The Discovery of Kv1.5 Blockers as a Case Study for the Application of Virtual Screening Approaches

Bernard Pirard,*,† Joachim Brendel,‡ and Stefan Peukert‡

Aventis Pharma Deutschland GmbH, A Company of the Sanofi-Aventis Group, Computational Chemistry, Medicinal Chemistry, Industrie Park Höchst, Building G878, D-65926 Frankfurt am Main, Germany

Received September 20, 2004

Different virtual screening techniques are available as alternatives to high throughput screening. These different techniques have been rarely used together on the same target. We had the opportunity to do so in order to discover novel blockers of the voltage-dependent potassium channel Kv1.5, a potential target for the treatment of atrial fibrillation. Our corporate database was searched, using a protein-based pharmacophore, derived from a homology model, as query. As a result, 244 molecules were screened in vitro, 19 of them (7.8%) were found to be active. Five of them, belonging to five different chemical classes, exhibited IC_{50} values under 10 μM. The performance of this structure-based virtual screening protocol has been compared with those of similarity and ligand-based pharmacophore searches. The analysis of the results supports the conventional wisdom of using as many virtual screening techniques as possible in order to maximize the chance of finding as many chemotypes as possible.

INTRODUCTION

Over the past few years, information-driven hit finding approaches such as virtual screening (VS) have emerged as an alternative to high throughput screening (HTS). Virtual screening encompasses a variety of computational techniques that aim to reduce a large collection of compounds to a short list of screening candidates.[1−3] The amount of information available within a given discovery project determines the choice of a computational technique. In many projects, ligand structures remain the only source of knowledge. This restricts the selection of computational methods to similarity searching, pharmacophore-based searches or classification techniques.[4−7] On the other hand the availability of a protein three-dimensional (3D) structure provides an opportunity for applying high throughput docking or protein-derived pharmacophore searching.[8−10]

Successful applications of VS are seldom reported in the literature because of the proprietary nature of the molecules identified.[11−15] In the handful of published success stories, researchers used either a single computational technique or a combination thereof. This combination forms a VS cascade which usually consists of fast and general filters, followed by more time-consuming and project specific ones, such as docking in a protein binding site.

Biochemical screening of the molecules passing these filters is expected to yield a higher hit rate than screening of a random selection. Therefore, the enrichment of the hit rate produced by VS is often considered as a figure of merit.[16] Other researchers have also used the number of different chemotypes discovered by a VS protocol for assessing its performance.[17,18] The identification of multiple lead series is a key element for a drug discovery project, since it enables the resolution of patent issues and provides a means to overcome chemical class-specific liabilities, such as poor pharmacokinetics and toxicity, which might stall a lead optimization program.

Herein, we report the application of a structure-based VS protocol which has led to the discovery of novel blockers of the Kv1.5 channel. This voltage-dependent potassium channel is regarded as a promising target for the development of novel atrial selective antiarrhythmics.[19] Then, we compare the performances of the structure-based VS protocol with those of ligand-based VS approaches that we have also applied in this project. These ligand-based approaches include two-dimensional (2D) similarity[20] and pharmacophore-based searches.[21] This comparative analysis is focused on hit rate and on the chemical novelty of the hits. To further assess the chemical novelty of the hits discovered through structure-based VS, we also consider competitors' Kv1.5 blockers. This analysis of ligands is based on molecular descriptors, frequently used for database mining. Although the structures of these new Kv1.5 blockers cannot be disclosed, the computational protocol for virtual screening and data analysis is described in detail. This case study illustrates the advantage of using several VS approaches nearly in parallel.

METHODS

Homology Modeling. Mammalian voltage dependent potassium (Kv) channels are tetramers of four identical subunits. Each subunit consists of six transmembrane domains (S1−S6), with S5 and S6 forming the conduction pore.[22] No experimentally determined 3D structure is currently available for any Kv channel. Therefore, we used the crystal structure of the bacterial KcsA channel[23] (Protein Data Bank code 1BL8) as the template for homology modeling of the human Kv1.5 pore domain. The KcsA structure, which

* Corresponding author phone: +49 (0) 69 305 83 768; fax: +49 (0) 69 33 13 99; e-mail bernard.pirard@aventis.com.
† Computational Chemistry.
‡ Medicinal Chemistry.

10.1021/ci0400011 CCC: $30.25
Published on Web 03/08/2005

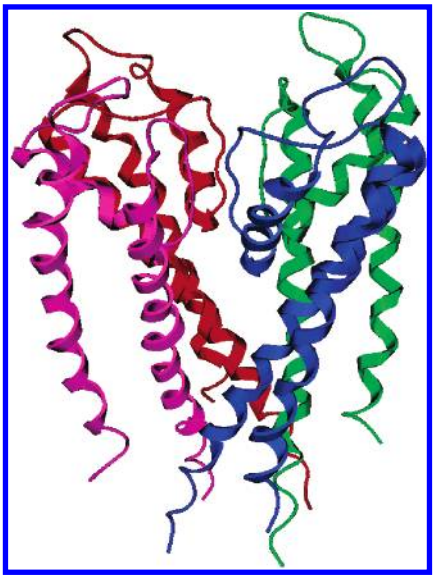

Figure 1. Crystal structure of the bacterial KcsA channel, used as template for homology modeling of the Kv1.5 pore domain. Each of the four subunits is displayed in a different color.

consists of a tetramer of S5−S6 segments (Figure 1), corresponds to the closed state of the channel. Previously, scientists at Merck used the KcsA structure as template for homology modeling of the HERG channel. This HERG homology model was used for docking known HERG blockers and was validated by site-directed mutagenesis.[24]

Each of the four subunits forming the Kv1.5 pore shares 30.2% of sequence identity with its KcsA counterpart. To start with, a homology model of a single subunit of the Kv1.5 pore was generated by using the SYBYL implementation of COMPOSER.[25] Then, the homology model of the Kv1.5 subunit was aligned on each of the KcsA subunits to produce a tetrameric arrangement. This structural alignment was performed by using the alignment routine implemented in the BIOPOLYMER routine of SYBYL.[25]

To refine the homology model of the Kv1.5 tetramer, we subsequently applied a two step minimization protocol. The first step was limited to the protein side chains. After convergence, the whole structure was minimized. At each step, the energy was evaluated, using the SYBYL implementation of the AMBER force field.[25] For these calculations, we employed a distance-dependent dielectric constant of 4r (r is the interatomic distance) and set the convergence criterion to 0.05 kJ $Å^{-1}$ mol^{-1}. Using a distance-dependent dielectric constant of 4r compensates for the lack of explicit solvation by implicitly damping the rather strong long-range intramolecular and intersubunits electrostatic interactions.[26]

The minimized homology model was submitted to a series of tests to determine its quality and internal consistency, using PROCHECK,[27−28] WHATCHECK[29] and MATCHMAKER.[25] 3D coordinates of the homology model can be obtained from the authors upon request.

Identification of a Putative Binding Site. The minimized homology model served as input for PASS[30,31] (Putative Active Site with Spheres), a binding site identification algorithm. This binding site identification procedure is based on purely geometrical criteria. Briefly, PASS coats a protein surface with an initial layer of spherical probes by calculating the locations at which a probe sphere may lie tangential to a triplet of protein atoms. Then, the probes that clash with the protein, or are not sufficiently buried or lie within 1 Å of a more buried probe, are discarded. Once the first layer of probes has been filtered, a second one is added and subsequently filtered. This sequence of coating-filtering is repeated until no newly added probe survives the filters. Finally, a ranked list of putative active site points is produced. Their positions correspond to the locations of the central probes occupying regions mainly populated by buried spheres. Their ranking is determined by their extent of burial.

Protein-Derived Pharmacophore. After visual inspection of the PASS output, we selected one potential binding site for further characterization by the GRID force field.[32] All the GRID calculations were carried out with a 1 Å grid size in a box enclosing this potential binding site. During the GRID calculations, the protein side chains were allowed to move (GRID directive Move = 1).

The molecular interaction fields (MIFs) were computed for three probes: the hydrophobic (DRY), the amide nitrogen (N1, H bond donor) and the carbonyl oxygen (O, H bond acceptor). Local minima were identified for the three MIFs. After visual inspection, we selected minima from the GRID energy maps. These selected minima served to define the features of a pharmacophoric query. We completed the definition of this pharmacophore query by adding an exclusion volume, defined by the protein atoms.

Virtual Screening. The protein-derived pharmacophore served as query to screen our corporate database, which approaches the size of the collection of large pharmaceutical companies.[33] UNITY[34] was used for this database search. Typically, we applied filters on the maximum number of rotatable bonds and on Lipinski's 'rule of five',[35] as available in the UNITY Flexsearch routine, before performing the search. The maximum number of rotatable bonds was set to 10, while the default values were considered for the 'rule of five' parameters. Partial match constraints, as implemented in UNITY, were also applied. A compound had to match at least three pharmacophoric features and at most twelve (i.e. all of the defined pharmacophoric features) to be considered as a hit. The hits of the database search were subsequently submitted to filters to discard duplicates, compounds with reactive groups[36] or those that had already been screened on Kv1.5. Finally, we selected representative compounds for in vitro screening, in conjunction with an experienced medicinal chemist. This chemist evaluated the compounds based on their synthetic feasibility and on their suitability for chemical optimization (opportunities for derivatization). This latter step is a common practice in drug discovery,[11] although this human intervention introduces some bias in the selection.[37] The protocol for the in vitro assay has been described elsewhere.[20]

Analysis of Screening Results. To assess the structural diversity of our most active hits (IC_{50} < 10 μM), we computed their pairwise similarities, based on UNITY fingerprints[34] and Feature trees descriptors.[38] Briefly, UNITY fingerprints encode the presence of 2D fragments in a binary format. Pairwise similarity based on UNITY fingerprints has been quantified by the Tanimoto coefficient. On the other hand, a Feature tree represents a molecule as a set of labeled nodes (i.e. a reduced graph) connected as in the chemical structure. Each node is labeled with a set of features representing physicochemical properties (hydrophobicity, H

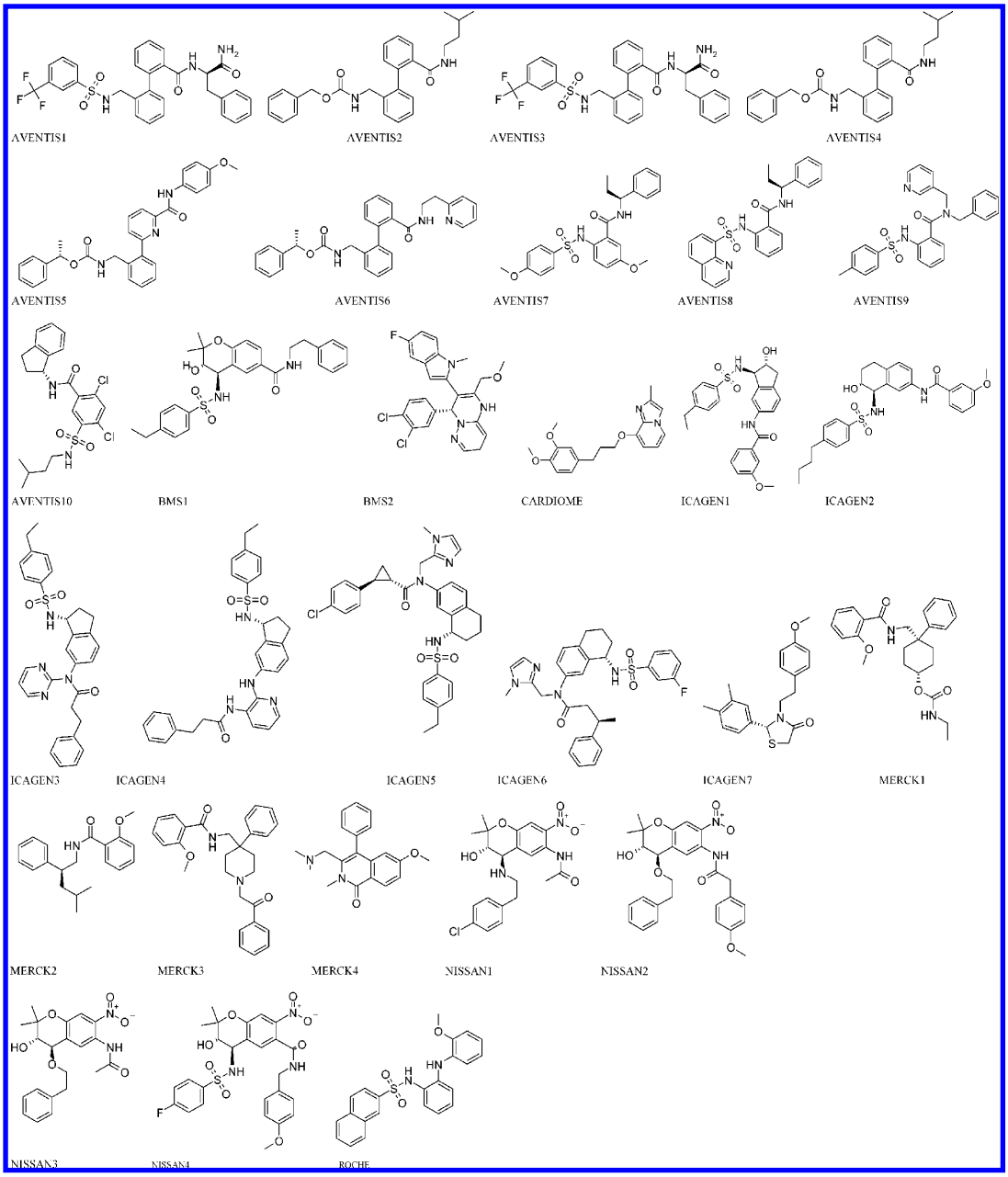

Figure 2. Chemical structures of reference Kv1.5 blockers.

bonding properties, etc.) of the region of the molecule described by the node. The pairwise molecular similarity is determined by matching of the corresponding trees. The matching procedure relates nodes in one tree with nodes in the second tree, preserving the molecular topology and maximizing a similarity score. This similarity score ranges from 0 (completely dissimilar structures) to 1 (identical structures). Feature trees which are less dependent on molecular topology than UNITY fingerprints can better gauge the pairwise similarity of structurally different compounds, binding alike to their target.[5] Similarity searches on Feature trees are therefore suitable for identifying molecules with novel scaffolds.

We were also interested in comparing the structures of our most active hits to those of other known Kv1.5 blockers. For this purpose, we built a database of reference Kv1.5 blockers, including in-house and competitors' compounds (Figure 2).[19] Some of the in-house compounds belong to chemical classes discovered using ligand-based virtual screening approaches. The comparative analysis of our most active hits with the reference Kv1.5 blockers involved computations of pairwise similarities based on UNITY fingerprints and Feature trees.

We also considered two additional types of descriptors: pharmacophore-based correlation vectors, known as CATS[39] descriptors, and VolSurf[40] descriptors. Briefly, CATS descriptors encode as a 150-dimensional vector, the occurrence of each of the 15 possible pairs of five pharmacophoric features (H bond donor, H bond acceptor, lipophilic, negatively charged, and positively charged), separated by a given topological distance. Like the Feature trees, the CATS are two-dimensional descriptors. Recent applications of the CATS descriptors include focused library design[41] and the identification of novel chemotypes through similarity searching.[39,42] For this work, we used our own implementation of the CATS methodology, which combines elements of the SPL and PERL scripting languages. These vectors were generated for our most active hits and for the compounds

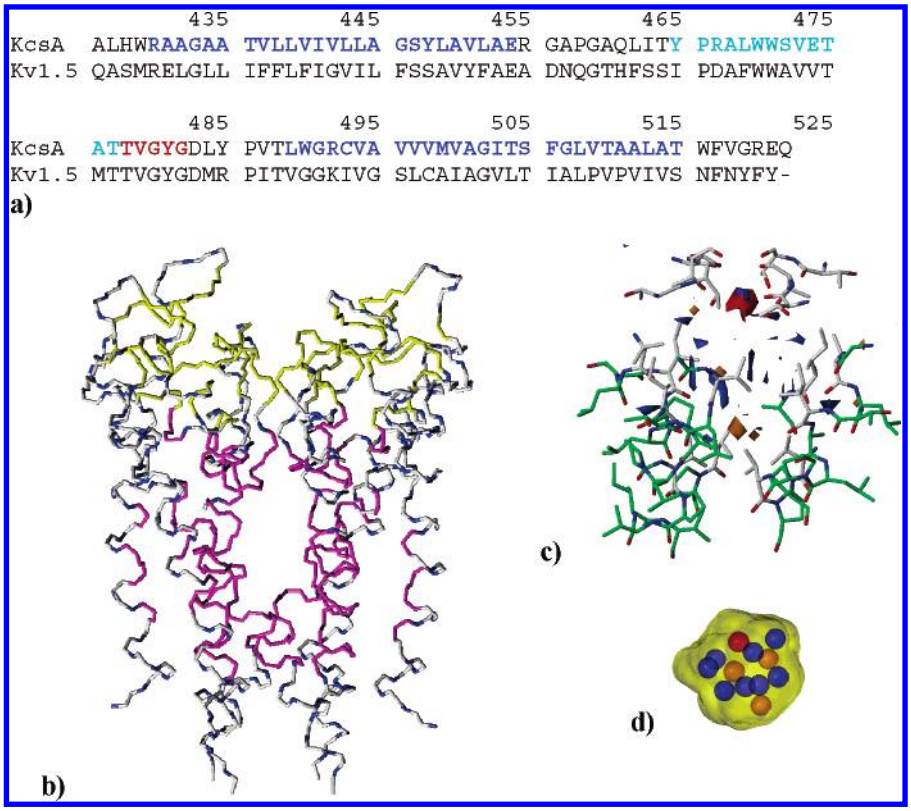

Figure 3. (a) Sequence alignment of one subunit of the bacterial KcsA and human Kv1.5 pore forming domains. KcsA residues forming the S5 and S6 transmembrane domains, the selectivity filter (GYG sequence motif) and the pore helix are highlighted in bold and displayed in blue, red and cyan, respectively. (b) Backbone representation of the homology model of the Kv1.5 pore. The O atoms have been omitted for clarity. The backbone of the amino acids forming the extracellular site is colored in yellow, while the backbone of the amino acids forming the internal site is in magenta. (c) Amino acids lining the internal site. The C atoms are displayed in gray for the amino acids identified as essential for the activity of Kv1.5 blockers[49] and in green for the other amino acids, the N and O atoms are colored in blue and red, respectively. GRID *iso*-contours are also displayed in blue, red and orange for the H bond donor (-5 kcal mol^{-1}), acceptor (-5 kcal mol^{-1}) and hydrophobic (-0.8 kcal mol^{-1}) probes respectively. (d) Protein-based pharmacophore derived from the GRID analysis of the internal site. The color convention is the same as for the GRID *iso*-contours. The yellow region corresponds to the excluded volume defined by the protein.

shown in Figure 2 and were subsequently analyzed by Principal Component Analysis (PCA), as implemented in the 4.5 release of GOLPE.[43] In PCA,[44] the original data matrix is approximated in terms of the product of two smaller matrices, the scores and the loadings matrices. The scores matrix gives a simplified picture of the objects (Kv1.5 blockers in this case), described in terms of their projection onto the principal components (PCs). The loadings matrix contains the PCs, which are linear combinations of the original variables (CATS). The first PC describes the maximum variance along all possible directions, the second component the next largest variation among all directions orthogonal to the first one, and so on. To find the minimum number of components necessary for data reproduction within residual errors, the components are added step by step to the model. Analysis of the score plots gives information about the clustering of the objects, while the loading plots contain information about the relative contributions of the variables to the PCs.

In contrast to the descriptors mentioned above, VolSurf descriptors are derived from the 3D structure.[40] The VolSurf approach starts from the molecular interaction fields (MIF), computed by the program GRID[32] with different probes and condenses the large body of information contained in the MIF into a few numerical descriptors. More precisely, VolSurf computes the volume and surface of the regions enclosing values of energies of interaction under given cutoff values, as well as other descriptors of their spatial distribution. VolSurf descriptors are useful for describing physical and pharmacokinetic properties such as solubility and intestinal absorption.[45] We generated the 3D structures with CORINA[46] and subsequently minimized them with the SYBYL 6.8 implementation of the MMF94s force field.[25] The minimized structure served as input for VolSurf3.0,[47] which generated 150 descriptors from the MIFs, obtained with the hydrophobic DRY, the water and the carbonyl O probes. These 150 descriptors were also submitted to a PCA.

RESULTS

Homology Modeling. Figure 3a shows the sequence alignment of one subunit of the KcsA and Kv1.5 pore domains. The selectivity filter, characterized by the sequence motif TVGYG, is conserved.[23] The homology model of the Kv1.5 pore is based on this sequence alignment.

After minimization, the homology model (Figure 3b) was submitted to several tests. First of all, least-squares fitting of the Cα atoms of the model and template structures yielded a root-mean-square standard deviation of 0.37 Å. Geometrical analyses and other tests showed that the model and template were of similar quality. Most of the amino acids (86.2%, compared to 78.8% for KcsA) of the Kv1.5 pore belong to the core of the Ramanchandran plot. Results from the geometrical analyses are available as Supporting Information.

Identification of a Putative Binding Site − Protein-Derived Pharmacophore. The minimized homology model of the Kv1.5 pore served as input for the binding site detection algorithm PASS. This resulted in the identification of nine putative binding sites. However, only two of them were deemed relevant, since the remaining putative sites were found in regions, where the Kv1.5 structure is incomplete. The first of these two sites (Figure 3b) is large, fairly flat and located on the extracellular surface of the pore and corresponds to the binding site for toxins, described for Kv channels.[48] The second site (Figure 3b-c) is located underneath the selectivity filter. Subsequent site-directed mutagenesis experiments confirmed that residues forming the internal site were important for the Kv1.5 blocking activity of our compounds.[49] Previously, other researchers had shown that some of the amino acids found in the same region of the HERG K^+ channel were essential for the interaction with small molecule blockers.[24] Therefore, we decided to focus our virtual screening efforts on the second of these two sites. In the rest of this paper, we refer to this site as the internal site.

The side chains of Thr477, Thr478, Val503, Ile506, Ala507, and Val510 define the internal binding site of Kv1.5 (Figure 3c). In contrast to the corresponding site of the HERG K^+ channel, there is no aromatic side chain lining the internal site of Kv1.5. [24]

We further characterized the internal site with the GRID force field and after visual inspection, selected three minima for the DRY probe (hydrophobic interactions), eight for the N1 probe (H bond donor) and one for the O probe (H bond acceptor). These twelve minima were then combined with an excluded volume, defined by the walls of the internal site, which resulted in the protein-based pharmacophoric query shown in Figure 3d.

Virtual Screening. The protein-derived pharmacophore served as the query to search our corporate database. The database search yielded 3102 virtual hits. None of the known ligand was retrieved by this structure-based VS protocol. The bisaryl compounds (AVENTIS1−6, Figure 2) were excluded because of their size, while the other chemical series such as AVENTIS7−10 (Figure 2) were not yet registered in the database at that time. Application of the sequence of filters, described in the Methods section, gave approximately 1000 hits for visual inspection by an experienced medicinal chemist. This step identified molecules that were difficult to synthesize or molecules that offered limited synthetic opportunities for future chemical optimization. Removal of those compounds resulted in a final list of 244 compounds being selected for in vitro screening.

Analysis of Screening Results. These 244 compounds were screened in vitro, as described elsewhere.[20] This yielded 19 actives (hit rate of 7.8%). Five of them showed an IC_{50} under 10 μM, with the best compound having an IC_{50} of 0.9 μM. The IC_{50} values of the five most active compounds are given in Table 1.

As mentioned in the Introduction, the number of different chemotypes identified by a given VS protocol can be considered as another metric to assess its performance. Herein, we computed the similarity matrices on UNITY fingerprints and on Feature trees for the five most active compounds. All of these compounds have a Tanimoto similarity, based on UNITY fingerprints, much lower than 0.85 (Table 2), which corresponds to the usual threshold for considering two compounds as analogues.[50] We were also interested in finding whether the structural diversity of the hits resulted from the power of the VS protocol or from the structural diversity of the screened database. Therefore, the five most active hits (Table 1) were used as queries for similarity searches based on UNITY fingerprints within our corporate collection. The minimum Tanimoto similarity was set to 0.70. These similarity searches resulted in the identification of 76, 124, 426, 776 and 67 analogues for compounds 1−5 (Table 1). None of these analogues was picked by the VS protocol. We can conclude that the structural diversity of the hits is not due to the structural diversity of the screened database.

On the other hand, matching of Feature trees, which can detect similarities between different chemical classes,[5] yields similarity values above 0.70 with only one pair of compounds showing similarity values above 0.85 (Table 3).

To further assess the novelty of the five most active hits from structure-based VS, we computed their pairwise similarities based on UNITY fingerprints and Feature trees, with a set of 29 reference Kv1.5 blockers (Figure 2). Each of the 29 × 5 comparisons yielded Tanimoto values based on UNITY fingerprints lower than 0.85 (Figure 4a). Therefore, these five hits are not close analogues of any of the reference Kv1.5 blockers. As expected, similarity values on Feature trees are higher, with nearly 70% of all the comparisons showing similarity between 0.7 and 0.8 (Figure 4b). For most of our similarity searches based on Feature trees, we usually obtain the best enrichment in actives with a similarity threshold in the range 0.85 to 0.9, depending on the compounds. Rarey and Stahl also suggested a target similarity of 0.9 after comparing the distribution of pairwise similarities based on Feature trees among a set of 54 dopamine D4 antagonists and between a subset of 7528 compounds from the World Drug Index (WDI) and the 54 D4 antagonists.[51] The distribution of pairwise similarities

Table 1. IC_{50} Values of the Five Most Active Hits

compound	IC_{50} (μM)
1	7.7
2	6.8
3	7.9
4	4.1
5	0.9

Table 2. Pairwise Tanimoto Similarity Values Based on UNITY Fingerprints for the Five Most Active Hits

compound	1	2	3	4	5
1	1.00	0.25	0.28	0.32	0.29
2		1.00	0.25	0.28	0.26
3			1.00	0.54	0.51
4				1.00	0.43
5					1.00

Table 3. Similarity Values Based on Feature Trees for the Five Most Active Hits

compound	1	2	3	4	5
1	1.00	0.80	0.81	0.72	0.81
2		1.00	0.76	0.74	0.77
3			1.00	0.72	0.91
4				1.00	0.72
5					1.00

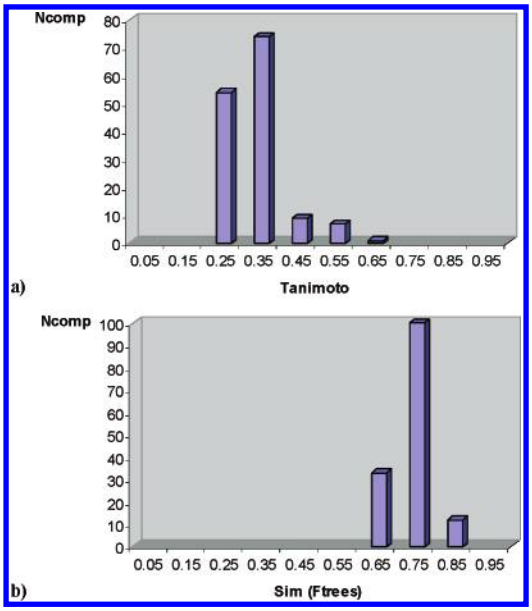

Figure 4. (a) Distribution of Tanimoto coefficients based on UNITY fingerprints. (b) Distribution of the similarities based on Feature trees. Both distributions result from the computations of pairwise similarities between the five most active hits from structure-based VS with the 29 reference Kv1.5 blockers, presented in Figure 2.

between the WDI subset and the D4 antagonist shows a maximum around 0.7. In other words, none of the five most active hits would have been retrieved by a similarity search based on Feature trees with a similarity threshold of 0.9 and any of the 29 reference compounds (Figure 2) as query. At a similarity level of 0.85, only hits 1 and 4 (Table 1) would have been identified, when using ICAGEN7 as query. Tables listing the pairwise similarities plotted in Figure 4 are available as Supporting Information.

In addition, we also performed a PCA on a 34 compound (29 reference compounds and five hits from structure-based VS) by 150 CATS matrix. This led to a three-component model, explaining 69.5% of the variance. These three PCs account for 39.5, 19.5 and 10.5% of the variance, respectively. The largest contributions to PC1 and PC2 loadings arise from pairs of lipophilic features separated by five to 10 bonds, while several pairs of lipophilic-H bond donor and lipophilic-H bond acceptor exhibit large contributions to PC3 loadings. This observation is consistent with the pharmacophore model published for three series of Aventis compounds.[20,21] Four of the five most potent hits, identified by structure-based VS, occupy the central region of the PC1−PC2 scores plot (Figure 5). Most of the reference Kv1.5 blockers tend to populate the central region of the PC1−PC2 scores plot. On the other hand, PC2 separates compound 3 from the other compounds. This separation arises from the presence in compound 3 of a basic substituent, which is not found in any of the other compounds. It is also noteworthy that our bisaryl class of compounds tends to group a little bit outside the central region (Figure 5). Finally, we also performed a PCA on the matrix of VolSurf descriptors. This yielded a two-component model, explaining 44.9% of the variance. These two PCs account for 32.6 and 12.3% of the variance. The largest contributions to PC1 and PC2 loadings arise from descriptors derived from the MIFs produced by the polar (water and carbonyl O) and hydrophobic GRID probes, respectively. Three of the five most potent hits mentioned above tend to cluster in a weakly populated region of the PC1−PC2 scores plot, while the

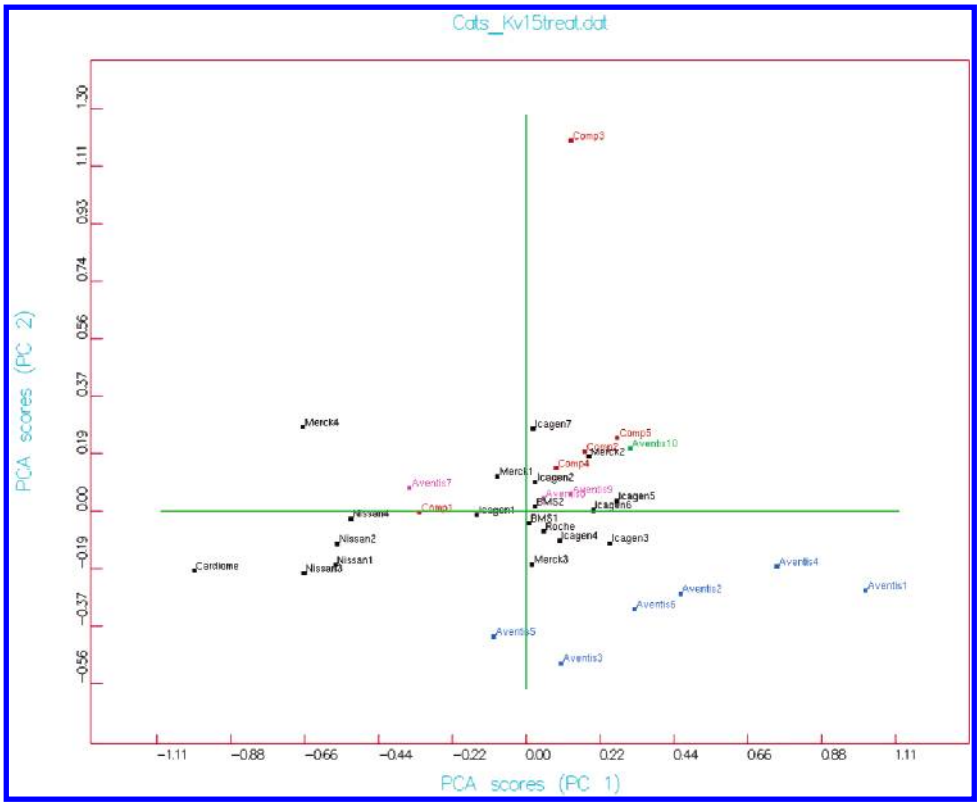

Figure 5. Scores plot obtained for the PCA on CATS descriptors for 34 compounds (the most active hits and the reference Kv1.5 blockers). The most active hits are displayed in red, the competitors' compounds in black, the other Aventis compounds in blue (bisaryls), magenta (anthranilic amides) or green (sulfonyl benzamides).

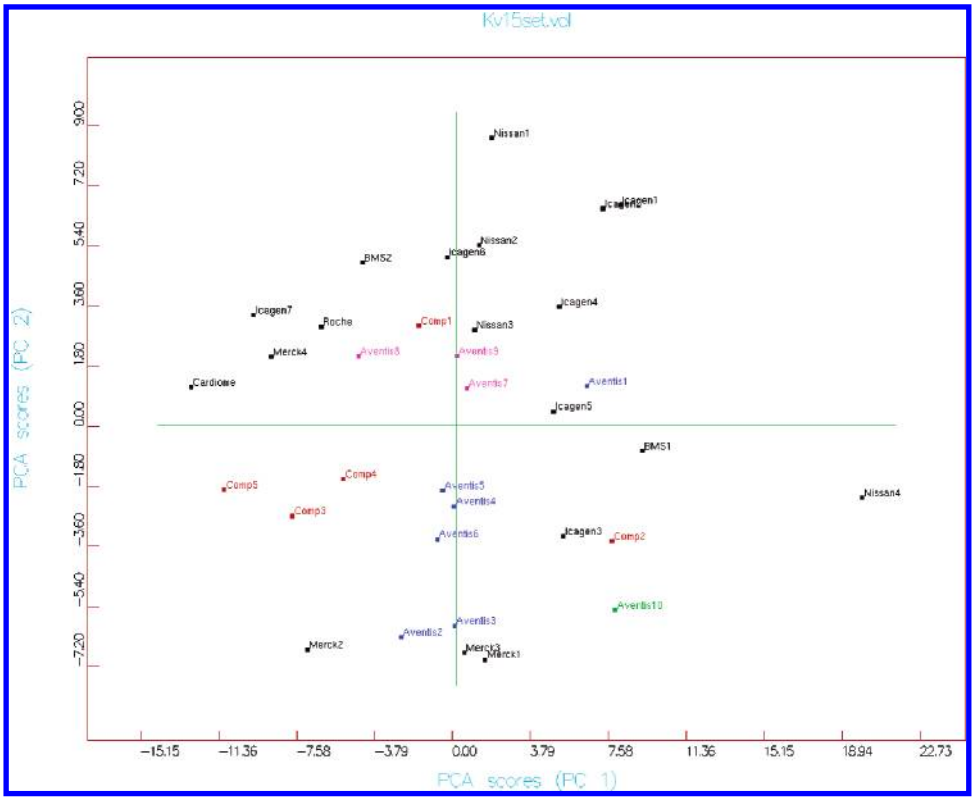

Figure 6. Scores plot obtained for the PCA on VolSurf descriptors for 34 compounds (the most active hits and the reference Kv1.5 blockers). The color code is the same as for Figure 5.

Table 4. Performance of Ligand- and Sturcture-Based Virtual Screening for Kv1.5

method	Ncomp[a]	Nact[b]	hit rate (%)	best IC_{50} (μM)[c]	Nclass[d]
2D similarity search[20]	75	2	2.7	7.7	2
ligand-based pharmacophore[21]	18	1	5.5	5.6	1
protein-based pharmacophore	244	19	7.8	0.9	5

[a] Number of compounds selected for in vitro screening. [b] Number of active compounds. [c] IC_{50} values of the most active compound. [d] Number of chemical classes identified.

remaining two hits belong to two quadrants of the plot (Figure 6). These two other quadrants are fairly well populated by Kv1.5 blockers of the reference set. In other words, some of the hits identified by structure-based VS might exhibit different physical properties. At this point, we should also mention that loading profiles for the most explanatory components of the PCA on CATS and VolSurf descriptors are available as Supporting Information.

Earlier in this project, we also ran 2D similarity searches based on UNITY fingerprints and ligand-derived 3D pharmacophore searches within the Aventis compound collection. As previously reported, these ligand-based VS efforts also led to the identification of Kv1.5 blockers.[20,21] In Table 4, we compare the performances of these ligand-based VS approaches with those of the structure-based VS protocol. Clearly, the structure-based VS protocol yielded a higher hit rate, more active compounds and more chemotypes than the ligand-based approaches. It is also noteworthy that there was no overlap between the hit lists identified by ligand- and structure-based approaches. The reasons for this absence of overlap are mentioned in the paragraph about the results of the structure-based approach.

To summarize, application of the structure-based VS protocol described in the Methods section has resulted in the identification of five novel classes of Kv1.5 blockers with IC_{50} values in the micromolar range. Some of these compounds might also exhibit different physicochemical and pharmacokinetic profiles, compared to previously known Kv1.5 blockers.

DISCUSSION AND CONCLUSIONS

In this paper, we have applied a structure-based VS protocol to identify novel Kv1.5 blockers and compared its performance with ligand-based VS approaches. Virtual screening has played a key role in our lead finding strategies, since no HTS assay was available for this target.

The structure-based VS protocol relies on a homology model of the Kv1.5 pore. The quality of a homology model depends on the level of sequence identity between the template of known 3D structure and the protein to be modeled.[52] In this case, the level of sequence identity between the target and the template (30.2%) is close to the threshold of 30% sequence identity. The alignment errors increase rapidly below this threshold.[53] However, Evers and Klebe reported a successful VS protocol based on a homology model of the neurokinin-1 receptor, which shares only 27% sequence identity with the transmembrane region of the bovine rhodopsin receptor, used as template.[54] Furthermore, after geometrical analysis, the quality of the Kv1.5 model was deemed sufficient to initiate structure-based VS efforts. We used the homology model of the Kv1.5 pore for two purposes. First of all, the homology model served as input for a binding site detection algorithm. This led to the identification of two relevant putative binding sites: a large and shallow site located on the extracellular surface and an

internal cavity underneath the selectivity filter. Mutational studies have confirmed the relevance of this internal binding site.[24,49] Therefore, we decided to focus our VS efforts on this second site, referred to as the internal site. Subsequently, we further characterized the internal site with the GRID force field in order to derive a protein-based pharmacophore. This pharmacophore served as query for database searching.

In contrast to other recent successful applications of VS using homology models, we did not perform any high throughput docking.[54−56] In our opinion, the evaluation of high throughput docking results would have been difficult, since no experimental 3D structure is available for an ion channel:small molecule complex and no evaluation of the performances of scoring functions is available for this target class. Instead, we preferred to apply softer filters, which led to the selection of 244 compounds for in vitro screening. As a result, 19 compounds were found to be active (7.8% hit rate), with five of them showing an IC_{50} value lower than 10 μM.

The hit rate is not the only criterion to assess the effectiveness of a VS approach, since the hit rate can be biased by the presence of close analogues in the hit list. To investigate this possibility, the pairwise similarities based on UNITY fingerprints were computed for the five most active hits. The Tanimoto similarities given in Table 2 indicate that none of these five compounds are close analogues. Subsequent similarity searches in the same database yielded for each of the five most active hits, a significant number of analogues, which were not identified by the structure-based VS approach. Therefore, the structural diversity of the hits is not a feature of the screened database. Since the possibility of false negatives cannot be excluded, we also suggest that substructure and 2D similarity searching could be used to expand out a single hit to molecules in the same chemical series. Recent work on the evaluation of reduced graph descriptors for similarity searching led to the same conclusion.[57] Furthermore, using the same similarity metrics (Figure 4), we found that none of these five hits was a close analogue of any member of a set of representative Kv1.5 blockers (Figure 2). On the other hand, four of these compounds clustered with known Kv1.5 blockers in the space of topological pharmacophore descriptors (Figure 5). This is an important result as it shows that this structure-based VS protocol can identify Kv1.5 blockers with significantly different molecular backbones (scaffold hopping). The simultaneous availability of multiple chemical series represents a key asset for a drug discovery project, since it provides different opportunities for synthetic medicinal chemistry and possible back-ups.

The structure-based VS protocol described in this work clearly outperformed ligand-based VS approaches applied in this project, in terms of both the hit rate and the number of chemotypes identified. Obviously, the effectiveness of different VS approaches is project dependent. Recently, Clark and colleagues obtained for a G-protein-coupled receptor, higher hit rates with ligand-based methods than with a protein-derived pharmacophore.[11] Overall, these observations support the value of employing multiple VS approaches in a given project.[6] These VS methodologies have not only different strengths and shortcomings, but they also consider different starting points. In particular, the ligand-based pharmacophore consists of three hydrophobic points that are consistent with the structure−activity relationships within two series of compounds. On the other hand, the protein-derived pharmacophore records all the potential interaction features of a given conformational state of a binding site. Different starting points are likely to yield different hit series, which could not be uncovered by a single approach.

In summary, a structure-based VS protocol, using a homology model of the Kv1.5 pore, domain led to the identification of five novel classes of Kv1.5 blockers with IC_{50} values under 10 μM. In this project, the structure-based strategy proved to be more effective than two ligand-based VS methods, in terms of hit rate and number of identified chemotypes. However, ligand-based VS approaches enabled the identification of compounds which have not been found by the structure-based protocol. All these different approaches can be seen as complementary. It should also be stressed that the ligand-based pharmacophore and structure-based searches were run in parallel. These provided medicinal chemists with different starting points for optimization. Some of these optimization efforts have already been reported, and further medicinal chemistry work on these series will be reported in due course.

ACKNOWLEDGMENT

Thank you to Annett Fengler for assistance with the homology modeling, to Andrea Brüggemann for measuring the biological activities, and to Hans Matter and Clemens Giegerich for the implementation of the CATS descriptors.

Supporting Information Available: Ramachandran plots for the template and the homology model, tables presenting the results of the geometric analysis for the template and the homology model as well as similarity values between the hits and the reference compounds, and loading profiles for the most significant PCs from the PCA on CATS and VolSurf descriptors. This material is available free of charge via the Internet at http://pubs.acs.org.

REFERENCES AND NOTES

(1) Bajorath, J. Integration of Virtual and High-Throughput Screening. *Nat. Rev. Drug Discovery* **2002**, *1*, 882−894.
(2) Langer, T.; Hoffmann, R. D. Virtual Screening: an Effective Tool for Lead Structure Discovery? *Curr. Pharm. Des.* **2001**, *7*, 509−527.
(3) Walters, W. P.; Stahl, M. T.; Murcko, M. A. Virtual Screening − an Overview. *Drug Discovery Today* **1998**, *3*, 160−178.
(4) Stahura, F. L.; Bajorath, J. Virtual Screening Methods that Complement HTS. *Comb. Chem. High Throughput Screening* **2004**, *7*, 259−269.
(5) Lengauer, T.; Lemmen, C.; Rarey, M.; Zimmermann, M. Novel Technologies for Virtual Screening. *Drug Discovery Today* **2004**, *9*, 27−34.
(6) Sheridan, R. P.; Kearsley, S. K. Why Do we Need so many Chemical Similarity Search Methods? *Drug Discovery Today* **2002**, *7*, 903−910.
(7) Bajorath, J. Selected Concepts and Investigations in Compound Classification, Molecular Descriptor Analysis, and Virtual Screening. *J. Chem. Inf. Comput. Sci.* **2001**, *41*, 233−245.
(8) Shoichet, B. K.; McGovern, S. L.; Wei, B.; Irwin, J. J. Lead Discovery Using Molecular Docking. *Curr. Opin. Chem. Biol.* **2002**, *6*, 439−446.
(9) Lyne, P. D. Structure-Based Virtual Screening: an Overview. *Drug Discovery Today* **2002**, *7*, 1047−1055.
(10) Schneider, G.; Boehm, H.-J. Virtual Screening and Fast Automated Docking Methods. *Drug Discovery Today* **2002**, *7*, 64−70.
(11) Clark, D. E.; Higgs, C.; Wren, S. P.; Dyke, H. J.; Wong, M.; Dennis, N.; Lockey, P. M.; Roach, A. G. A Virtual Screening Approach to Finding Novel and Potent Antagonists at the Melanin-Concentrating Hormone 1 Receptor. *J. Med. Chem.* **2004**, *47*, 3962−3971.
(12) Lyne, P. D.; Kenny, P. W.; Cosgrove, D. A.; Deng, C.; Zabludoff, S.; Wendoloski, J. J.; Ashwell, S. Identification of Compounds with

Nanomolar Binding Affinity for Checkpoint Kinase-1 Using Knowledge-Based Virtual Screening. *J. Med. Chem.* **2004**, *47*, 8, 1962−1968.
(13) Rastelli, G.; Pacchioni, S.; Sirawaraporn, W.; Sirawaraporn, R.; Parenti, M.; Marco, D.; Ferrari, A. M. Docking and Database Screening Reveal New Classes of Plasmodium Falciparum Dihydrofolate Reductase Inhibitors. *J. Med. Chem.* **2003**, *46*, 2834−2845.
(14) Brenk, R.; Naerum, L.; Graedler, U.; Gerber, H.-D.; Garcia, G. A.; Reuter, K.; Stubbs, M. T.; Klebe, G. Virtual Screening for Submicromolar Leads of tRNA-guanine Transglycosylase Based on a New Unexpected Binding Mode Detected by Crystal Structure Analysis. *J. Med. Chem.* **2003**, *46*, 1133−1143.
(15) Grüneberg, S.; Stubbs, M. T.; Klebe, G. Successful Virtual Screening for Novel Inhibitors of Human Carbonic Anhydrase: Strategy and Experimental Confirmation. *J. Med. Chem.* **2002**, *45*, 3588−3602.
(16) Doman, T. N.; McGovern, S. L.; Witherbee, B. J.; Kasten, T. P.; Kurumbail, R.; Stallings, W. C.; Connolly, D. T.; Shoichet, B. K. Molecular Docking and High-Throughput Screening for Novel Inhibitors of Protein Tyrosine Phosphatase-1B. *J. Med. Chem.* **2002**, *45*, 2213−2221.
(17) Evensen, E.; Eksterowicz, J. E.; Stanton, R. V.; Oshiro, C.; Grootenhuis, P. D. J.; Bradley, E. K. Comparing Performance of Computational Tools for Combinatorial Library Design. *J. Med. Chem.* **2003**, *46*, 5125−5128.
(18) Good, A. C.; Cheney, D. L.; Sitkoff, D. F.; Tokarski, J. S.; Stouch, T. R.; Bassolino, D. A.; Krystek, S. R.; Li, Y.; Mason, J. S.; Perkins, T. D. J. Analysis and Optimization of Structure-based Virtual Screening Protocols. 2. Examination of Docked Ligand Orientation Sampling Methodology: Mapping a Pharmacophore for Success. *J. Mol. Graph. Mod.* **2003**, *22*, 31−40.
(19) Brendel, J.; Peukert, S. Blockers of the Kv1.5 Channel for the Treatment of Atrial Arrhythmias. *Curr. Med. Chem.-Cardiovascular Hematological Agents* **2003**, *1*, 273−287.
(20) Peukert, S.; Brendel, J.; Pirard, B.; Brüggemann, A.; Below, P.; Kleemann, H.-W.; Hemmerle, H.; Schmidt, W. Identification, Synthesis, and Activity of Novel Blockers of the Voltage-Gated Potassium Channel Kv1.5. *J. Med. Chem.* **2003**, *46*, 486-498.
(21) Peukert, S.; Brendel, J.; Pirard, B.; Strübing, C.; Kleemann, H.-W.; Böhme, T.; Hemmerle, H. Pharmacophore-Based Search, Synthesis, and Biological Evaluation of Anthranilic Amides as Novel Blockers of the Kv1.5 Channel. *Bioorg. Med. Chem. Lett.* **2004**, *14*, 2823−2827.
(22) Yellen, G. The Voltage-Gated Potassium Channels and their Relatives. *Nature* **2002**, *419*, 35−42.
(23) Doyle, D. A.; Morais Cabral, J.; Pfuetzner, R. A.; Kuo, A.; Gulbis, J. M.; Cohen, S. L.; Chait, B. T.; MacKinnon, R. The Structure of the Potassium Channel: Molecular Basis of K^+ Conduction and Selectivity. *Science* **1998**, *280*, 69−77.
(24) Mitcheson, J. S.; Chen, J.; Lin, M.; Culberson, C.; Sanguinetti, M. A Structural Basis for Drug-Induced Long QT Syndrome. *Proc. Natl. Acad. Sci. U.S.A.* **2000**, *97*, 12329−12333.
(25) SYBYL6.8; Tripos Inc., St. Louis, MO, 2001 (www.tripos.com).
(26) Weiner, S. J.; Kollman, P. A.; Case, D. A.; Singh, U. C.; Ghio, C.; Alagona, G.; Profeta, S. Jr.; Weiner, P. A New Force Field for Molecular Mechanical Simulation of Nucleic Acids and Proteins. *J. Am. Chem. Soc.* **1984**, *106*, 765−784.
(27) Laskowski, R. A.; MacArthur, M. W.; Moss, D. S.; Thornton, J. M. PROCHECK: a Program to Check the Stereochemical Quality of Protein Structures. *J. Appl. Crystallogr.* **1993**, *26*, 283−291.
(28) Morris, A. L.; MacArthur, M. W.; Hutchison, E. G.; Thornton, J. M. Stereochemical Quality of Protein Structure Coordinates. *Proteins* **1992**, *12*, 345−364.
(29) Hooft, R. W. W.; Vriend, G.; Sander, C.; Abola, E. E. Errors in Protein Structures. *Nature* **1996**, *381*, 272−272.
(30) Brady, G. P.; Stouten, P. F. W. Fast Prediction and Visualization of Protein Binding Pockets with PASS. *J. Comput.-Aided Mol. Des.* **2000**, *14*, 383−401.
(31) Available on the Web: http://www.ccl.net/cca/software/UNIX/pass/overview.shtml.
(32) GRID version 19; Molecular Discovery Ltd., Oxford, UK, 2001 (http://www.moldiscovery.com/).
(33) Bleicher, K. H.; Böhm, H.-J.; Müller, K.; Alanine, A. Hit and Lead generation: beyond High-Throughput Screening. *Nat. Rev. Drug Discovery* **2003**, *2*, 369−377.
(34) UNITY; Tripos Inc., St. Louis, MO, 2001 (www.tripos.com).
(35) Lipinski, C. A.; Lombardo, F.; Dominy, B. W.; Feeney, P. J. Experimental and Computational Approaches to Estimate Solubility and Permeability in Drug Discovery and Development Settings. *Adv. Drug Deliv. Rev.* **1997**, *23*, 3−25.
(36) Rishton, G. M. Nonleadlikeness and Leadlikeness in Biochemical Screening. *Drug Discovery Today* **2003**, *8*, 86−96.
(37) Lajiness, M. S.; Maggiora, G. M.; Shanmugasundaram, V. Assessment of the Consistency of Medicinal Chemists in Reviewing Sets of Compounds. *J. Med. Chem.* **2004**, *47*, 4891−4896.
(38) Rarey, M.; Dixon, J. S. Feature Trees: a New Molecular Similarity Measure Based on Tree Matching. *J. Comput.-Aided Mol. Des.* **1998**, *12*, 471−490.
(39) Schneider, G.; Neidhart, W.; Giller, T.; Schmid, G. ×e3Scaffold-Hopping“ by Topological Pharmacophore Search: a Contribution to Virtual Screening. *Angew. Chem., Int. Ed. Engl.* **1999**, *38*, 2894−2896.
(40) Cruciani, G.; Pastor, M.; Guba, W. VolSurf: a New Tool for the Pharmacokinetic Optimization of Lead Compounds. *Eur. J. Pharm. Sci.* **2000**, *11* (Suppl. 2), S29−S39.
(41) Schneider, G.; Nettekoven, M. Ligand-Based Combinatorial Design of Selective Purinergic Receptor (A_{2A}) Antagonists Using Self-Organizing Maps. *J. Comb. Chem.* **2003**, *5*, 233−237.
(42) Naerum, L.; Nørskov-Lauritsen, L.; Olesen, P. H. Scaffold Hopping and Optimization towards Libraries of Glycogen Synthase Kinase-3 Inhibitors. *Bioorg. Med. Chem. Lett.* **2002**, *12*, 1525−1528.
(43) GOLPE Version 4.5; MIA S.r.l, Perugia, Italy, 2004 (http://www.miasrl.com/index.htm).
(44) Franke, R.; Gruska, A. Principle Component and Factor Analysis. In *Chemometric Methods in Molecular Design*, *Methods and Principle in Medicinal Chemistry*; van de Waterbeemd, H., Ed.; VCH Publishers: Weinheim, 1995; Vol. 2, p 113.
(45) Cruciani, G.; Pastor, M.; Mannhold, R. Suitability of Molecular Descriptors for Database Mining. A Comparative Analysis. *J. Med. Chem.* **2002**, *45*, 2685−2694.
(46) CORINA; Molecular Networks GmbH, Erlangen, Germany (http://www.mol-net.de/).
(47) VolSurf3.0; Molecular Discovery Ltd., Oxford, UK, 2003 (http://www.moldiscovery.com/).
(48) Pongs, O. Pharmacology of Voltage-Gated Potassium Channels. In *Perspectives in Drug Discovery and Design*; Darbon, H., Sabatier, J.-M., Eds.; Kluwer/Escom: Dordrecht, 1999; Vols. 15/16, p 215.
(49) Decher, N.; Pirard, B.; Bundis, F., Peukert, S.; Baringhaus, K.-H.; Busch, A. E.; Steinmeyer, K.; Sanguinetti, M. C. Molecular Basis for Kv1.5 Channel Block. *J. Biol. Chem.* **2004**, *279*, 394−400.
(50) Martin, Y. C.; Kofron, J. L.; Traphagen, L. M. Do Structurally Similar Molecules Have Similar Biological Activity? *J. Med. Chem.* **2002**, *45*, 4350−4358.
(51) Rarey, M.; Stahl, M. Similarity Searching in Large Combinatorial Chemistry Spaces. *J. Comput.-Aided Mol. Des.* **2001**, *15*, 497−520.
(52) Hillisch, A.; Pineda, L. P.; Hilgenfeld, R. Utility of Homology Models in the Drug Discovery Process. *Drug Discovery Today* **2004**, *9*, 659−669.
(53) Baker, D.; Sali, A. Protein Structure Prediction and Structural Genomics. *Science* **2001**, *294*, 93−96.
(54) Evers, A.; Klebe, G. Ligand-Supported Homology Modeling of G-Protein-Coupled Receptor Sites: Models Sufficient for Successful Virtual Screening. *Angew. Chem., Int. Ed.* **2004**, *43*, 248−251.
(55) Liu, H.; Li, Y.; Song, M.; Tan, X.; Cheng, F.; Zheng, S.; Shen, J.; Luo, X.; Ji, R.; Yue, J.; Hu, G.; Jiang, H.; Chen, K. Structure-Based Discovery of Potassium Channel Blockers from Natural Products: Virtual Screening and Electrophysiological Assay Testing. *Chem. Biol.* **2003**, *10*, 1103−1113.
(56) Bissantz, C.; Bernard, P.; Hibert, M.; Rognan, D. Protein-Based Virtual Screening of Chemical Databases II. Are Homology Models of G-Protein Coupled Receptors Suitable Targets? *Proteins: Struct., Funct., Genet.* **2003**, *50*, 5−25.
(57) Harper, G.; Bravi, G. S.; Pickett, S. D.; Hussain, J.; Green, D. V. S. The Reduced Graph Descriptor in Virtual Screening and Data-Driven Clustering of High-Throughput Screening Data. *J. Chem. Inf. Comput. Sci.* **2004**, *44*, 2145−2156.

CI0400011