DOI: 10.1002/cmdc.200800301

Virtual Screening and Biological Characterization of Novel Histone Arginine Methyltransferase PRMT1 Inhibitors

Ralf Heinke,[a] Astrid Spannhoff,[b, c] Rene Meier,[a] Patrick Trojer,[d] Ingo Bauer,[e] Manfred Jung,[b] and Wolfgang Sippl*[a]

Lysine and arginine methyltransferases participate in the post-translational modification of histones and regulate key cellular functions. Protein arginine methyltransferase 1 (PRMT1) has been identified as an essential component of mixed lineage leukemia (MLL) oncogenic complexes, revealing its potential as a novel therapeutic target in human cancer. The first potent arginine methyltransferase inhibitors were recently discovered by random- and target-based screening approaches. Herein we report virtual and biological screening for novel inhibitors of PRMT1. Structure-based virtual screening (VS) of the Chembridge database composed of 328 000 molecules was performed with a combination of ligand- and target-based in silico approaches. Nine inhibitors were identified from the top-scored docking solutions; these were experimentally tested using human PRMT1 and an antibody-based assay with a time-resolved fluorescence readout. Among several aromatic amines, an aliphatic amine and an amide were also found to be active in the micromolar range.

Introduction

Histones are subject to a host of various posttranslational modifications that set a pattern of molecular signals for the access of transcription factors toward DNA, the so-called histone code.[1–3] The histone N-termini, which are rich in arginine and lysine residues, protrude from the histone octamers and undergo many types of posttranslational modifications such as acetylation, methylation, phosphorylation, ubiquitinylation, and sumoylation.[3] The extent of chromatin condensation is regulated in part through these modifications, which have an impact on gene transcription and the maintenance of altered transcription after cell division.[2] Histone acetylation and DNA methylation are epigenetic modifications that have been linked closely to the pathology of human cancers, and inhibitors of both enzyme classes for clinical use are close at hand. In contrast, other fields of epigenetics still lack similarly thorough knowledge. This is especially true for the group of histone methyltransferases and their inhibitors. Because connections between histone methylation patterns and cancer progression have been recognized, histone methyltransferases represent promising targets for future cancer treatment.

Histone methylation has been reported to play an important role in the regulation of gene expression patterns. In contrast to histone acetylation, histone methylation does not alter the charge of the histone tail, but influences mainly the hydrophobicity of histones and their affinity for certain proteins such as transcription factors. Histone lysine and arginine methyltransferases are known not only to methylate histone proteins, they also catalyze the methylation of many proteins that are involved in signal transduction and cell proliferation.[4] For example, arginine methylation of the histone acetyltransferase CBP by the transferase CARM1 contributes to its co-activating effects.[5] Arginine methylation also affects various proteins involved in RNA maturation[6,7] and DNA repair.[8] The latter may pose questions on the safety of arginine methyltransferase inhibitors, as it has been shown that the function of the DNA damage control protein MRE 11 is dependent on arginine methylation.[9]

Protein arginine methylation is carried out by two classes of enzymes: type I methyltransferases and type II methyltransferases (PRMTs). Type II enzymes catalyze the formation of symmetric ω-N_G,N'_G-dimethylarginine tails, whereas the formation of asymmetric ω-N_G,N_G-dimethylarginine tails is maintained by type I enzymes. PRMTs 1–4, 6, and 8 belong to the type I class, and PRMTs 5, 7, and 9 are type II enzymes. PRMT1 is the predominant type I PRMT in mammalian cells, accounting for 85% of cellular PRMT activity. PRMT1 is a necessary component for oncogenic transformation induced by a mixed lineage leuke-

[a] *R. Heinke, R. Meier, Prof. Dr. W. Sippl*
Universität Halle-Wittenberg
Department of Pharmaceutical Chemistry
Wolfgang-Langenbeckstraße 4, 06120 Halle/Saale (Germany)
Fax: (+49) 345-5527355
E-mail: Wolfgang.sippl@pharmazie.uni-halle.de

[b] *Dr. A. Spannhoff, Prof. Dr. M. Jung*
Institute of Pharmaceutical Sciences
Albert-Ludwigs-Universität Freiburg
Albertstraße 25, 79104 Freiburg (Germany)

[c] *Dr. A. Spannhoff*
current address: The Virginia Harris Cockrell Cancer Research Center at The University of Texas M. D. Anderson Cancer Center
1808 Park Road 1C, Smithville, Texas 78957 (USA)

[d] *Dr. P. Trojer*
Department of Biochemistry, Robert Wood Johnson Medical School
University of Medicine and Dentistry New Jersey
Piscataway, NJ 08854 (USA)

[e] *Dr. I. Bauer*
Division of Molecular Biology, Biocenter
Innsbruck Medical University (Austria)

mia (MLL) complex.[10] The direct fusion of MLL with PRMT1 enhances the self-renewal of primary hematopoietic cells. Conversely, specific knockdown of PRMT1 expression suppresses MLL-mediated transformation. The study uncovered an essential function of PRMTs in oncogenesis and revealed their potential as novel therapeutic targets in human cancer. Therefore, therapy with small-molecule PRMT1 inhibitors represents a promising approach to new cancer treatments.

Detailed structure–activity relationships have been set up for HDAC inhibitors,[3,11,12] whereas much less is known in the area of histone methyltransferase inhibitors. Besides cofactor analogues such as sinefungin,[12] there are only a few inhibitors of lysine[13] and arginine[14–17] methyltransferases available. High-throughput screening led to the discovery of new active compounds for both target classes. The PRMT inhibitor AMI-1[7] (Figure 1) has shown the potential to block transcriptional activation of androgen and estrogen receptors by agonists.

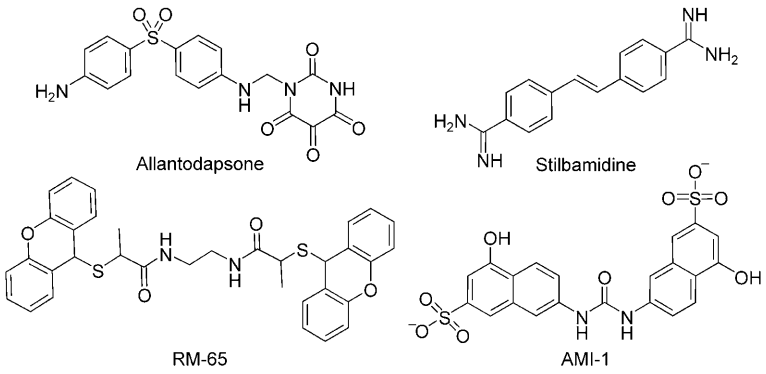

Figure 1. Molecular structures of known PRMT1 inhibitors.

In the last decade, virtual screening (VS) has emerged as a powerful tool for the identification of novel and diverse lead structures.[18] VS can be regarded as a complementary approach to high-throughput screening (HTS); it allows the discovery of novel bioactive molecules from very large compound collections by using either the information about a set of active ligands or the protein binding pocket. By applying a combination of pharmacophore- and target-based approaches, we have successfully employed the technique in the search of novel leads for several targets.[14,15,19,20]

We recently identified new PRMT1 inhibitors by VS of the National Cancer Institute (NCI) diversity set for substrate-competitive compounds. The PRMT1 inhibitors, allantodapsone and stilbamidine, were indeed characterized in competition experiments as being substrate competitive.[14] In this study, the search for novel PRMT1 inhibitors was pursued with the Chembridge compound collection containing approximately 328000 molecules. A combination of pharmacophore search and target-based VS was conducted based on a model of human PRMT1 complexed with *S*-adenosylhomocysteine (SAH). Compounds that were top-ranked by the docking program and that passed the pharmacophore model, which was generated on the basis of previously identified hits, were tested in a biochemical PRMT1 assay. Potential inhibitors were identified that are able to block the substrate pocket of PRMT1. These compounds are promising anticancer drug candidates, particularly for androgen- and estrogen-dependent tumors.

Results and Discussion

X-ray crystal structure information for human PRMT1 (hPRMT1) is not yet available, but structural data are available for the homologous rat proteins PRMT1[21] (sequence identity 95.1%) and PRMT3[22] (sequence identity 47.2%). However, the rat PRMT1 crystal structure is not suitable as a target structure for VS, as it was obtained at nonphysiological pH (pH 4.7; maximum enzymatic activity at pH 8.5). In addition, not all residues of the substrate binding site have been resolved in the PRMT1 crystal structure. The N-terminal region (residues 1–40), including a helical section (helix αX) proximal to the binding pocket, is missing in the PRMT1 structure, but was resolved for the PRMT3 protein (Figure 2). Similar data were recently reported for the homologous PRMT4 (also known as CARM1).[23] In the absence of the bound cofactor, the PRMT4 crystal structure revealed that amino acid residues 144–154, which compose the N-terminal helix αX, are not observed in the electron density map and are therefore probably disordered. In contrast, helix αX was observed in the crystal structure of PRMT4 (as well as in PRMT3) co-crystallized with the cofactor analogue SAH. These results indicate that the helix is flexible, allowing the cofactor to enter the binding site.[24]

Recently, we reported the generation of a homology model for the active form of hPRMT1 on the basis of the rat PRMT3 X-ray structure.[14] The high sequence identity between hPRMT1 and PRMT3 (47% overall, 95% at the binding pocket) resulted in high structural similarity between the hPRMT1 model and the PRMT3 crystal structure (RMSD: 1.8 Å, Cα atoms). The positions of the cofactor and substrate binding

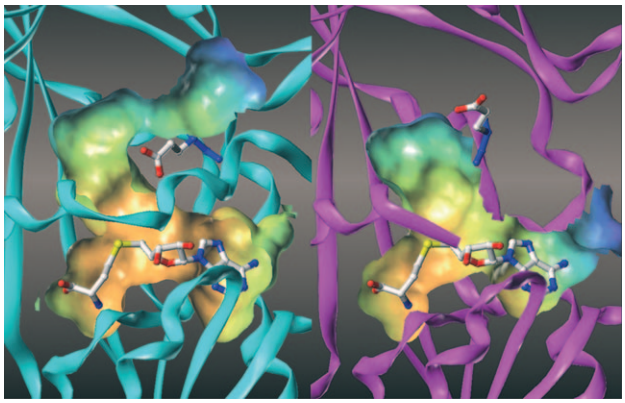

Figure 2. Comparison of the binding pockets observed in the rat PRMT3 (left) and rat PRMT1 (right) crystal structures. The co-crystallized SAH and Glu 161, which adopts a different conformation in the two proteins, are displayed. The Conolly molecular surface is colored according the cavity depth.

pocket are particularly well conserved and show the same architecture as observed in the PRMT3 crystal structure. The substrate binding surface of hPRMT1 is expected to be acidic, because most substrates for PRMT1 contain one or multiple arginine residues. The substrate binding pocket is cone shaped with an opening of ~8×6 Å and a depth of ~8 Å. The pocket accommodates the arginine side chain of the substrate so that the terminal amino group reaches the deeply buried *S*-adenosylmethionine (SAM) cofactor. The residues lining the substrate pocket are negatively charged at the bottom (Glu 152 and Glu 161) and hydrophobic in the middle (Tyr 147, Tyr 156, Phe 160). This fits well with the asymmetric polarity of an arginine residue with three hydrophobic methylene groups and the basic guanidine group.

The hPRMT1 model was successfully used to virtually screen the NCI diversity set containing 1990 structurally diverse molecules. By this approach two potent and cell-permeable PRMT1 inhibitors (stilbamidine and allantodapsone, Figure 1) were identified.[14] The interaction of allantodapsone with the hPRMT1 substrate binding site is shown in Figure 3. A

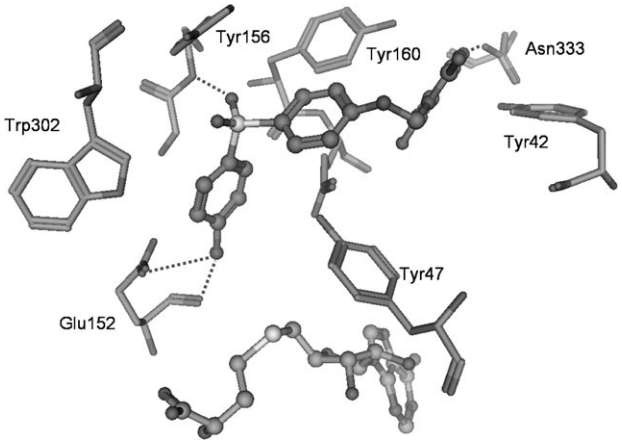

Figure 3. Predicted binding mode of allantodapsone (upper structure in ball-and-stick representation) at hPRMT1. Hydrogen bonds are indicated by dashed lines.

common feature of the active inhibitors is the interaction (hydrogen bond) between a basic or polar group and the acidic residue Glu 152 that mimics the interaction of the Arg guanidine group of the substrate peptide. The most potent compound in vitro, allantodapsone, shows van der Waals interactions with several aromatic residues of the binding pocket (Tyr 47, Tyr 156, Trp 302), and hydrogen bonds with the backbone of Tyr 156 and the amide side chain of Asn 333 are formed.

Based on the successful VS of the diversity subset of the NCI we expanded our search for novel PRMT inhibitors to a larger compound collection. Therefore, we considered the Chembridge compound collection (Chembridge Corporation, San Diego, CA, USA), containing 328 000 molecules for our present work.

Pharmacophore search

As we were interested in novel lead-like PRMT inhibitors, we filtered the Chembridge database for molecules with a molecular weight <400, a topological polar surface area (TPSA) <150 Å^2, and at least one nitrogen atom. This resulted in 189 000 molecules, which were converted into a multi-conformer library consisting of 19 million conformations. This library was virtually screened using a structure-based pharmacophore model. The pharmacophore model was generated on the basis of the PRMT1–allantodapsone interaction model. We used the LigandScout software (Inteligand, Vienna, Austria) to derive the pharmacophore from the PRMT1–inhibitor complex. LigandScout is a program for developing structure-based pharmacophore models based on a given protein–ligand complex. The derived pharmacophore, shown in Figure 4, contains one hydrogen bond donor (interaction with Glu 152 and the backbone CO group of Glu 152) and one hydrogen bond acceptor function (interaction with Tyr 156), as well as two hydrophobic/aromatic features. In addition, we used one included volume and 23 excluded volumes on the basis of the allantodapsone structure and the surrounding amino acid residues. LigandScout generates the excluded volumes on the basis of the Cα atoms. The pharmacophore was stored in LigandScout and exported to the MOE software.

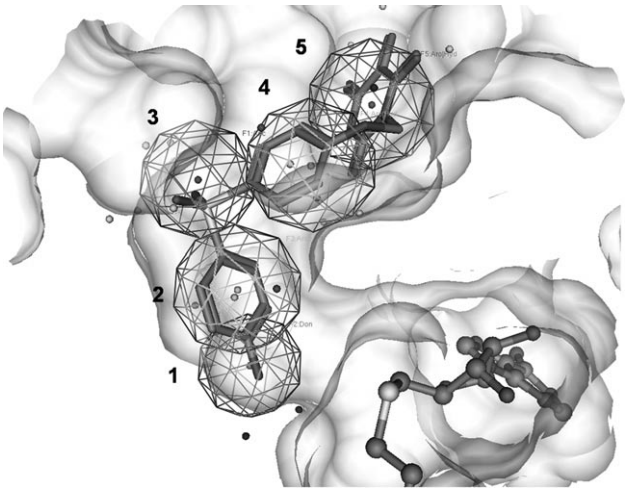

Figure 4. Structure-based PRMT1 pharmacophore generated on the basis of allantodapsone and the LigandScout program. The five pharmacophoric features are shown as spheres along the allantodapsone structure (1: hydrogen bond donor, 3: hydrogen bond acceptor, 2 and 4: aromatic features, 5: included volume feature). The Conolly molecular surface of the binding pocket is displayed as well as the cofactor analogue SAH (below).

In the subsequent pharmacophore search, the screened compounds were required to match at least four out of the five essential pharmacophoric features. In general, the number of identified hits that match the pharmacophore model depends on the size of the pharmacophoric features. We empirically modified the radius of the spheres to a value of 1.3 Å for the hydrogen bond features and to a value of 1.7 Å for the aromatic and hydrophobic features. The screening resulted in

6232 compounds (3.3%) that passed the LigandScout pharmacophore. Because it is likely that not all of the 6232 pharmacophore hits are PRMT1 inhibitors, we further filtered the compound selection by carrying out a docking study.

Ligand docking

The 6232 compounds were subsequently docked into the substrate binding pocket of PRMT1. Because we were interested in substrate-competitive inhibitors, the cofactor binding pocket was blocked during docking by the analogue SAH; thus the cofactor was regarded as a part of the protein during the docking process. Due to the fact that docking scores are not good discriminators between true actives and false positives we focused not only on the derived GoldScores but used additional filters for selecting potential candidates for biological testing. Using the SVL script language within MOE we filtered the docking output for appropriate poses that showed a similar binding mode as the known PRMT1 inhibitors allantodapsone and stilbamidine. We calculated the distances to amino acids of the binding pocket which were already identified as essential binding partners (i.e. Glu 152 and Tyr 156). Among the top-ranked 100 docking poses that fulfilled the distance constraints (< 3.5 Å) we manually selected nine compounds from clusters of similar molecules (Figure 5). The nine compounds were tested afterward in an in vitro assay for their ability to inhibit hPRMT1.

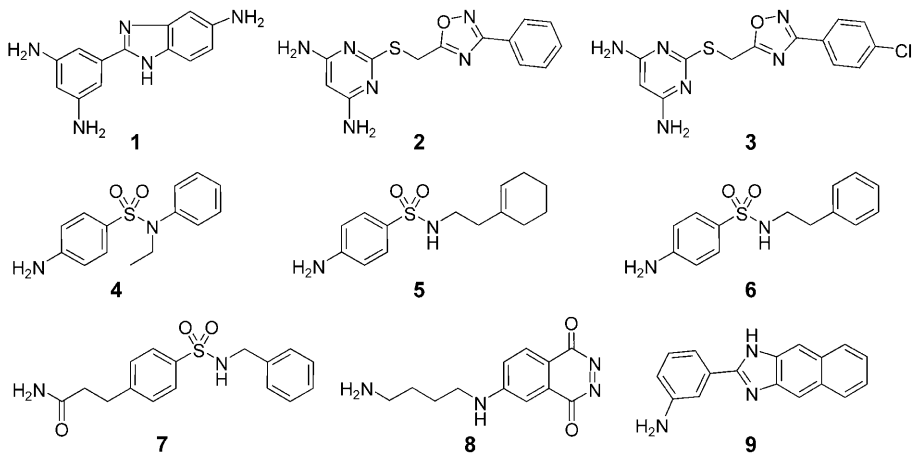

Figure 5. Molecular structures of PRMT1 inhibitors identified using target-based VS.

Enzyme inhibition studies

For the measurement of in vitro methyltransferase activity we used our recently developed assay employing a biotinylated oligopeptide that contains the amino acids 1–21 of human histone H4 as the methylation substrate for hPRMT1. The peptide was immobilized on streptavidin-coated microtiter plates. Methylation was recognized by using an antibody against dimethyl-Arg 3 on H4. Treatment with a europium-labeled secondary antibody and final measurement of time-resolved fluorescence (TRF) was used for the detection of the methylation level.[14] Table 1 shows the IC_{50} values of the identified compounds for the human enzyme. AMI-1 was used as a reference (see Table 1) and similar inhibition values as reported previously[23] were observed. Among the nine identified inhibitors, compounds **5**, **6**, **7**, and **9** are the most active (IC_{50} range: 12.75–16.93 μM). The diaminobenzene and diaminopyrimidine derivatives **1**, **2**, and **3** are less active (IC_{50} range: 21.88–36.74 μM). The smaller sulfonamide **4** and the aliphatic amine **8** are also less active inhibitors.

Table 1. Inhibition of hPRMT1.

Compound	$IC_{50} \pm SE$ [μM]
6000016 (**1**)	33.81 ± 16.56
7112201 (**2**)	36.74 ± 10.89
7155176 (**3**)	21.88 ± 2.35
7736382 (**4**)	29.91 ± 1.70
6689772 (**5**)	16.93 ± 3.23
7280948 (**6**)	12.75 ± 2.88
7789734 (**7**)	15.32 ± 4.83
5784982 (**8**)	33.63 ± 8.26
5756663 (**9**)	14.34 ± 1.27
AMI-1	1.20 ± 0.50[a]
Allantodapsone	1.70 ± 0.3[a]

[a] Data from Ref. [14].

Analysis of PRMT1–inhibitor complexes

Of the nine compounds assayed, the aniline derivative **6** is the most potent hPRMT1 inhibitor (Table 1). The docking results show that the *p*-aminophenylsulfonamide fragment of **4**, **5**, and **6** adopts a similar orientation in the binding pocket to that of the corresponding element in the allantodapsone structure. The dapsone fragment of allantodapsone as well as the *p*-aminophenylsulfonamide group of the novel inhibitors show two hydrogen bonds to the side chain and the backbone of Glu 152 (Figure 6). In the case of the more potent inhibitors **5** (Figure 6A) and **6** (Figure 6B) the alkylbenzene and alkylcyclohexene rings interact with aromatic residues at the binding pocket (Tyr 147 and Tyr 160), whereas the less potent and smaller analogue **4** is not able to interact with these aromatic residues. Interestingly, the sulfonamide substructure was also identified as a binding motif in the structurally diverse amide **7** (Figure 6C). However, the sulfonamide group of **7** is not involved in a hydrogen bond as observed for

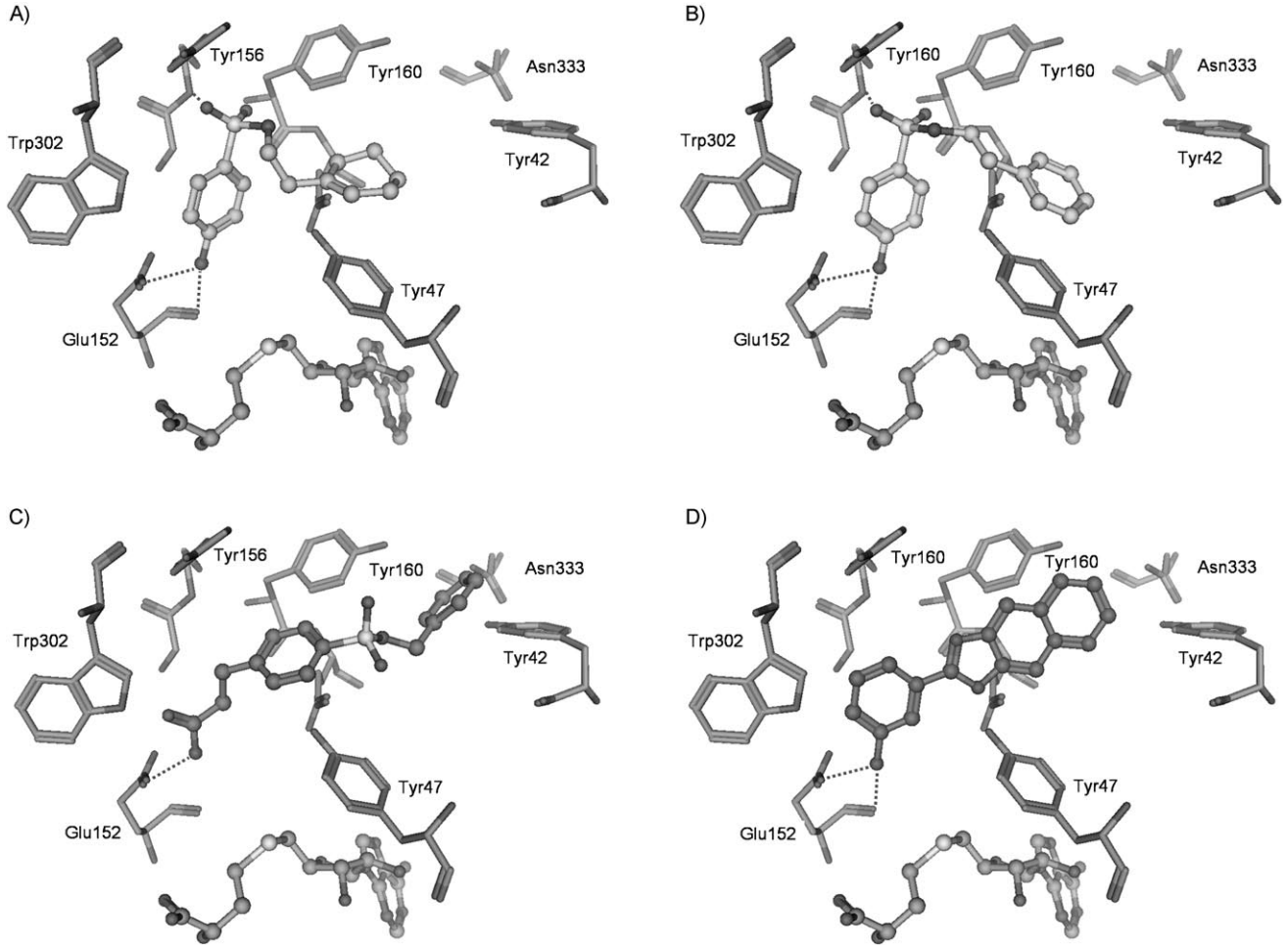

Figure 6. Docking results for the four most active inhibitors: A) **5**, B) **6**, C) **7**, and D) **9**, showing the interaction at the hPRMT1 substrate binding site. The cofactor analogue SAH is displayed in each case. Hydrogen bonds are indicated by dashed lines.

the inhibitors **5** and **6**. In the case of the aromatic amine derivatives (**1**, **2**, **3**, and **9**) compound **9** (Figure 6D) with a single amino group in the meta position of the aromatic ring shows the most potent inhibitory activity. The aromatic nitrogen atom donates hydrogen bonds to the side chain and backbone CO group of Glu 152 as observed for allantodapsone. Compound **9** also shows van der Waals interactions with the two aromatic residues Tyr 147 and Tyr 160 as detected for the other potent inhibitors. Therefore, along with the hydrogen bond to Glu 152, van der Waals interactions with the two aromatic residues also seems to be a prerequisite for high inhibitory activity. Interestingly, the guanidine group of the Arg substrate can be mimicked not only by basic amino groups but also by an amide, representing an appropriate binding motif (Figure 6C). Based on the docking poses obtained, a qualitative interpretation of the structure–activity relationships was possible, whereas no significant correlation was obtained between docking scores and activity data.

To analyze the structural similarity between the retrieved Chembridge compounds and the known inhibitor allantodapsone used for pharmacophore generation, we calculated similarity indices on the basis of MACCS keys and graph-3-point pharmacophore fingerprints. The Tanimoto similarity coefficients are listed in Table 2 for the nine novel inhibitors. Using the MACCS keys, the Tanimoto coefficients are all <0.60 (graph-3-point pharmacophore <0.4), indicating the low structural similarity between the nine novel inhibitors and allanto-

Table 2. Structural similarity between the novel inhibitors and the reference compound allantodapsone.

Compound	MACCS Keys[a]	GpiDAPH3[a]
6000016 (**1**)	0.28	0.33
7112201 (**2**)	0.47	0.32
7155176 (**3**)	0.45	0.30
7736382 (**4**)	0.54	0.36
6689772 (**5**)	0.48	0.27
7280948 (**6**)	0.54	0.27
7789734 (**7**)	0.57	0.37
5784982 (**8**)	0.43	0.29
5756663 (**9**)	0.25	0.27

[a] Tanimoto coefficients between the seven hit structures and the most similar analogues were calculated on the basis of MACCS keys and graph-3-point pharmacophore (GpiDAPH3) fingerprints.

dapsone. Thus, although we applied a clearly defined pharmacophore for the VS, the identified hits represent different chemotypes.

In order to estimate the enrichment rate of the VS we carried out two experiments. First, we randomly selected a subset of 500 compounds from the NCI diversity set and used it as a set of decoys for the VS. We docked the randomly selected compounds and the nine identified active inhibitors into the hPRMT1 model using Gold and the same settings as described before. The enrichment curve that we obtained from analyzing the GoldScores is shown in Figure 7. All nine inhibitors were identified among the top 30% of the dataset, resulting in an enrichment rate of 16.7. Ligand enrichment among top-ranking hits is a key metric of VS. There is a big debate in the literature on how to avoid bias, so that enrichment is not attributable to simple differences of gross features.[25] Because the main focus of the work described herein was the discovery of novel PRMT1 inhibitors and not an analysis of various VS settings, we did not compare different decoy sets or docking programs. To assess an exact hit rate and enrichment factor, biological testing of all compounds would be necessary. However, as we only had a limited amount of hPRMT1 protein available it was not possible to screen a large number of compounds. Nonetheless, in order to determine whether the results we obtained would be different from those attained from random sampling, five compounds (Figure 8) were chosen at random, and their ability to inhibit hPRMT1 was also tested. None of the randomly chosen compounds from the NCI diversity set showed any inhibitory effect at 40 μM, whereas all nine compounds chosen by VS did exhibit inhibition <35 μM, suggesting that a significant enrichment had occurred. In our analysis of the obtained docking ranks (Table 3) and the interaction of the five inactive NCI compounds (not shown) we observed that none of the compounds is among the top 20% of the ranking list and that most of the molecules did not show an interaction with the active site Glu 152.

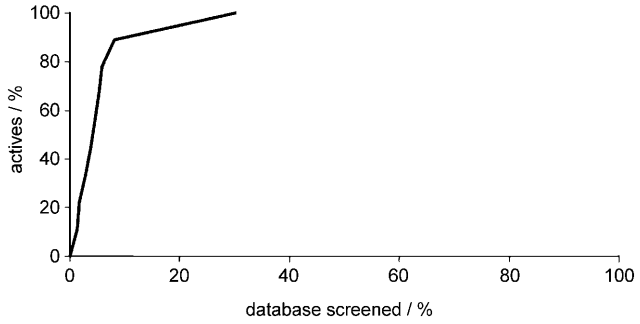

Figure 7. Enrichment curve obtained for the PRMT1 VS of the subset of the NCI diversity set and the nine active inhibitors.

Table 3. Docking rank of the five inactive compounds of the NCI diversity set (1990 compounds) observed in the VS.

Compound	VS Rank
NSC 25141	104
NSC 143099	224
NSC 97845	231
NSC 134754	248
NSC 118176	473

Discussion

In an ideal situation, virtual screening is carried out with a high-resolution structure of a ligand-bound protein. However, the crystal structure of a protein of interest is often not available, has low resolution, or has not been crystallized in the ligand-bound form. In such cases structural protein models can be generated to get a detailed representation of the binding site. Depending on the degree of sequence similarity to the known structure, a structural model can contain errors in both side chain and backbone conformation. In the case of a rigid protein docking approach, structural errors significantly influence the accuracy of VS.[26] Despite these limitations, most examples of successful VS have been accomplished without consideration of protein flexibility. This may be largely due to the fortunate result of relatively minor induced-fit adaptations of proteins upon ligand binding.[27]

It is surprising that homology models often provide significant enrichment of actives against a background of decoys. A variety of successful VS studies based

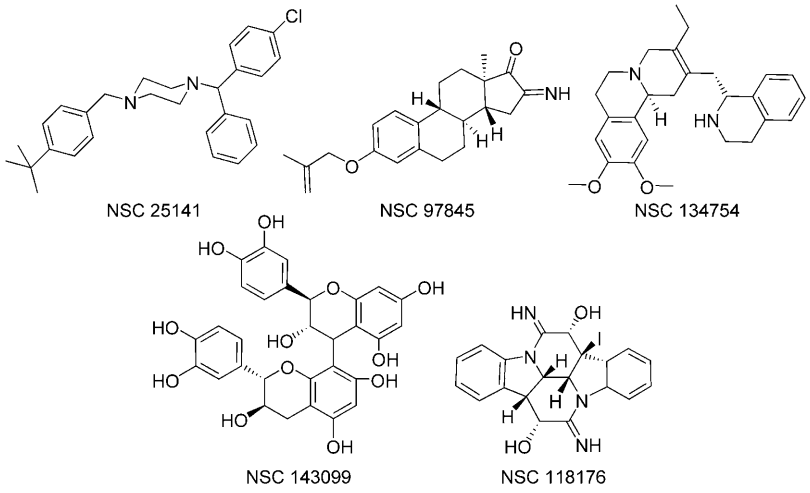

Figure 8. Molecular structures of five randomly selected compounds of the NCI diversity set that showed no inhibitory activity.

on protein models have been published in the last few years.[19,28–30] In a retrospective study, Gilson and co-workers[31] analyzed five drug targets for which the crystal structure was available. The VS study showed that some homology models (based on a template with a sequence identity $>30\%$) yielded enrichment equal to or even greater than that obtained with the crystal structure of the target. They also observed that there is no smooth correlation between enrichment rate and sequence identity of the template. It is generally difficult to predict a priori whether docking into a homology model yields reliable or misleading geometries.

The sequence identities between hPRMT1 and the template structure PRMT3 (as well as PRMT4) are in the range of 45–50% (and $>90\%$ considering only the residues of the binding pocket). Therefore, it can be expected that the structural similarity between hPRMT1 and PRMT3 is high. The recently reported crystal structures of PRMT4 (in the apo and SAM-bound forms) show low RMSD values when superimposed onto the hPRMT1 model as well as onto the rat PRMT3 crystal structure. Superimposing PRMT4 (residues 140–480) on the other PRMT structures gives RMSD in the range of 1.4–1.8 Å between corresponding 296 Cα atoms. Due to the high similarity between the individual PRMTs and the fact that the conformation of the missing helix αX in the rat PRMT1 structure could be modeled from the crystal structures of cofactor-bound PRMT3 and PRMT4, the PRMT1 model represents a suitable target structure for virtual screening.

A recently published perspective on protein flexibility in drug design focused on the pitfalls and limitations of target-based virtual screening.[32] Using flexibility as a criterion, a classification of three types of proteins was suggested: 1) "rigid" proteins, in which ligand-induced changes are limited to relatively small side chain rearrangements, 2) flexible proteins, for which relatively large movements around "hinge points" or at active site loops occur upon ligand binding with concomitant side chain motion, and 3) intrinsically unstable proteins, where conformation is not defined until ligand binding.

The analysis of the available crystal structures of PRMTs revealed that they belong to the class of "rigid" proteins. Only helix αX shows some flexibility; in the rat PRMT1 structure (with bound SAH) and the apo form of PRMT4, helix αX could not be resolved. In the SAH-bound form of PRMT3 and PRMT4, the helix is positioned like a lid over the binding pocket, thereby stabilizing the enzyme–cofactor interaction. The high hit rate that we observed for the selected VS hits may be the result of the limited flexibility of the cofactor-bound form of PRMT1.

Conclusions

In summary, the Chembridge database consisting of some 328000 compounds was screened against a homology model of human PRMT1 to identify novel drug-like inhibitors. Nine inhibitors were identified by the hierarchical VS, and all compounds showed activity below 35 μM. Besides the successful identification of novel PRMT1 inhibitors, we analyzed the enrichment of the VS procedure by screening a subset of the NCI diversity set. The random selection of five compounds from the NCI set that were all ranked as poorly binding compounds and the subsequent proof of their inactivity underscores the value of the virtual screening procedure. The active hits identified in the present study can be optimized into potent selective PRMT1 inhibitors. This in turn will be helpful to elucidate the therapeutic potential of this new class of potential drugs, particularly in the treatment of hormone-dependent cancer. The obtained results also demonstrate the unique ability of VS to significantly increase hit rates and to complement biological screening.

Experimental Section

Computational methods

All calculations were performed on a Pentium IV 2.2 GHz based Linux cluster (20 CPUs). The Gold software package was used for docking, whereas the calculation of all molecular descriptors and the analysis of the docking results were carried out in MOE2007.09 (Chemical Computing Group).

Database filtration

The 3D structures of the Chembridge database were obtained from the ZINC homepage[33] and were imported as Mol2 files into the MOE program. The following filters were applied to decrease the number of compounds for the present study: M_r between 250 and 400 Da, TPSA <150 Å, and at least one nitrogen atom. This resulted in ~189000 compounds for which we generated a multi-conformer library containing about 19 million conformations using the Omega program (OpenEye Software).

For estimation of the enrichment of the applied VS procedure, the NCI diversity set was used to generate a dataset of decoys. From the 1990 molecules, 500 were randomly selected and transformed into 3D structures using the Omega program.

Pharmacophore-based screening

The multi-conformer database was screened with a pharmacophore model that we generated on the basis of a previously identified PRMT1 inhibitor (allantodapsone) and by using LigandScout software. LigandScout is a program for developing structure-based pharmacophore models based on protein–ligand structures. The pharmacophore (Figure 4) contains one hydrogen bond donor and one hydrogen bond acceptor function, two aromatic features, and one included volume feature. In addition, we used 23 excluded volumes on the basis of the surrounding amino acid residues. LigandScout generates the excluded volumes on the basis of the Cα atoms. More details about the software can be found elsewhere.[34] The pharmacophore was exported as PH4 file and imported into the MOE program. Applying the MOE pharmacophore search module we screened the derived multi-conformer database of the Chembridge subset. Molecules that fulfilled at least four of the five pharmacophore features were retrieved. The pharmacophore screen resulted in 6232 hits which were further considered for the molecular docking study.

Protein modeling

The crystal structures of rat PRMT1 (PDB code: 1OR8) and rat PRMT3 (PDB code: 1F3L) were acquired from the Protein Data Bank. The rat PRMT1 X-ray crystal structure is not suitable as a target structure for VS: the crystal structure of rat PRMT1 was obtained at nonphysiological pH (pH 4.7; maximum activity at pH 8.5), and an important helical section near the binding pocket was not resolved in the structure. An active and complete form of a rat PRMT is available only for subtype 3 (1F3L). A homology model for human PRMT1 was generated with the rat PRMT3 (1F3L) X-ray structure as template. The model generation was described previously.[14] The sequence identity between the individual enzymes is sufficiently high for this approach (human PRMT1 and PRMT3, 47.2% overall sequence identity; human PRMT1 and human PRMT4, 31.2% identity; human PRMT3 and human PRMT4, 20.6% sequence identity). The residues within the binding pocket are especially highly conserved among the various PRMTs.

We suggest that the PRMT1 homology model possesses high accuracy. The sequence identities between PRMT1 and PRMT3 as well as PRMT4 are in the range between 45–50% (and $>90\%$ when considering only the residues of the binding pocket). Therefore, it is very likely that the geometrical deviation between model and actual structure is <2 Å. The recently reported crystal structures of PRMT4 (in the apo and cofactor-bound forms) show low RMSD values when superimposed onto the PRMT1 model as well as onto the PRMT3 crystal structure.[24] Superimposing PRMT4 (residues 140–480) on the other PRMT structures gives RMSD in the range of 1.4–1.8 Å between corresponding 296 Cα atoms. Due to the high similarity between the individual PRMTs and the fact that we know the conformation of the missing helix αX from the crystal structures of PRMT3–SAH and PRMT4–SAH, we propose that the PRMT1 model is highly accurate and a suitable target structure for virtual screening.

Ligand docking

The docking study was performed using the standard speed-up mode in Gold 3.2.[35] The default Gold parameters for these settings were used except the 'number of allowed ligand bumps' was set to 3. Gold was run to save up to 10 top-ranked docking solutions for the ligands. Docking was carried out to obtain a population of possible conformations and orientations for these inhibitors at the putative active site. GoldScore was chosen as fitness function because it has outperformed other scoring functions in our previous studies.[14]

Inhibitors

Compounds **1**–**9** were purchased from Chembridge. Compound **11** and the five randomly selected compounds shown in Figure 8 were obtained from the Office of the Associate Director, Developmental Therapeutics Program, Division of Cancer Treatment and Diagnosis, National Cancer Institute, USA. More information is publicly available at http://dtp.nci.nih.gov/branches/dscb/diversity_explanation.html.

Enzymes

hPRMT1 was expressed and purified as published.[36]

In vitro methyltransferase assay

The assays were performed in streptavidin-coated 96-well microplates (Nunc). After each incubation step, six washing steps were performed to remove the excess reagents. (Tecan Columbus plate washer; 300 μL per step in the overflow modus; washing solution: 100 mM Tris, 0.1% Tween 20, pH 7.5). The biotinylated substrate (residues 1–21 of human histone H4, Upstate) was initially bound to the streptavidin-coated surface. The biotinylated substrate (20 pmol) in a 200-μL solution of 100 mM Tris, 0.1% Tween 20, pH 7.5 was applied to each well. The microplates were incubated for 1 h at room temperature with slow agitation. In the next incubation step, the bound substrate was turned over by hPRMT1 methyltransferase activity. Pre-incubation of hPRMT1 (5 μL), inhibitor solution (5 μL), and Tris buffer (10 μL, 15 mM, pH 8.5) at room temperature is recommended for good reproducibility. The pre-incubation step was performed in 600-μL reaction tubes (Eppendorf). Control reactions (Enzyme control and substrate control) were performed accordingly, with buffer in place of inhibitor. To estimate the influence of DMSO, 5 μL DMSO were added instead of inhibitor solution. After 5 min, each incubation mixture was transferred to its well, and buffer (15 mM Tris buffer, pH 8.5) was added to obtain a final assay volume of 200 μL. To start the enzymatic reaction, 10 μL of an 800 μM SAM solution in 15 mM Tris buffer (pH 8.5) was added to each well except for the negative control, for which buffer alone (15 mM Tris, pH 8.5) was added. The microplate was kept at 30 °C for 1 h with slow agitation. The degree of methylation was detected by a primary rabbit IgG antibody (anti-H4R3, Upstate); a 1:1000 dilution (200 μL) in 100 mM Tris containing 0.1% Tween 20 and 0.5% BSA (protease free), pH 7.5, was added to each well. The microplate was incubated for 1 h at room temperature with slow agitation. Afterward, a europium-labeled secondary antibody (anti-rabbit IgG, PerkinElmer) was used for the readout; a 1:5000 dilution (200 μL) in 100 mM Tris containing 0.1% Tween 20 and 0.5% BSA (protease free), pH 7.5, was applied to each well. The microplate was incubated for 1 h at room temperature with slow agitation. Finally, the europium label was cleaved off by the addition of a proprietary enhancement/chelating cocktail (DELFIA Enhancement Solution, PerkinElmer); 200 μL were added to each well, and after revolving the plate for 5 min at slow agitation, a final time-resolved fluorescence measurement ($\lambda_{ex}=340$ nm, $\lambda_{em}=615$ nm, performed on a BMG Polarstar plate reader) gave the readout for quantitation of the enzymatic turnover.

Acknowledgements

Funding by the Deutsche Krebshilfe (Grant No. 107898) is gratefully acknowledged.

Keywords: drug design · histones · methyltransferases · prmt1 · virtual screening

[1] T. Jenuwein, C. D. Allis, *Science* **2001**, *293*, 1074–1080.
[2] M. Biel, V. Wascholowski, A. Giannis, *Angew. Chem.* **2005**, *117*, 3248–3280; *Angew. Chem. Int. Ed.* **2005**, *44*, 3186–3216.
[3] S. Schäfer, M. Jung, *Arch. Pharm.* **2005**, *338*, 347–357.
[4] M. T. Bedford, S. Richard, *Mol. Cell* **2005**, *18*, 263–272.
[5] M. Chevillard-Briet, D. Trouche, L. Vandel, *EMBO J.* **2002**, *21*, 5457–5466.
[6] H. Brahms, L. Meheus, V. de Brabandere, U. Fischer, R. Luhrmann, *RNA* **2001**, *7*, 1531–1542.
[7] D. Cheng, N. Yadav, R. W. King, M. S. Swanson, E. J. Weinstein, M. T. Bedford, *J. Biol. Chem.* **2004**, *279*, 23892–23899.

[8] U. Dery, Y. Coulombe, A. Rodrigue, A. Stasiak, S. Richard, J. Y. Masson, *Mol. Cell. Biol.* **2008**, *28*, 3058–3069.
[9] F. M. Boisvert, U. Dery, J. Y. Masson, S. Richard, *Genes Dev.* **2005**, *19*, 671–676.
[10] N. Cheung, L. C. Chan, A. Thompson, M. L. Cleary, C. W. Eric So, *Nat. Cell Biol.* **2007**, *9*, 1208–1215.
[11] T. A. Miller, D. J. Witter, S. Belvedere, *J. Med. Chem.* **2003**, *46*, 5097–5116.
[12] M. Vedel, F. Lawrence, M. Robert-Gero, E. Lederer, *Biochem. Biophys. Res. Commun.* **1978**, *85*, 371–376.
[13] D. Greiner, T. Bonaldi, R. Eskeland, E. Roemer, A. Imhof, *Nat. Chem. Biol.* **2005**, *1*, 143–145.
[14] A. Spannhoff, R. Heinke, I. Bauer, P. Trojer, E. Metzger, R. Gust, R. Schüle, G. Brosch, W. Sippl, M. Jung, *J. Med. Chem.* **2007**, *50*, 2319–2325.
[15] A. Spannhoff, R. Machmur, R. Heinke, P. Trojer, I. Bauer, G. Brosch, R. Schule, W. Hanefeld, W. Sippl, M. Jung, *Bioorg. Med. Chem. Lett.* **2007**, *17*, 4150–4153.
[16] R. Ragno, S. Simeoni, S. Castellano, C. Vicidomini, A. Mai, A. Caroli, A. Tramontano, C. Bonaccini, P. Trojer, I. Bauer, G. Brosch, G. Sbardella, *J. Med. Chem.* **2007**, *50*, 1241–1253.
[17] A. Mai, D. Cheng, M. T. Bedford, S. Valente, A. Nebbioso, A. Perrone, G. Brosch, G. Sbardella, F. De Bellis, M. Miceli, L. Altucci, *J. Med. Chem.* **2008**, *51*, 2279–2290.
[18] B. K. Shoichet, *Nature* **2004**, *432*, 862–865.
[19] B. Schlegel, C. Laggner, R. Meier, T. Langer, D. Schnell, R. Seifert, H. Stark, H. D. Höltje, W. Sippl, *J. Comput. Aided Mol. Des.* **2007**, *21*, 437–453.
[20] R. C. Neugebauer, U. Uchiechowska, R. Meier, H. Hruby, V. Valkov, E. Verdin, W. Sippl, M. Jung, *J. Med. Chem.* **2008**, *51*, 1203–1213.
[21] X. Zhang, X. Cheng, *Structure* **2003**, *11*, 509–520.
[22] X. Zhang, L. Zhou, X. Cheng, *EMBO J.* **2000**, *19*, 3509–3519.
[23] D. Cheng, J. Cote, S. Shaaban, M. T. Bedford, *Mol. Cell* **2007**, *25*, 71–83.
[24] N. Troffer-Charlier, V. Cura, P. Hassenboehler, D. Moras, J. Cavarelli, *EMBO J.* **2007**, *26*, 4391–4401.
[25] N. Huang, B. K. Shoichet, J. J. Irwin, *J. Med. Chem.* **2006**, *49*, 6789–6801.
[26] A. May, M. Zacharias, *J. Med. Chem.* **2008**, *51*, 3499–3506.
[27] E. S. Bolstad, A. C. Anderson, *Proteins* **2008**, *73*, 566–580.
[28] T. Noeske, A. Jirgensons, I. Starchenkovs, S. Renner, I. Jaunzeme, D. Trifanova, M. Hechenberger, T. Bauer, V. Kauss, C. G. Parsons, G. Schneider, T. Weil, *ChemMedChem* **2007**, *2*, 1763–1773.
[29] A. L. Parill, U. Echols, T. Nguyen, T. C. Pham, A. Hoeglung, D. L. Baker, *Bioorg. Med. Chem.* **2008**, *16*, 1784–1795.
[30] E. Kellenberger, J. Y. Springael, M. Parmentier, M. Hachet-Haas, J. L. Galzi, D. Rognan, *J. Med. Chem.* **2007**, *50*, 1294–1303.
[31] V. Kairys, M. X. Fernandes, M. K. Gilson, *J. Chem. Inf. Model.* **2006**, *46*, 365–379.
[32] P. Cozzini, G. E. Kellogg, F. Spyrakis, D. J. Abraham, G. Costantino, A. Emerson, F. Fanelli, H. Gohlke, L. A. Kuhn, G. M. Morris, M. Orozco, T. A. Pertinhez, M. Rizzi, C. A. Sotriffer, *J. Med. Chem.* **2008**, *51*, 6237–6255.
[33] J. J. Irwin, B. K. Shoichet, *J. Chem. Inf. Model.* **2005**, *45*, 177–182.
[34] G. Wolber, T. Langer, *J. Chem. Inf. Model.* **2005**, *45*, 160–169.
[35] G. Jones, P. Willet, R. C. Glen, A. R. Leach, R. Taylor, *J. Mol. Biol.* **1997**, *267*, 727–748.
[36] P. Trojer, M. Dangl, I. Bauer, S. Graessle, P. Loidl, G. Brosch, *Biochemistry* **2004**, *43*, 10834–10843.

Received: September 15, 2008
Revised: October 22, 2008
Published online on December 10, 2008

