Computational Validation of the Importance of Absolute Stereochemistry in Virtual Screening

Wesley H. Brooks,[†,⊥] Kenyon G. Daniel,[‡,⊥] Shen-Shu Sung,[†] and Wayne C. Guida*[,‡,§,||]

High Throughput Screening and Chemistry Core and Drug Discovery Program, H. Lee Moffitt Cancer Center & Research Institute, Tampa, Florida 33612, and Department of Chemistry and Center for Molecular Diversity in Drug Design, Discovery, and Delivery and Department of Interdisciplinary Oncology, University of South Florida, Tampa, Florida 33620

Received September 28, 2007

Consideration of stereochemistry early in the identification and optimization of lead compounds can improve the efficiency and efficacy of the drug discovery process and reduce the time spent on subsequent drug development. These improvements can result by focusing on specific enantiomers that have the desired potential therapeutic effect (eutomers), while removing from consideration enantiomers that may have no, or even undesirable, effects (distomers). A virtual screening campaign that correctly takes stereochemical information into account can, in theory, be utilized to provide information about the relative binding affinities of enantiomers. Thus, the proper enumeration of the relevant stereoisomers in general, and enantiomeric pairs in particular, of chiral compounds is crucial if one is to use virtual screening as an effective drug discovery tool. As is obvious, in cases where no stereochemical information is provided for chiral compounds in a 2D chemical database, then each possible stereoisomer should be generated for construction of the subsequent 3D database to be used for virtual screening. However, acute problems can arise in 3D database construction when relative stereochemistry is encoded in a 2D database for a chiral compound containing multiple stereogenic atoms but absolute stereochemistry is not implied. In this case, we report that generation of enantiomeric pairs is imperative in database development if one is to obtain accurate docking results. A study is described on the impact of the neglect of enantiomeric pairs on virtual screening using the human homolog of murine double minute 2 (MDM2) protein, the product of a proto-oncogene, as the target. Docking in MDM2 with GLIDE 4.0 was performed using the NCI Diversity Set 3D database and, for comparison, a set of enantiomers we created corresponding to mirror image structures of the single enantiomers of chiral compounds present in the NCI Diversity Set. Our results demonstrate that potential lead candidates may be overlooked when databases contain 3D structures representing only a single enantiomer of racemic chiral compounds.

INTRODUCTION

The importance of stereochemistry in biological molecules, such as proteins, is well established.[1−3] For example, in the tricarboxylic acid cycle (also known as the citric acid cycle or Krebs cycle), the enzyme aconitase uses the achiral molecule citric acid as a substrate to produce a single enantiomer of isocitric acid, which has two chiral centers. It does so by operating on the pro-R hydrogen of the pro-R arm of citric acid to first produce *cis*-aconitic acid via a dehydration reaction followed by stereospecific hydration of *cis*-aconitic acid to produce isocitric acid.[4] Likewise, it is well-known that the binding affinity of a chiral drug for its receptor will, generally, be different for its two mirror image forms (enantiomers).[5] Furthermore, it has become increasingly important in drug discovery and development to focus on a homochiral compound (as opposed to a racemic mixture) that has the desired therapeutic effect, referred to as the eutomer, while removing the distomer (the enantiomer that has no, or even an undesirable, effect). This focus on absolute stereochemistry has become more feasible in recent years as a result of improvements in asymmetric synthesis and chiral separation that allow facile production of homochiral compounds for in vitro and in vivo testing and subsequent drug development.[6] Furthermore, in 1992, the FDA issued a policy statement on the development of stereoisomeric drugs to give guidance to drug companies.[7] The emphasis of the FDA guidelines is that the absolute stereochemical composition, identity, strength, quality, and purity of drugs with chiral centers should be known for pharmacological, toxicological, and clinical studies and for the final product. The guidelines suggest that for chiral compounds the stereochemical composition and determination of the activity of each *enantiomer* should be established early in pharmacological studies since any initial studies may be meaningless if these aspects are not known. Finally, a number of the leading drugs developed in recent years are marketed as single enantiomers, such as Lipitor (Pfizer), Zocor (Merck), Plavix (Sanofi-Synthelabo), and Nexium (AstraZeneca).[8,9]

* Corresponding author e-mail: Wayne.Guida@moffitt.org.
† High Throughput Screening and Chemistry Core, H. Lee Moffitt Cancer Center & Research Institute.
‡ Drug Discovery Program, H. Lee Moffitt Cancer Center & Research Institute.
§ Department of Chemistry and Center for Molecular Diversity in Drug Design, Discovery, and Delivery, University of South Florida.
|| Department of Interdisciplinary Oncology, University of South Florida.
⊥ W.H.B. and K.G.D. contributed equally to this manuscript.

10.1021/ci700358r CCC: $40.75
Published on Web 02/12/2008

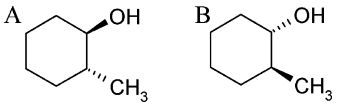

Figure 1. Example enantiomeric pair: A) 1R,2R-2-methylcyclohexanol and B) 1S,2S-2-methylcyclohexanol.

In fact, almost half of the drugs currently on the market are homochiral.[5] Thus, the interest in stereochemistry in drug development has been sparked not only by the desire to reduce potential adverse effects of distomers in racemic mixtures but also the need to meet regulatory requirements for testing and labeling as well as issues regarding patent protection.

Virtual screening has been used effectively for the identification of potential lead candidates for drug discovery,[10] and virtual screening protocols that correctly take stereochemical information into account can theoretically be used to provide information about the relative binding affinities of enantiomers. Thus, for chiral compounds, proper enumeration of the relevant stereoisomers in general, and pairs of enantiomers in particular, is key to the effective use of virtual screening as a drug discovery tool.

Generally, synthetic compounds are produced as racemic mixtures unless asymmetric synthesis or chiral separation has been performed. Owing to the difficulty and cost associated with both asymmetric synthesis and chiral separation to obtain enantiomerically pure material, when dealing with compound libraries of unknown chiral purity, it is prudent to assume the physical samples exist as racemic mixtures. Nonetheless, databases for virtual screening that represent compound libraries often do not contain both enantiomers of compounds that are likely to be synthesized as racemic mixtures. For example, the 3D databases currently available from the NCI's Developmental Therapeutics Program[11] do not contain pairs of enantiomers for the chiral compounds contained therein. Although commercially available software packages, like Corina,[12] or the Stereoplex module for Concord,[13] are capable of enumerating multiple stereoisomers of a given compound, a problem arises when *relative* stereochemistry is specified for a 2D structure undergoing transformation to 3D. It is a common convention, when using programs like ISISDraw[14] to generate MDL formatted files, to specify relative stereochemistry for diastereomers by drawing "up" (solid wedge) and "down" (hashed wedge) bonds in a 2D representation of a molecule. However, it is understood that such notation can, and often does, represent a racemic mixture (i.e., the presence of the other enantiomer is implied). For example, the chiral compound 2-methylcyclohexanol contains two chiral centers. Thus, there are four possible stereoisomers of this compound (two pairs of enantiomers). As depicted in Figure 1, structure A could be used to represent the *relative* stereochemistry of *trans*-2-methylcyclohexanol. It is implied that the other enantiomer (structure B) of this diastereomer of *cis*-2-methylcyclohexanol would be present in the racemic mixture that results when this compound is prepared under achiral conditions. 2D structure files relying upon MDL's sdf format, for example, have the capability to encode such relative stereochemistry. However, when using this type of 2D structural information to generate 3D structures, programs such as Corina,[12] when instructed to retain the stereochemistry encoded in the original 2D file, will quite appropriately maintain the *absolute* stereochemistry encoded in the structure file and fail to generate the opposite enantiomer. However, for commercially available compounds, chiral molecules are generally sold as racemic mixtures, and both enantiomers corresponding to a particular relative stereochemical representation should be contained in the final 3D file. Nonetheless, we are unaware of any program used for conversion of 2D structures to 3D that will generate enantiomeric pairs for compounds that have multiple chiral centers and have relative stereochemistry specified in the 2D structure file. Moreover, since enantiomers of a chiral compound might have different binding free energies for a given receptor, it is important in virtual screening that both enantiomers are docked to the receptor. In such a virtual screening campaign, it might be anticipated that the docking algorithms presently in use would be able to produce different docking scores for enantiomers that bind asymmetrically to a particular receptor.

The general question of how to handle chiral centers in molecules whose structures are encoded in 2D and 3D databases is a particularly important issue that can arise when one is preparing databases for virtual screening. Failure to include appropriate diastereomers and, relevant to this study, enantiomeric pairs of compounds with multiple chiral centers could significantly increase the number of false negatives and adversely affect enrichments. However, we submit that the many layers and nuances of the problem have not been adequately described in the literature.

In order to provide a frame of reference for our investigation, a short description of the problem follows. Frequently, compound libraries for virtual screening are provided as MDL sdf files. These files may or may not contain complete stereochemical information. Indeed, short of visual inspection of the file's contents, it can be extremely difficult to determine the extent of stereochemical information provided. The problem faced in virtual screening projects can be broken down into several different potential types. If the stereochemistry is indicated for a compound (e.g., either "solid wedge" or "hashed wedge" bonds or Cahn, Ingold, and Prelog R and S chirality is encoded in the structure file) for a compound with a single chiral center, this indicates an enantiomerically pure compound, and no further action needs to be taken (Figure 2A). However, if there are multiple chiral centers present and the stereochemistry is encoded, then this is most likely indicative of relative stereochemistry, and *the existence of the other enantiomer is implied but may not be included in the structure file*. In this case, the enantiomeric partner should be generated based on the relative stereochemical information provided (Figure 2B). We note that although it is possible that absolute stereochemistry is what is meant in this case, it is prudent to generate both enantiomeric partners for subsequent virtual screening. It is also possible that multiple stereocenters may be present but only one of them possesses stereochemical information. For these structures, all the variants for the unknown center(s) should be generated, but the single known center can remain constant since it can be assumed that the presence of stereochemical information for a single chiral center is indicative of absolute stereochemistry (Figure 2C). If no stereochemical information is provided in the structure file, then each of the 2^n possible stereoisomers should be generated (where n = the number of chiral centers) (Figure

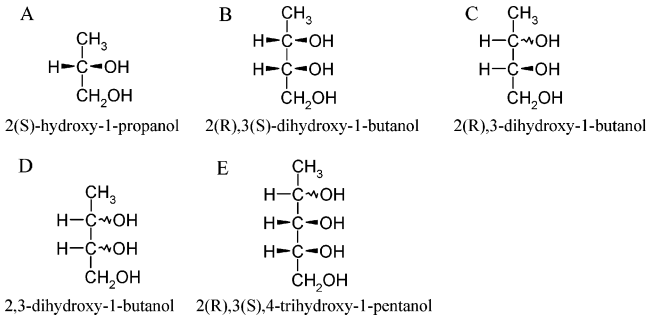

Figure 2. A series of polyhydroxy compounds illustrating different levels of stereochemical detail. A) A compound with a single chiral center that is noted: no further structures need to be generated. B) Multiple chiral centers with indicated stereochemistry: in this case the 2(S),3(R) enantiomer should be generated. C) Multiple chiral centers one of which possesses noted stereochemistry: in this case the single chiral center is considered to be absolute so only the 2(R),3(R) and the 2(R),3(S) structures should be generated. D) Multiple chiral centers, no stereochemistry indicated: in this case all possible configurations should be generated: 2(R),3(R) 2(S),3-(S), 2(R),3(S), 2(S),3(R). E) Multiple chiral centers with relative stereochemistry indicated for a subset of stereocenters: in this case each variant for the unknown center(s) should be generated and the enantiomeric partners for the entire set should be generated: 2(R),3(S),4(R) and 2(R),3(S),4(S) and 2(S),3(R),4(R) and 2(S),3(R),4(S).

2D). In some cases, multiple stereocenters may be present, but only partial stereochemical information may be provided (Figure 2E). Thus, while some of the chiral centers may show known *relative* stereochemistry, all of the stereoisomers for the unknown center(s) should be generated leading to a set of diastereomers, and then all the enantiomeric pairs of these diastereomers should be generated. These stereochemical issues can emerge during 3D structure generation if the application retains chirality based on provided structural information. Unfortunately, working with a 3D structure file does not avoid the issue. It is to the investigator's advantage to determine how the 3D coordinates were generated and how complete and accurate the stereochemical information is. It should be apparent that at a minimum enantiomeric pairs should be generated for molecules with multiple chiral centers and only relative stereochemistry indicated.

It may be of concern that 3D databases for virtual screening can grow exponentially and become unwieldy. It should be emphasized that in some cases this is unavoidable (such as the case in Figure 2D) but in others only an enantiomeric partner needs to be generated (Figure 2B). Rather than use the "generate every stereoisomer for every compound" approach for an entire database, only certain classes of compounds would fall into this category. Implementation of small programs or scripts (such as the ones we have utilized in this study) can readily identify the types of chiral compounds present in the file and automate the generation of enantiomeric pairs, whereas commonly available commercial applications can be used to generate diastereomers when needed. Restricting exponential growth to only certain classes of compounds could help avoid the exponential explosion of 3D chemical databases while eliminating as many false negatives as possible in subsequent virtual screening.

In the present work, we investigate whether enantiomeric pairs of molecules, when docked to the same receptor, are scored differently enough to have a significant impact on

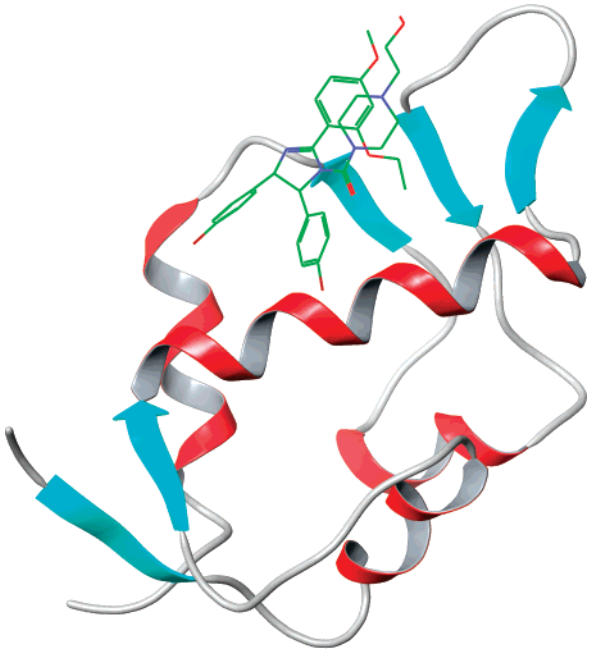

Figure 3. Structure of p53 binding domain of MDM2 from pdb 1RV1. Alpha helices are shown in red and beta sheets are shown in blue. Nutlin-2 is shown in its crystallographically determined position (with green carbons).

the virtual screening of 3D databases. In order to investigate the importance of absolute stereochemistry in virtual screening, we decided to use the human homolog of the murine double minute 2 protein (MDM2) as a test system. The proto-oncogene product, MDM2, is a putative E3 type ubiquitin ligase involved in the ubiquitin-proteasome system for degradation of p53, a well-studied tumor suppressor. It also functions in its own degradation as well as having its expression regulated by p53. This provides tight control of the levels of p53. In those tumors attributed to inactivation of p53, the majority results from either mutations of p53 or amplification of wild type MDM2. Moreover, MDM2 has recently been reported as targeting the retinoblastoma protein (Rb)[15] and the E2F1-DP1 transcription factor complex[16] as additional means by which MDM2 can affect control of cell growth. Therefore, MDM2 is a useful target for virtual screening of small molecule databases in the search for anticancer drugs since disruption of p53-MDM2 binding can release p53 to function in its role as a tumor suppressor.[17,18] The p53 binding site in the N-terminal domain of MDM2 is primarily a hydrophobic cleft formed by an α-helix on each side, a β sheet at each end, and two short α helices at the bottom (Figure 3). Residues 17−27 of p53, which form a short amphipathic α-helix, appear to be sufficient for p53 binding in which two aromatic (Phe19 and Trp23) and one aliphatic residue (Leu26) in this p53 transactivation domain form the primary contacts in the MDM2 binding pocket.[19] This suggests that small molecules of approximately 300 Daltons that mimic the combination of these three residues could bind in the MDM2 pocket and disrupt the binding of p53 to MDM2.[19]

There are several X-ray structures for human MDM2 available in the Protein DataBank (http://www.rcsb.org).[20] We chose to use the structure contained in the file 1RV1.pdb. Recently, a group of chiral *cis*-imidazoline analogs, the nutlins, were discovered via experimental screening that can

disrupt the interaction between MDM2 and p53, thereby releasing p53 to carry out its role as a transcription factor involved in cell cycle control, apoptosis, and cell senescence.[21,22] A racemic mixture of one of these compounds, nutlin-2, was cocrystallized with truncated human MDM2 (84 residues), and the structure of the complex is contained in the file 1RV1.[22] In this cocrystal structure, with a resolution of 2.30 Å, only a single enantiomer of nutlin-2 is observed to be bound within the p53-binding pocket of MDM2. We chose this structure for our study over others available in the Protein DataBank because it has the chiral molecule nutlin-2 bound and this system provides an excellent example of a small molecule for which only one enantiomer from a racemic mixture binds with sufficient affinity to cocrystallize with its protein receptor. Furthermore, the closely related analog, nutlin-3, was resolved into its individual enantiomers, which were tested separately for their ability to bind to MDM2 and one enantiomer was found to be significantly more potent than the other.

For our virtual screening study, we used the NCI Diversity Set 3D database as a representative collection of highly diverse small molecules that contains only a single enantiomer of each chiral compound. For each of these chiral compounds we generated the corresponding enantiomeric partner. We used Schrödinger's GLIDE 4.0 program to perform automated docking. GLIDE Scores (GScores) could then be compared to identify the predicted differences in binding free energies of the enantiomeric pairs and study their ranks relative to the entire NCI Diversity Set 3D database. It is worth noting that it is quite possible that some (or even many) of the chiral compounds in the NCI Diversity Set 3D database have incorrect stereochemistry.[11] Thus, the stereochemical information of the 1990 compounds of the NCI Diversity Set 3D database should be examined carefully prior to use in an actual virtual screening campaign. However, for our purpose the NCI Diversity Set fulfilled the requirement of a 3D database of highly diverse compounds containing a significant number of chiral molecules. Since our aim was not to perform a virtual screening campaign followed by experimental testing, the NCI Diversity Set 3D database was suitable for examining the role of stereochemistry, and especially the importance of missing enantiomeric pairs, in virtual screening.

MATERIALS AND METHODS

Hardware. Molecular modeling and virtual screening were performed using a Dell Precision 490 workstation running Fedora Core 5 with dual Intel Xeon 3.20 GHz processors, 4 Gb RAM, 120 Gb hard drive, and nVidia Quadro FX 4500 graphics card with hardware stereo.

Software. Schrödinger's Maestro 7.5 was used as the primary graphical user interface (GUI).[23] Schrödinger's LigPrep 2.0[23] was used for processing of the NCI Diversity Set 3D database, downloaded from NCI's DTP Web site.[11] During that process, LigPrep was instructed to maintain the stereochemistry encoded in the original sdf file. It is noteworthy that LigPrep, like Corina, when instructed to retain stereochemistry encoded in a structure file, will generate enantiomeric pairs for molecules with one chiral center since no stereochemical descriptors should be present in the file unless the compound is known to be enantiomerically pure. On the other hand, LigPrep, like Corina, will generate only one enantiomer for molecules with two or more chiral centers for which relative stereochemistry is encoded in a 2D structure file. Swiss PBD Viewer[24] and Schrödinger's MacroModel 9.1[23] were used in the preparation of the protein structure. Schrödinger's GLIDE 4.0[23] was used for the generation of grid files and the virtual screening. PyMol from DeLano Scientific, along with Maestro, was used for graphical presentation of our results in the figures.[25]

Data. The NCI Diversity Set 3D database, consisting of the structures of 1990 compounds, was used as our diverse small molecule database. Schrödinger's LigPrep was used to generate alternative tautomers, ring conformations, and ionization states, which increased the number of 3D structures to 2967. Conversion from sdf file format to Maestro format (.mae extension) was also performed using LigPrep. The enzyme structure of 1RV1.pdb,[22] which contains the crystal structure of a single nutlin-2 enantiomer bound to truncated human MDM2, was used for preliminary docking to determine whether GLIDE could qualitatively discriminate between the known difference in binding affinity of the nutlin-2 enantiomers. In order to prepare the protein for use in virtual screening, the pdb file was first opened in Swiss PDB Viewer, which automatically corrects any missing residues due to incomplete electron density. The corrected structure was then imported into Schrödinger's Maestro. Water molecules were removed, hydrogens were added, and the protein was then prepared for use in docking studies by processing it with Schrödinger's protein preparation utility. This preparation process minimizes the protein's potential energy gradient through a series of constrained energy minimizations to an rmsd of 0.30 Å using the OPLS-2001 force field.

Small Molecule Selection. Enantiomers of chiral compounds were created from the LigPrep processed NCI Diversity Set 3D database (and contained in a Maestro formatted file) using a Perl script that we created. There were 967 structures extracted from the NCI Diversity Set 3D database, of which 923 were successfully converted into enantiomeric pairs (the remaining compounds were predominantly meso-compounds). These isomers were appended to the 2967 NCI Diversity Set structures. The nutlin-2 enantiomeric pair was added to the resulting data set as a control since the binding mode of one enantiomer to MDM2 is known.[22] This comprehensive file was then used in a GLIDE docking run.

Docking. For all docking studies we employed Schrödinger's GLIDE 4.0 SP using the default settings.

Analysis. The GLIDE 4.0 docking scores (GScores) provide estimates of relative binding free energies for each docked structure. These GScores were calculated and used to rank order the structures relative to each other.

RESULTS AND DISCUSSION

MDM2 and Nutlin. The nutlins, mentioned previously, are a class of compounds identified from experimental screening of a library of compounds against MDM2.[22] The nutlins, synthesized as racemates, displaced a recombinant p53 protein with median inhibitory concentration (IC_{50}) values in the range of 100−300 nM.[22] One of the nutlins (nutlin-3) was separated on chiral columns. One of its

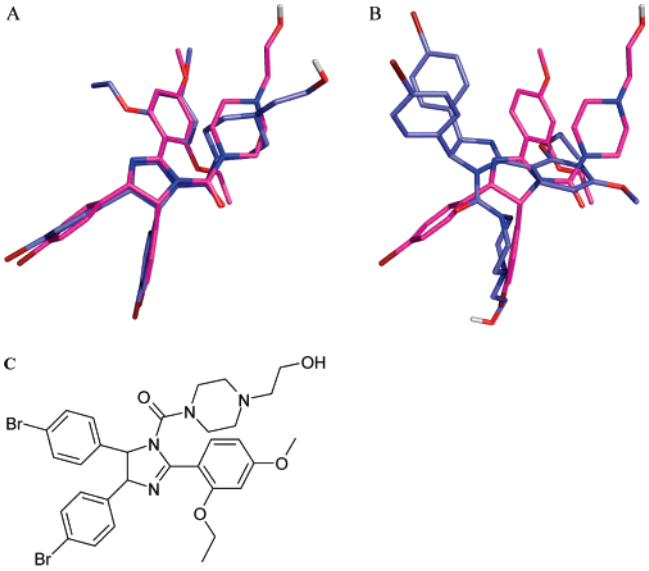

Figure 4. Comparison of the crystallographically determined pose of nutlin-2 with the GLIDE 4.0 SP docking result for nutlin-2 enantiomer-a vs enantiomer-b. A) The pose observed using Glide 4.0 for self-docking of nutlin-2 (enantiomer-b) to the p53 binding domain of MDM2 (shown in blue) vs the cystallographically observed pose (shown in purple; superimposition rmsd=2.41 Å). B) The GLIDE docking pose (blue) found for enantiomer-b of nutlin-2 docked to the p53 binding domain of MDM2 vs the cystallographically observed pose of nutlin-2 (purple; superimposition rmsd=7.51 Å). C) 2D structure of nutlin-2.

enantiomers (labeled enantiomer-a) was found to bind 150 times more tightly to MDM2 than the other enantiomer (enantiomer-b).[22] In preliminary studies, we compared the cocrystal structure of the nutlin-2 enantiomer that binds to truncated MDM2, as determined cystallographically,[22] with the pose of each separate nutlin-2 enantiomer docked to MDM2 using GLIDE 4.0 SP. The nutlin-2 enantiomer that cocrystallized with MDM2 presumably has the same absolute stereochemistry as nutlin-3 enantiomer-a, and we have labeled this nutlin-2 stereoisomer enantiomer-a as well. Thus, enantiomer-a of nutlin-2 docked with a GScore of −7.22 kcal/mol, whereas enantiomer-b docked with a score −4.93 kcal/mol. Enantiomer-a of nutlin-2 docked with a pose very similar to the one observed crystallographically (rmsd: 2.41 Å, Figure 4A). However, enantiomer-b docked with a completely different pose (rmsd: 7.51 Å, Figure 4B). Due to the differences in GScore and pose as a result of this docking study, we were encouraged that GLIDE 4.0 would be able to adequately distinguish the binding of enantiomers to stereoselective targets.

GLIDE Docking with the NCI Diversity Set. In order to explore the importance of including enantiomeric pairs of chiral compounds suggested by the nutlin-2 results as previously described, we used the NCI Diversity Set 3D database. We performed virtual screening using Schrödinger's GLIDE 4.0 using an input file of 3892 structures, which was composed of 2967 NCI Diversity Set structures (76.2%), our 923 enantiomers (23.7%), and our nutlin controls. After virtual screening, we examined the rankings and GScores of the top 100 and top 10% of unique best scoring compounds and the best scoring enantiomeric partner, discarding duplicates comprised of worse ranking alternate structures such as tautomers and ionization states to avoid dilution of the analysis set (Table 1).

In the top ranking 100 compounds, 41 contain chiral centers. Of these compounds, 33 had only one partner within the top 100 compounds. This suggests the possibility of decreasing this pool of potential leads by 33% if only one member of the enantiomeric pair, the "wrong" member, had been considered. Assuming an equal probability of either the "right" or "wrong" enantiomer appearing in the database used for virtual screening, over 15% of the potential leads would have been excluded as false negatives. Likewise, extending the analysis to the top 10% of compounds with the best GLIDE score (239), the percent of chiral compounds with only one partner in the analysis window remained high as compared to the top 100 (26% vs 33%).

It is noteworthy that the difference in ranking can be dramatic. For example, nutlin-2's rank changed, based on the enantiomer considered, from rank 11 (for enantiomer-a) to rank 2304 (enantiomer-b). We note that the enantiomer with rank 11 (enantiomer-a) is the same one that is observed experimentally to bind to MDM2. Frequently, a very similar trend could be observed among the NCI compounds as well (Table 1). These results strongly suggest that a significant number of potential leads (as many as ∼15% in the present case) could be missed when enantiomeric pairs are not considered during screening of the NCI Diversity Set. In fact, three of the top ten chiral hits were compounds generated by our script underscoring the significance of specifically including enantiomeric partners.

Finally, it is worth mentioning that in those cases where the absolute stereochemistry of active compounds is known, enantiomers of these known actives make a perfect set of decoy compounds (if the target is enantioselective) since they have identical physical properties to their bioactive enantiomeric partner except for the rotation of plane polarized light. Such enantiomers would provide a useful addition to decoy sets like the Directory of Useful Decoys[26] that is built on a similar premise.

CONCLUSION

This study describes the results of virtual screening with human MDM2, a proto-oncogene product, using the NCI Diversity Set 3D database augmented with missing enantiomeric partners of chiral compounds. Our results show that one can obtain significant differences in ranking based on GLIDE scores for the individual enantiomers of numerous compounds in the database. Furthermore, our study demonstrates that potential lead candidates may be overlooked when databases containing 3D structures representing only a single enantiomer of racemic chiral compounds are used for virtual screening. The results we have obtained clearly establish that potential improvements in lead generation are likely to accompany a comprehensive approach to the inclusion of stereoisomers in databases used for virtual screening. Thus, several examples have been obtained where the enantiomer included in the original database would not have been considered for further analysis after virtual screening based on typical cutoff criteria (i.e., using the top 10% by rank of the compounds with the best docking score for further study), whereas the enantiomer we generated showed a significant improvement in docking score that would have justified its further investigation as a potential lead candidate. Our data suggest that in a standard virtual screening protocol, in which

Table 1. Ranking Results for Top 10% of NCI Diversity Set Compounds with Chiral Centers from GLIDE SP[a]

compound	rank	enantiomer rank	compound	rank	enantiomer rank
372127	2	594	51070	125	69
121912:Enant	8	741	372280:Enant	127	488
231643	9	298	363744:Enant	131	301
84093:Enant	10	112	158383:Enant	132	564
nutlin2 (enantiomer-a)	11	2304	17777:Enant	133	181
34924	13	89	112671	136	254
71426:Enant	14	1130	138756	139	462
84094	16	159	363935:Enant	140	207
116709	18	274	328087	141	144
114414	20	111	328087:Enant	144	141
1012	22	212	661755:Enant	145	464
48630	23	1263	1013:Enant	146	319
92412:Enant	26	108	170992:Enant	148	1349
143099:Enant	33	1012	372294	149	764
633406:Enant	35	575	17354	150	91
122405	38	49	6176	152	484
372074	40	77	84094:Enant	159	16
96948	41	103	83633	160	90
13480:Enant	44	294	150316	162	444
122405:Enant	49	38	289336:Enant	163	519
191260:Enant	57	295	3323:Enant	164	198
117268:Enant	61	1292	26273	166	1248
321578:Enant	62	99	25431:Enant	169	368
24047	64	123	105017	170	370
51070:Enant	69	125	634396	171	551
364164	71	472	135371:Enant	180	404
123526:Enant	73	106	17777	181	133
141566	74	292	186057	182	655
372074:Enant	77	40	369294	184	251
24048	79	110	639174	188	689
186066	81	731	321568:Enant	189	692
401077	83	647	309883	191	241
58255:Enant	84	347	331973:Enant	192	337
208758	86	217	3323	198	164
34924:Enant	89	13	39207	201	1073
83633:Enant	90	160	23925	202	114
17354:Enant	91	150	134755:Enant	204	1411
153625:Enant	94	210	309874:Enant	205	958
121860:Enant	96	778	363935	207	140
150117:Enant	98	121	43513	208	1100
321578	99	62	42448	209	478
2052:Enant	101	214	153625	210	94
96948:Enant	103	41	1012:Enant	212	22
121847	105	1497	2052	214	101
123526	106	73	27959:Enant	215	272
92412	108	26	17387	216	492
201579	109	229	208758:Enant	217	86
24048:Enant	110	79	368270	218	607
114414:Enant	111	20	201863	220	559
84093	112	10	2561:Enant	225	534
322661:Enant	113	323	13512	227	535
23925:Enant	114	202	201579:Enant	229	109
351520	115	265	255980	230	1455
117614:Enant	118	787	119847:Enant	231	309
150117	121	98	305787:Enant	233	1381
24047:Enant	123	63	169534	234	1869
44102	124	760	319440:Enant	235	342

[a] NSC numbers are listed for the NCI Diversity Set compounds. Enantiomers that were generated and not part of the original NCI Diversity Set 3D database are noted by the extension "Enant". Total unique enantiomers in the top 100 = 41 (41%). Enantiomers that have only 1 member in the top 100 = 33 (33%).Total unique enantiomers in the top 10% (239) = 114 (48%). Enantiomers that have only 1 member in the top 10% = 62 (26%).

the top ranking compounds are selected for experimental testing, a significant fraction of the potential leads could be lost if enantiomeric pairs are not considered.

The purpose of this report is to establish that thorough and proper representation of chiral structures in virtual screening databases should receive as much consideration as ionic states, ring conformations, and tautomeric states. Lack of proper representation of stereochemistry can become particularly insidious when the physical compound is known to be a racemic mixture, but only an arbitrary member of each enantiomeric pair that comprises the mixture is included in the 3D database used for virtual screening. Our data demonstrate that both members of enantiomeric pairs should be included during virtual screening to decrease potential false negatives. We believe that this inclusion can significantly enhance the ability of the docking code employed for virtual screening to generate an enriched set of top scoring compounds within a chemical library for lead identification.

Investigators should be aware that the available 3D data sets in common use (e.g., NCI Diversity Set) are likely to be incomplete with respect to stereochemical information and almost certainly lacking enantiomeric pairs for racemic compounds containing multiple chiral centers. Therefore, prior to using a 3D structure file for virtual screening, investigators would benefit by examining the initial structure file for completeness of stereochemical information, in general, and generating enantiomeric pairs for chiral compounds for which such pairs are lacking.

ACKNOWLEDGMENT

This research was supported, in part, by grant P01 CA-94000 from the National Cancer Institute, Bethesda, MD. The 1990 compound NCI Diversity Set 3D database was obtained from the Developmental Therapeutics Program, Division of Cancer Treatment and Diagnosis, National Cancer Institute, 6130 Executive Blvd., Room 8020 Rockville, MD 20852. This work has been supported, in part, by the High Throughput Screening and Chemistry Core Facilities at the H. Lee Moffitt Cancer Center and Research Institute.

REFERENCES AND NOTES

(1) Caldwell, J. The importance of stereochemistry in drug action and disposition. *J. Clin. Pharmacol*. **1992**, *32*, 925−9.
(2) Ariens, E. J. Stereochemistry, a basis for sophisticated nonsense in pharmacokinetics and clinical pharmacology. *Eur. J. Clin. Pharmacol.* **1984**, *26*, 663−8.
(3) Hutt, A. J.; O'Grady, J. Drug chirality: a consideration of the significance of the stereochemistry of antimicrobial agents. *J. Antimicrob. Chemother*. **1996**, *37*, 7−32.
(4) Garrett, R. H.; Grisham, C. M. *Biochemistry*, 3rd ed.; Thompson Brooks/Cole: Belmont, 2005; p 1086.
(5) Beroza, P.; Suto, M. J. Designing chiral libraries for drug discovery. *Drug Discovery Today* **2000**, *5*, 364−372.
(6) Thayer, A. M. Centering on Chirality. *Chem. Eng. News* **2007**, 11−19.
(7) FDA. *FDA's policy statement for the development of new stereoisomeric drugs*. http://www.fda.gov/cder/guidance/stereo.htm (accessed Nov 20, 2007).
(8) Thayer, A. M. Trial Separations. *Chem Eng. News* **2005**, 49.
(9) Thayer, A. M. Chiral Catalysis. *Chem Eng. News* **2005**, *83*, 40.
(10) Ghosh, S.; Nie, A.; An, J.; Huang, Z. Structure-based virtual screening of chemical libraries for drug discovery. *Curr. Opin. Chem. Biol.* **2006**, *10*, 194−202.
(11) NCI. *Developmental Therapeutics Program*. http://dtp.nci.nih.gov/ (accessed Nov 20, 2007).
(12) Sadowski, J.; Gasteiger, J.; Klebe, G. Comparison of Automatic Three-Dimensional Model Builders Using 639 X-Ray Structures. *J. Chem. Inf. Comput. Sci*. **1994**, 1000−1008.
(13) Pearlman, R. Rapid Generation of High Quality Approximate 3-dimension Molecular Structures. *Chem. Des. Auto News* **1987**, *2*.
(14) *MDL/ISIS Draw*; MDL Information Systems Inc.: San Leandro, CA, 2002
(15) Xiao, Z. X.; Chen, J.; Levine, A. J.; Modjtahedi, N.; Xing, J.; Sellers, W. R.; Livingston, D. M. Interaction between the retinoblastoma protein and the oncoprotein MDM2. *Nature* **1995**, *375*, 694−8.
(16) Martin, K.; Trouche, D.; Hagemeier, C.; Sorensen, T. S.; La Thangue, N. B.; Kouzarides, T. Stimulation of E2F1/DP1 transcriptional activity by MDM2 oncoprotein. *Nature* **1995**, *375*, 691−4.
(17) Zheleva, D. I.; McInnes, C.; Baxter, C.; Gibson, D.; Maccallum, D.; Powers, H.; Duncan, K.; Bailey, K.; Cummings, L.; Thomas, M.; Wang, S.; Turner, N.; Uhrinova, S.; Barlow, P.; Taylor, P.; Walkinshaw, M.; Lane, D.; Fischer, P. Bisarylsulfonamides - novel small molecule inhibitors of p53-Mdm2 interaction. *Proc. Am. Assoc. Cancer Res.* **2004**, 5552.
(18) Lu, Y.; Nikolovska-Coleska, Z.; Fang, X.; Gao, W.; Shangary, S.; Qiu, S.; Qin, D.; Wang, S. Discovery of a nanomolar inhibitor of the human murine double minute 2 (MDM2)-p53 interaction through an integrated, virtual database screening strategy. *J. Med. Chem.* **2006**, *49*, 3759−62.
(19) Kussie, P. H.; Gorina, S.; Marechal, V.; Elenbaas, B.; Moreau, J.; Levine, A. J.; Pavletich, N. P. Structure of the MDM2 oncoprotein bound to the p53 tumor suppressor transactivation domain. *Science* **1996**, *274*, 948−53.
(20) Berman, H. M.; Westbrook, J.; Feng, Z.; Gilliland, G.; Bhat, T. N.; Weissig, H.; Shindyalov, I. N. and Bourne, P. E. The Protein Data Bank. *Nucleic Acids Res*. **2000**, *28*, 235−42.
(21) Klein, C.; Vassilev, L. T. Targeting the p53-MDM2 interaction to treat cancer. *Br. J. Cancer* **2004**, *91*, 1415−9.
(22) Vassilev, L. T.; Vu, B. T.; Graves, B.; Carvajal, D.; Podlaski, F.; Filipovic, Z.; Kong, N.; Kammlott, U.; Lukacs, C.; Klein, C.; Fotouhi, N.; Liu, E. A. In vivo activation of the p53 pathway by small-molecule antagonists of MDM2. *Science* **2004**, *303*, 844−8.
(23) *Schrödinger Softwawre Suite;* Schrodinger LLC: New York, 2006.
(24) Guex, N.; Peitsch, M. C. SWISS-MODEL and the Swiss-PdbViewer: an environment for comparative protein modeling. *Electrophoresis* **1997**, *18*, 2714−23.
(25) DeLano, W. L. *The PyMOL Molecular Graphics System*; DeLano Scientific LLC: Palo Alto, CA, 2002.
(26) Shoichet, B. K. *DUD − A Directory of Useful Decoys*. http://dud.docking.org/ (accessed Nov 20, 2007).

CI700358R