. doi: 10.3389/fphys.2014.00420
- Bento, A. P., Gaulton, A., Hersey, A., Bellis, L. J., Chambers, J., Davies, M., et al. (2014). The ChEMBL bioactivity database: an update. *Nucleic Acids Res.* 42, D1083–D1090. doi: 10.1093/nar/gkt1031
- Chen, C. Y. (2011). TCM Database@Taiwan: the world's largest traditional chinese medicine database for drug screening *in silico*. *PLoS ONE* 6:e15939. doi: 10.1371/journal.pone.0015939
- D'Amore, C., Di Leva, F. S., Sepe, V., Renga, B., Del Gaudio, C., D'Auria, M. V., et al. (2014). Design, synthesis, and biological evaluation of potent dual agonists of nuclear and membrane bile acid receptors. *J. Med. Chem.* 57, 937–954. doi: 10.1021/jm401873d
- Evans, K. A., Budzik, B. W., Ross, S. A., Wisnoski, D. D., Jin, J., Rivero, R. A., et al. (2009). Discovery of 3-aryl-4-isoxazolecarboxamides as TGR5 receptor agonists. *J. Med. Chem.* 52, 7962–7965. doi: 10.1021/jm901434t
- Evers, A., Hessler, G., Matter, H., and Klabunde, T. (2005). Virtual screening of biogenic amine-binding G-protein coupled receptors: comparative evaluation of protein- and ligand-based virtual screening protocols. *J. Med. Chem.* 48, 5448–5465. doi: 10.1021/jm050090o
- Fiorucci, S., Distrutti, E., Ricci, P., Giuliano, V., Donini, A., and Baldelli, F. (2014). Targeting FXR in cholestasis: hype or hope. *Expert Opin. Ther. Targets* 18, 1449–1459. doi: 10.1517/14728222.2014.956087
- Floreani, A., and Mangini, C. (2018). Primary biliary cholangitis: old and novel therapy. *Eur. J. Intern. Med.* 47, 1–5. doi: 10.1016/j.ejim.2017.06.020
- Friedrich, N. O., de Bruyn Kops, C., Flachsenberg, F., Sommer, K., Rarey, M., and Kirchmair, J. (2017). Benchmarking commercial conformer ensemble generators. *J. Chem. Inf. Model.* 57, 2719–2728. doi: 10.1021/acs.jcim.7b00505
- Genet, C., Strehle, A., Schmidt, C., Boudjelal, G., Lobstein, A., Schoonjans, K., et al. (2010). Structure-activity relationship study of betulinic acid, a novel and selective TGR5 agonist, and its synthetic derivatives: potential impact in diabetes. *J. Med. Chem.* 53, 178–190. doi: 10.1021/jm900872z
- Gertzen, C. G., Spomer, L., Smits, S. H., Haussinger, D., Keitel, V., and Gohlke, H. (2015). Mutational mapping of the transmembrane binding site of the G-protein coupled receptor TGR5 and binding mode prediction of TGR5 agonists. *Eur. J. Med. Chem.* 104, 57–72. doi: 10.1016/j.ejmech.2015.09.024

study shows that the two models in combination are qualified for their application in the future assessments of a molecules' GPBAR1 activating profile, in particular for the assessment of NPs, as the models comprise scaffolds that are widespread in nature. This is particularly helpful for increasing our insight into the molecular mechanism of traditionally used herbal remedies with complex compositions of secondary metabolites. A fast appraisal of their pharmacological profile can give direction and fast-forward research (i.e., pinpointing most promising constituents), alongside reducing expenses.

## AUTHOR CONTRIBUTIONS

JR planned and supervised the study. BK and JK created the pharmacophore models under supervision of TL and JR. BK and JK performed the virtual screening and hit selection along with UG. BK and AL conducted biological experiments and analyzed the data under supervision of VD. The manuscript was written with contributions of all authors. All authors have given approval to the final version of the manuscript.

## ACKNOWLEDGMENTS

The authors thank OpenEye for providing the ROCS software free of charge.

## SUPPLEMENTARY MATERIAL

The Supplementary Material for this article can be found online at: <https://www.frontiersin.org/articles/10.3389/fchem.2018.00242/full#supplementary-material>

- Grienke, U., Kaserer, T., Pfluger, F., Mair, C. E., Langer, T., Schuster, D., et al. (2015). Accessing biological actions of *Ganoderma* secondary metabolites by *in silico* profiling. *Phytochemistry* 114, 114–124. doi: 10.1016/j.phytochem.2014.10.010
- Ha, H., Debnath, B., Odde, S., Bensman, T., Ho, H., Beringer, P. M., et al. (2015). Discovery of novel CXCR2 inhibitors using ligand-based pharmacophore models. *J. Chem. Inf. Model.* 55, 1720–1738. doi: 10.1021/acs.jcim.5b00181
- Hawkins, P. C., Skillman, A. G., and Nicholls, A. (2007). Comparison of shape-matching and docking as virtual screening tools. *J. Med. Chem.* 50, 74–82. doi: 10.1021/jm0603365
- Herbert, M. R., Siegel, D. L., Staszewski, L., Cayanan, C., Banerjee, U., Dharni, S., et al. (2010). Synthesis and SAR of 2-aryl-3-aminomethylquinolines as agonists of the bile acid receptor TGR5. *Bioorg. Med. Chem. Lett.* 20, 5718–5721. doi: 10.1016/j.bmcl.2010.08.014
- Hodge, R. J., and Nunez, D. J. (2016). Therapeutic potential of Takeda-G-protein-receptor-5 (TGR5) agonists. Hope or hype? *Diabetes Obes. Metab.* 18, 439–443. doi: 10.1111/dom.12636
- Jensen, D. D., Godfrey, C. B., Niklas, C., Canals, M., Kocan, M., Poole, D. P., et al. (2013). The bile acid receptor TGR5 does not interact with beta-arrestins or traffic to endosomes but transmits sustained signals from plasma membrane rafts. *J. Biol. Chem.* 288, 22942–22960. doi: 10.1074/jbc.M113.455774
- Kawamata, Y., Fujii, R., Hosoya, M., Harada, M., Yoshida, H., Miwa, M., et al. (2003). A G protein-coupled receptor responsive to bile acids. *J. Biol. Chem.* 278, 9435–9440. doi: 10.1074/jbc.M209706200
- Keitel, V., Donner, M., Winandy, S., Kubitz, R., and Haussinger, D. (2008). Expression and function of the bile acid receptor TGR5 in Kupffer cells. *Biochem. Biophys. Res. Commun.* 372, 78–84. doi: 10.1016/j.bbrc.2008.04.171
- Keitel, V., Görg, B., Bidmon, H. J., Zemtsova, I., Spomer, L., Zilles, K., et al. (2010). The bile acid receptor TGR5 (Gpbar-1) acts as a neurosteroid receptor in brain. *Glia* 58, 1794–1805. doi: 10.1002/glia.21049
- Keitel, V., and Haussinger, D. (2011). TGR5 in the biliary tree. *Dig. Dis.* 29, 45–47. doi: 10.1159/000324127
- Keitel, V., Reinehr, R., Gatsios, P., Rupprecht, C., Gorg, B., Selbach, O., et al. (2007). The G-protein coupled bile salt receptor TGR5 is expressed in liver sinusoidal endothelial cells. *Hepatology* 45, 695–704. doi: 10.1002/hep.21458
- Kratz, J. M., Mair, C. E., Oettl, S. K., Saxena, P., Scheel, O., Schuster, D., et al. (2016). hERG channel blocking ipecac alkaloids identified by combined *in silico-in vitro* screening. *Planta Med.* 82, 1009–1015. doi: 10.1055/s-0042-105572
- Kumar, D. P., Asgharpour, A., Mirshahi, F., Park, S. H., Liu, S., Imai, Y., et al. (2016). Activation of transmembrane bile acid receptor TGR5 modulates pancreatic islet alpha cells to promote glucose homeostasis. *J. Biol. Chem.* 291, 6626–6640. doi: 10.1074/jbc.M115.699504
- Kumar, D. P., Rajagopal, S., Mahavadi, S., Mirshahi, F., Grider, J. R., Murthy, K. S., et al. (2012). Activation of transmembrane bile acid receptor TGR5 stimulates insulin secretion in pancreatic beta cells. *Biochem. Biophys. Res. Commun.* 427, 600–605. doi: 10.1016/j.bbrc.2012.09.104
- Ladurner, A., Zehl, M., Grienke, U., Hofstadler, C., Faur, N., Pereira, F. C., et al. (2017). Allspice and clove as source of triterpene acids activating the G protein-coupled bile acid receptor TGR5. *Front. Pharmacol.* 8:468. doi: 10.3389/fphar.2017.00468
- Larsson, J., Gottfries, J., Muresan, S., and Backlund, A. (2007). ChemGPS-NP: tuned for navigation in biologically relevant chemical space. *J. Nat. Prod.* 70, 789–794. doi: 10.1021/np070002y
- Lavoie, B., Balemba, O. B., Godfrey, C., Watson, C. A., Vassileva, G., Corvera, C. U., et al. (2010). Hydrophobic bile salts inhibit gallbladder smooth muscle function via stimulation of GPBAR1 receptors and activation of K(ATP) channels. *J. Physiol.* 588(Pt 17), 3295–3305. doi: 10.1113/jphysiol.2010.192146
- Li, T., Holmstrom, S. R., Kir, S., Umetani, M., Schmidt, D. R., Kliever, S. A., et al. (2011). The G protein-coupled bile acid receptor, TGR5, stimulates gallbladder filling. *Mol. Endocrinol.* 25, 1066–1071. doi: 10.1210/me.2010-0460
- Londregan, A. T., Piotrowski, D. W., Futatsugi, K., Warmus, J. S., Boehm, M., Carpino, P. A., et al. (2013). Discovery of 5-phenoxy-1,3-dimethyl-1H-pyrazole-4-carboxamides as potent agonists of TGR5 via sequential combinatorial libraries. *Bioorg. Med. Chem. Lett.* 23, 1407–1411. doi: 10.1016/j.bmcl.2012.12.076
- Macchiarulo, A., Gioiello, A., Thomas, C., Pols, T. W., Nuti, R., Ferrari, C., et al. (2013). Probing the binding site of bile acids in TGR5. *ACS Med. Chem. Lett.* 4, 1158–1162. doi: 10.1021/ml400247k
- Mair, C. E., Grienke, U., Wilhelm, A., Urban, E., Zehl, M., Schmidtke, M., et al. (2018). Anti-influenza triterpene saponins from the bark of *Burkea africana*. *J. Nat. Prod.* 81, 515–523. doi: 10.1021/acs.jnatprod.7b00774
- Martin, R. E., Bissantz, C., Gavelle, O., Kuratli, C., Dehmlow, H., Richter, H. G., et al. (2013). 2-Phenoxy-nicotinamides are potent agonists at the bile acid receptor GPBAR1 (TGR5). *ChemMedChem* 8, 569–576. doi: 10.1002/cmdc.201200474
- Maruyama, T., Miyamoto, Y., Nakamura, T., Tamai, Y., Okada, H., Sugiyama, E., et al. (2002). Identification of membrane-type receptor for bile acids (M-BAR). *Biochem. Biophys. Res. Commun.* 298, 714–719. doi: 10.1016/S0006-291X(02)02550-0
- Mysinger, M. M., Carchia, M., Irwin, J. J., and Shoichet, B. K. (2012). Directory of useful decoys, enhanced (DUD-E): better ligands and decoys for better benchmarking. *J. Med. Chem.* 55, 6582–6594. doi: 10.1021/jm300687e
- OpenEye (2016). *ROCS Information*. Available online at: <http://www.eyesopen.com/rocs> (Accessed Jun 8, 2016).
- Park, E. J., Ahn, Y. G., Jung, S. H., Bang, H. J., Kim, M., Hong, D. J., et al. (2014). Discovery of novel pyrimidine and malonamide derivatives as TGR5 agonists. *Bioorg. Med. Chem. Lett.* 24, 4271–4275. doi: 10.1016/j.bmcl.2014.07.026
- Peeters, M. C., van Westen, G. J., Li, Q., and Ijzerman, A. P. (2011). Importance of the extracellular loops in G protein-coupled receptors for ligand recognition and receptor activation. *Trends Pharmacol. Sci.* 32, 35–42. doi: 10.1016/j.tips.2010.10.001
- Pellicciari, R., Gioiello, A., Macchiarulo, A., Thomas, C., Rosatelli, E., Natalini, B., et al. (2009). Discovery of 6 $\alpha$ -ethyl-23(S)-methylcholic acid (S-EMCA, INT-777) as a potent and selective agonist for the TGR5 receptor, a novel target for diabetes. *J. Med. Chem.* 52, 7958–7961. doi: 10.1021/jm901390p
- Pols, T. W., Nomura, M., Harach, T., Lo Sasso, G., Oosterveer, M. H., Thomas, C., et al. (2011). TGR5 activation inhibits atherosclerosis by reducing macrophage inflammation and lipid loading. *Cell Metab.* 14, 747–757. doi: 10.1016/j.cmet.2011.11.006
- Rollinger, J. M., Haupt, S., Stuppner, H., and Langer, T. (2004). Combining ethnopharmacology and virtual screening for lead structure discovery: COX-inhibitors as application example. *J. Chem. Inf. Comput. Sci.* 44, 480–488. doi: 10.1021/ci030031o
- Rollinger, J. M., Steindl, T. M., Schuster, D., Kirchmair, J., Anrain, K., Ellmerer, E. P., et al. (2008). Structure-based virtual screening for the discovery of natural inhibitors for human rhinovirus coat protein. *J. Med. Chem.* 51, 842–851. doi: 10.1021/jm701494b
- Sato, H., Genet, C., Strehle, A., Thomas, C., Lobstein, A., Wagner, A., et al. (2007). Anti-hyperglycemic activity of a TGR5 agonist isolated from *Olea europaea*. *Biochem. Biophys. Res. Commun.* 362, 793–798. doi: 10.1016/j.bbrc.2007.06.130
- Sato, H., Macchiarulo, A., Thomas, C., Gioiello, A., Une, M., Hofmann, A. F., et al. (2008). Novel potent and selective bile acid derivatives as TGR5 agonists: biological screening, structure-activity relationships, and molecular modeling studies. *J. Med. Chem.* 51, 1831–1841. doi: 10.1021/jm7015864
- Schuster, D. (2010). 3D pharmacophores as tools for activity profiling. *Drug Discov. Today Tech.* 7, e205–e211. doi: 10.1016/j.ddtec.2010.11.006
- Schuster, D., Laggner, C., Steindl, T. M., Paluszczak, A., Hartmann, R. W., and Langer, T. (2006). Pharmacophore modeling and *in silico* screening for new P450 19 (aromatase) inhibitors. *J. Chem. Inf. Model.* 46, 1301–1311. doi: 10.1021/ci050237k
- Schuster, D., Waltenberger, B., Kirchmair, J., Distinto, S., Markt, P., Stuppner, H., et al. (2010). Predicting cyclooxygenase inhibition by three-dimensional pharmacophoric profiling. Part I: model generation, validation and applicability in ethnopharmacology. *Mol. Inform.* 29, 75–86. doi: 10.1002/minf.200900071
- Seidel, T., Ibis, G., Bendix, F., and Wolber, G. (2010). Strategies for 3D pharmacophore-based virtual screening. *Drug Discov. Today Tech.* 7, e221–e228. doi: 10.1016/j.ddtec.2010.11.004
- Vaidehi, N., Bhattacharya, S., and Larsen, A. B. (2014). Structure and dynamics of G-protein coupled receptors. *Adv. Exp. Med. Biol.* 796, 37–54. doi: 10.1007/978-94-007-7423-0\_3
- Valli, M., dos Santos, R. N., Figueira, L. D., Nakajima, C. H., Castro-Gamboa, I., Andricopulo, A. D., et al. (2013). Development of a natural

- products database from the biodiversity of Brazil. *J. Nat. Prod.* 76, 439–444. doi: 10.1021/np3006875
- Vassileva, G., Golovko, A., Markowitz, L., Abbondanzo, S. J., Zeng, M., Yang, S., et al. (2006). Targeted deletion of Gpbar1 protects mice from cholesterol gallstone formation. *Biochem. J.* 398, 423–430. doi: 10.1042/BJ20060537
- Velazquez-Villegas, L. A., Perino, A., Lemos, V., Zietak, M., Nomura, M., Pols, T. W. H., et al. (2018). TGR5 signalling promotes mitochondrial fission and beige remodelling of white adipose tissue. *Nat. Commun.* 9:245. doi: 10.1038/s41467-017-02068-0
- Waltenberger, B., Schuster, D., Paramapojn, S., Gritsanapan, W., Wolber, G., Rollinger, J. M., et al. (2011). Predicting cyclooxygenase inhibition by three-dimensional pharmacophoric profiling. Part II: identification of enzyme inhibitors from Prasapjai, a Thai traditional medicine. *Phytomedicine* 18, 119–133. doi: 10.1016/j.phymed.2010.08.002
- Watanabe, M., Houten, S. M., Matak, C., Christoffolete, M. A., Kim, B. W., Sato, H., et al. (2006). Bile acids induce energy expenditure by promoting intracellular thyroid hormone activation. *Nature* 439, 484–489. doi: 10.1038/nature04330
- Wolber, G., and Langer, T. (2005). LigandScout: 3-D pharmacophores derived from protein-bound ligands and their use as virtual screening filters. *J. Chem. Inf. Model.* 45, 160–169. doi: 10.1021/ci049885e
- Woods, S. C., and D'Alessio, D. A. (2008). Central control of body weight and appetite. *J. Clin. Endocrinol. Metab.* 93(11 Suppl. 1), S37–S50. doi: 10.1210/jc.2008-1630
- Yu, D. D., Sousa, K. M., Mattern, D. L., Wagner, J., Fu, X., Vaidehi, N., et al. (2015). Stereoselective synthesis, biological evaluation, and modeling of novel bile acid-derived G-protein coupled bile acid receptor 1 (GP-BAR1, TGR5) agonists. *Bioorg. Med. Chem.* 23, 1613–1628. doi: 10.1016/j.bmc.2015.01.048

**Conflict of Interest Statement:** The authors declare that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

Copyright © 2018 Kirchweger, Kratz, Ladurner, Grienke, Langer, Dirsch and Rollinger. This is an open-access article distributed under the terms of the Creative Commons Attribution License (CC BY). The use, distribution or reproduction in other forums is permitted, provided the original author(s) and the copyright owner(s) are credited and that the original publication in this journal is cited, in accordance with accepted academic practice. No use, distribution or reproduction is permitted which does not comply with these terms.