Bioorganic & Medicinal Chemistry 17 (2009) 3169–3176

Contents lists available at ScienceDirect

Bioorganic & Medicinal Chemistry

journal homepage: www.elsevier.com/locate/bmc

The identification of novel PLC-γ inhibitors using virtual high throughput screening

Jóhannes Reynisson *, William Court, Ciaran O'Neill, James Day, Lisa Patterson, Edward McDonald, Paul Workman, Matilda Katan, Suzanne A. Eccles *

Cancer Research UK Centre for Cancer Therapeutics, The Institute of Cancer Research, 15 Cotswold Road, Sutton, Surrey SM2 5NG, UK

ARTICLE INFO

Article history:
Received 28 October 2008
Revised 25 February 2009
Accepted 25 February 2009
Available online 3 March 2009

Keywords:
Virtual screening
PLC-protein
Oncology
Docking
Drug discovery

ABSTRACT

Phospholipase C-γ (PLC-γ) has been identified as a possible biological target for anticancer drug therapy but suitable inhibitors are lacking. Therefore, in order to identify active compounds (hits) virtual high throughput screening was performed. The crystal structure of the PLC-δ isoform was used as a model docking scaffold since no crystallographic data are available on its γ counterpart. A pilot screen was performed using $\sim 9.2 \times 10^4$ compounds, where the robustness of the methodology was tested. This was followed by the main screening effort where $\sim 4.4 \times 10^5$ compounds were used. In both cases, plausible compounds were identified (virtual hits) and a selection of these was experimentally tested. The most potent compounds were in the single digit micro-molar range as determined from the biochemical (Flashplate) assay. This translated into ~15 μM in a functional assay in cells. About 30% of the virtual hits showed activity against PLC-γ ($IC_{50} < 50$ μM).

1. Introduction

Phosphoinositide specific-phospholipase C (pi-PLC) is a membrane bound protein which hydrolyses phosphatidylinositol 4,5-diphosphate (PIP_2) to diacylglycerol and inositol 1,4,5-triphosphate.[1,2] Diacylglycerol activates phospholipid-dependent protein serine/threonine kinase and protein kinase C, and inositol 1,4,5-triphosphate promotes release of Ca^{2+} from intracellular stores.[1] Together, these important secondary messengers mediate downstream cell motility and proliferation.[1] It is thought that the former signalling pathway (leading to cell motility) is of particular significance in tumour progression, which suggests that it may be a potential target for therapy. Motility of tumour cells is a key component of invasion and dissemination, a major causes of morbidity and death.[3,4] There are six subfamilies belonging to the mammalian pi-PLC family, which are classified as PLC-β (β1–β4), PLC-γ (γ1 and γ2), PLC-δ (δ1–δ4), PLC-ε, PLC-ζ and PLC-η.[5–7] The PLC-γ subfamily is divided into PLC-γ1, which is present in most cell types, and the minor isoform PLC-γ2, which is restricted to hematopoietic cells.[1,7,8] The deletion of the PLC-γ1 gene in mice results in early embryonic death revealing its importance in cell proliferation and differentiation. Deletion of the PLC-γ2 gene results in impairment of immune function.[7–10] The hydrolysis of PIP_2 by PLC-γ is linked to cytoskeletal reorganisation necessary for cell polarisation and in cell adhesion signalling which implicates a wider role for the enzyme in integrin related cell motility.[2,11] Therefore, PLC-γ is involved in two key processes; namely, growth factor induced cell motility and cell adhesion, which are key to tumour development. Growth factor- and cytokine-induced cell motility occurs primarily during organogenesis, inflammation and wound repair in which inhibition of PLC-γ may be more specific for tumour cells.[12]

A range of different molecular inhibitors are known for the PLC enzymes such as steroid analogues,[13,14] peptides[15] and various natural products.[16,17] Unfortunately, none of these inhibitors are suitable for drug development, thus the need exists to identify molecular entities for this purpose.

In this paper, we describe the search for suitable starting compounds (hits) for drug development using virtual high throughput screening (vHTS) methodology for the PLC-γ enzyme. The success of the vHTS method in discovering chemically tractable hit compounds encouraged us to implement this technique.[18–26] This was possible due to the availability of crystal structure on the binding site of PLC-δ,[27] which was used as a model. This is justifiable because the region with the highest amino acid sequence similarity between the isoforms is contained within the substrate binding pocket.[28] The virtual screening was performed with compound collections acquired from the ZINC website.[29] The prediction

* Corresponding authors at present address: Department of Chemistry and Auckland Bioengineering Institute, The University of Auckland, Private Bag 92019, Auckland 1142, New Zealand (J.R.); Tumour Biology and Metastasis, Cancer Research UK Centre for Cancer Therapeutics, The Institute of Cancer Research, McElwain Laboratories, 15 Cotswold Road, Belmont, Sutton, Surrey SM2 5NG, UK (S.A.E.).
E-mail addresses: j.reynisson@auckland.ac.nz (J. Reynisson), sue.eccles@icr.ac.uk (S.A. Eccles).

0968-0896/$ - see front matter © 2009 Elsevier Ltd. All rights reserved.
doi:10.1016/j.bmc.2009.02.049

power of the GOLD (Genetic Optimisation for Ligand Docking)[30] docking algorithm was implemented. Its capabilities are well documented in the literature, and of the docking algorithms available GOLD is generally considered to give the most reliable results.[30–33]

2. Computational and experimental methods

2.1. Virtual screening

Three commercially available compound collections were downloaded from the ZINC[29] web site totalling $\sim 5.3 \times 10^5$ molecules. Two relatively small compound collections of $\sim 5.3 \times 10^4$ and $\sim 3.9 \times 10^4$ molecules were used as a pilot screen to verify the method employed. A large collection of $\sim 4.4 \times 10^5$ was used for the main screening effort. Multiple representations of each compound were included, for example, protonation states between pH 5.0 and 9.5 and tautomeric forms except for the initial pilot screen where single representation of each molecule was used. The structures were docked to the PLC-δ crystal structure (1DJX, resolution 2.30 Å)[27], which was obtained from the Protein Data Bank (PDB).[34] The QikProp 6.0 program[35] was used to prepare the crystal structure for docking, that is, hydrogen atoms were added, the co-crystallised ligand (inositol 1,4,5-triphosphate) was removed as well as crystallographic water molecules. The centre of the binding pocket was defined as the position of the Ca^{2+} ion ($x = 126.257$, $y = 38.394$, $z = 22.370$) with 10 Å radius. The basic amino acids lysine and arginine were defined as protonated. Furthermore, aspartic and glutamic acids were assumed to be deprotonated. The GOLDSCORE (GS)[30] and CHEMSCORE (CS)[36] algorithms were implemented to validate the predicted binding modes and relative energies of the ligands with GOLD 2.2 for the pilot screen and GOLD 3.0 for the main screen.[30] Library screening settings were used in conjunction with 10 docking runs per ligand. The candidates emerging from the consensus scoring and filtering (see Fig. 1) were re-docked with the standard default settings in conjunction with maximum 50 docking runs per ligand. It has been shown that the standard default setting in the GOLD software generates ligand poses closer to the parent X-ray pictures rather than the library screening setting.[37] This is important for the visualisation step in the screening cascade. For the molecular descriptors such as Log*S* (*S* = water solubility) and Log*P* the program CAChe was used.[38–42]

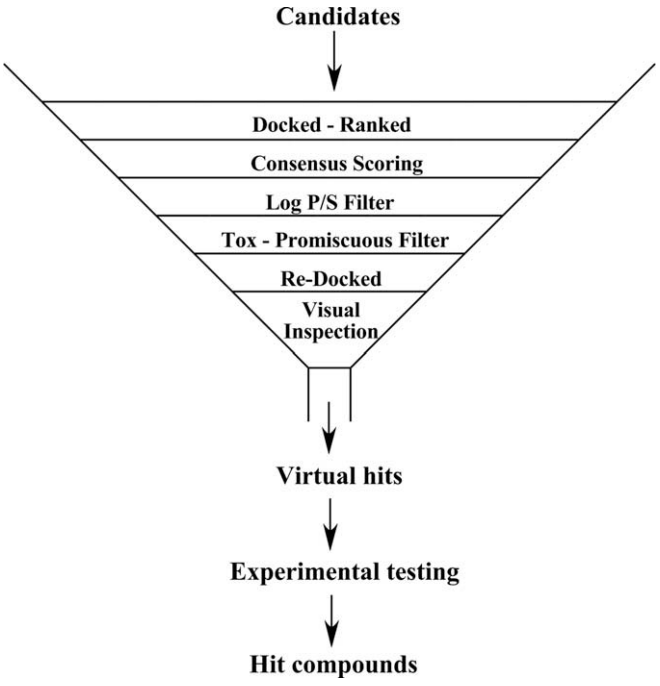


Figure 1. The screening cascade implemented in this work.

2.2. Enzyme preparation

A recombinant baculovirus encoding full length, histidine tagged, PLC-γ2 was used to infect Sf-9 insect cell shaker cultures grown in SF-900 II SFM media (Invitrogen). These were grown to a density of around 2×10^6 cells per mL and infected with baculovirus at a multiplicity of infection greater than one. Infected cultures were harvested at ~72 h post infection by centrifugation (2000*g*, 20 min). Cell pellets were resuspended in lysis buffer (50 mM Hepes pH 7.4, 300 mM NaCl, 5% glycerol, 1 mM NaF, 1 mM Na_3VO_4, 1 mM β-glycerophosphate and 1 × complete EDTA-free protease inhibitors (Roche)) and lysed by sonication. The cell lysate was clarified by centrifugation (25,000*g*, 45 min) and filtered through a 0.45 μm pore size system. The cell lysates were passed over a Talon IMAC resin (Takara Bio Europe/Clontech) packed into a Tricorn column (GE Healthcare) at 0.5 mL/min (a ration of 1 mL resin per litre of insect cell culture was used). The column was then washed with 20 column volumes of buffer (50 mM Hepes pH 7.4, 300 mM NaCl, 5% glycerol). The protein was eluted with buffer (50 mM Hepes pH 7.4, 300 mM NaCl, 250 mM imidazole, 5% glycerol). Selected eluates were concentrated by ultra filtration and dialysed against 20 mM Tris pH 7.5, 200 mM NaCl, 1 mM EDTA, 2 mM DTT, 5% glycerol.

2.3. ^{3}H-PIP$_2$ flashplate biochemical assay

The compounds were dissolved in DMSO with 10 μL aliquots which were added to the assay resulting in DMSO of 5% v/v concentration. PLC-γ2 lipase activity was measured using Flashplate Plus technology® (PerkinElmer) which incorporates a solid scintillant to which an artificial phospholipid membrane layer is covalently bound. The substrate (^{3}H-PIP$_2$) is then bound to the phospholipid layer and the proximity of the isotope to the scintillant is detected. Hydrolysis of the substrate by the enzyme removes ^{3}H-inositol and the signal is lost.[43] The plates were coated by incubation for 24 h with ^{3}H-PIP$_2$ (0.1 μCi/well) in 200 μL coating buffer (NEN) at room temperature then the excess radiolabel was discarded. The enzyme and inhibitor or DMSO in assay buffer (50 mM Tris–HCl, 200 mM NaCl, 100 μM $CaCl_2$, 100 μM EGTA, 100 μM spermine, and 0.5 mM deoxycholate, pH 7.4 ± 0.2), were added to the wells (200 μL final volume) and the assay followed using a Topcount™ scintillation counter (Perkin Elmer) at room temperature, typically for 30 min. Inhibition of enzyme activity was measured by comparing CPM (counts per minute) of sample wells with control wells.

2.4. Calcium release assay

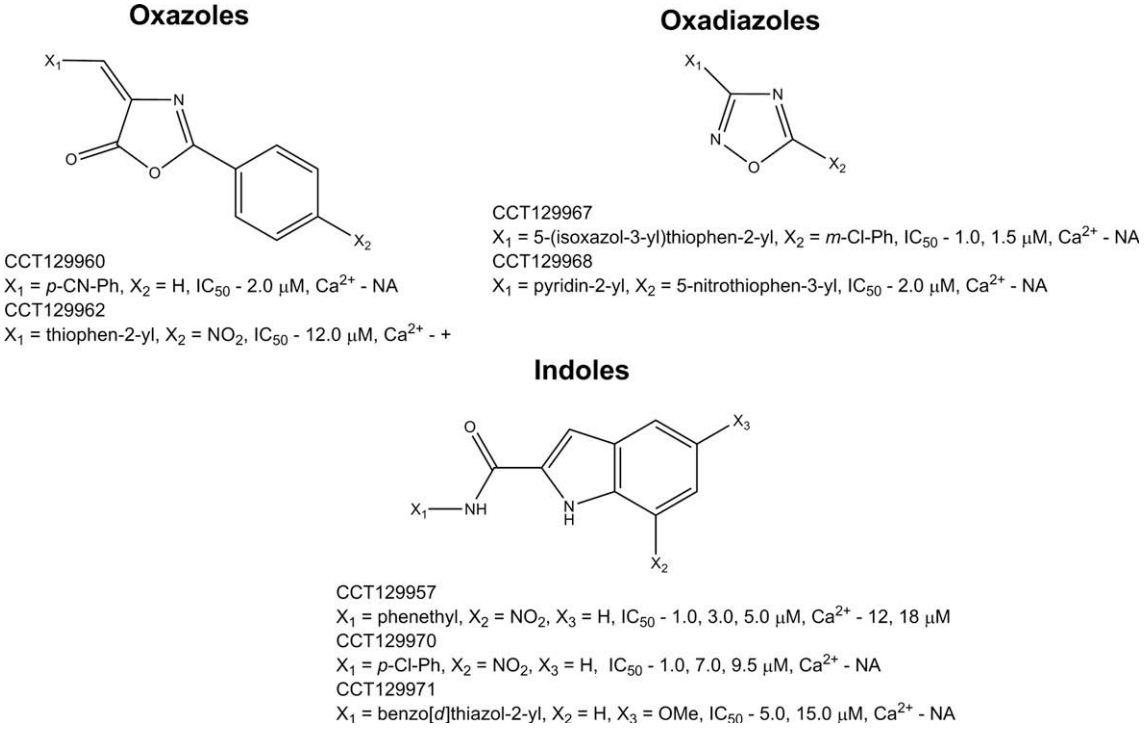
The assay performed was a modification of the method published in the Calcium 3 assay kit (Molecular Devices Corp). Briefly, 4×10^5 cells of the squamous carcinoma cell line A431 were seeded into each well of a 96 well black-walled plate in growth medium and left overnight in an incubator at 37 °C, 5% CO_2. The following day the medium was removed from the plate and replaced with a 50 μL doubling dilution of each compound in Hepes buffered saline (HBSS) starting at 100 μM. A working dilution of the calcium dye was prepared as outlined in the protocol and probenecid added to give a final concentration of 2.5 mM. A 50 μL aliquot was added to each well of the test plate which was then incubated at 37 °C, 5% CO_2 for 45 min. Then the plate was placed in a Fluorescence Imaging Plate Reader (FLIPR) at 37 °C. Each well of the plate was then stimulated with 50 μL of EGF at 400 ng/mL and the FLIPR set to acquire readings every second at

510–570 nm emission for 120 s following excitation at 488 nm. Any activity of the compound that resulted in a decrease in calcium release at 50 μM was scored as + and at 25 μM ++. In some cases the quantitative GC_{50} (Germination half-inhibition Concentration) values were determined when deemed appropriate.

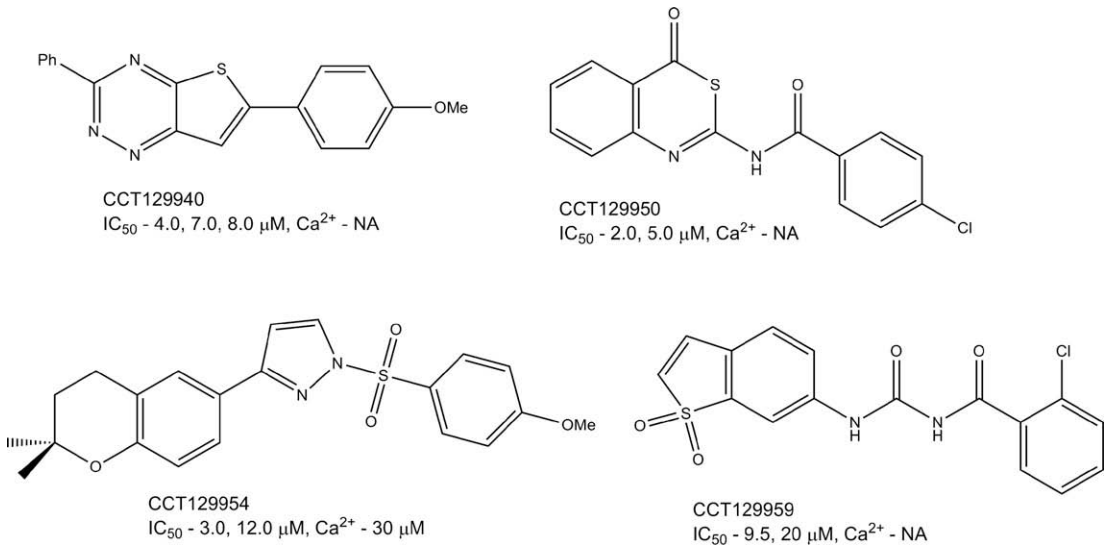
3. Results

3.1. Pilot screen

The screening cascade used in this project is shown in Figure 1. The first step was to dock and rank all of the compounds ($\sim 9.2 \times 10^4$) using GOLD- and CHEMSCORE; cut-off filters were implemented based on the relative score. Furthermore, all compounds showing no hydrogen-bonding activity were disqualified. Only a single representation of each molecule was used. At this stage consensus scoring[44–47] was executed, that is, compounds that were within the top ~5% for both fitness functions were carried forward. This brought the number of candidates to <1% of the total number of docked compounds. The Log*P* and Log*S* values were calculated in order to eliminate water insoluble and excessively lipophilic compounds. The limits were set at Log*P* < 4.5 and Log*S* > −5.5. Furthermore, after filtration the remaining candidates were visually inspected for promiscuous molecular species (frequent hitters) and compounds containing toxic moieties. Many such molecular entities are known and documented in the literature.[48–52] The



Scheme 1. The three chemical families identified in the pilot screen according to the results from the ^{3}H-PIP$_2$ assay. The IC_{50} are the results of the ^{3}H-PIP$_2$ assay and Ca^{2+} are the results of the calcium release cell-based assay (NA = not active). Often the measurements were repeated in which case all the results are depicted.



Scheme 2. Miscellaneous compounds found in the pilot screen. The IC_{50} are the results of the ^{3}H-PIP$_2$ assay and Ca^{2+} are the results of the calcium release cell-based assay (NA = not active). Often the measurements were repeated in which case all the results are shown.

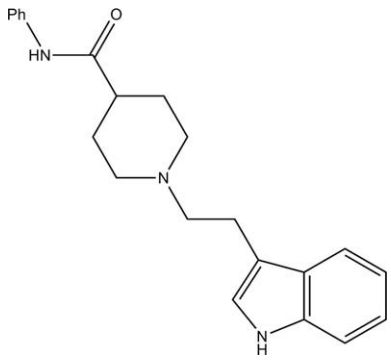


Figure 2. The molecular structure of Idoramin, which is a marketed drug structurally related to CCT129957.

two last steps brought down the number of candidates to ~0.3% of the compound collections screened. At this point the molecules were re-docked using a more rigorous theoretical approach (see computational methods). In addition, multiple representations of each compound were included. The docked configurations of the candidate compounds were visually examined and disqualified according to whether CHEM- and GOLDSCORE predicted a similar pose; whether lipophilic moieties protrude into the water phase and hydrogen-bonding pattern. This procedure reduced the number of candidates to only ~0.05% (50 compounds) of the total number of molecules screened and these were defined as virtual hits. Thirty-seven of these were commercially available and tested against PLC-γ2 using the ^{3}H-PIP$_2$ assay. Eleven of the compounds were considered active (IC_{50} < 50 μM). This represented ~30% of the tested compounds, which was a very acceptable result. No solubility problems were encountered for these compounds in the ^{3}H-PIP$_2$ assay. Three chemical families were identified, that is, oxazole-, oxadiazole- and indole-derivatives and they are presented in Scheme 1. Additionally, a group of miscellaneous compounds was defined, which are depicted in Scheme 2. Some of the compounds were quite potent with measured activity in the low single digit micro-molar region such as CCT129950, CCT129957 and CCT129967.

Twenty-six of the compounds were tested in the calcium release cell-based assay and five of them had activity <50 μM, which represented ~19% of the compounds. The most potent molecule investigated was CCT129957 of the indole family with a GC_{50} of ~15 μM (see Scheme 1). The second most potent compound was CCT129954, which is a singleton and has a measured GC_{50} of ~30 μM (see Scheme 2).

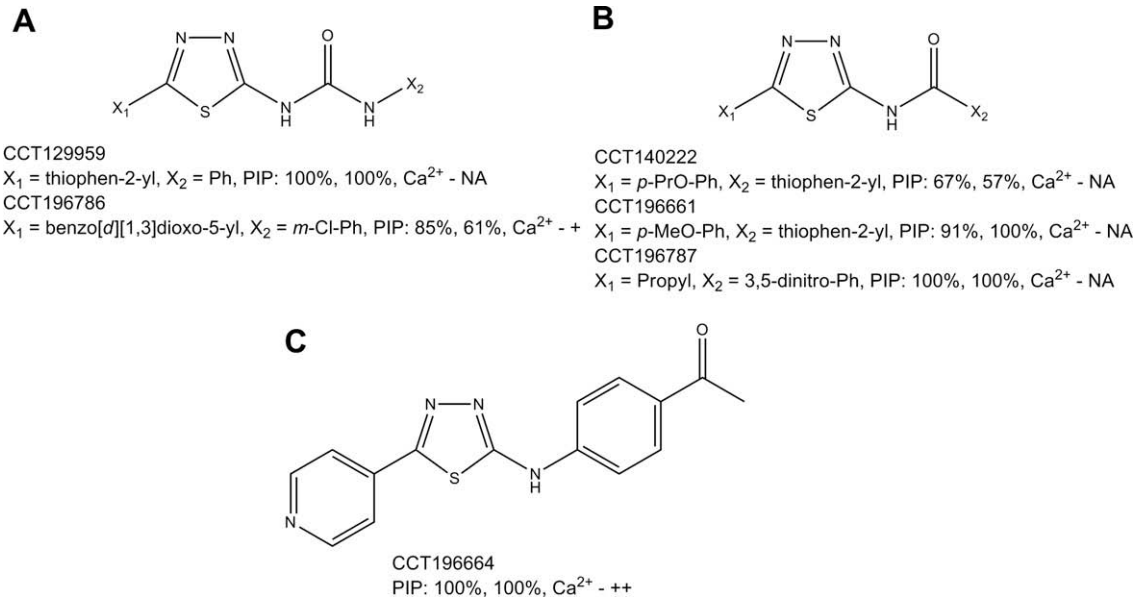
There are two compounds that are clearly the most interesting for drug development emerging from this pilot screen, that is, CCT129954 and CCT129957. The latter has better potency in both assays and therefore can be considered the *best* compound identified. Interestingly, CCT129957 has a close structural relative marketed as Doralese® or Baratol® with the active ingredient Indoramin. It is used to treat enlarged prostate glands and hypertension.[53,54] The molecular structure of this drug is shown in Figure 2.

3.2. Main screen

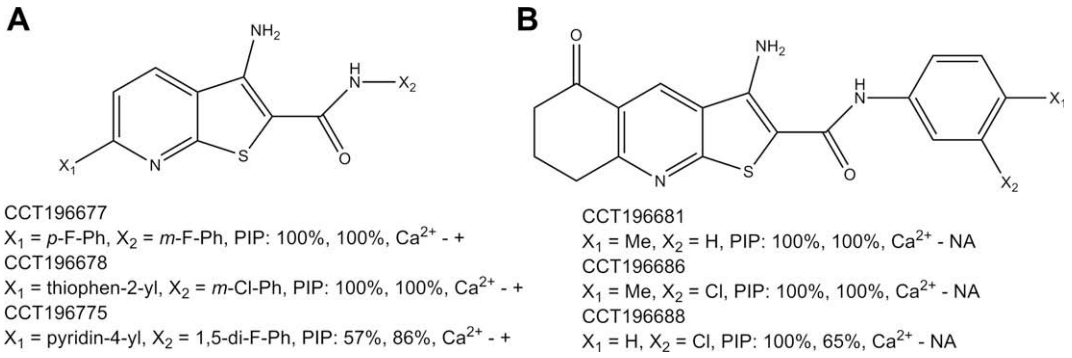
For the main screen, an attempt was made to improve the process by analysing the results of the pilot screen. An important difference between the pilot and main screens is that multiple representations of the compounds were used in both docking stages of the latter. 4.4×10^5 Compounds were represented by 6.3×10^5 molecular species. First, cut-off filters were used based on the predicted affinity data. Second, all ligands showing no hydrogen-bonding were excluded. These two steps reduce the number of candidates using CHEMSCORE to 7.4% of the total collection and 22.8% for GOLDSCORE. Consensus scoring was the third step and it resulted in reduction of candidates to ~2.0% of the total collection for which the Log*P* and Log*S* were calculated. All the compounds with Log*P* > 5 and Log*S* < −6 were not considered further, which led to a reduction of candidates to ~0.8%. All of these were visually inspected for frequent hitters and compounds containing toxic moieties leading to a reduction to ~0.7%. In order to reduce the number of candidates even further, molecular descriptors were used based on the eleven compounds showing activity in the ^{3}H-PIP$_2$ assay from the pilot screen. All of the hits had an elongated shape and their mean globularity was calculated as 0.823 (complete sphere = 1) with the standard deviation of $\sigma = 0.013$ (max = 0.843 and min = 0.799). Based on this analysis all the candidates with globularity between 0.84 and 0.80 were taken forward and the rest discarded. Furthermore, the following molecular descriptors were employed in the same way as the globularity descriptor: MW, Log*P*, dipole moment, rotatable bonds, number of hydrogen bond donors and acceptors. The limits established with this type of statistical analysis were: MW < 400, Log*P* 0–4.5, dipole moment, 0–12.5, rotatable bonds <7, HbA 2–10 and HbD 0–5. The application of these limits reduced the number of candidates to ~0.2%. Also, a more stringent toxicity, promiscuity and reactivity moiety inspection was implemented and it brought the number of compounds considered to ~0.14%. The remaining compounds were re-docked (1046 molecular derivatives). The docked compounds were visually inspected using the same criteria as in the pilot screen. After the elimination process 210 molecular species remained representing 171 compounds, which is 0.04% of the total collection screened.

Eighty-three of these were experimentally tested using the ^{3}H-PIP$_2$ biochemical assay format at 50 μM and two measurements were made for each compound. Only semi-quantitative measurements were obtained for these molecules. Thirty-one of the virtual hits gave 100% inhibition of PLC-γ activity for both measurements, which was ~38% of the compounds tested. This means that their IC_{50} was lower than 50 μM. Eleven of the candidates gave 100% inhibition for one of the measurements and a high inhibition (>60%) in the second measurement, representing ~13% of tested compounds. Here it can be argued that quantitative measurements would lead to an IC_{50} lower than 50 μM, perhaps as low as ~30 μM. Sixteen ligands gave inhibition between 99% and 50% for both measurements (~20%) and they can be roughly estimated at IC_{50} ~50 μM. The remainder of the candidates (24) showed little or no inhibition. For the Ca^{2+} release cell-based assay the same 83 compounds were tested and 32 of these were active (39%), two were marginally active and 49 compounds were inactive (59%). Seventeen of the compounds had solubility issues in the Ca^{2+} release assay but this was solved either by heating, sonicating and vortexing or a combination of these. Only one compound was completely insoluble.

Three chemical families were identified. The first family contained a 1,3,4-thiadiazole ring moiety in the centre of the compounds as shown in Scheme 3. Six compounds were represented in this family and they were all active according to the ^{3}H-PIP$_2$ hydrolysis assay. However, only two compounds showed activity in the cell-based Ca^{2+} assay. The basic structure of this family are three aryl-rings arranged in a linear fashion with 1,3,4-thiadiazole in the centre. These can be split into two groups and a singleton depending on the type of linker between the 1,3,4-thiadiazoles and the aryl groups (X_2) as shown in Scheme 3. Group A has a urea linker, group B an amide linker and finally the singleton (C) has the shortest linker consisting of an amine moiety. Interestingly, the singleton has the best overall affinity in this chemical family.



Scheme 3. The 1,3,4-thiadiazole family. This family is split into two groups and one singleton. PIP is the percentage of inhibition at 50 μM for two measurements using the ^{3}H-PIP$_2$ assay. Ca^{2+} is the result of the cell-based assay. Any activity of the compound that resulted in a decrease in calcium release at 50 μM was scored as + and at 25 μM ++. (NA—not active).

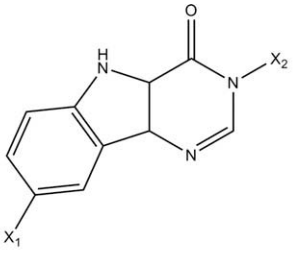


Scheme 4. The thieno[2,3-*b*]pyridine-3-amine family. PIP is the percentage of inhibition at 50 μM for two measurements using the ^{3}H-PIP$_2$ assay and Ca^{2+} is the result of the cell-based assay. Any activity of the compound that resulted in a decrease in calcium release at 50 μM was scored as + and at 25 μM ++. (NA means not active).

Six compounds were found in the thieno[2,3-*b*]pyridine-3-amine family and were split into two groups as depicted in Scheme 4. First, the structures on the left side (A) have a substituent on the C6 (X_1), in all cases an aryl ring. The X_2 substituent depicts a phenyl moiety substituted with fluorine or chlorine. All of these compounds were active in the ^{3}H-PIP$_2$ hydrolysis assay and also showed moderate activity in the cell-based Ca^{2+} assay. The second group (depicted in B) has also a thieno[2,3-*b*]pyridine-3-amine core but in addition a cyclohexanone moiety (3-amino-7,8-dihydrothieno[2,3-*b*]quinolin-5(6*H*)-one). Three derivatives were found with methyl and chlorine substituents on the *meta*- and *para*- positions on the phenyl moiety. None of these compounds were active in the cell-based assay.

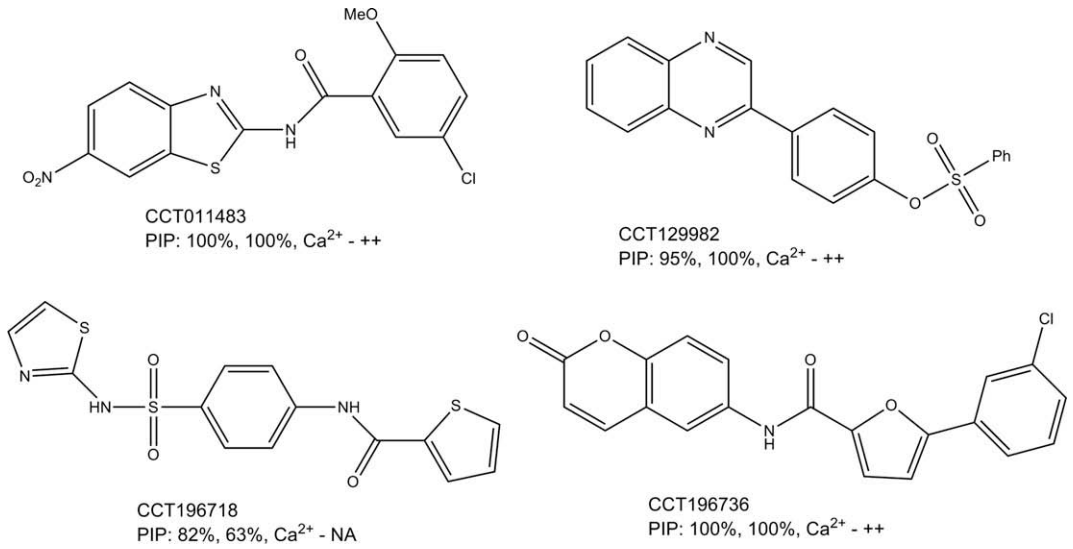
The third chemical family identified had a 3*H*-pyrimido[5,4-*b*]indol-4(5*H*)-one core with mainly bromine substituted at the C8 position (X_1) but also meth- and ethoxy as shown in Scheme 5. At the X_2 site various ring systems occur with an ethyl linker except in one case the aliphatic linker is composed of one carbon atom (CCT196797). All of the compounds are active against PLC-γ according to the ^{3}H-PIP$_2$ assay, however only one compound (CCT196720) displayed a moderate activity in the cell- based assay.

In Scheme 6, miscellaneous compounds are shown. These were only a few of the active compounds identified but give an idea of the types of molecules found. They all contained aromatic and aryl

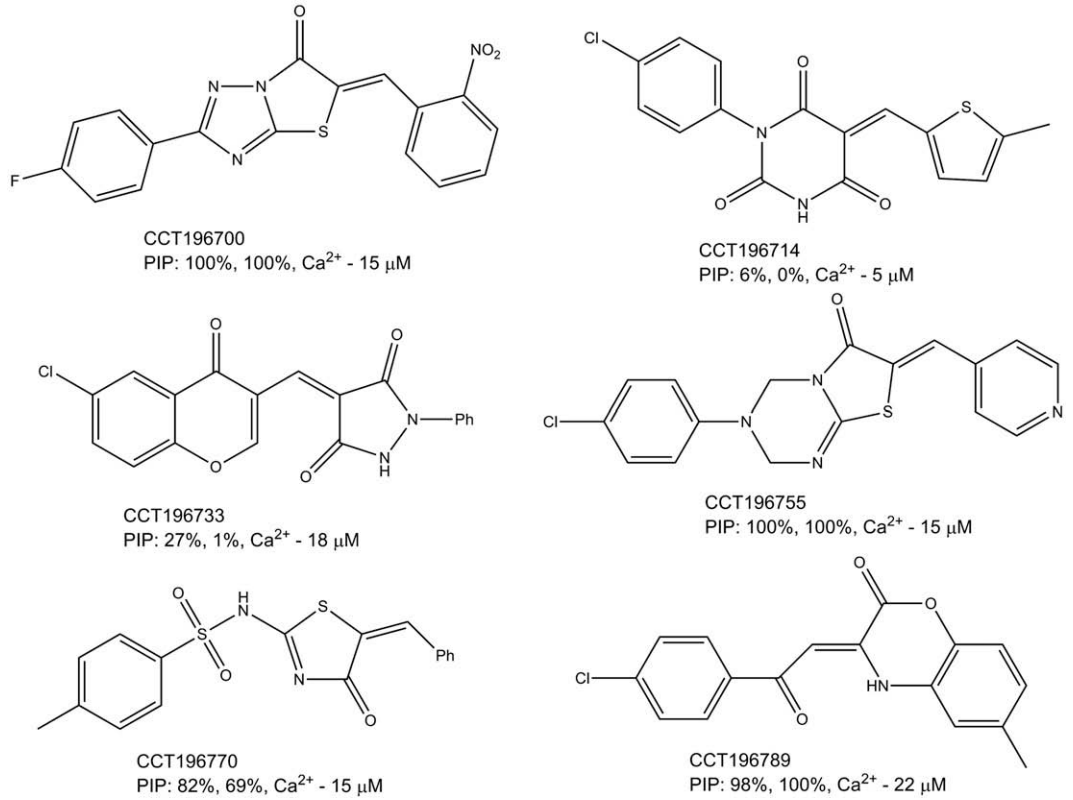


CCT196720
X1 = MeO, X2 = indol-3-yl-ethyl, PIP: 89%, 62%, Ca2+ - +
CCT196722
X1 = Br, X2 = 2-cyclohexenyl-ethyl, PIP: 100%, 100%, Ca2+ - NA
CCT196768
X1 = Br, X2 = phenethyl, PIP: 83%, 100%, Ca2+ - NA
CCT196777
X1 = Br, X2 = p-F-phenethyl, PIP: 100%, 100%, Ca2+ - NA
CCT196797
X1 = EtO, X2 = benzyl, PIP: 76%, 89%, Ca2+ - NA

Scheme 5. The 3*H*-pyrimido[5,4-*b*]indol-4(5*H*)-one group. PIP is the percentage of inhibition at 50 μM for two measurements using the ^{3}H-PIP$_2$ assay and Ca^{2+} is the result of the cell-based assay. Any activity of the compound that resulted in a decrease in calcium release at 50 μM was scored as + and at 25 μM ++. (NA means not active).



Scheme 6. Miscellaneous hit compounds. The four molecules depict the chemical variety of compounds found in the main screen. PIP is the percentage of inhibition at 50 μM for two measurements and Ca^{2+} is the result of the cell-based assay. Any activity of the compound that resulted in a decrease in calcium release at 50 μM was scored as + and at 25 μM ++.



Scheme 7. The most active compounds according to the Ca^{2+} cell-based assay results and their quantitative GC_{50} determination.

ring systems and were elongated in shape like the rest of the compounds with measured activity against the PLC-γ enzyme.

Six of the compounds tested were sufficiently potent to merit a quantitative GC_{50} determination using the cell-based assay and the results are shown in Scheme 7.

Two of the compounds (CCT196714 and CCT196733) demonstrate good potency according to the cell-based assay. They, however, interact weakly or not at all with PLC-γ. This means that the potency in the cell-based assay were not due to the inhibition of PLC-γ but some other mechanism(s). The rest of the compounds depicted in Scheme 7 were potent in both assays.

Most of the experimental data generated for the compounds identified in the main screen are semi-quantitative. Undoubtedly, it is much better to have the full quantitative IC_{50} and GC_{50} for all of the virtual hits generated; however, it was not always possible to obtain the data due to material restrictions and other factors.

3.3. Binding to PLC

Modelling of the most interesting ligands was undertaken again using the PLC-δ isoform as a model. CCT129957 was chosen since it was the most potent compound measured both in the biochemical and cell-based assays (see Scheme 1). Figure 3 depicts this ligand docked in the binding site. Hydrogen bonding is predicted (green lines, Fig. 3A) between the indole moiety and the amino acids Glu341, Lys438 and Arg549. In all of these cases, the side chains of the amino acids were involved in the hydrogen bonding mostly with the $-NO_2$ moiety in the indole ring. Also, the carboxyl group in the linker is predicted to interact with the Ca^{2+} cation in the binding pocket. The phenyl group on the left side sat in a lipophilic pocket formed of various amino acids as depicted in Figure 3B. In general, the predicted binding mode of CCT129957 displayed a robust hydrogen bond pattern as well as a lipophilic contact, that is, everything a potent ligand needs as is reflected in the assay results. Thus, the conclusion can be reached that the binding mode was plausible.

Several other inhibitors were also modelled and the same pattern emerged as for the CCT129957 ligand, that is, a hydrogen bonding network with the Glu341, Lys438 and Arg549 amino acids and a lipophilic contact in the pocket on the left side as depicted in Figure 3B.

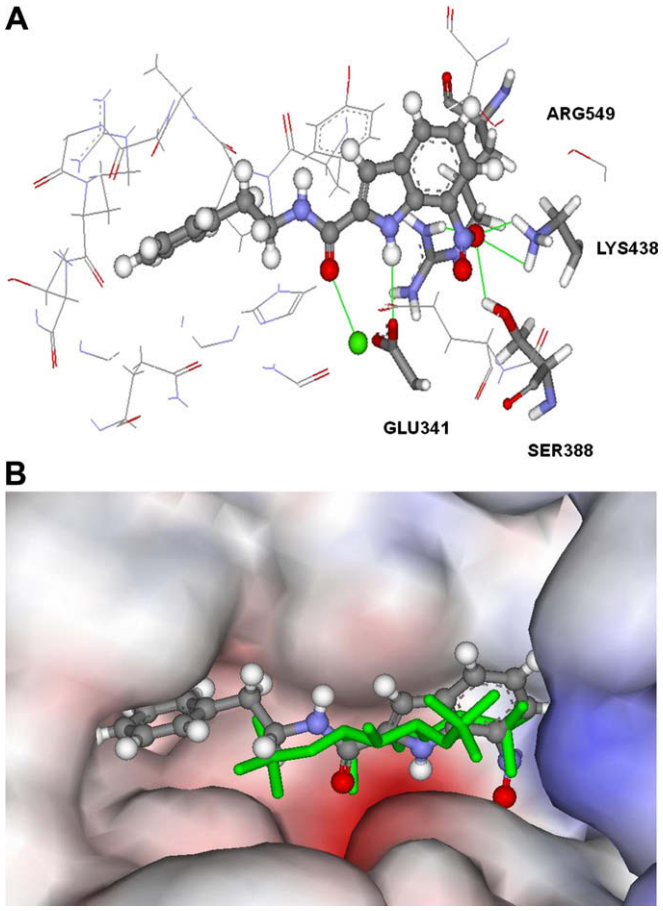


Figure 3. The docked configuration of CTT129957 in the binding site of PLC-δ. (A) Hydrogen bonds are depicted in green between the amino acids Glu341, Lys438 and Arg549. Furthermore, the Ca^{2+} is shown to interact with the carboxyl group of the amide moiety of the ligand. (B) CTT129957 shown in the binding pocket with the protein surface rendered and overlain with the co-crystallised inositol 1,4,5-triphosphate shown in green. The phenyl group of CTT129957 is predicted to occupy a lipophilic cavity as depicted on the left side of the figure. A good overlap with the inositol is predicted. Red depicts a positive partial charge on the surface, blue negative and grey neutral/lipophilic.

4. Discussion

In general, this virtual screening campaign was successful since active chemical entities were found in the context of a drug discovery programme. Virtual screening projects usually find 5–35% of the virtual hits to be active.[55] In this study ~30% of the compounds were reasonably active, which is at the higher end of reported results from similar work. A considerable strength of this work is that not only biochemical data were produced but also a cell-based assay was used, which lends more credibility to the results. However, there are still many aspects of the approach used here that can be improved. First, the sequence of the elimination of candidates can be made more efficient by installing the Log*P*/*S* filters as the first step (see Fig. 1). These calculations require very little computational power and performing them early in the screening cascade would accelerate the whole process. The sequence shown in Figure 1 was used for technical reasons, that is, how easily various software packages could be integrated into the cascade. Moving the Log*P*/*S* step forward would accelerate the process but in principle the same results should be obtained. Also, automating and moving the elimination of promiscuous moieties forward would make the process more streamlined.

Another way of improving the screening cascade would be to use more fitness functions in the consensus scoring step. Considerable evidence exists to support the application of this step,[44,45,47,56] and Wang et al.[46] found that three to four fitness functions gave the optimum results. Thus adding one or two more fitness functions to the elimination process can theoretically improve it.

Commercial compound collections tend to be lipophilic[57] and it is important to eliminate these chemical species from the molecular candidates. We had some problems with compound insolubility but it was minimised by installing the Log*P*/*S* filters. It was possible to set more stringent limits, but this would reduce the probability of finding tractable hit compounds.

Many of the experimentally confirmed hits contain chemical moieties that are considered undesirable due to their ability to perturb assays or are toxic. As shown in Figure 1 and described in this study, it was attempted to eliminate these molecular species in the screening cascade but it was not completely successful. In particular, Michael acceptors[49] incorporated into ring systems slipped through. However, according to the experimental data many of the inactive compounds were also Michael acceptors (data not shown). Also, cyclohexanone is a moiety linked to cytotoxicity and therefore could give false positives in cell-based assays.[49,58,59] Interestingly, the subgroup B of the thieno[2,3-*b*]pyridine-3-amine family (Scheme 4) contains a cyclohexanone moiety and none of them was active in the Ca^{2+} release cell-based assay. Furthermore, compounds containing moieties like nitro-aryls, which are linked to mutagenic and carcinogenic properties (see e.g., Refs. 60–62 and references therein) can often be substituted with trifluoromethyl- or methylsulphone—groups due to their similar electron negativity properties.[63] Interestingly, our modelling studies revealed that when the nitro- group was substituted with a chloro-group, good binding was achieved for the CCT129957 derivative. Therefore, it can be concluded that elimination of undesirable moieties should be used with caution since otherwise valuable regions of chemical space are unduly ignored. Finally, in a recent study it was found that 26% of marketed drug compounds contain a moiety reported to perturb biochemical and/or biological assays.[64]

5. Conclusions

Even though many improvements can be made to the process used in this work, it produced a wealth of compounds active against PLC-γ, some being chemically interesting for drug develop-

ment. Compounds with low micro-molar potency were confirmed using a biochemical assay and verified in a cell-based functional assay. Roughly 30% of the compounds tested experimentally showed relatively strong affinity to PLC-γ, which is a very good return for a virtual high throughput screen.

References and notes

1. Rhee, S. G. *Annu. Rev. Biochem.* **2001**, *70*, 281.
2. Wells, A.; Grandis, J. R. *Clin. Exp. Met.* **2003**, *20*, 285.
3. Eccles, S. A. *Curr. Opin. Genet. Dev.* **2005**, *15*, 77.
4. Eccles, S. A.; Box, C.; Court, W. *Biotechnol. Annu. Rev.* **2005**, *11*, 391.
5. Rhee, S. G.; Choi, K. D. *J. Biol. Chem.* **1992**, *267*, 12393.
6. Saunders, C. M.; Larman, M. G.; Parrington, J.; Cox, L. J.; Royse, J.; Blayney, L. M.; Swann, K.; Lai, F. A. *Development* **2002**, *129*, 3533.
7. Sekiya, F.; Poulin, B.; Kim, Y. J.; Rhee, S. G. *J. Biol. Chem.* **2004**, *279*, 32181–32190.
8. Wilde, J. I.; Watson, S. P. *Cell. Signalling* **2001**, *13*, 691.
9. Carpenter, G.; Ji, Q. S. *Exp. Cell Res.* **1999**, *253*, 15.
10. Wang, D. M.; Feng, J.; Wen, R. R.; Marine, J. C.; Sangster, M. Y.; Parganas, E.; Hoffmeyer, A.; Jackson, C. W.; Cleveland, J. L.; Murray, P. J.; Ihle, J. N. *Immunity* **2000**, *13*, 25.
11. Jones, N. P.; Peak, J.; Brader, S.; Eccles, S. A.; Katan, M. *J. Cell. Sci.* **2005**, *118*, 2695.
12. Wells, N. *Adv. Cancer Res.* **2000**, *78*, 31.
13. Xie, W.; Peng, H.; Zalkow, L. H.; Li, Y. H.; Zhu, C.; Powis, G.; Kunkel, M. *Bioorg. Med. Chem.* **2000**, *8*, 699.
14. Xie, W.; Peng, H. R.; Kim, D. I.; Kunkel, M.; Powis, G.; Zalkow, L. H. *Bioorg. Med. Chem.* **2001**, *9*, 1073.
15. Homma, M. K.; Yamasaki, M.; Ohmi, S.; Homma, Y. *J. Biochem.* **1997**, *122*, 738.
16. Lee, J. S.; Yang, M. Y.; Yeo, H.; Kim, J. B.; Lee, H. S.; Ahn, J. S. *Bioorg. Med. Chem. Lett.* **1999**, *9*, 1429.
17. Aoki, M.; Itezonzo, Y.; Shirai, H.; Nakayama, N.; Sakai, A.; Tanaka, Y.; Yamaguchi, A.; Shimma, N.; Yokose, K.; Seto, H. *Tetrahedron Lett.* **1991**, *32*, 4737.
18. Doman, T. N.; McGroven, S. L.; Witherbee, B. J.; Kasten, T. P.; Kurumbail, R.; Stallings, W. C.; Connolly, D. T.; Shoichet, B. K. *J. Med. Chem.* **2002**, *45*, 2213.
19. Singh, J.; Chuaqui, C. E.; Boriack-Sjodin, P. A.; Lee, W.-C.; Pontz, T.; Corbley, M. J.; Cheung, H. K.; Arduini, R. M.; Mead, J. N.; Newman, M. N.; Papadatos, J. L.; Bowes, S.; Josiah, S.; Ling, L. E. *Bioorg. Med. Chem. Lett.* **2003**, *13*, 4355.
20. Vangrevelinghe, E.; Zimmermann, K.; Schoepfer, J.; Portmann, R.; Fabbro, D.; Furet, P. *J. Med. Chem.* **2003**, *46*, 2656.
21. Lyne, P. D.; Kenny, P. W.; Cosgrove, D. A.; Deng, C.; Zabuldoff, S.; Wendolosky, J. J.; Ashwell, S. *J. Med. Chem.* **2004**, *47*, 1962.
22. Huang, N.; Nagarsekar, A.; Xia, G.; Hayashi, J.; MacKerell, A. D. J. *J. Med. Chem.* **2004**, *47*, 3502.
23. Forino, M.; Jung, D.; Easton, J. B.; Houghton, P. J.; Pellecchia, M. *J. Med. Chem.* **2005**, *48*, 2278.
24. Toledo-Sherman, L.; Deretey, E.; Slon-Usakiewicz, J. J.; Ng, W.; Dai, J.; Foster, J. E.; Redden, P. R.; Uger, M. D.; Liao, L. C.; Pasternak, A.; Reid, N. *J. Med. Chem.* **2005**, *48*, 3221.
25. Besong, G. E.; Bostock, J. M.; Stubbings, W.; Chopra, I.; Roper, D. I.; Lloyd, A. J.; Fishwick, C. W. G.; Johnson, A. P. *Angew. Chem., Int. Ed.* **2005**, *44*, 6403.
26. Polgár, T.; Baki, A.; Szendrei, G. I.; Keseru, G. M. *J. Med. Chem.* **2005**, *48*, 7946.
27. Essen, L. O.; Perisic, O.; Katan, M.; Wu, Y.; Roberts, M. F.; Williams, R. L. *Biochemistry* **1997**, *36*, 1704.
28. Ellis, M. V.; James, S. R.; Perisic, O.; Downes, C. P.; Williams, R. L.; Katan, M. *J. Biol. Chem.* **1998**, *273*, 11650.
29. Irwin, J. J.; Shoichet, B. K. *J. Chem. Inf. Model.* **2005**, *45*, 177.
30. Jones, G.; Willet, P.; Glen, R. C.; Leach, A. R.; Taylor, R. *J. Mol. Biol.* **1997**, *267*, 727.
31. Nissink, J. W. M.; Murray, C.; Hartshorn, M.; Verdonk, M. L.; Cole, J. C.; Taylor, R. *Proteins* **2002**, *49*, 457.
32. Kontoyianni, M.; McClellan, L. M.; Sokol, G. S. *J. Med. Chem.* **2004**, *47*, 558.
33. Verdonk, M. L.; Chessari, G.; Cole, J. C.; Hartshorn, M.; Murray, C. W.; Nissink, J. W. M.; Taylor, R. D.; Taylor, R. *J. Med. Chem.* **2005**, *48*, 6504.
34. Berman, H. M.; Westbrook, J.; Feng, Z.; Gilliland, G.; Bhat, T. N.; Weissig, H.; Shindyalov, I. N.; Bourne, P. E. *Nucleic Acids Res.* **2000**, *28*, 235.
35. CACHE 6.1, Fijitsu Limited 2000–2003, Oxford Molecular, 1989–2000.
36. Eldridge, M. D.; Murray, C.; Auton, T. R.; Paolini, G. V.; Mee, P. M. *J. Comput. Aided Mol. Des.* **1997**, *11*, 425.
37. Thomas, M. P.; McInnes, C.; Fischer, P. M. *J. Med. Chem.* **2006**, *49*, 92.
38. QIKPROP 2.1.103, Schrödinger Inc., New York, 2003.
39. Duffy, E. M.; Jorgensen, W. L. *J. Am. Chem. Soc.* **2000**, *122*, 2878.
40. Jorgensen, W. L.; Duffy, E. M. *Bioorg. Med. Chem. Lett.* **2000**, *10*, 1155.
41. Jorgensen, W. L.; Duffy, E. M. *Adv. Drug Delivery Rev.* **2002**, *54*, 355.
42. Ioakimidis, L.; Thoukydidis, L.; Naeem, S.; Mirza, A.; Reynisson, J. *QSAR Comb. Sci.* **2008**, *27*, 445.
43. Katan, M.; Rodriguez, R.; Matsuda, M.; Newbatt, Y.; Aherne, W. *Adv. Enzyme Regul.* **2003**, *43*, 77.
44. Charifson, P. S.; Corkery, J. J.; Murcko, M. A.; Walters, W. P. *J. Med. Chem.* **1999**, *42*, 5100.
45. Clark, R. D.; Strizhev, A.; Leonard, J. M.; Blake, J. F.; Matthew, J. B. *J. Mol. Graphics Modell.* **2002**, *20*, 281.
46. Wang, R.; Wang, S. *J. Chem. Inf. Comput. Sci.* **2001**, *41*, 1422.
47. Stahl, M.; Rarey, M. *J. Med. Chem.* **2001**, *44*, 1035.
48. Rishton, G. M. *DDT* **2003**, *8*, 86.
49. Rishton, G. M. *DDT* **1997**, *2*, 382.
50. Roche, O.; Schneider, P.; Zuegge, J.; Guba, W.; Kansy, M.; Alanine, A.; Bleicher, K.; Danel, F.; Gutknecht, E. M.; Rogers-Evans, M.; Neidhart, W.; Stalder, H.; Dillon, M.; Sjögren, E.; Fotouhi, N.; Gillespie, P.; Goodnow, R.; Harris, W.; Jones, P.; Taniguchi, M.; Tsujii, S.; von der Saal, W.; Zimmermann, G.; Schneider, G. *J. Med. Chem.* **2002**, *45*, 137.
51. McGovern, S. L.; Caselli, E.; Grigorieff, N.; Schoichet, B. K. *J. Med. Chem.* **2002**, *45*, 1712.
52. McGovern, S. L.; Helfand, B. T.; Feng, B.; Schoichet, B. K. *J. Med. Chem.* **2003**, *46*, 4265.
53. http://www.netdoctor.co.uk/.
54. http://www.patient.co.uk/.
55. Shoichet, B. K. *Nature* **2004**, *432*, 862.
56. Yang, J. M.; Chen, Y. F.; Shen, T. W.; Kristal, B. S.; Hsu, D. F. *J. Chem. Inf. Model.* **2005**, *45*, 1134.
57. Baurin, N.; Richardson, B. C.; Chen, I.; Foloppe, N.; Potter, A.; Jordan, A.; Roughley, S.; Parratt, M.; Greany, P.; Morley, D.; Hubbard, R. E. *J. Chem. Inf. Comput. Sci.* **2004**, *44*, 643.
58. Oprea, T. I.; Bologa, C.; Olah, M. In *Virtual Screening in Drug Discovery*; Alvarez, J., Shoichet, B. K., Eds.; Compound Selection for Virtual Screening; Taylor & Francis: London, 2005. Chapter 4, pp 89–106.
59. Arnold, L. A.; Estébanez-Perpiñá, E.; Togashi, M.; Jouravel, N.; Shelat, A.; McReynolds, A. C.; Mar, E.; Nguyen, P.; Baxter, J. D.; Fletterick, R. J.; Webb, P.; Guy, R. K. *J. Biol. Chem.* **2005**, *280*, 43048–43055.
60. Iwata, N.; Fukuhara, K.; Suzuki, K.; Miyata, N.; Takahashi, A. *Chem. Biol. Interact.* **1992**, *85*, 187.
61. Josephy, P. D.; Summerscales, J.; DeBruin, L. S.; Schlaeger, C.; Ho, J. *Biol. Chem.* **2002**, *383*, 977.
62. Arlt, V. M.; Schmeiser, H. H.; Osborne, M. R.; Kawanishi, M.; Kanno, T.; Yagi, T.; Phillips, D. H.; Takamura-Enya, T. *Int. J. Cancer* **2006**, *118*, 2139.
63. Reynisson, J.; McDonald, E. *J. Comput. Aided Mol. Des.* **2004**, *18*, 421.
64. Axerio-Cilies, P.; Castañeda, I. P.; Mirza, A.; Reynisson, J. *Eur. J. Med. Chem.* **2009**, *44*, 1128.