

Spherical Harmonics Coefficients for Ligand-Based Virtual Screening of Cyclooxygenase Inhibitors

Quan Wang¹, Kerstin Birod², Carlo Angioni², Sabine Grösch², Tim Geppert³, Petra Schneider³, Matthias Rupp^{4,9}, Gisbert Schneider^{3,*}

1 Frankfurt Institute for Advanced Studies (FIAS), Goethe University, Frankfurt, Germany, **2** Institute for Clinical Pharmacology, Goethe University, Frankfurt, Germany, **3** Institute of Pharmaceutical Sciences, Swiss Federal Institute of Technology (ETH), Zürich, Switzerland, **4** Machine Learning Group, Technical University, Berlin, Germany

Abstract

Background: Molecular descriptors are essential for many applications in computational chemistry, such as ligand-based similarity searching. Spherical harmonics have previously been suggested as comprehensive descriptors of molecular structure and properties. We investigate a spherical harmonics descriptor for shape-based virtual screening.

Methodology/Principal Findings: We introduce and validate a partially rotation-invariant three-dimensional molecular shape descriptor based on the norm of spherical harmonics expansion coefficients. Using this molecular representation, we parameterize molecular surfaces, i.e., isosurfaces of spatial molecular property distributions. We validate the shape descriptor in a comprehensive retrospective virtual screening experiment. In a prospective study, we virtually screen a large compound library for cyclooxygenase inhibitors, using a self-organizing map as a pre-filter and the shape descriptor for candidate prioritization.

Conclusions/Significance: 12 compounds were tested *in vitro* for direct enzyme inhibition and in a whole blood assay. Active compounds containing a triazole scaffold were identified as direct cyclooxygenase-1 inhibitors. This outcome corroborates the usefulness of spherical harmonics for representation of molecular shape in virtual screening of large compound collections. The combination of pharmacophore and shape-based filtering of screening candidates proved to be a straightforward approach to finding novel bioactive chemotypes with minimal experimental effort.

Citation: Wang Q, Birod K, Angioni C, Grösch S, Geppert T, et al. (2011) Spherical Harmonics Coefficients for Ligand-Based Virtual Screening of Cyclooxygenase Inhibitors. PLoS ONE 6(7): e21554. doi:10.1371/journal.pone.0021554

Editor: Paul Wrede, Charité-Universitätsmedizin Berlin, Germany

Received: November 19, 2010; **Accepted:** June 3, 2011; **Published:** July 27, 2011

Copyright: © 2011 Wang et al. This is an open-access article distributed under the terms of the Creative Commons Attribution License, which permits unrestricted use, distribution, and reproduction in any medium, provided the original author and source are credited.

Funding: The authors received an academic MOE software license from Chemical Computing Group Inc., Montreal, Canada. M.R. acknowledges partial support from DFG grant MU 987/4-2 and the FP7-ICT programme of the European Community under the PASCAL2 network of excellence, ICT-216886. The funders had no role in study design, data collection and analysis, decision to publish, or preparation of the manuscript.

Competing Interests: The authors received an academic MOE software license from Chemical Computing Group Inc., Montreal, Canada. There are no patents, products in development or marketed products to declare. This does not alter the authors' adherence to all the PLoS ONE policies on sharing data and materials, as detailed online in the guide for authors.

* E-mail: gisbert.schneider@pharma.ethz.ch

9 These authors contributed equally to this work.

‡ Current address: Institute for Pure and Applied Mathematics (IPAM), University of California, Los Angeles, Los Angeles, California, United States of America

Introduction

Ligand-based virtual screening [1,2], quantitative structure-property and structure-activity relationships [3,4], and other concepts in computational medicinal chemistry are based on the similarity principle [5], which states that (structurally) similar compounds generally exhibit similar properties. Such methods require quantitative representations of molecules, usually in the form of chemical descriptors, i. e., computable numerical attributes in vector form [6].

Numerous molecular 3D-descriptors and alignment methods have been proposed. Examples include CoMFA (comparative molecular field analysis) [7], Randic molecular profiles [8], 3D-MoRSE code (3D-molecule representation of structures based on electron diffraction) [9], invariant moments and radial scanning and integration [10], radial distribution function descriptors [11], WHIM (weighted holistic invariant molecular descriptors) [12], length-to-breadth ratios [13], USR (ultrafast shape recognition,

based on statistical moments) [14], ROCS (rapid overlay of chemical structures, based on Gaussian densities) [15], VolSurf (volumes and surfaces of 3D molecular fields) [16], GETAWAY (geometry, topology, and atom weights assembly) [17], and shrink-wrap surfaces [18], to name just a few prominent representatives.

In computer graphics, several methods exist for the more general problem of comparing arbitrary 3D objects [19,20], including distribution-based shape histograms [21], the D2 shape descriptor [22], and, the scaling index method [23]; the view-based methods of extended Gaussian images [24], and the light field descriptor [25]; the surface decomposition-based methods of Zernike moments [26], REXT (radialized spherical extent function) [27], and spherical harmonics descriptors [28].

Spherical harmonics have been used in cheminformatics as a global feature-based parametrization method of molecular shape [28–38]. Their attractive properties with regard to rotations make them an intuitive and convenient choice as basis functions when searching in a rotational space [31]. A review article by

Venkatraman et al. [38] highlights applications of spherical harmonics to protein structure comparison, ligand binding site similarity, protein-protein docking, and virtual screening. Jakobi et al. [37] use spherical harmonics in their ParaFrag approach to derive 3D pharmacophores of molecular fragments. Recently, Ritchie and co-workers have applied the ParaSurf and ParaFit methodologies [32,33] (Cepos InSilico Ltd., Erlangen, Germany) in a virtual screening study on the directory of useful decoys (DUD) data set [39], which motivates 3D shape-property combinations specifically for flexible ligands [40]. The DUD data set was also used in a comparative analysis of the performance of various shape descriptors alone and in combination with property and pharmacophore features [41]. See the section on related methods for further discussion of spherical harmonics approaches.

In this work, we introduce a partially rotation-invariant descriptor of molecular shape based on spherical harmonics decomposition coefficients. The idea is to decompose the molecular surface using spherical harmonics and to use the norm of the decomposition coefficients as a description of molecular shape. In this, we take advantage of the fact that the norm of the coefficients does not change under rotation around the z -axis, which we align to the primary axis of the molecule. We retrospectively evaluate our descriptor, and prospectively apply it to screen for novel inhibitors of the enzymes cyclooxygenase-1 (COX-1) and cyclooxygenase-2 (COX-2). Particular focus is on the practical application of the virtual screening technique as an evaluation of its actual suitability for early-phase drug discovery.

Materials and Methods

Spherical harmonics

Let $l \in \{0, 1, 2, \dots\}$, let $|m| \leq l$, let (θ, ϕ) indicate spherical coordinates, and let P_l^m denote the Legendre polynomials [42]. The spherical harmonics [43] Y_l^m of order l (frequency, angular quantum number) and degree m (azimuthal quantum number),

$$Y_l^m(\theta, \phi) = \sqrt{\frac{2l+1}{4\pi} \frac{(l-m)!}{(l+m)!}} P_l^m(\cos \theta) e^{im\phi}, \quad (1)$$

form an orthonormal (with respect to integration over the unit sphere) and complete set of basis functions (Fig. 1). They are solutions to Laplace's differential equation $\nabla^2 \Psi = 0$ in spherical coordinates [44].

Any square-integrable spherical function $f(\theta, \phi)$ can be decomposed as

$$f(\theta, \phi) = \sum_{l=0}^{\infty} \sum_{m=-l}^l c_l^m Y_l^m(\theta, \phi), \quad (2)$$

with complex coefficients c_l^m . The spherical harmonics decomposition can be viewed as a generalization of the Fourier decomposition to three dimensions [45].

The coefficients c_l^m of an harmonic expansion can be found using the orthonormality property. Multiplying each side of Eq. 2 by the complex conjugate $\overline{Y_l^m}(\theta, \phi)$ and integrating over the sphere yields

$$c_l^m = \int_{\theta=0}^{\pi} \int_{\phi=0}^{2\pi} f(\theta, \phi) \overline{Y_l^m}(\theta, \phi) \sin(\theta) d\phi d\theta. \quad (3)$$

Small values of l correspond to low frequencies, and describe the overall low-resolution shape; higher values of l add finer, high-

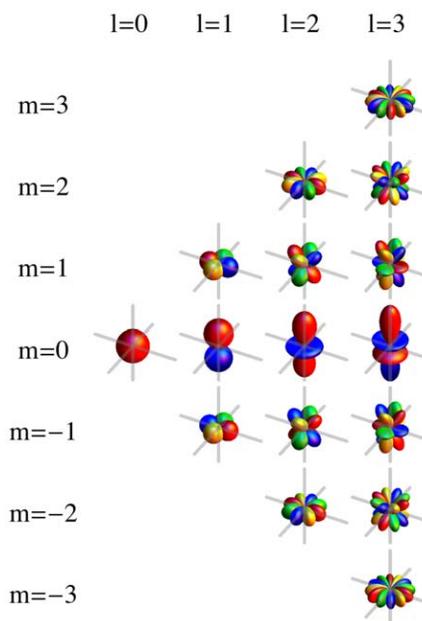


Figure 1. Spherical harmonics by order $l \leq 3$ (columns, left to right) and degree $m \in \{-1, \dots, l\}$ (rows, bottom to top). Shown are negative real (blue), positive real (red), negative imaginary (green), and positive imaginary (yellow) parts of Y_l^m . doi:10.1371/journal.pone.0021554.g001

frequency detail. The coefficients are unique, and can therefore be used as feature vectors for shape description.

Rotation of a molecule (its shape function f) changes the coefficients. A conventional solution is to define a canonical orientation of the molecule. For the purpose of shape comparisons, this implies an alignment of the compared molecules, with all associated problems and computational requirements. As an alternative, we use a partial orientation in conjunction with certain rotational invariance properties of the coefficients.

Descriptor definition

Let $\mathbf{X} = (\vec{x}_1, \dots, \vec{x}_n)^T \in \mathbb{R}^{n \times 3}$ denote the Cartesian coordinates of points $\vec{x}_1, \dots, \vec{x}_n \in \mathbb{R}^3$ sampled from a molecular surface. We assume that the surface is “star-like” (single-valued) in the sense that rays radiating outward from the molecule's origin intersect the surface only once (this is more of an issue for proteins; as argued elsewhere [35], small molecules are little, if at all, affected). Let $\mathbf{B} \in \mathbb{R}^{n \times k}$ denote the spherical harmonics basis functions Y_l^m evaluated at $\vec{x}_1, \dots, \vec{x}_n$, $0 \leq l \leq L$, $-l \leq m \leq l$, with $L \in \mathbb{N}$ the maximum order used and $k = \sum_{l=0}^L \sum_{m=-l}^l 1 = (L+1)^2$ the number of basis functions. The sampled molecular surface \mathbf{X} can be reconstructed using a matrix $\mathbf{C} \in \mathbb{R}^{k \times 3}$ of coefficients as $\mathbf{X} = \mathbf{B}\mathbf{C}$. The coefficient matrix is given by $\mathbf{C} = \mathbf{B}^{-1}\mathbf{X}$, where \mathbf{B}^{-1} denotes the pseudo-inverse [46] of \mathbf{B} .

Lemma. The p -norm of the rows of \mathbf{C} does not change under rotation around the z -axis (polar axis, change in ϕ).

Proof It is sufficient to consider a single coefficient, i.e., $n = k = 1$, and $\vec{c} = b^{-1} \vec{x} = (x, y, z)/b$. Here, $\vec{x} = (x, y, z) \in \mathbb{R}^3$ is a sampled surface point, $\vec{c} \in \mathbb{R}^3$ are coefficients, and $b \in \mathbb{R}$ is the spherical harmonics basis function Y_l^m . From Eq. 1, it is clear that if ϕ changes, only the part $\exp(im\phi)$ of the spherical harmonic basis function $Y_l^m(\theta, \phi)$ changes, while the rest of Y_l^m stays constant. Thus, $b = \exp(im\phi)d$ for some constant $d \in \mathbb{R}$ depending only on m, l, θ , but not on ϕ . Since $|\exp(im\phi)| = 1$, and thus $|b| = |d|$,

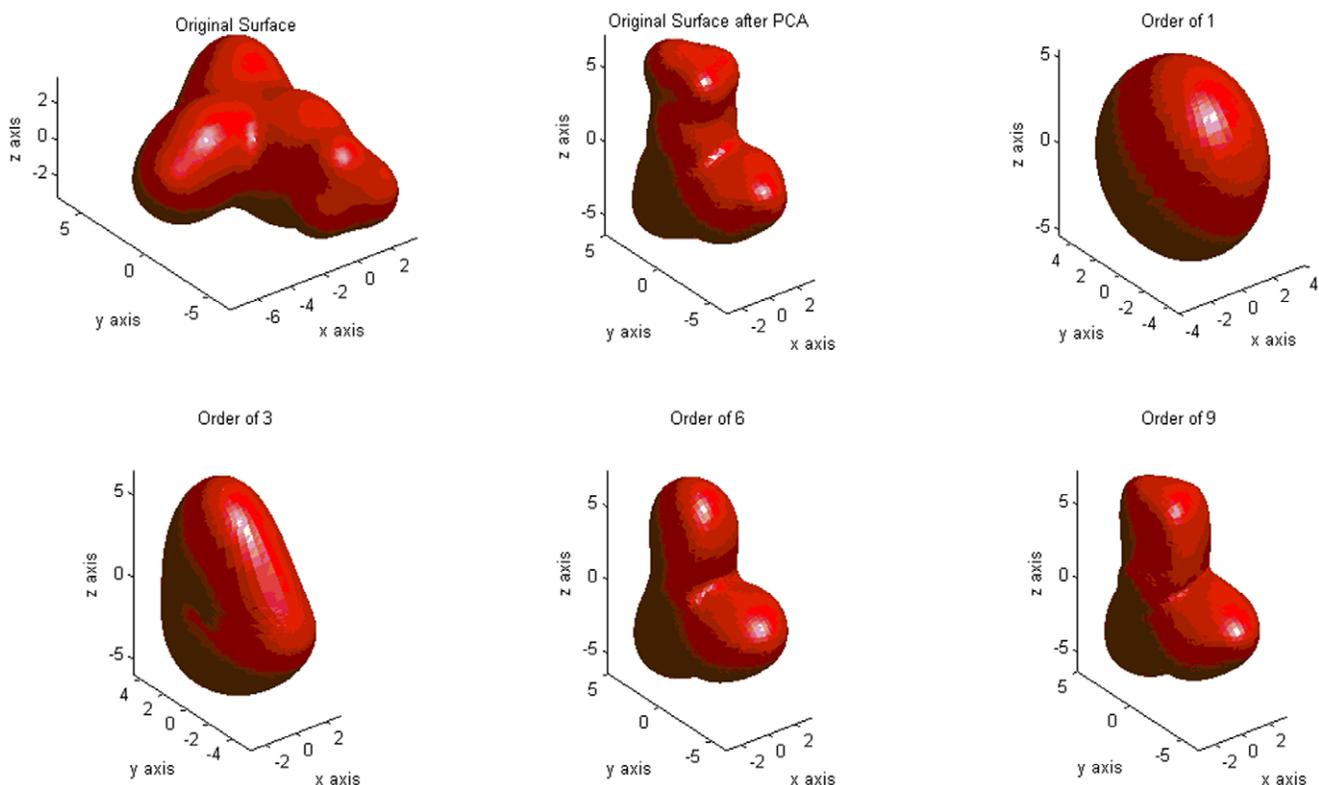


Figure 2. Surface reconstruction using spherical harmonics. Shown are the original surface (top left), the surface after alignment to the z -axis (top middle), and reconstructions using spherical harmonics of order l up to 1 (top right), 3 (bottom left), 6 (bottom middle), and 9 (bottom right). doi:10.1371/journal.pone.0021554.g002

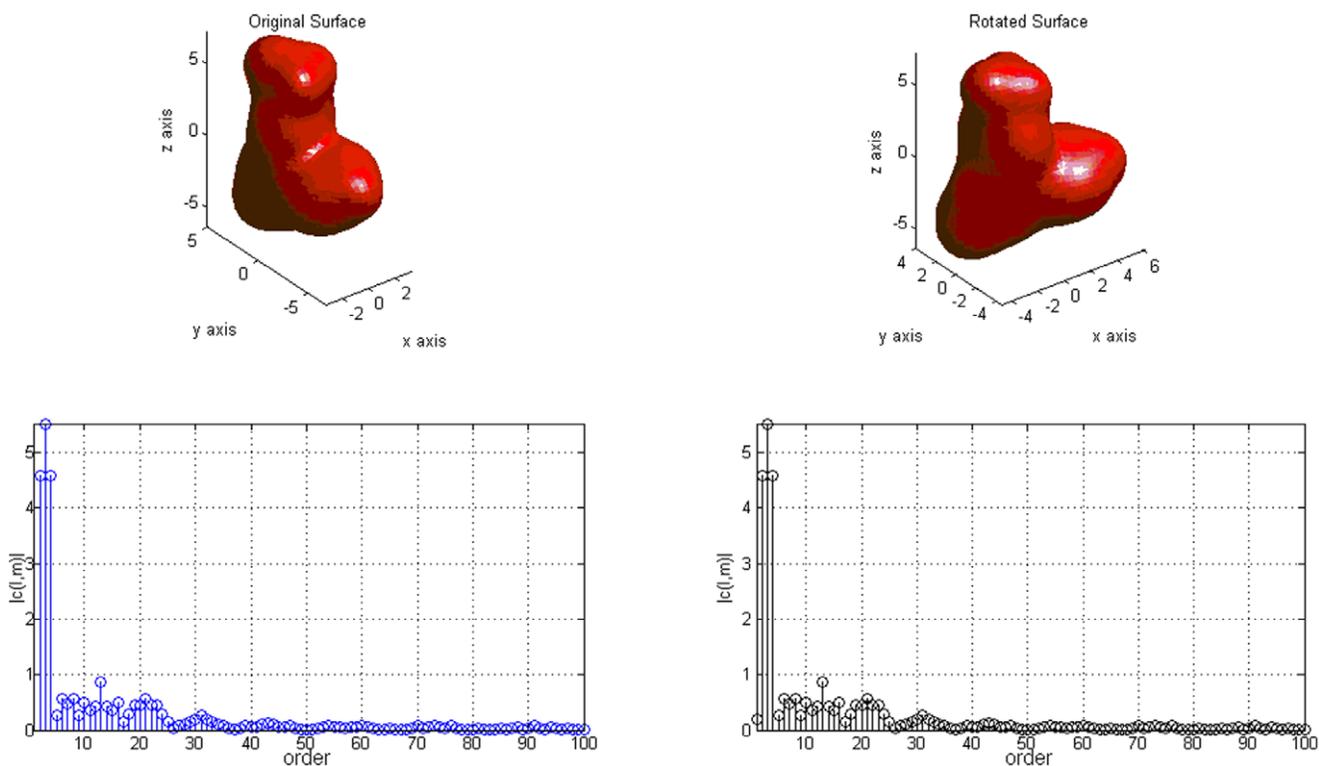


Figure 3. Spherical harmonics decomposition coefficients c_l^m of a molecular surface for $0 \leq l \leq 9$. The original (top left) and the rotated (top right) surfaces yield coefficients with identical norm (bottom), up to numerical noise (differences were below $2 \cdot 10^{-13}$). doi:10.1371/journal.pone.0021554.g003

$$\|\vec{c}\|_p = \left(\left| \frac{x}{b} \right|^p + \left| \frac{y}{b} \right|^p + \left| \frac{z}{b} \right|^p \right)^{1/p} = \frac{1}{|d|} \|\vec{x}\|_p. \quad (4)$$

After a rotation around the z -axis (a change in ϕ), the same holds for the rotated point $\vec{x}' = (x', y', z')$ and its coefficient \vec{c}' , i.e., $\|\vec{c}'\|_p = |d|^{-1} \|\vec{x}'\|_p$. Since the rotation matrix is unitary, $\|\vec{x}\|_p = \|\vec{x}'\|_p$, and it follows that $\|\vec{c}\|_p = \|\vec{c}'\|_p$.

Our spherical harmonics descriptor is the k -component vector

$$(\|\vec{c}_0^0\|_p, \|\vec{c}_1^{-1}\|_p, \|\vec{c}_1^0\|_p, \|\vec{c}_1^1\|_p, \|\vec{c}_2^{-2}\|_p, \dots, \|\vec{c}_L^L\|_p)$$

of the norms of the coefficients. It is a description of molecular shape that is invariant to rotations around the z -axis.

Before spherical harmonics decomposition, we place molecules into a common frame of reference by translating their center of gravity to the coordinate system origin and by aligning their first principle component (the direction of maximum variance as given by principle component analysis [47]) with the z -axis. In other words, we align molecules according to their longest spatial extent, and then apply our descriptor which is invariant to rotations around the z -axis.

Descriptor computation

Gaussian contact surfaces [48] of all compounds were computed using MOE (Molecular Operating Environment, version 2009.10, Chemical Computing Group Inc., Montreal, Canada, www.chemcomp.com). Spherical harmonics decomposition was then carried out on the vertices of these surfaces, giving approximate coefficients [49]. To limit computational expense, we truncated spherical harmonics expansions after order $L=9$. The resulting $k=(9+1)^2=100$ decomposition coefficients were sufficient to represent fine molecular detail and approximately reconstruct the original molecular surfaces (Fig. 2). The partial rotational invariance of the coefficient norms $\|\vec{c}^m\|$ is demonstrated in Fig. 3. Computation was done in Matlab (The MathWorks, version R2007a, www.mathworks.com), partly based on code by Dr. Andrew Hanna (University of East Anglia, United Kingdom, www.cmp.uea.ac.uk/~aih). Average computing time was <3 seconds per compound, which is acceptable for medium-sized libraries but will require speed-up for high-throughput virtual screening.

Related methods

Spherical harmonics have been widely used in cheminformatics as a global feature-based parametrization method of molecular shape [28–38]. Most current approaches, including ours, use the center of gravity as the center of the spherical harmonics decomposition. Molecular surface sampling can be done by sampling iso-probability surfaces of molecular property densities. One aspect in which methods differ is the way they deal with rotations in 3D space.

Ritchie and Kemp [31] apply the rotational property of spherical harmonics (a rotation of the surface can be simulated by rotating the expansion coefficients) to maximize the pairwise superposition of two molecules. The software ParaSurf superposes molecules using a brute-force rotational search over the three Euler rotation angles [50]. In a recent publication, Cai et al. [36] use a similar approach to obtain the minimal root-mean-square distance between a ligand molecule and a target protein. In these related studies, molecular surfaces were rotated by transforming their expansion coefficients.

Standard orientation of compounds prior to spherical harmonics decomposition was proposed by Morris et al. [34]. Their work registered molecules and binding pockets in a standard frame by translating their center of mass to the coordinate origin and aligning their variance-covariance matrix to the axes of the coordinate system. They then use the coefficients of a real spherical harmonics expansion to describe and compare the molecular shape of binding pockets and ligands. This approach aligns molecules to minimize rotation-dependent differences in the coefficients.

Rotation-invariant spherical harmonics descriptors were applied by Kazhdan et al. [28] and Mavridis et al. [35,51], using the fact that expansion coefficients of the same order l transform among themselves to construct rotationally invariant spherical harmonics coefficients $\|\sum_{m=-l}^l c_l^m\|_2$. In their approach, coefficients of the same order l are binned together, thereby losing information contained in the individual degrees m , but gaining complete rotational invariance.

In this work, we combine partial orientation of the molecules with the magnitude of the expansion coefficients as a partially rotation-invariant shape descriptor. Our proposed descriptor retains more information than the spherical harmonics descriptors by Kazhdan et al. [28] and Mavridis et al. [35,51] in the sense that coefficients within the same order are not summed up, but kept. Compared with standard orientation methods, our descriptor is potentially less susceptible to problems in the orientation step than most others because only the first (and most stable) principle component is used for orientation.

Retrospective evaluation

For retrospective validation, we ranked the compounds in a database according to their similarity to a reference compound, as measured by Euclidean distance and our descriptor. Two conceptually different collections of reference data were used, the DUD data set (release 2, from <http://dud.docking.org/r2>, unmodified data) [39], and the COBRA data set (version 10.3, 11 244 compounds annotated with activity on a total of 677 individual macromolecular targets) [52]. COBRA 10.3 contains 168 COX-2 inhibitors.

Gaussian contact surfaces were generated with the MOE 2009.10 (Molecular Operating Environment, Chemical Computing Group Inc., Montreal, Canada, www.chemcomp.com) GaussianSurface function, with parameter pos set to 'aPos a', rad 'dock_aRadius a', nearpos 'aPos a', neardist '5', maxMb '1', and fuzzy '0'. All other parameters were kept at their default values. Virtual screening experiments in COBRA were carried out using a single conformation generated by CORINA (version 2007, Molecular Networks GmbH, Erlangen, Germany, www.molecular-networks.com).

We used the selective COX-2 inhibitor SC-558 and the non-selective inhibitor indomethacin as queries for ligand-based similarity searching, with the conformations extracted from the crystal structure (protein data bank [53] identifiers (PDB ID) 6cox [54] and 4cox [54]). Enrichment factors [55], receiver operating characteristic curves (ROC curves [56]), and the area under these curves (ROC AUC) were used as performance measures.

Prospective virtual screening

We screened the ChemBridge compound pool (457 226 compounds, ChemBridge Corp., San Diego, USA, www.chembridge.com) for potential COX ligands using a single CORINA conformer query as in the retrospective screening. The database was preprocessed using the “washing” procedure in MOE (protonation of strong bases and de-protonation of strong acids; all other parameters were kept at their default values).

To reduce computational effort and allow for pharmacophore feature-based compound ranking, the screening compound pool was pre-filtered using a self-organizing map (SOM [57]) trained on the ChemBridge collection and 275 COX-1 and COX-2 inhibitors from the COBRA database. SOM topology was toroidal with 20×20 neurons (1 200 molecules per neuron on average); compounds were represented using the 150-dimensional CATS2D topological pharmacophore descriptor [58,59] and compared using the Manhattan distance. The initial width of the Gaussian neighborhood function was set to 5; training was terminated after $5 \cdot 10^6$ steps (using each compound 10 times on average). We used the MOLMAP software tool for SOM generation [60].

After pre-filtering, 21 950 compounds of the ChemBridge database that were similar to the COX inhibitors from the COBRA database were retained for virtual screening using our spherical harmonics shape descriptor. Two potent COX inhibitors served as reference molecules (queries; Fig. 4). All parameters were set to the values used in retrospective virtual screening. The spherical harmonics descriptor was calculated for the 21 950 retained molecules and for the two reference molecules.

Enzyme inhibition assay

Inhibition of COX-1 (ovine) and COX-2 (human recombinant) activity was measured using a COX inhibitor screening assay kit (Cayman Chemicals, Ann Arbor, MI, USA, www.caymanchem.com), according to the manufacturer's protocol. SC-560, a selective COX-1 inhibitor, and celecoxib, a selective COX-2 inhibitor, served as positive controls. The COX inhibitor screening assay directly measures the amount of prostaglandins PGE_2 , PGD_2 and $PGF_{2\alpha}$ produced by $SnCl_2$ reduction of COX-derived PGH_2 . In addition to this protocol, the amounts of prostaglandins were quantified by LC-MS/MS analysis as described previously [61].

Whole blood assay

COX-1 whole blood assay. One-milliliter heparinized human blood samples were incubated with $4 \mu\text{l}$ test substance (in DMSO) or $4 \mu\text{l}$ DMSO (control) for 10 min at 37°C . After this, thrombocyte aggregation was stimulated by addition of calcium ionophore A23187 ($50 \mu\text{M}$) for 30 min at 37°C . Plasma was separated by centrifugation for 20 min at 2000 rpm, 4°C and kept at -80°C until assayed for TXB_2 by LC-MS/MS (see below).

COX-2 whole blood assay. For the determination of COX-2 activity, 1 ml of heparinized human blood was incubated at 37°C with $10 \mu\text{l}$ of acetylsalicylic acid (1 mg ml^{-1} in PBS), $4 \mu\text{l}$ DMSO or $4 \mu\text{l}$ inhibitor (in DMSO) for 10 min. After this, $2 \mu\text{l}$ of LPS (5 mg ml^{-1} in DMSO) was added and incubated for 24 hr at 37°C . The reaction was terminated by quickly chilling on ice. Plasma was separated by centrifuging (20 min, 2000 rpm, 4°C), stored at -80°C until analysis of prostaglandins by LC-MS/MS within two weeks.

Extraction procedure. $250 \mu\text{l}$ plasma was incubated with $600 \mu\text{l}$ $45 \text{ mM H}_3\text{PO}_4$, $100 \mu\text{l}$ 0.15 M EDTA , $10 \mu\text{l}$ BHT (butylated hydroxytoluene, 2 mg ml^{-1}), $20 \mu\text{l}$ MeOH, $20 \mu\text{l}$ internal standard PGE_2-d4 (25 ng ml^{-1}), PGD_2-d4 (25 ng ml^{-1}), $PGF_{2\alpha}-d4$ (10 ng ml^{-1}), $6k$ $PGF-1 \alpha-d4$ (10 ng ml^{-1}), TXB_2-d4 (25 ng ml^{-1}) for 1 min, and passed through a ABS ELUT-Nexus cartridge (Varian, Darmstadt, Germany) preconditioned with methanol (1 ml), followed by distilled water (1 ml). The cartridge was washed with distilled water (1 ml) and 30% MeOH (1 ml). PGE_2 , $PGF_{2\alpha}$, $6-keto-PGF_{1\alpha}$ and TXB_2 were eluted with hexane-ethylacetate-isopropranolol (30:65:5, v/v, 1 ml). After vaporating the solvent under nitrogen atmosphere, the residue was

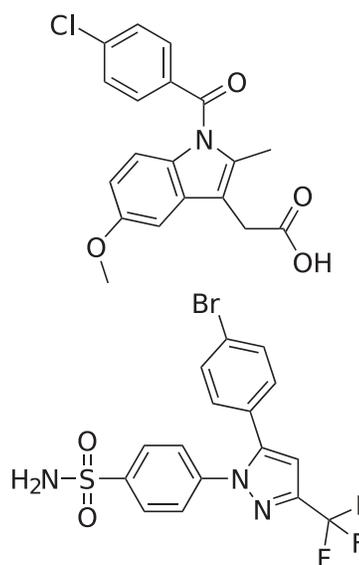


Figure 4. Reference COX inhibitors used for prospective screening with the shape descriptor. Indomethacin (top, PDB ID 4cox), a non-selective COX inhibitor, and, SC-558 (bottom, PDB ID 6cox), a selective COX-2 inhibitor.

doi:10.1371/journal.pone.0021554.g004

reconstituted in $50 \mu\text{l}$ acetonitrile / formic acid. PGE_2 concentrations were quantified by means of a validated LC-MS/MS assay described previously [61]. The lower limit of quantification was 10 pg ml^{-1} .

Results and Discussion

We validated our spherical harmonics (SpH) descriptor in a retrospective setting (statistical validation on known data), and in a prospective study to obtain biochemical confirmation of our model.

Retrospective evaluation

As a first analysis, we used the DUD compound collection for a preliminary comparison of selected shape- and structure-based virtual screening methods. ROC AUC [62] values were computed

Table 1. Results (ROC AUC) of retrospective virtual screening of DUD data set for COX-2 inhibitors.

Method	Single query		DUD COX-2	
	SC558	Indomethacin	average \pm std.dev.	
CATS2D	0.76	0.49	$0.63 \pm$	0.14
LIQUID v.1	0.59	0.60	$0.58 \pm$	0.10
LIQUID v.2	0.80	0.61	$0.74 \pm$	0.12
PRPS	0.83	0.15	$0.76 \pm$	0.19
SpH	0.91	0.84	$0.86 \pm$	0.12
ShaEP	0.89	0.79	$0.61 \pm$	0.03^a
ROCS	n.d.	n.d.	0.68^a	

^aValues reproduced from the original study by Vainio *et al.* [68], with ShaEP and ROCS run in 'shape-only' mode.

DUD COX-2 data: 426 actives, 13 289 decoys. Query SC558 has PDB ID 6cox; indomethacin has PDB ID 4cox. n.d. = not determined.

doi:10.1371/journal.pone.0021554.t001

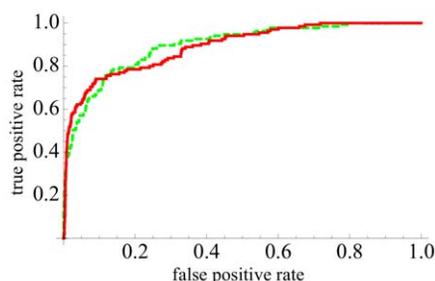


Figure 5. Receiver operating characteristic (ROC) curves for virtual screening by ranking against the COX-2 ligand SC-558 (PDB ID 6cox). Shown are curves for shapelets (solid red line), and spherical harmonics descriptor (dashed green line). doi:10.1371/journal.pone.0021554.g005

for each of the methods compared. ROC AUC values lie in the interval $[0,1]$, with values closer to 1 indicating higher enrichment of actives in a ranked list of compounds. The analysis was limited to the original COX-2 data from DUD (426 actives, 13 289 decoys). We did not perform exhaustive comparative analyses of virtual screening performance or focus on ‘early recognition’ of actives [63,64], as the primary purpose of this study was to determine whether our SpH descriptor might be a useful shape-based filtering criterion for COX inhibitors. Retrospective screening was restricted to COX-2, our original target.

Table 1 summarizes the results obtained for CATS2D (topological pharmacophore descriptor [58]), LIQUID (three-dimensional pharmacophore descriptor using Gaussian feature

points; v1: hydrogen-bond donors, hydrogen-bond acceptors, lipophilic [65]; v2: additional aromatic, positive and negative charge features (manuscript in preparation)), PRPS (Gaussian pseudoreceptor model [66,67]), ShaEP (field-based subgraph matching [68]), and ROCS (Gaussian shape model [15,69]). For the DUD COX-2 data, ROC AUC values indicate better than random performance for all methods. SpH yielded an average of 0.86, which compares to Ritchie’s ParaFit spherical harmonics descriptor (note that the ParaFit ROC AUC value is not given in the original publication; we estimated it from graphical material provided in the article’s supplementary material [41]). Among the tested methods, SpH performed best for the selective COX-2 inhibitor SC-558 (Fig. 4) yielding a ROC AUC = 0.91. Notably, high values were also obtained for indomethacin (Fig. 4), a non-selective COX inhibitor (COX-1 IC_{50} = 18 nM; COX-2 IC_{50} = 26 nM) [70]. Apparently, only the PRPS pseudoreceptor model distinguished between the selective (ROC AUC = 0.83) and the non-selective (ROC AUC = 0.15) query.

In contrast to DUD (unmodified data), the COBRA database contains only druglike bioactive compounds. Ranking of the COBRA database with SC-558 as query resulted in an enrichment factor (computed for the first percentile) of 23. We compared this result to those obtained by the shapelets [71] method from our group, using the same version of the COBRA database and the same reference structure. The shapelets shape-only virtual screening method achieved a comparable enrichment factor of 24. ROC curves are presented in Fig. 5 (numbers for shapelets refer to COBRA version 8.4 containing 8 311 compounds including 136 COX-2 inhibitors).

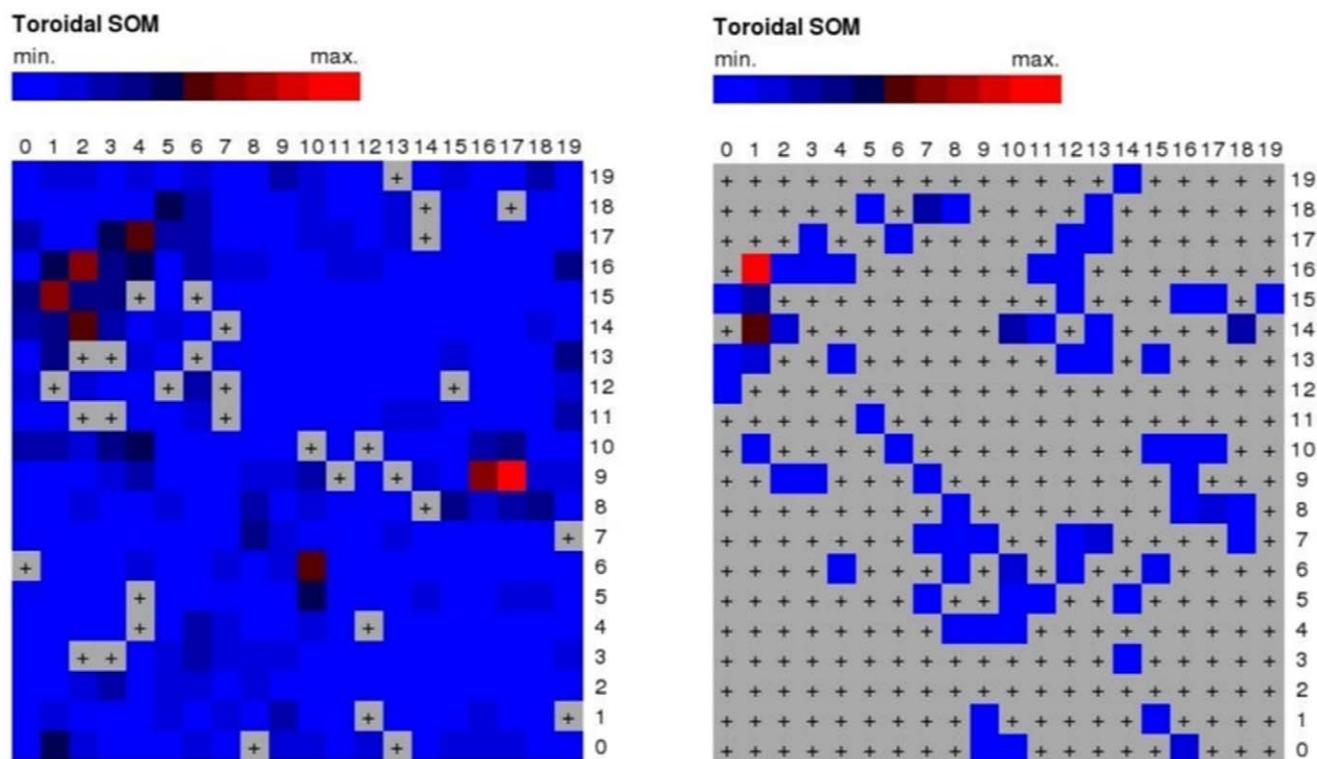


Figure 6. Self-organizing map projection of the ChemBridge database in CATS topological pharmacophore space, using a 20×20 toroidal grid. Colors correspond to the number of compounds clustered, separately scaled for each plot, with + indicating empty neurons. The left panel presents the distribution of the 457 226 compounds from the ChemBridge database, the right panel shows the 275 COX ligands from the COBRA database. doi:10.1371/journal.pone.0021554.g006

In summary, our spherical harmonics coefficients-based approach SpH achieves notable enrichment of actives and seems suitable for COX-2 inhibitor retrieval. This outcome is in agreement with the study of shape-based virtual screening approaches by Ritchie et al. [41], who report high hit rates for COX-2 using shape descriptors. We conclude that spherical harmonics-based decomposition of molecular shape captures structural features that are relevant for virtual screening. Due to the limited number of published prospective applications [72], it seems premature to render any conclusion regarding certain implementation preferences or 'best-in-class' spherical harmonics methods. To further assess our SpH approach, we performed a prospective study using SpH in a virtual screening cascade with the aim to identify new COX inhibitors.

Prospective virtual screening

Virtual screening. We used a SOM to pre-select potential COX inhibitors from the screening compound pool. The SOM (Fig. 6) of COX activity islands contains six neurons with more than three ligands (neurons (1,16), (1,14), (1,15), (7,18), (18,14), (10,14) with 49, 25, 15, 14, 12, 11 ligands, respectively). We selected all compounds from the ChemBridge database contained in these neurons, 21 950 in total (5% of the pool).

In the second virtual screening step, SpH was used for shape-based filtering. Two reference molecules (SC-558 and indomethacin; Fig. 4) resulted in two ranked lists of the pre-filtered ChemBridge compounds. 10 duplicates were found among the 50 top-ranking compounds from the two lists (20% overlap). In total, 12 compounds were selected by visual inspection, preferring potentially new scaffolds ('cherry-picking', Fig. 7), and submitted

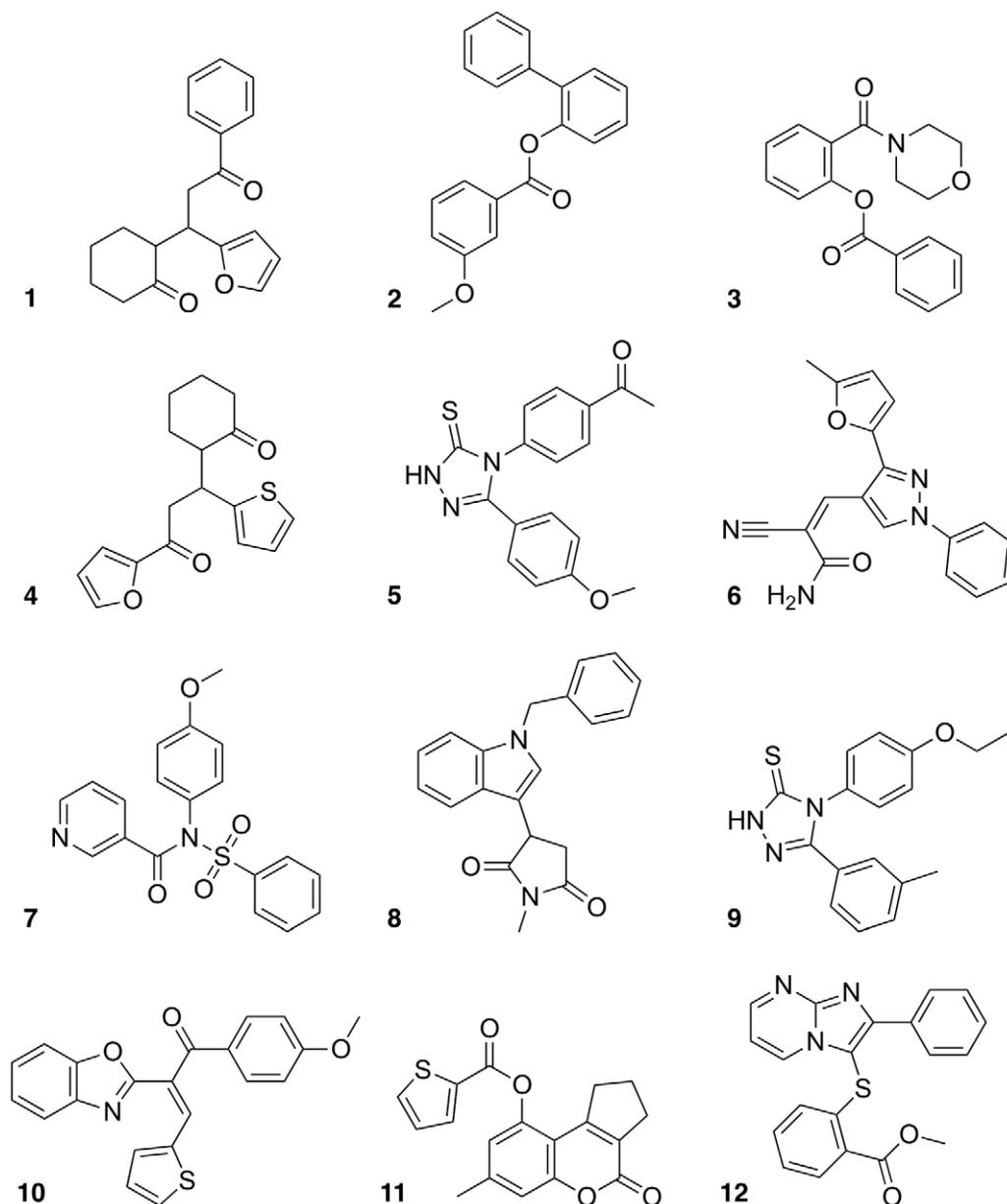


Figure 7. Compounds selected for the COX inhibition assay.

doi:10.1371/journal.pone.0021554.g007

for activity determination in a direct enzyme inhibition and a whole blood assay.

Assay results

We determined the COX-inhibitory activity of 12 compounds by performing a commercially available competitive COX-inhibition assay using purified COX-1 (ovine) and COX-2 (human recombinant) enzymes. Compounds **5** and **9** inhibit COX-1 in a concentration dependent manner (Fig. 8 and Table 2). At 100 μM compounds **5** and **9** inhibit COX-1 activity to 72% and 89%, respectively. Both compounds have only marginal effects on COX-2-activity at concentrations up to 100 μM . All other substances have no effect on COX-1 or COX-2 activity in this *in vitro* assay. While this outcome supports our general virtual screening approach, we failed to retrieve COX-2 inhibitors. This might be a consequence of using the selective COX-2 inhibitor SC-558 in combination with the non-selective COX inhibitor indomethacin as queries for the spherical harmonics shape filter. Apparently, the COX activity island on the SOM and SpH consensus filtering eliminated COX-2 specific features. It is also possible that there were no hitherto unidentified COX-2 ligands in the compound pool.

In the whole blood assay (Fig. 9, Table 3), compounds **5** and **9** are less effective, with maximum COX-1 inhibition of about 30% and no COX-2 inhibitory efficacy. Interestingly, in this assay, compounds **6**, **10**, **2** and **8** inhibit TXB_2 production in a concentration dependent manner up to 72%, 72% and 61% at 100 μM , respectively. Compounds **6** and **10** have only marginal inhibitory potency on PGE_2 production, which points to selective

COX-1 inhibitors *in vivo*. Compound **2** also inhibits PGE_2 production comparable to TXB_2 , indicating that this compound is a COX-unselective inhibitor. In contrast, substance **8** increases the PGE_2 amount in a concentration dependent manner, which argues for an activator of PGE_2 production in the cellular context. All other compounds show only very weak or no effect on PGE_2 production.

The inhibitory data obtained from the whole blood assay might be meaningful for further hit optimization. Compounds that are active in this assay are not snatched away by binding to serum albumin, but cross the cell membrane and overcome possible interactions with cellular substances or enzymes. This could explain why compounds **5** and **9** are active in the enzyme assay, but inactive in the whole blood assay. In contrast, compounds **6**, **10**, **2** and **8**, which were more active in the whole blood assay, possibly interact with the arachidonic acid pathway in other ways than direct inhibition of COX-1 or COX-2. Also, these compounds might be metabolized by cellular enzymes to more active derivatives, but this hypothesis needs to be tested by further experiments. Compound **8** is of special interest, as it induces PGE_2 production up to 322%. This increase could be due to an activation of enzyme activity, possibly by binding to the “inactive” monomer of the COX-homodimer complex [73,74], or, due to an enhancement of COX-2 protein, either by transcriptional or post-transcriptional mechanisms.

As a preliminary novelty check, similarity searches were performed using SciFinder Web (2010-10-21) for data retrieval from the CAS database (Chemical Abstracts Service, Columbus, Ohio, USA; www.cas.org). For none of the actives any reference to COX inhibition was found, and only for compound **9** substructure

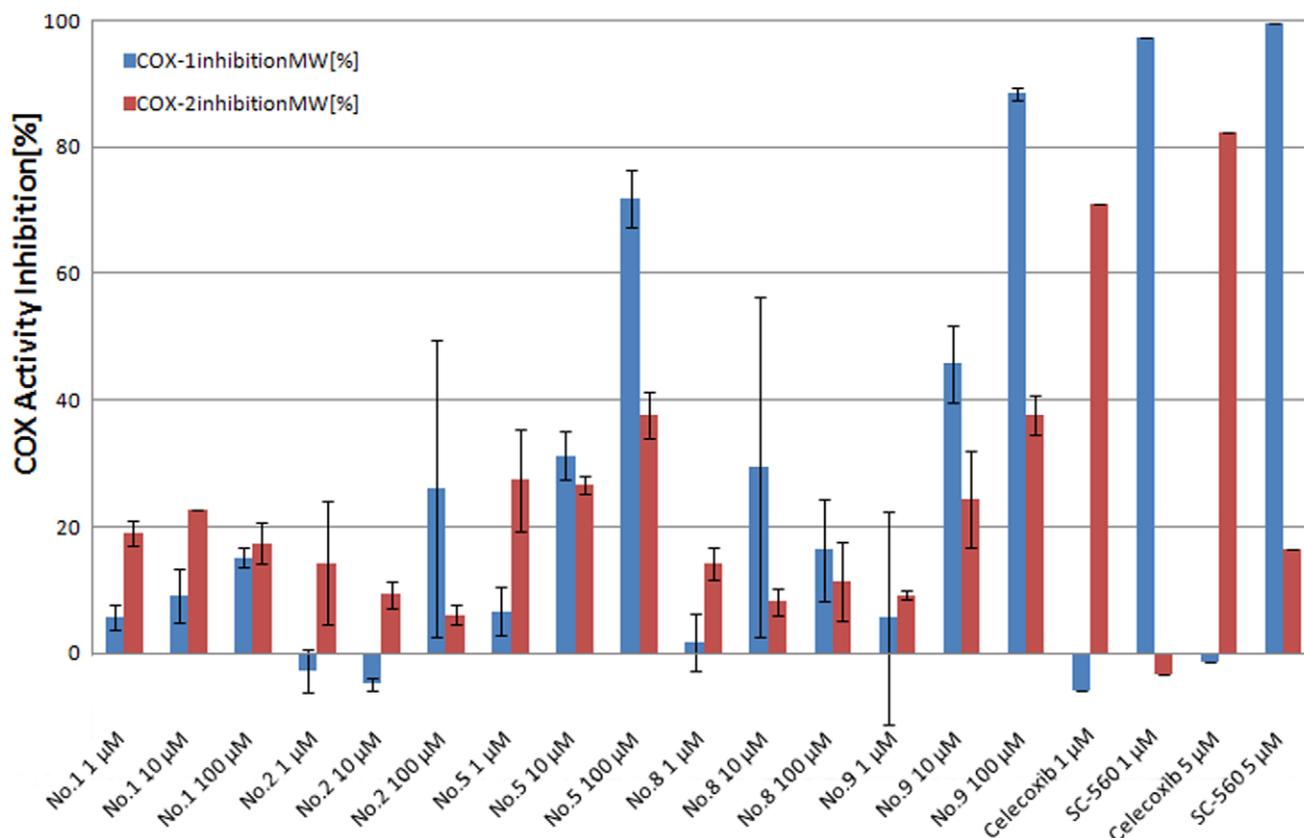


Figure 8. COX-2 inhibition *in vitro* assay results. Shown are COX-1 (blue) and COX-2 (red) inhibition. Celecoxib and SC-560 are known inhibitors selective for COX-2 and COX-1, respectively. doi:10.1371/journal.pone.0021554.g008

Table 2. Results of in vitro enzyme inhibition assay tests.

ID	COX-1 [%]						COX-2 [%]					
	1 μ M		10 μ M		100 μ M		1 μ M		10 μ M		100 μ M	
1'	6 \pm	2.0	9 \pm	4.2	15 \pm	1.6	19 \pm	2.1	23 \pm	12.7 †	17 \pm	3.2
2*'	-3 \pm	3.5	-5 \pm	1.0	26 \pm	23.5	14 \pm	9.7	9 \pm	2.2	6 \pm	1.6
3	8 \pm	1.3	12 \pm	3.5	10 \pm	5.1	19 \pm	0.7	10 \pm	1.9	14 \pm	0.1
4	6 \pm	10.7	-6 \pm	3.8	-2 \pm	2.8	3 \pm	3.2	3 \pm	7.4	-1 \pm	9.9
5'	7\pm	3.9	31\pm	3.8	72\pm	4.5	27\pm	8.1	27\pm	1.4	38\pm	3.7
6*	16 \pm	4.7	18 \pm	9.3	15 \pm	1.6	19 \pm	68.3 †	0 \pm	10.0	12 \pm	2.9
7	2 \pm	4.5	4 \pm	8.4	7 \pm	62.6 †	20 \pm	9.4	13 \pm	8.8	15 \pm	0.3
8*'	2 \pm	4.5	29 \pm	26.8	16 \pm	8.0	14 \pm	2.6	8 \pm	2.1	11 \pm	6.2
9'	6\pm	16.8	46\pm	6.1	89\pm	1.0	9\pm	0.7	24\pm	7.6	38\pm	3.1
10*	-7 \pm	1.6	-10 \pm	8.5	7 \pm	2.8	2 \pm	5.5	-2 \pm	2.7	-1 \pm	3.7
11	11 \pm	12.6	5 \pm	0.6	1 \pm	1.5	5 \pm	2.9	4 \pm	4.2	0 \pm	1.1
12	18 \pm	4.0	8 \pm	4.5	14 \pm	1.5	15 \pm	0.1	15 \pm	7.5	19 \pm	2.2

Primed (') compounds are shown in Fig. 8, starred (*) compounds are shown in Fig. 9. Discussed compounds are shown in bold face.

† For these three measurements, high standard deviations were observed. This could be due to solubility problems, impurities, protein degradation, or other unspecific effects. The corresponding measurements should be treated carefully.

doi:10.1371/journal.pone.0021554.t002

matches (lacking the *meta* methyl group) were retrieved with regard to bioactivities other than COX inhibition. It is therefore reasonable to conclude that COX inhibition by compounds **5** and **9** represents a novel finding resulting from our study. We did not perform additional analytical investigations of compound

integrity and purity other than those provided by the compound supplier. Therefore, we cannot exclude that the activities measured in the assays might be partially owed to decomposition or oxidation products. Analog compound design and testing will be mandatory.

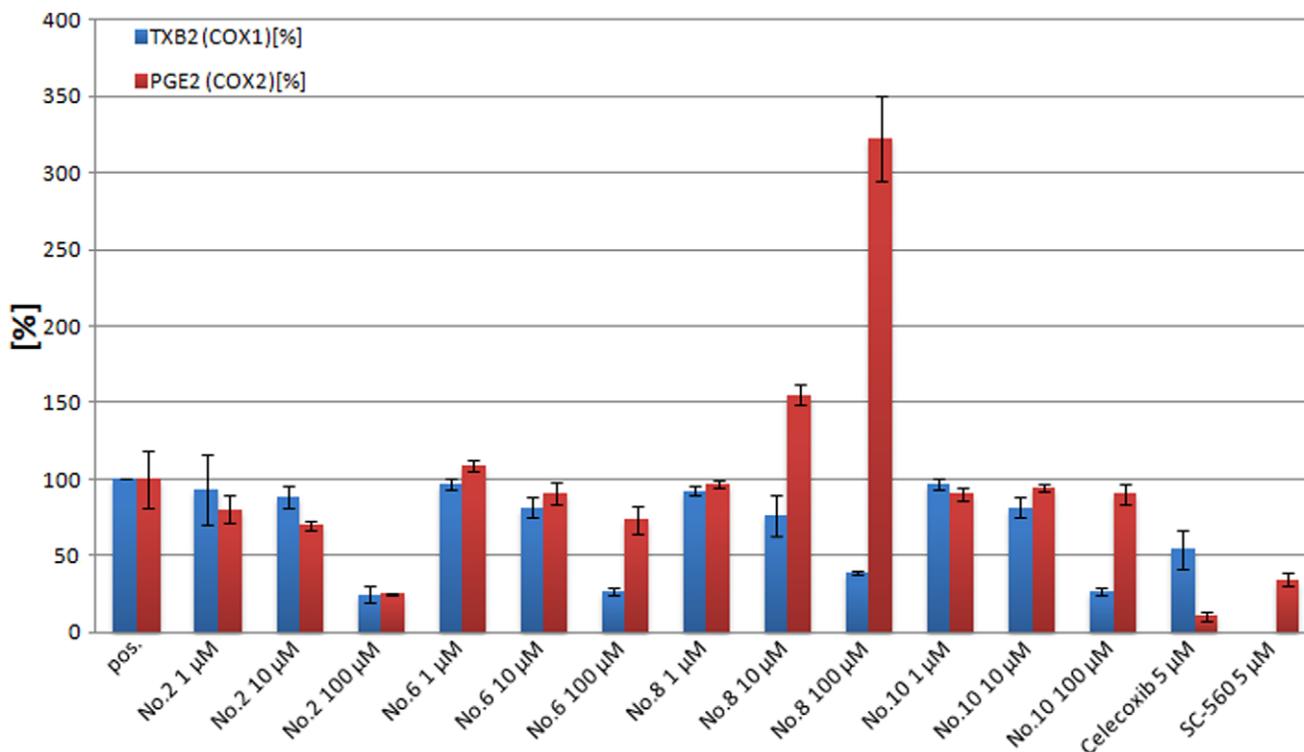


Figure 9. COX-2 inhibition whole blood assay results. Shown are TXB₂ (blue, indicative of COX-1 activity) and PGE₂ (red, indicative of COX-2 activity) amounts relative to the control (DMSO). Celecoxib and SC-560 are known inhibitors selective for COX-2 and COX-1, respectively.
doi:10.1371/journal.pone.0021554.g009

Table 3. Results of whole blood assay tests.

ID	TXB ₂ [%]						PGE ₂ [%]					
	1 μM		10 μM		100 μM		1 μM		10 μM		100 μM	
1'	122±	15.0	94±	3.4	81±	1.2	105±	4.9	112±	8.4	97±	19.9
2*'	93±	22.7	88±	6.9	25±	5.3	80±	9.0	70±	3.1	25±	0.5
3	107±	0.0	107±	2.8	91±	2.3	93±	15.8	91±	8.9	88±	34.5
4	102±	3.5	91±	4.4	85±	1.8	110±	12.3	111±	16.5	105±	22.0
5'	100±	0.9	98±	1.3	71±	1.3	105±	7.9	118±	6.9	106±	33.8
6*	97±	3.8	82±	6.6	28±	2.4	109±	3.6	91±	7.4	74±	9.1
7	102±	6.3	89±	12.1	77±	5.8	99±	13.7	86±	11.9	93±	21.2
8*'	93±	2.8	77±	13.3	39±	1.2	97±	2.5	155±	6.7	322±	27.5
9'	86±	6.9	71±	2.3	67±	12.1	73±	6.1	84±	9.8	89±	3.1
10*	97±	3.8	82±	6.6	28±	2.4	91±	4.2	94±	2.9	90±	6.3
11	92±	1.8	87±	2.7	60±	1.9	105±	2.1	98±	9.2	69±	6.7
12	100±	1.3	85±	4.4	80±	0.9	102±	10.6	120±	9.4	113±	14.5

TXB₂ is indicative of COX-1 activity, PGE₂ of COX-2 activity relative to DMSO control. Primed (') compounds are shown in Fig. 8, starred (*) compounds are shown in Fig. 9. Discussed compounds are shown in bold face.

doi:10.1371/journal.pone.0021554.t003

Conclusions

We presented a favorable retrospective evaluation of the SpH approach using COX-2 data from the DUD collection, and in a first prospective application demonstrated the usefulness of the descriptor in combination with a self-organizing map for retrieving bioactive ligands from a large compound pool. Although we did not retrieve a potent COX-2 inhibitor, which is likely owed to the setup of the virtual screening cascade, two novel COX-1 inhibitors were discovered. Future research will have to focus on mathematical descriptions of molecular shape that allow for enzyme subtype-selective ligand screening.

We introduced the magnitude of spherical harmonics coefficients as a partially rotation-invariant descriptor of molecular shape. In retrospective validation on the DUD dataset, the performance (as estimated by ROC AUC) of our shape-only

method was comparable to other shape-based similarity searching methods. Results show that the magnitude of spherical harmonics decomposition coefficients can be used to describe molecular shape in a partially rotation-invariant way, resulting in a notable enrichment of active compounds in virtual and real screening studies. The combination of pharmacophore filtering by a self-organizing map and shape-filtering by spherical harmonics descriptors might be a useful two-step virtual screening protocol for hit retrieval from large screening compound collections.

Author Contributions

Conceived and designed the experiments: QW SG MR GS. Performed the experiments: QW KB CA TG MR GS. Analyzed the data: QW SG PS MR GS. Contributed reagents/materials/analysis tools: PS SG. Wrote the paper: QW MR GS.

References

- Böhm HJ, Schneider G, eds. (2000) Virtual Screening for Bioactive Molecules. Weinheim, Germany: Wiley-VCH.
- Douguet D (2008) Ligand-based approaches in virtual screening. *Curr Comput Aided Drug Des* 4: 180–190.
- Jurs P (2003) Quantitative structure-property relationships. In: Gasteiger J, ed. *Handbook of Chemoinformatics: From Data to Knowledge*, Wiley, volume 3, chapter 1.2. pp 1314–1135.
- Kubinyi H (2003) QSAR in drug design. In: Gasteiger J, ed. *Handbook of Chemoinformatics: From Data to Knowledge*, Wiley, volume 4, chapter 4.2. pp 1532–1554.
- Johnson M, Maggiora G, eds. *Concepts and Applications of Molecular Similarity*. New York: Wiley.
- Rupp M, Schneider P, Schneider G (2009) Distance phenomena in high-dimensional chemical descriptor spaces: Consequences for similarity-based approaches. *J Comput Chem* 30: 2285–2296.
- Cramer III R, Patterson D, Bunce J (1988) Comparative molecular field analysis (CoMFA). 1. Effect of shape on binding of steroids to carrier proteins. *J Am Chem Soc* 110: 5959–5967.
- Randic M, Kleiner A, Alba LD (1994) Distance/distance matrices. *J Chem Inf Comput Sci* 34: 277–286.
- Schuur J, Gasteiger J (1997) Infrared spectra simulation of substituted benzene derivatives on the basis of a 3D structure representation. *Anal Chem* 69: 2398–2405.
- Robinson D, Barlow T, Richards G (1997) The utilization of reduced dimensional representations of molecular structure for rapid molecular similarity calculations. *J Chem Inf Comput Sci* 37: 943–950.
- Hemmer M, Steinhauer V, Gasteiger J (1999) Deriving the 3D structure of organic molecules from their infrared spectra. *Vib Spectros* 19: 151–164.
- Gramatica P, Corradi M, Consonni V (2000) Modelling and prediction of soil sorption coefficients of non-ionic organic pesticides by molecular descriptors. *Chemosphere* 41: 763–777.
- Todeschini R, Consonni V (2000) *Handbook of Molecular Descriptors*. Weinheim, Germany: Wiley-VCH, first edition.
- Ballester PJ, Richards WG (2007) Ultrafast shape recognition to search compound databases for similar molecular shapes. *J Comput Chem* 28: 1711–1723.
- Grant A, Gallardo A, Pickup B (1996) A fast method of molecular shape comparison: A simple application of a Gaussian description of molecular shape. *J Comput Chem* 17: 1653–1666.
- Cruciani G, Crivori P, Carrupt PA, Testa B (2000) Molecular fields in quantitative structure-permeation relationships: the VolSurf approach. *J Mol Struct* 503: 17–30.
- Consonni V, Todeschini R, Pavan M, Gramatica P (2002) Structure / response correlations and similarity / diversity analysis by GETAWAY descriptors. 2. Application of the novel 3D molecular descriptors to QSAR/QSPR studies. *J Chem Inf Comput Sci* 42: 693–705.
- Van Drie JH (1997) “shrink-wrap” surfaces: A new method for incorporating shape into pharmacophoric 3D database searching. *J Chem Inf Comput Sci* 37: 38–42.
- Shilane P, Min P, Kazhdan M, Funkhouser T (2004) The Princeton shape benchmark. In: *Proceedings of the International Conference on Shape Modeling & Applications (SMI 2004)*, Genova, Italy, June 7–9 IEEE Computer Society. pp 167–178.
- Iyer N, Jayanti S, Lou K, Kalyanaraman Y, Ramani K (2005) Three-dimensional shape searching: State-of-the-art review and future trends. *Comput Aided Des* 37: 509–530.
- Ankerst M, Kastenmüller G, Kriegel HP, Seidl T (1999) 3D shape histograms for similarity search and classification in spatial databases. In: Güting R,

- Papadias D, Lochovsky F, eds. Proceedings of the 6th International Symposium on Spatial Databases (SSD 1999), Hong Kong, China, July 20–23 Springer. pp 207–228.
22. Osada R, Funkhouser T, Chazelle B, Dobkin D (2001) Matching 3D models with shape distributions. In: Proceedings of the 3rd International Conference on Shape Modeling & Applications (SMI 2001), Genova, Italy, May 7–11 IEEE Computer Society. pp 154–166.
 23. Jamitzky F, Stark R, Bunk W, Thalhammer S, R ath C, et al. (2001) Scaling-index method as an image processing tool in scanning-probe microscopy. *Ultramicroscopy* 86: 241–246.
 24. Horn B (1984) Extended Gaussian images. *Proc IEEE* 72: 1671–1686.
 25. Chen DY, Tian XP, Shen YT, Ouhyoung M (2003) On visual similarity based 3D model retrieval. *Comput Graph Forum* 22: 223–232.
 26. Novotni M, Klein R (2003) 3D Zernike descriptors for content based shape retrieval. In: Proceedings of the 8th ACM Symposium on Solid modeling and Applications (SM 2003), Seattle, Washington, USA, June 16–20 Association for Computing Machinery. pp 216–225.
 27. Vranic D (2003) An improvement of rotation invariant 3D-shape based on functions on concentric spheres. In: Proceedings of the International Conference on Image Processing (ICIP 2003), Barcelona, Spain, September 14–17 IEEE Computer Society, volume 3. pp 757–760.
 28. Kazhdan M, Funkhouser T, Rusinkiewicz S (2003) Rotation invariant spherical harmonic representation of 3D shape descriptors. In: Kobbelt L, Schr oder P, Hoppe H, eds. Proceedings of the 1st Eurographics Symposium on Geometry Processing (SGP 2003), Aachen, Germany, June 23–25 Association for Computing Machinery. pp 167–175.
 29. Max N, Getzoff E (1988) Spherical harmonic molecular surfaces. *IEEE Comput Graph Appl* 8: 42–50.
 30. Duncan B, Olson A (1993) Approximation and characterization of molecular surfaces. *Biopolymers* 33: 219–229.
 31. Ritchie D, Kemp G (1999) Fast computation, rotation, and comparison of low resolution spherical harmonic molecular surfaces. *J Comput Chem* 20: 383–395.
 32. Ritchie DW, Kemp GJL (2000) Protein docking using spherical polar fourier correlations. *Proteins: Struct Funct Bioinf* 39: 178–194.
 33. Lin JH, Clark T (2005) An analytical, variable resolution, complete description of static molecules and their intermolecular binding properties. *J Chem Inf Model* 45: 1010–1016.
 34. Morris R, Najmanovich R, Kahraman A, Thornton J (2005) Real spherical harmonic expansion coefficients as 3D shape descriptors for protein binding pocket and ligand comparisons. *Bioinformatics* 21: 2347–2355.
 35. Mavridis L, Hudson B, Ritchie D (2007) Toward high throughput 3D virtual screening using spherical harmonic surface representations. *J Chem Inf Model* 47: 1787–1796.
 36. Cai W, Xu J, Shao X, Leroux V, Beaudrait A, et al. (2008) SHEF: A vHTS geometrical filter using coefficients of spherical harmonic molecular surfaces. *J Mol Model* 14: 393–401.
 37. Jakobi AJ, Mauser H, Clark T (2008) Parafraq—an approach for surface-based similarity comparison of molecular fragments. *J Mol Model* 14: 547–558.
 38. Venkatraman V, Sael L, Kihara D (2009) Potential for protein surface shape analysis using spherical harmonics and 3D Zernike descriptors. *Cell Biochem Biophys* 54: 23–32.
 39. Huang N, Shoichet B, Irwin J (2006) Benchmarking sets for molecular docking. *J Med Chem* 49: 6789–6801.
 40. P erez-Nu eno VI, Venkatraman V, Mavridis L, Clark T, Ritchie DW (2011) Using spherical harmonic surface property representations for ligand-based virtual screening. *Mol Inf* 30: 151–159.
 41. Venkatraman V, P erez-Nu eno VI, Mavridis L, Ritchie DW (2010) Comprehensive comparison of ligand-based virtual screening tools against the DUD data set reveals limitations of current 3D methods. *J Chem Inf Model* 50: 2079–2093.
 42. Abramowitz M, Stegun I (1972) *Handbook of Mathematical Functions*. New York: Dover.
 43. Press W, Teukolsky S, Vetterling W, Flannery B (2007) *Numerical Recipes. The Art of Scientific Computing*. Cambridge: Cambridge University Press, third edition.
 44. Vilenkin NY (1968) *Special Functions and the Theory of Group Representations*, volume 22 of *Translations of Mathematical Monographs*. Washington DC: American Mathematical Society.
 45. Funkhouser T, Min P, Kazhdan M, Chen J, Halderman A, et al. (2003) A search engine for 3D models. *ACM Trans Graph* 22: 83–105.
 46. Ben-Israel A, Greville T (2003) *Generalized Inverses. Theory and Applications* Springer, second edition.
 47. Jolliffe I (2004) *Principle Component Analysis*. New York: Springer, second edition.
 48. Grant A, Pickup B, Nicholls A (2001) A smooth permittivity function for Poisson-Boltzmann solvation methods. *J Comput Chem* 22: 608–640.
 49. Brechb uhler C, Gerig G, K ubler O (1995) Parametrization of closed surfaces for 3-D shape description. *Comput Vis Image Understand* 61: 154–170.
 50. Clark T (2010) *ParaSurf 10 User Manual* CePos InSilico Ltd., The Old Vicarage, 132 Bedford Road, Kempston, United Kingdom.
 51. Mavridis L, Ritchie DW (2010) 3D-blast: 3D protein structure alignment, comparison, and classification using spherical polar fourier correlations. In: Proceedings of the 15th Pacific Symposium on Biocomputing (PSB 2010), Maui, Hawaii, USA, January 3–7. pp 281–292.
 52. Schneider P, Schneider G (2003) Collection of bioactive reference compounds for focused library design. *QSAR Comb Sci* 22: 713–718.
 53. Berman H, Westbrook J, Feng Z, Gilliland G, Bhat T, et al. (2000) The protein data bank. *Nucleic Acids Res* 28: 235–242.
 54. Kurumbail R, Stevens A, Gierse J, McDonald J, Stegeman R, et al. (1996) Structural basis for selective inhibition of cyclooxygenase-2 by anti-inflammatory agents. *Nature* 384: 644–648.
 55. Hawkins P, Warren G, Skillman G, Nicholls A (2008) How to do an evaluation: pitfalls and traps. *J Comput Aided Mol Des* 22: 179–190.
 56. Jain A, Nicholls A (2008) Recommendations for evaluation of computational methods. *J Comput Aided Mol Des* 22: 133–139.
 57. Kohonen T (2001) *Self-Organizing Maps*. New York: Springer, third edition.
 58. Schneider G, Neidhart W, Giller T, Schmid G (1999) “Scaffold-hopping” by topological pharmacophore search: A contribution to virtual screening. *Angew Chem Int Ed* 38: 2894–2896.
 59. Fechner U, Schneider G (2004) Optimization of a pharmacophore-based correlation vector descriptor for similarity searching. *QSAR Comb Sci* 23: 19–22.
 60. Schneider G, Wrede P (1998) Artificial neural networks for computer-based molecular design. *Progr Biophys Mol Biol* 70: 175–222.
 61. Schmidt R, Coste O, Geisslinger G (2005) LC-MS/MS-analysis of prostaglandin E2 and D2 in microdialysis samples of rats. *J Chrom B* 826: 188–197.
 62. Fawcett T (2006) An introduction to ROC analysis. *Pattern Recogn Lett* 27: 861–874.
 63. Mackey MD, Melville JL (2009) Better than random? The chemotype enrichment problem. *J Chem Inf Model* 49: 1154–1162.
 64. Truchon JF, Bayly C (2007) Evaluating virtual screening methods: Good and bad metrics for the “early recognition” problem. *J Chem Inf Model* 47: 488–508.
 65. Tanrikulu Y, Nietert M, Scheffer U, Proschak E, Grabowski K, et al. (2007) Scaffold hopping by “fuzzy” pharmacophores and its application to RNA targets. *Chem Bio Chem* 8: 1932–1936.
 66. Tanrikulu Y, Schneider G (2008) Pseudoreceptor models in drug design: Bridging ligand- and receptor-based virtual screening. *Nat Rev Drug Discov* 7: 667–677.
 67. Tanrikulu Y, Proschak E, Werner T, Geppert T, Todoroff N, et al. (2009) Homology model adjustment and ligand screening with a pseudoreceptor of the human histamine H4 receptor. *Chem Med-Chem* 4: 820–827.
 68. Vainio MJ, Puranen JS, Johnson MS (2009) ShaEP: Molecular overlay based on shape and electrostatic potential. *J Chem Inf Model* 49: 492–502.
 69. Rush III TS, Grant JA, Mosyak L, Nicholls A (2005) A shape-based 3-d scaffold hopping method and its application to a bacterial protein–protein interaction. *J Med Chem* 48: 1489–1495.
 70. Riendeau D, Percival MD, Boyce S, Brideau C, Charleson S, et al. (1997) Biochemical and pharmacological profile of a tetrasubstituted furanone as a highly selective COX-2 inhibitor. *Br J Pharmacol* 121: 105–117.
 71. Proschak E, Rupp M, Derksen S, Schneider G (2008) Shapelets: Possibilities and limitations of shape-based virtual screening. *J Comput Chem* 29: 108–114.
 72. Ripphausen P, Nisius B, Bajorath J (2011) State-of-the-art in ligand-based virtual screening. *Drug Discov Today* 16: 372–376.
 73. Yuan C, Rieke CJ, Rimon G, Wingard BA, Smith WL (2006) Partnering between monomers of cyclooxygenase-2 homodimers. *Proc Natl Acad Sci USA* 103: 6142–6147.
 74. Vecchio AJ, Simmons DM, Malkowski MG (2010) Structural basis of fatty acid substrate binding to cyclooxygenase-2. *J Biol Chem* 285: 22152–22163.