

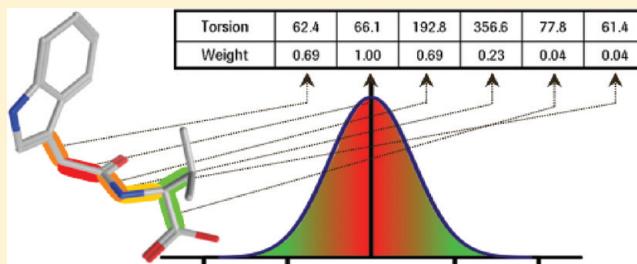
TFD: Torsion Fingerprints As a New Measure To Compare Small Molecule Conformations

Tanja Schulz-Gasch,[†] Christin Schärfer,[‡] Wolfgang Guba,[†] and Matthias Rarey^{*,‡}

[†]Pharmaceutical Research, Chemistry Technologies & Innovation, F. Hoffmann-La Roche Ltd., CH-4070 Basel, Switzerland, and

[‡]Center for Bioinformatics, University of Hamburg, Bundesstrasse 43, D-20146 Hamburg, Germany

ABSTRACT: Advantages like intuitive interpretation, objectivity, general applicability, and its easy, automated calculation make the rmsd (root-mean-squared deviation) the measure of choice for the investigation of the accuracy of conformational model generators. For comparing conformations of a single molecule this is a clearly superior method. Single molecule analysis is, however, a rare scenario. Typically, conformations are generated for huge corporate or external vendor databases of high diversity which are then further investigated with high-throughput computational methods like docking or pharmacophore searching, in virtual screening campaigns. Representative subsets for accuracy investigations of computational methods need to mimic this diversity. Averaged rmsd values over these data sets are frequently used to assess the accuracy of the methods. There are, however, significant weaknesses in rmsd comparisons for such kind of data sets. The interpretation is for example no longer intuitive because what can be expected in terms of good or bad rmsd values crucially depends on the data set composition like size or number of rotatable bonds of the underlying molecules. Further, rmsd lacks normalization which might result in very high averaged rmsd values for highly flexible molecules and thus might completely skew results. We have developed a novel measure to compare conformations of molecules called Torsion Fingerprint Deviation (TFD). It extracts, weights, and compares Torsion Fingerprints from a query molecule and generated conformations under consideration of acyclic bonds as well as ring systems. TFD is alignment-free and overcomes major limitations of rmsd while retaining its advantages.



INTRODUCTION

The root-mean-square deviation (rmsd) is frequently used to measure accuracy in very different areas of research like meteorology, economics, or life sciences like chemoinformatics, bioinformatics, etc. It aggregates differences (residuals) between predicted or estimated values and observed values in a single value to judge the predictive power of an underlying model. The present work challenges the suitability of the rmsd for the comparison of conformations from druglike, small molecules and presents a new measure to overcome major limitations of rmsd in this field of research.

Virtual (database) screening (VS)^{1,2} for hit identification is a common procedure in the drug design process of pharmaceutical or biotechnical companies. The (diverse) usage of rmsd within the VS process is state-of-the-art. VS methods like docking,^{3–5} shape matching,⁶ or pharmacophore searching⁷ strongly depend on high quality small molecule conformations. Conformations are either pregenerated^{8,9} (external conformational model) or sampled during the virtual screening process (internal conformational model).^{10–12} Structural deviations between the bioactive reference conformation, e.g. derived from a complex X-ray structure, and the computationally generated conformation are measured through relative rmsd values. This requires rigid-body translation and rotation of the generated conformation on top of the reference structure such that the distances between corresponding atom pairs are minimal. Thus,

with respect to rmsd, the virtual screening results can only be as good as the applied conformational model. Accuracy of VS results are measured through absolute rmsd values. The position and orientation of the small molecule is determined by the VS algorithm, and no further alignment-step is involved. Absolute rmsd thus includes a relative rmsd contribution from the conformational deviation plus a VS algorithm-dependent offset. Absolute rmsd in docking studies for example evaluates how accurate a given conformational model and the applied docking engine can reproduce an experimentally observed complex structure.

We want to strongly emphasize this difference between absolute and relative rmsd since this is not always done appropriately in literature. Often identical rmsd cutoffs are applied in publications on the evaluation of docking engines or conformation generators in order to discriminate between close or distant docking poses on the one hand or conformations on the other hand. We are not aware of a docking evaluation study that investigates to which extent relative rmsd contributes to and effects absolute rmsd and how much the offset between the two is.

rmsd for conformations or poses of single molecules can be calculated in order to assess the quality of conformation

Received: May 24, 2011

Published: June 6, 2012



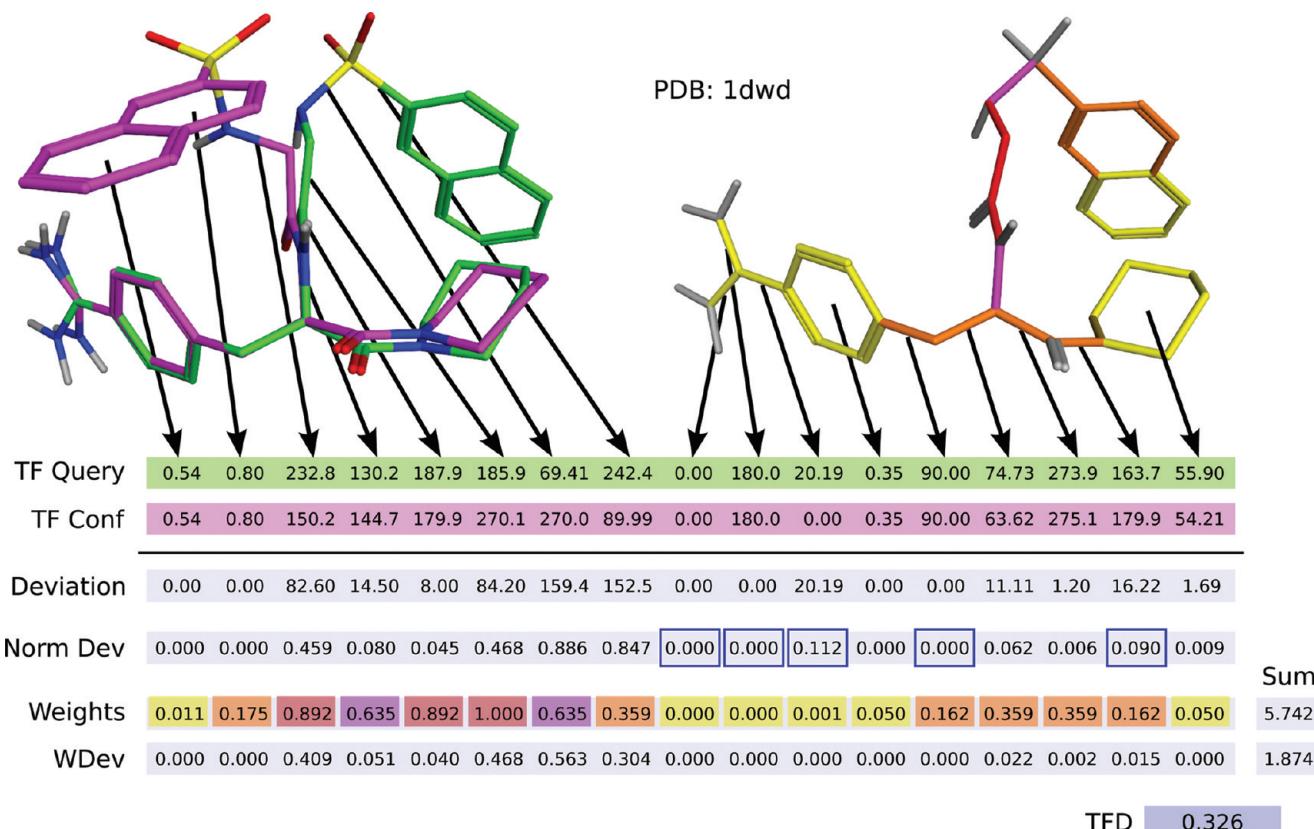


Figure 1. Example TFD calculation for a conformation of the ligand from PDB complex 1dwd: “TF Query” depicts the Torsion Fingerprint of the green query structure and “TF Conf” the Torsion Fingerprint of the pink conformation structure. Torsion angle values in each fingerprint are given in degrees. The structure on the right is colored according to the Gaussian weighting scheme. “Norm Dev” lists the normalized deviations, and “WDev” lists the weighted normalized deviations. The blue rectangles highlight cases in which symmetry is involved.

generators or VS tools toward a given target or a structural class.^{13–15} The information content of such an approach is however limited since it is not known if this is applicable for major or even minor structural modifications of the small molecules or observed flexibility in the binding site. rmsd is more frequently used in evaluation studies that try to mimic high-throughput virtual screening campaigns where conformations for molecules of huge and diverse corporate or external vendor databases are generated and are then postprocessed via for example docking or pharmacophore searching.⁷ In order to provide robust defaults and methods for VS toward diverse targets and small molecules the evaluation databases need to mimic this structural diversity. Averaged rmsd values over these data sets are used to assess the accuracy of the methods and significantly weaken the meaning of rmsd in structure comparisons because it is no longer intuitive to interpret the results. What can be expected in terms of good or bad rmsd values crucially depends on the data set composition like size or number of rotatable bonds. A further disadvantage of rmsd is its lack of normalization. For example highly flexible molecules might have very high averaged rmsd values which might completely skew results. On the other hand conformations or random placements of small and globular molecules can easily result in good rmsd values and lead to some kind of artificial “beautification” of VS results.

The computational chemistry community is aware of the above-described issues with rmsd for structure comparison.¹³ Several attempts to replace rmsd with improved methods were made, but so far no method is really established because

improvements often came along with the reduction of obvious advantages for rmsd calculations. CAD (contact area difference)¹⁶ and T-Analyst¹⁷ for example are dealing with conformations of protein structures only. IBAC (Interactions-Based Accuracy Classification)¹⁸ is analyzing the presence of key interactions between ligand and the protein and is thus limited to structure-based (ligand+protein) approaches. It can also not easily be automated. RSR (real space R-factor)¹⁹ tries to overcome the issue with model vs model comparison and instead uses the primary experimental data of electron density for the ligand and compare it to the expected, calculated density for the model pose. RSR is limited to structure-based approaches, and it is not normalized to a constant range, e.g. a pose with RSR of 1 is as good as the crystallographers refined model, whereas anything greater than 1 would fit less well. GARD (Generally Applicable Replacement of rmsd)²⁰ is another attempt to replace rmsd and would work for structure- and ligand-based methods. GARD calculates alignment scores between the atoms of a reference structure and the atoms of one of its conformations. The alignment score of two atoms is based on their distance and weighted according to the relative importance of the atom to binding. GARD is normalized to a value between 0.0 (worst pose) and 1.0 (best pose). The weighting scheme for functional groups was derived from regression analysis of commonly occurring functional groups and the free energy upon binding. Statistics are biased toward the data set composition, e.g. contributions of charged groups turned out to be much higher than for polar or unpolar groups. This of cause might be misleading if a charged group is a

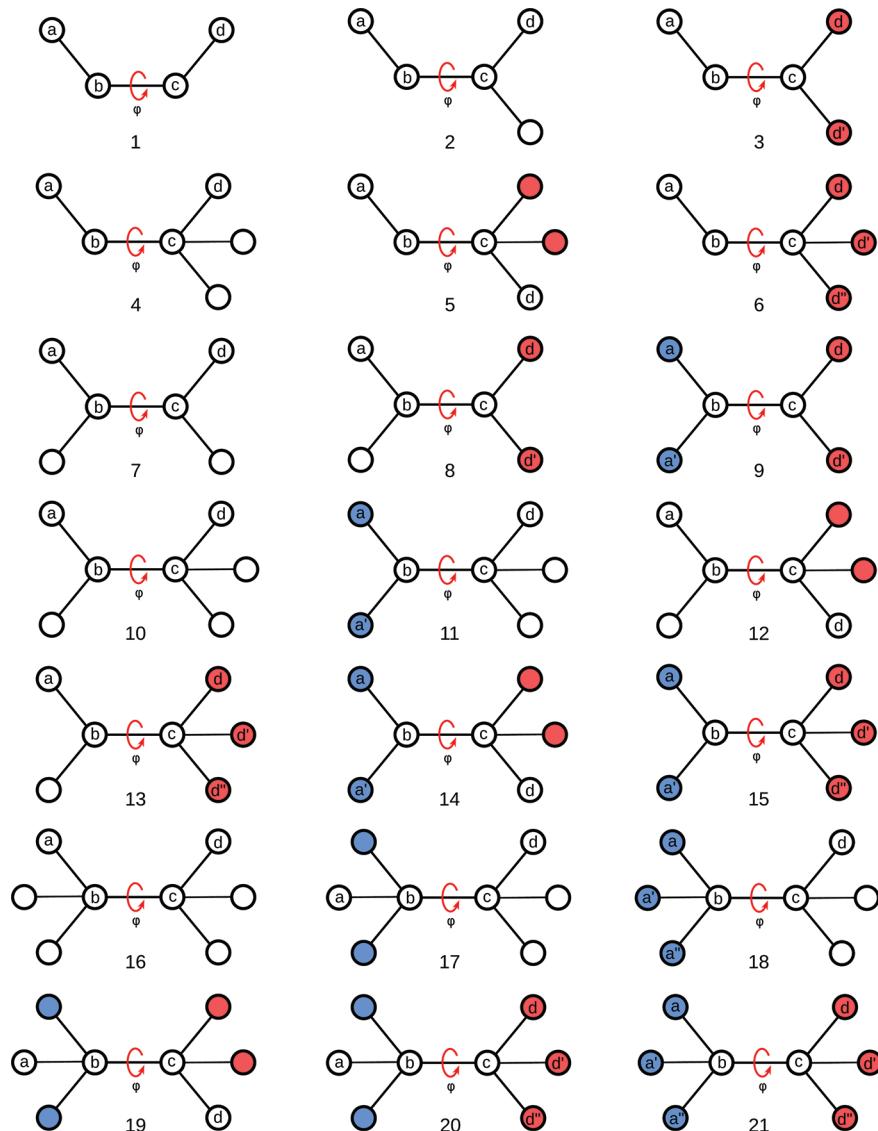


Figure 2. Case distinction in TF calculation: φ : torsion angle, white circles: atoms, red and blue circles: symmetric atoms. See Table 1 for the torsion angle calculation in each case.

solubilizing group of the ligand and not interacting with the protein. Interpretation is thus not always intuitive and – depending on the composition of the data set – might get highly biased.

In the present work we want to introduce a novel measure to compare different conformations of a molecule which can be used for example to assess the quality of conformation generators. It retains all the fundamental advantages of rmsd, and it is further normalized as well as applicable and informative for a broad range of data sets. Details of the method are described in the methods section of this publication. We will show examples of diverse molecules to demonstrate intuitivity of TFD and how it compares to relative rmsd. We have further selected a well-established conformational model (OMEGA 2.4.1)⁹ for conformation generation and a published data set²¹ to show the superior behavior of averaged TFD values vs averaged rmsd values. We will also show this for certain subsets of the data set and the dependence of the subset composition. We want to point out that our novel measure is not meant to replace absolute rmsd as a measure for docking accuracy.

MATERIALS AND METHODS

General. The Torsion Fingerprint (TF) of a molecule is a set of numeric values representing conformational states of all its bonds and rings. The TFD between a molecule in a reference conformation and one of its predicted conformations is calculated as follows: 1. The difference of the (pseudo-) torsion angle (torsion deviation) is calculated for each bond and ring system. 2. Each deviation is normalized to a number between 0 (no deviation) and 1 (maximal deviation). 3. A Gaussian weighting scheme is applied to the normalized deviations, assuring that deviations at topologically central bonds or rings have more influence on the TFD than deviations at terminal bonds or rings. 4. Finally the sum of the weighted deviations is normalized with respect to the Gaussian weights to obtain the TFD. An example TFD calculation is shown in Figure 1.

TF Calculation. The calculation of the Torsion Fingerprint of a molecule is done in two steps. In the first step the torsion angle of each bond is calculated, excluding ring bonds and bonds to terminal atoms. In this step, the reference atoms

Table 1. Case Distinction in TF Calculation^a

case	torsion angle calculation	maximal deviation
1	$T(a,b,c,d)$	180°
2	$T(a,b,c,d)$. Store d	180°
3	$\min\{T(a,b,c,d), T(a,b,c,d')\}$ if atom c is N trigonal pyramidal: $T(a,b,c,LP)$	90°
4	$T(a,b,c,d)$. Store d	180°
5	$T(a,b,c,d)$	180°
6	$\min\{T(a,b,c,d), T(a,b,c,d'), T(a,b,c,d'')\}$	60°
7	$T(a,b,c,d)$. Store a,d	180°
8	$\min\{T(a,b,c,d), T(a,b,c,d')\}$. Store a if atom c is N trigonal pyramidal: $T(a,b,c,LP)$. Store a	90°
9	$\min\{T(a,b,c,d), T(a,b,c,d'), T(a',b,c,d), T(a',b,c,d')\}$ if atom c is N trigonal pyramidal: $\min\{T(a,b,c,LP), T(a',b,c,LP)\}$	90°
10	$\min\{T(a,b,c,d), T(a',b,c,d)\}$. Store a,d	180°
11	$\min\{T(a,b,c,d), T(a',b,c,d)\}$. Store d if atom b is N trigonal pyramidal: $T(LP,b,c,LP)$. Store d	90°
12	$T(a,b,c,d)$. Store a	180°
13	$\min\{T(a,b,c,d), T(a,b,c,d'), T(a,b,c,d'')\}$. Store a	60°
14	$\min\{T(a,b,c,d), T(a',b,c,d)\}$ Atom b is N trigonal pyramidal: $T(LP,b,c,d)$	90°
15	$\min\{T(a,b,c,d), T(a,b,c,d'), T(a,b,c,d''), T(a',b,c,d), T(a',b,c,d'), T(a',b,c,d'')\}$ if atom b is N trigonal pyramidal: $\min\{T(LP,b,c,d), T(LP,b,c,d'), T(LP,b,c,d'')\}$	30°
16	$T(a,b,c,d)$. Store a, d	180°
17	$T(a,b,c,d)$. Store d	180°
18	$\min\{T(a,b,c,d), T(a',b,c,d), T(a'',b,c,d)\}$. Store d	60°
19	$T(a,b,c,d)$	180°
20	$\min\{T(a,b,c,d), T(a,b,c,d'), T(a,b,c,d''), T(a',b,c,d), T(a',b,c,d'), T(a',b,c,d''), T(a'',b,c,d), T(a'',b,c,d'), T(a'',b,c,d'')\}$	60°
21	$\min\{T(a,b,c,d), T(a,b,c,d'), T(a,b,c,d''), T(a',b,c,d), T(a',b,c,d'), T(a',b,c,d''), T(a'',b,c,d), T(a'',b,c,d'), T(a'',b,c,d'')\}$	60°

^aThe numbers for each case (column 1) refer to the illustrations in Figure 2. $T(a,b,c,d)$ denotes a torsion angle between atoms a, b, c , and d with b, c adjacent to the rotatable bond, a and d reference atoms. LP denotes lone pair. Maximal deviation means the maximal possible deviation for this torsion angle in two different conformations assuming ideal bond angle geometry.

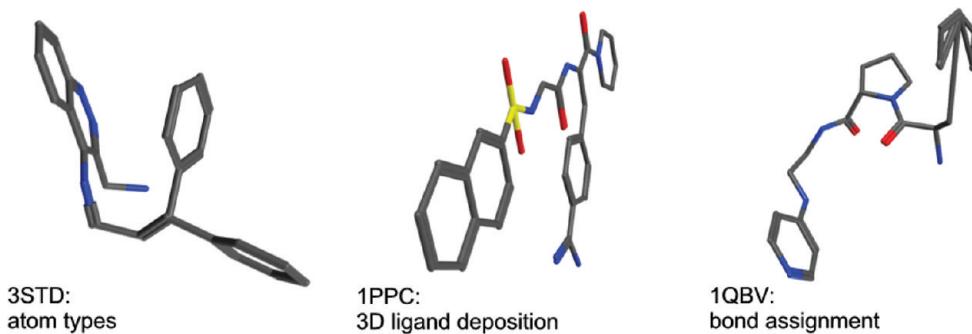


Figure 3. Three examples of ligands that were excluded from the data set because of errors in PDB structure deposition.

between which the torsion angle is measured have to be uniquely defined. Furthermore, molecular symmetry has to be taken into account. To address these issues for each bond, we count the atoms adjacent to those connected by the bond. The easiest case is the one in which there is only one atom on each side (see case 1 in Figure 2 and Table 1). Then we can use these two atoms as a reference for the bond to calculate the torsion angle. As long as no symmetric substituents occur, the torsion angle can be computed this way picking two arbitrary adjacent atoms for reference. Note that – assuming the bond angles are constant – the torsion angle difference is independent from the selected reference atoms. Ignoring rotation direction, the maximum torsion angle difference is 180 degrees. As soon as symmetry is involved, various cases

have to be distinguished. If two symmetric atoms are connected with planar-trigonal geometry to an atom of the bond considered (see case 3 in Figure 2 and Table 1), we select the atom resulting in the minimal torsion angle as reference. In this case the maximal possible deviation is 90 degrees. All further cases as well as the scenario of two atoms connected in pyramidal geometry are summarized in Figure 2 and Table 1. In some cases there are a lot of extra torsion angle calculations. Case 21 in Figure 2 and Table 1 for example uses 9 calculations. If atoms b and c in this case have a perfect tetrahedral geometry, the 9 calculations can be reduced to 3. However, as a perfect geometry is not always given in molecules we included the extra calculations to make sure the torsion angles are always

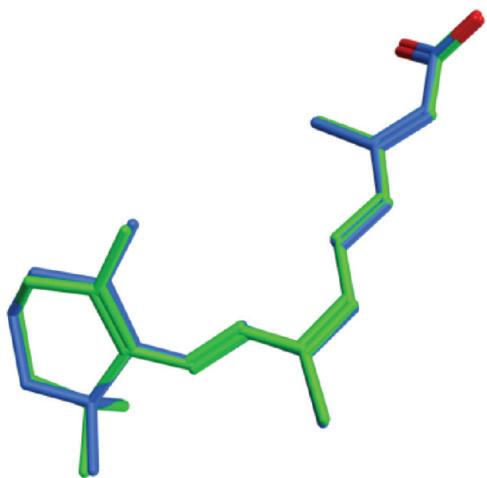


Figure 4. Ligands from PDB codes 1fm6 and 1fm9. Relative rmsd between both is $<0.2 \text{ \AA}$.

calculated properly. In case 21 this means that we need all 9 values for the correct torsion angle calculation.

In order to also take ring conformations into account, a compatible measure has to be developed for rings. The set of rings considered here is defined to be the set of relevant cycles as calculated with the method of Vismara.²² For each ring the torsion angles of all its ring bonds are calculated, using

neighboring atoms belonging to this ring as a reference. The sum of the absolute values of the torsion angles divided by the ring size is taken as a torsion angle equivalent for the whole ring. The last step of the ring torsion calculation is to decide about its maximal deviation. We used the following Gaussian function for rings of size 3 to 14

$$\maxDev(\text{ring size}) = 180e^{-0.025(\text{ring size}-14)^2} \quad (1)$$

The maximal deviation for rings of size 14 or greater is 180 degrees.

Topological Weighting. The weighting assures that deviations at topologically central bonds or rings have more influence on the final TFD than deviations at terminal bonds or rings. The first step is to identify the central bond of a molecule. We start by calculating the shortest paths $\delta(b_1, b_2)$ between all pairs of bonds (b_1, b_2) (excluding hydrogen atoms) using the Floyd-Warshall algorithm.²³ The central bond c is defined as the bond with minimal standard deviation over all shortest paths to the other bonds. A Gaussian function is used to transfer distance values into normalized weights

$$w(b) = e^{-\beta(\delta(b, c))^2} \quad (2)$$

After inspection of initial test cases we decided that a bond half the way from the central bond to the farthest atom of a molecule should get 10% of the maximal weight. We therefore

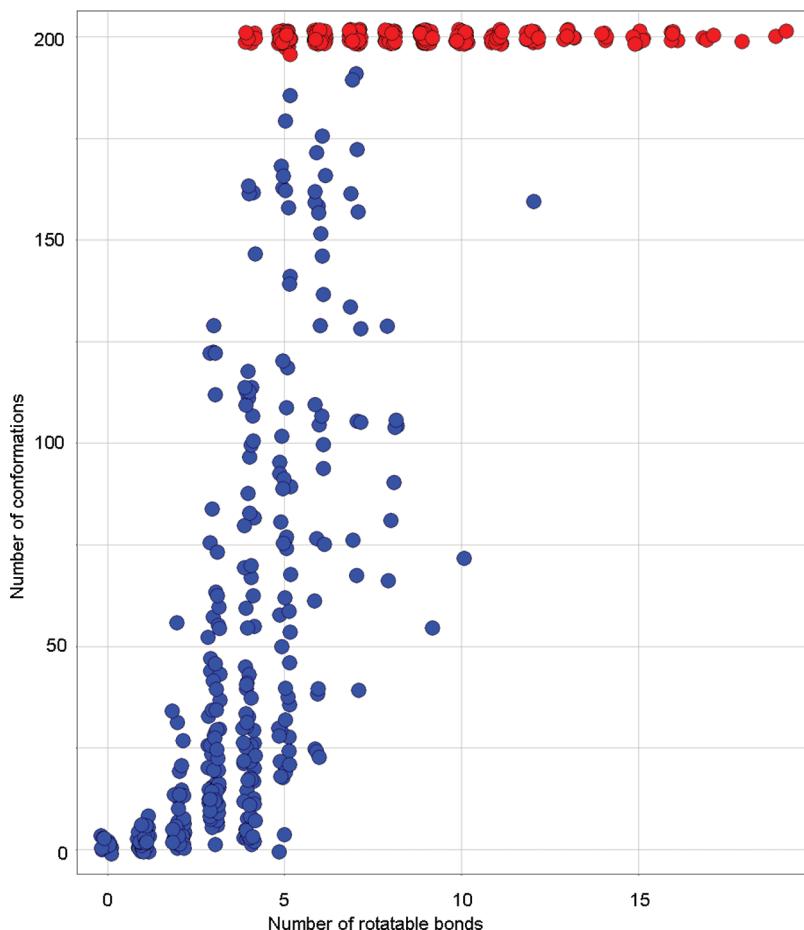


Figure 5. Dependence of conformational sampling from number of rotatable bonds. Some degree of jittering has been applied to improve the visibility of population. Red compounds have reached the maximum number of conformations which by default is limited to 200.

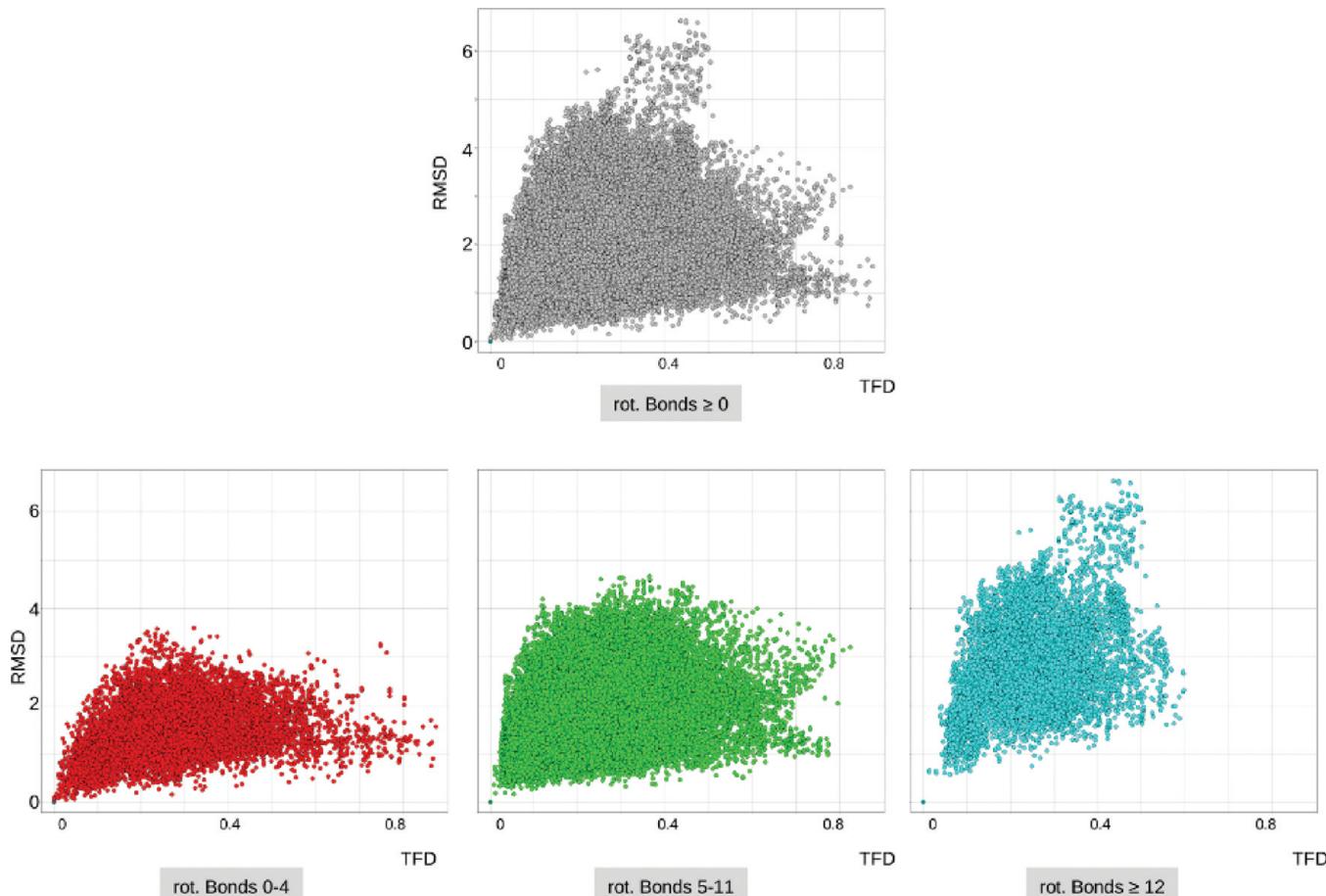


Figure 6. TFD and relative rmsd for all generated conformations of the data set are hardly correlated, $R^2 = 0.1$ (top). There are however trends in the data if subsets depending on the number of rotatable bonds are defined. Less flexible molecules (left) have a less steep slope than the more flexible molecules.

adjusted β such that $w((\delta_{max})/2) = 0.1$. δ_{max} denotes the length of the longest shortest path.

On condition that a ring torsion counts as much as ringsize/2 acyclic bonds the weight for a ring is calculated by dividing the sum of the normalized weights of the ring bonds by 2.

TFD Calculation. To calculate the TFD between two TFDs we start by calculating the absolute deviation between each torsion angle pair, measured in degrees. The deviations are then divided by 180 degrees (the maximal possible deviation) to normalize them to a number between 0 (no deviation) and 1 (maximal deviation). In cases in which symmetry is involved, the maximal normalized deviation of 1 is not reached, e.g. if the maximal possible deviation is 90 degrees the maximal possible normalized deviation is 0.5. For the ring torsions the absolute deviations are normalized by dividing them by their maximal possible deviation ($maxDev$) as calculated by eq 1. Each normalized deviation is then multiplied by its corresponding Gaussian weight. The weighted deviations are summed up and divided by the overall sum of the weights to obtain the final TFD.

TFD values range from 0.0 (no deviation) to 1.0 (maximal deviation). If two conformations are compared and none of their torsion angles and ring torsions deviate, the TFD value will be 0.0. By contrast the TFD value will be 1.0 if for every torsion angle and ring torsions the deviation is maximal. A TFD value of 0.5 could mean several things, for example that there is a maximal deviation of 180 degrees at the central bond or that

the torsions angles of two relatively central bonds deviate by 90 degrees. A TFD value of around 0.3 could indicate that there is a maximal deviation at a terminal bond of a molecule or that there are smaller deviations at several rotatable bonds of a molecule.

Data Set. In order to be unbiased toward the underlying work on TFD we used a data set from literature.²¹ The data set contains structures from three different publications^{21,24,25} and is described as diverse in terms of target and ligand classes. The 667 ligands were extracted from the PDB²⁶ and inspected with MOE.⁷ 40 ligands show unreasonably high ligand strain mainly because of errors in PDB structure deposition and were excluded from the final data set. Three examples of “strained” molecules are shown in Figure 3.

Although it is claimed that the data set is diverse toward targets and ligand classes, there is redundancy in there, e.g. there are more than 20 couples and triples of identical SMILES (1ghw, 1ghx, 1ghy), some for the same target and no significant change in conformation (see for example Figure 4). Those structures were kept for consistency.

Conformation Generation. Conformations were generated with OMEGA 2.4.1 and default settings. OMEGA failed to generate conformations for 23 ligands which also were excluded from the final data set (604 molecules remaining). For approximately 50% of the compounds the limit for the maximum number of conformations is reached, and conformational sampling is thus not fully exhaustive (Figure 5). Figure 5

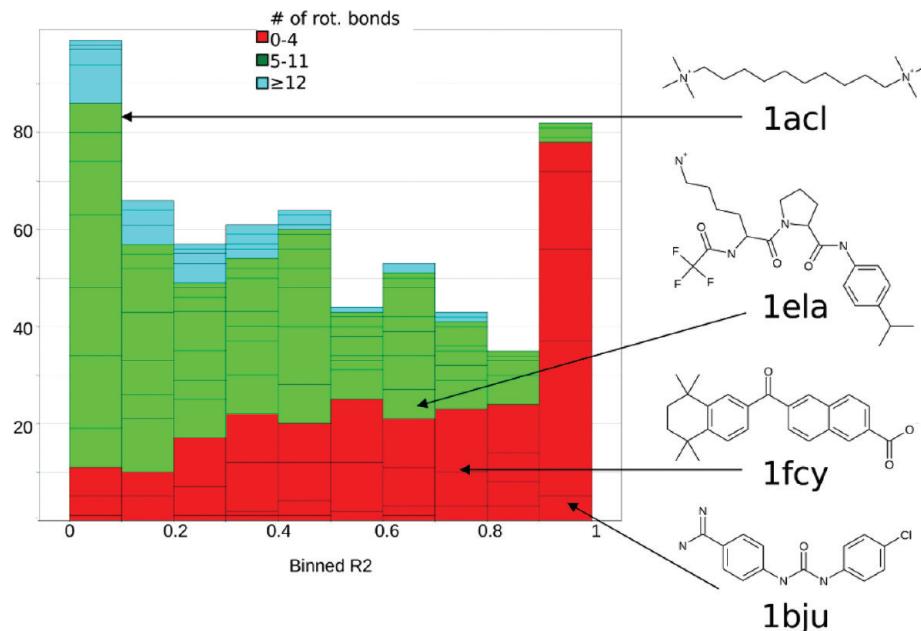


Figure 7. Binned R^2 for conformations of single molecules from the data set. Color-coding is based on the number of rotatable bonds. Molecules with few rotatable bonds have a higher tendency to be correlated with rmsd, while the correlation for more flexible molecules is reduced.

shows the dependence of conformational sampling from the number of rotatable bonds. For some molecules the limit in the generated number of conformations is already reached with 4 rotatable bonds. In total, approximately 71,500 conformations were generated for the 604 molecules and further analyzed. It should be emphasized, that this work is not about judging or optimizing the conformation generator but about disclosing effects on rmsd calculations and their suitability in evaluation studies of conformational models. For every molecule that is not exhaustively sampled the likelihood increases that the conformation which would be closest to the bioactive one is rejected resulting in an increased average relative rmsd.

Data Analysis. All data preparation, structure analysis, compound superimposition, descriptor, and rmsd calculations have been performed with MOE 2010.10.

ROCS Shape- and Color-Tanimoto As a Measure To Compare Conformations. ROCS (Rapid Overlay of Chemical Structures)⁶ is a commercial application from Openeye for fast shape comparisons to support the Virtual Screening process. It is based on the idea that molecules have similar shape if their volumes overlay well, and volume mismatch can be considered as a measure of dissimilarity. Besides shape matching ROCS includes a chemical force field. This can be used to measure chemical complementarity and to refine shape-based superpositions based on chemical similarity. Scoring can be either based on shape (TanimotoShape), color (TanimotoColor), or a combination of the two (TanimotoCombo). Scores are normalized to 1 for TanimotoShape and TanimotoColor and 2 for TanimotoCombo respectively. Higher Tanimoto values indicate better matching overlays. ROCS is also suited for comparing conformations of a molecule. In the present work we are using ROCS scores (ROCS 3.1.0) to confirm the superior behavior of TFD compared to rmsd. For in-house Virtual Screening campaigns we usually define a cutoff for TanimotoCombo of 1.5. Contributions of TanimotoShape and TanimotoColor should both be >0.6 in order to ensure a sufficient degree of shape and color complementarity to the query. Since TFD does not

contain a TanimotoColor counterpart in order to account for chemical similarity we are comparing TFD results to TanimotoShape where usually a cutoff of >0.75 is applied.

RESULTS AND DISCUSSION

TFD Evaluation Strategy. TFD was designed to overcome major drawbacks and limitations of rmsd for the analysis of

Table 2. Comparison of Normalized TFD Values against RMSD Values for Three Differently Sized Molecules: Small, Medium, and Large^a

molecule	torsion fingerprint	TFD	rmsd
small 1	(60.00 180.00 60.00)	0	0
medium 1	(60.00 180.00 180.00 180.00 60.00)	0	0
large 1	(60.00 180.00 180.00 180.00 180.00 180.00 60.00)	0	0
small 2	(60.00 270.00 60.00)	0.50	0.54
medium 2	(60.00 270.00 90.00 90.00 60.00)	0.50	1.02
large 2	(60.00 270.00 270.00 270.00 270.00 270.00 60.00)	0.50	0.78
small 3	(60.00 0.00 60.00)	0.99	0.92
medium 3	(60.00 0.00 0.00 0.00 60.00)	0.99	2.05
large 3	(60.00 0.00 0.00 0.00 0.00 60.00)	0.99	3.08

^aTFD and RMSD values are calculated using the first molecule as a reference for each group.

conformation ensembles. In the present evaluation study we thus calculate TFD and rmsd values between computationally generated conformations and the corresponding bioactive conformation. In our study we applied a cutoff of 0.2 for TFD. Conformations with TFD < 0.2 typically are close to the bioactive conformation and are therefore likely to result in a hit in later VS campaigns. The assessment and comparison of the different measures over a whole data set as well as for selected specific examples should demonstrate the superior behavior of TFD. The specific examples should also demonstrate that TFD mimics how an experienced user would classify conformations

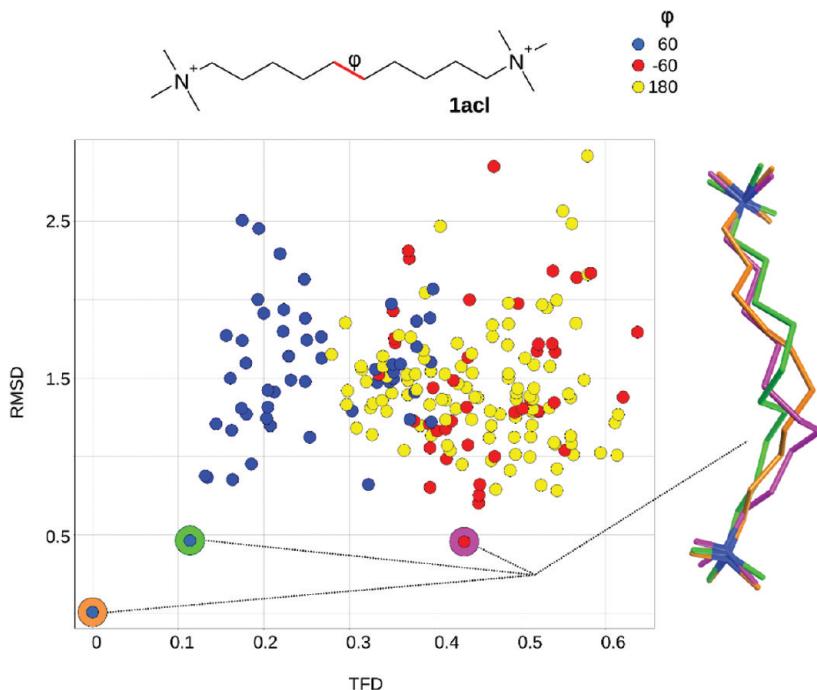


Figure 8. Behavior of TFD and rmsd for conformations of the ligand of PDB complex 1acl. Color-coding of spheres is by the TFD contribution of the central (highlighted in red) rotatable bond. $R^2 = 0.00$.

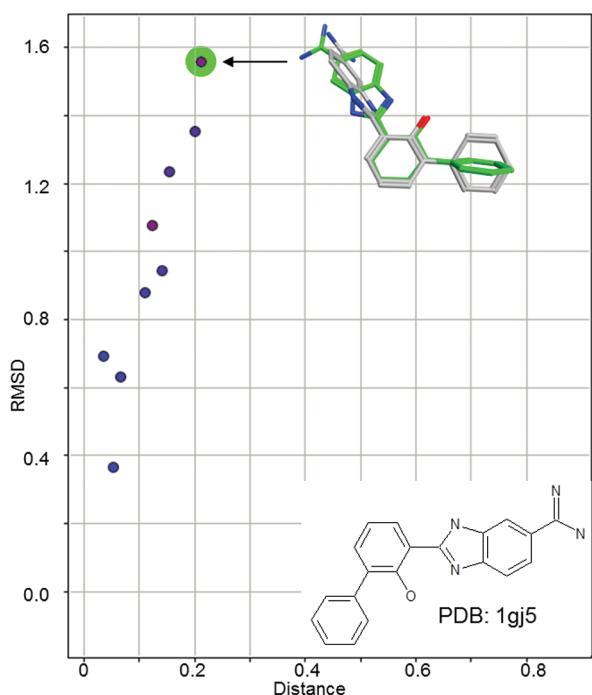


Figure 9. Behavior of TFD and rmsd for conformations of the ligand from PDB complex 1gj5. Color-coding of spheres is by TanimotoShape (1 - blue, 0 - red). A selected overlay with the bioactive conformation is shown. $R^2 = 0.84$.

as suitable or unsuitable for their later usage in VS approaches. They should also exemplify that for the chosen TFD cutoff conformations typically have no major steric violations or mismatching interaction patterns when compared to the bioactive conformation and that these are thus valid and crucial conformations within a conformational ensemble for the successful application in later stage VS approaches. In order to

have another independent measure of conformational fit we also compare TFD results to TanimotoShape obtained with the commercial software ROCS.

General Behavior of TFD. Relative rmsd and TFD for the overall generated conformations within the data set are not correlated ($R^2 = 0.1$). As it can be seen from Figure 6 conformations with a very low TFD (<0.2) exceed rmsd values of 4 Å and conformations with a very high TFD (>0.8) might have very good rmsd values below 1 Å. A high level of correlation between the two different measures is not necessarily desired because this would only reflect redundancy of the different methods. The bottom scatterplots in Figure 6 show however that there are trends compared to the fully random behavior of the top scatterplot if the data set is clustered in three sets depending on the number of rotatable bonds. For the less flexible molecules with <5 rotatable bonds a trendline intercepting at 0 has a slope of 4.6, for compounds with a medium number of rotatable bonds (5–11) the slope increases to 7.9 and for compounds with ≥ 12 rotatable bonds the slope reaches a maximum of 10.7. The scatterplots in Figure 6, especially the bottom right scatterplot for the subset with ≥ 12 rotatable bonds, also show that TFD values do not necessarily reach the maximum value of 1. This is because the conformation generation is knowledge-based, and some conformations with maximal torsional deviation will not be generated because they would be conformationally unfavorable (a specific example on this will be discussed in the following section). Other reasons why certain conformations are not observed could be filtering because of internal steric strain or clustering of similar conformations to increase the diversity of conformations in the final ensemble. An averaged R^2 for conformations of single entries rather than for the whole data set increases R^2 to 0.5. As it can be seen from the R^2 histogram in Figure 7 which is color-coded by sets defined by the number of rotatable bonds in a molecule the level of correlation increases for the less flexible molecules. However, like for

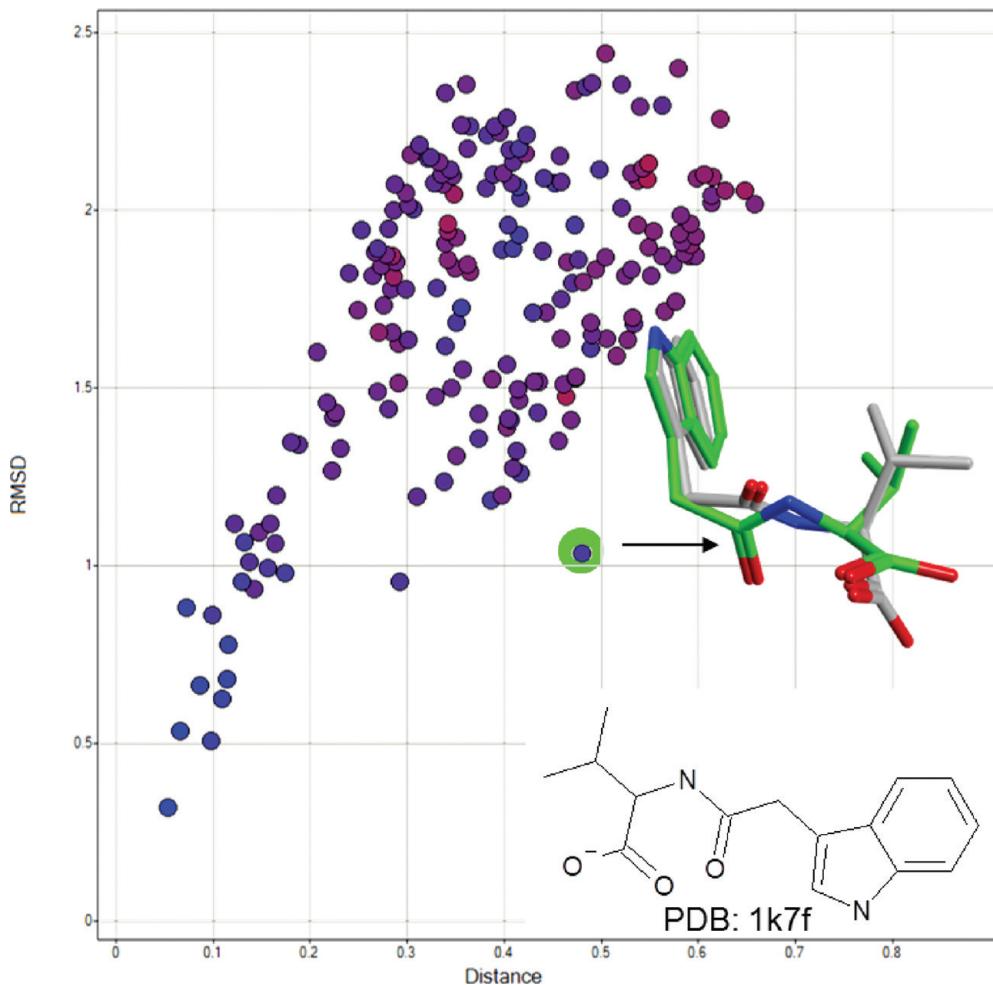


Figure 10. Behavior of TFD and rmsd for conformations of the ligand from PDB complex 1k7f. Color-coding of spheres is by TanimotoShape (1 - blue, 0 - red). A selected overlay with the bioactive conformation is shown. $R^2 = 0.38$.

Figure 6 it is necessary to look at specific, chemically diverse examples in order to be able to interpret and judge the quality and suitability of TFD. This detailed analysis of single molecules, especially the extreme cases with strongly oppositional behavior of the two measures, is shown in the following section.

Normalization of TFD. To show that — in contrast to rmsd values — normalized TFD values are comparable between diverse molecules we generated 3 carbon chains of variable length: small (3 bonds), medium (5 bonds), and large (7 bonds). The two terminal bonds are always considered as nonrotatable, while all other bonds are considered as rotatable. We generated 3 conformations for each carbon chain so that the rotatable bonds of the second conformations deviate by half the maximal possible deviation, and the rotatable bonds of the third conformations deviate maximally from the rotatable bonds of the first conformations. TFD and rmsd values are then calculated for each group by using the first conformation as a reference. Table 2 shows the Torsion Fingerprints, TFD, and rmsd for each group. TFD values for all second conformations are the same, whereas the rmsd varies between the three differently sized molecules. For the third conformations there is an even larger rmsd variation than for the second conformations, ranging from about 0.9 Å up to over 3 Å, whereas the TFD again stays the same.

TFD Behavior for Structurally Diverse Molecules.

Figures 8 to 11 show the behavior of TFD vs rmsd for four diverse molecules from the data set. The ligand of PDB complex 1acl shown in Figure 8 is highly flexible. It contains two charged terminal ammonium substituents linked by an aliphatic -decyl (C10) chain. For this example there is no correlation between rmsd and TFD ($R^2 = 0.0$). The central torsion of the 1acl ligand is 67 degrees (C—C—C—C). During the knowledge-based conformation generation with OMEGA the matching torsion rule from the torsion library for this SMARTS pattern is the generic pattern sp₃-sp₃ (rotatable bond between two sp₃ centers). For this pattern torsions of +60, -60, and 180 degrees are adjusted. When coloring the entries in Figure 8 by the TFD contribution of this central rotatable bond (highlighted in red) one can see the three different subsets of ligand conformations of 1acl. The blue subset is closest to the bioactive conformation; it has a central torsion of 60 degrees. The small deviation of only 7 degrees from the bioactive conformation results in a small TFD contribution (0.05 before weighting). The yellow and orange subsets are the more distant orientations with -60 (deviation 127 degrees) and 180 degrees (deviation 113 degrees). Their unweighted contributions to TFD are 0.62 and 0.71, respectively. Since we are applying a Gaussian weighting scheme, a violation of this central torsion is largely contributing, and those conformations do not end up below a TFD cutoff of 0.2. There is also an

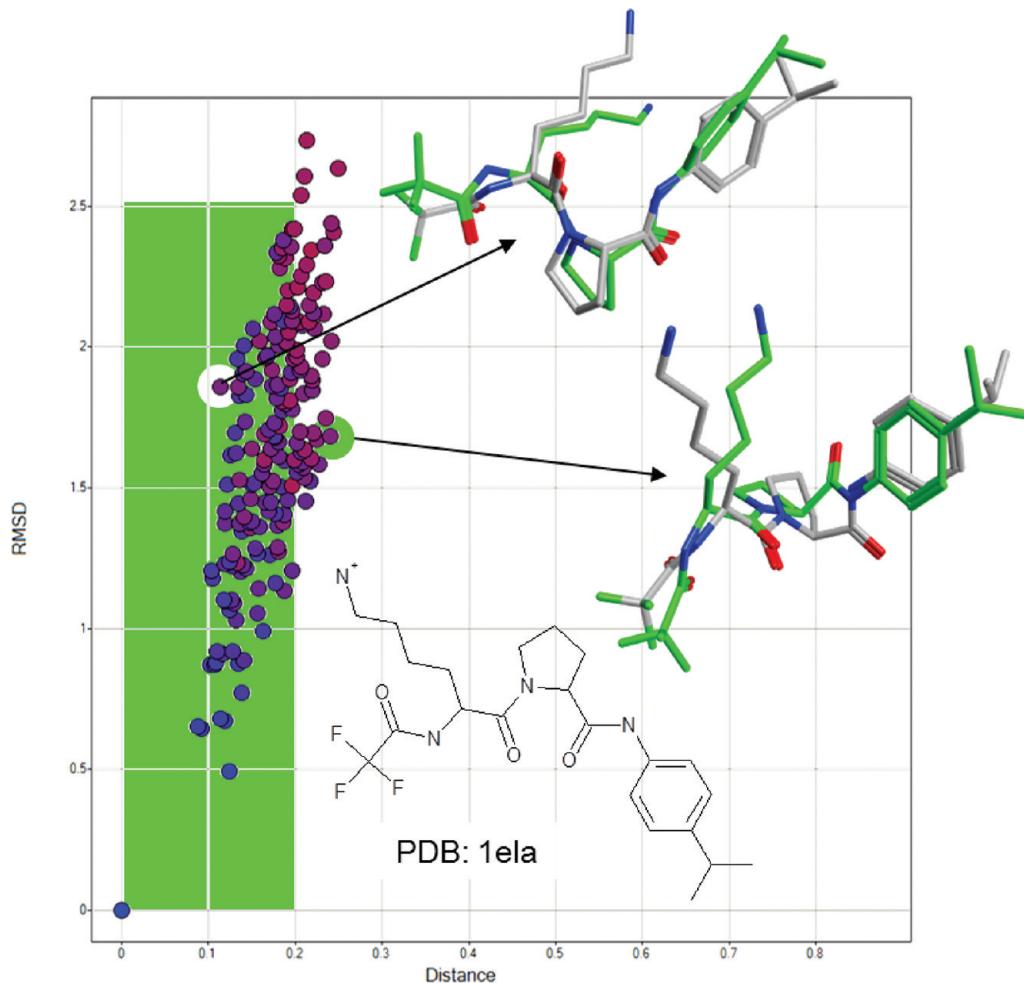


Figure 11. Behavior of TFD and rmsd for conformations of the ligand from PDB complex 1ela. Color-coding of spheres is by TanimotoShape (1 - blue, 0 - red). Two selected overlays with the bioactive conformation are shown. $R^2 = 0.68$.

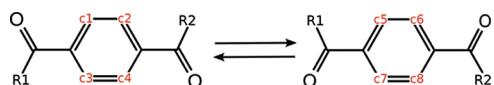


Figure 12. Because only one mapping of atoms between these two conformations was used, the TFD will be close to 1 instead of 0.

overlay shown of the conformation with the best TFD (green) and the bioactive conformation (orange). Since the charged components of the molecule do nicely match and the part in between links them properly, this is a valid conformation in a conformational ensemble. Also overlaid is a conformation with a similar rmsd but a TFD > 0.4 (magenta). As it can be seen the overlay of this conformation is just as good as the green structure and as it is indicated by the very similar rmsd values. rmsd is outperforming TFD in this example. Because of the atomic contributions rmsd is able to capture if one torsion deviation is balanced by one or more other torsion deviations in a molecule that bring substituents back into the position of the bioactive conformation, whereas the torsion contributions in TFD are contributing independently and any kind of communication between torsions is not taken into consideration.

The ligand of PDB complex 1gj5 shown in Figure 9 belongs to the compound class that is compact and rmsd values never exceed 2 Å. Default settings of the conformation generator produce 9 conformations for the ligand of 1gj5. Relative rmsd

nicely correlates with TFD ($R^2 = 0.84$). A relative rmsd cutoff of 2 Å for conformations classified as “close to the bioactive conformation” has often been applied^{15,21} in publications. It has however also been realized that one has to discriminate between the applied cutoff for absolute and relative rmsd. Thus, in some publications the relative rmsd cutoff has been reduced to 1.5 Å or even below.²⁷ A < 2 Å cutoff for 1gj5 classifies all conformations as “close” conformations. Even a more stringent cutoff of 1.5 Å still classifies 8 out of 9 conformations as “close” conformations. With our experience-based TFD cutoff of < 0.2 for “close” conformations we also found 8 “close” conformations. One conformation slightly exceeds this cutoff of < 0.2 and is shown in Figure 9 (gray) superimposed on the bioactive conformation taken from the X-ray complex structure (green). If one considers the pi-interactions of the terminal substituents as well as the directed interactions via hydrogen bonding of the amidine group as crucial for molecular recognition it makes sense that this conformation is neither classified by rmsd nor TFD as “close” to the bioactive conformation.

In contrast to the ligand of PDB complex 1gj5 from Figure 9 the ligand of PDB complex 1k7f (Figure 10) is more flexible such that the maximum number of generated conformations is reached, and rmsd values span a range from below 0.5 up to 2.5 Å. Like for the previous example rmsd and TFD are correlated ($R^2 = 0.38$). The level of correlation is higher in the region below the usually applied cutoffs for both measures. A rmsd

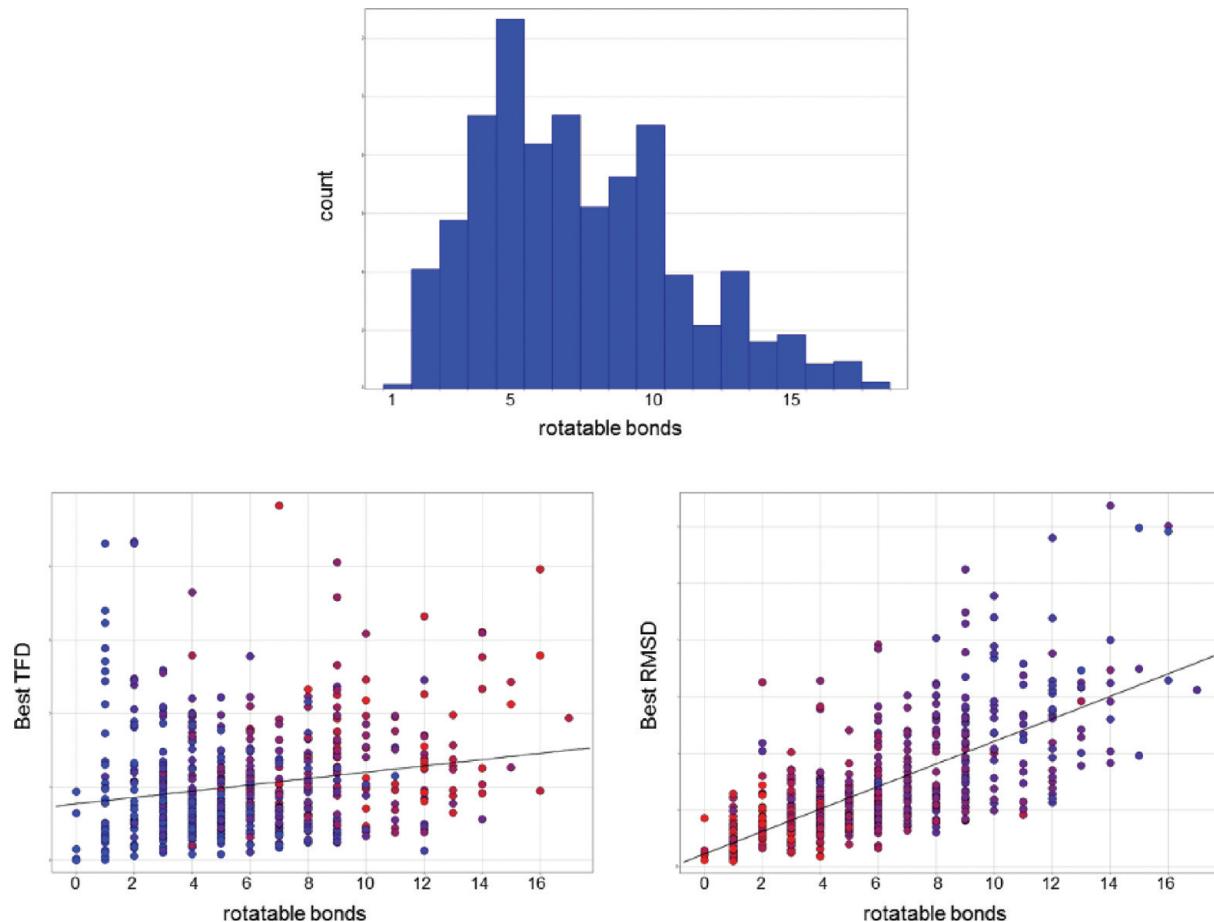


Figure 13. Top - Distribution of rotatable bonds in the data set. The two bottom scatterplots show best TFD (bottom left) and best rmsd (bottom right) values for every entry in the data set against the number of rotatable bonds. TFD is not correlated ($R^2 = 0.04$); rmsd ($R^2 = 0.47$) in contrast is highly correlated. Color-coding is by the number of heavy atoms to show the strong correlation between the number of rotatable bonds and heavy atoms.

cutoff of <2 Å classifies 136 conformations as close to the bioactive conformation; for a cutoff of <1.5 Å there are still 50 close conformations. In contrast a TFD cutoff of 0.2 classifies fewer, only 22 conformations, as close. In Figure 10 an overlay of a molecule with a low rmsd (1 Å) but a high TFD (0.5) to the bioactive conformation (green) is shown. The central amide group is oriented such that the hydrogen bonding directions substantially differ from those of the bioactive conformation. There is nevertheless a good shape match which is reflected in a TanimotoShape of 0.83. Overall, it is very unlikely that the molecule would bind to its target tryptophan synthase in this conformation. The other way round however there is no conformation with rmsd >1.5 Å that TFD would classify as close.

As already mentioned before the ligand of PDB complex 1k7f can also be used to demonstrate why the maximum TFD of 1 is not reached. The reason for this is the knowledge-based conformation generation that prevents conformationally unfavorable structures to be generated. The ligand of PDB complex 1k7f has a central, high-weighted amide bond. The maximum deviation of 180 degrees would convert the favorable trans amide to a cis amide. Since this rule is captured in the underlying knowledge base of the conformation generator the maximal TFD contribution of this bond and thus the maximal TFD of 1 is not observed in this study.

This section is completed by showing the TFD behavior of the highly flexible ligand of PDB complex 1ela (Figure 11). Like for the previous example the maximum allowed number of conformations has been generated. rmsd and TFD are again highly correlated ($R^2 = 0.68$). In Figure 11 two different alignments with the bioactive conformation (green) are shown. Both conformations are in a very similar rmsd range of acceptable conformations but clearly separated through TFD, one in the accepted the other in the rejected TFD area. Visual inspection shows that for the example on the bottom right (conformation distant from the bioactive conformation) the central amide group is oriented in a way that directionality of interactions are not in line with the ones from the bioactive conformation. Since this is a central torsion the deviation has a high weight and a strong contribution to TFD. On the other hand the overlay on the top right shows that the amide orientation and thus the central torsion is correct. The interaction pattern could properly be formed within the binding site. Coarse minimization of the conformation within the binding site however shows a correct interaction pattern and confirms the TFD classification.

TFD Limitations. As it has already been shown in Figure 8 application of TFD is limited for highly flexible molecules with many adjacent, only generic torsion patterns, for example sp₃-sp₃. The high contribution of deviating, central torsions and the lacking consideration of rebalancing effects of other torsions

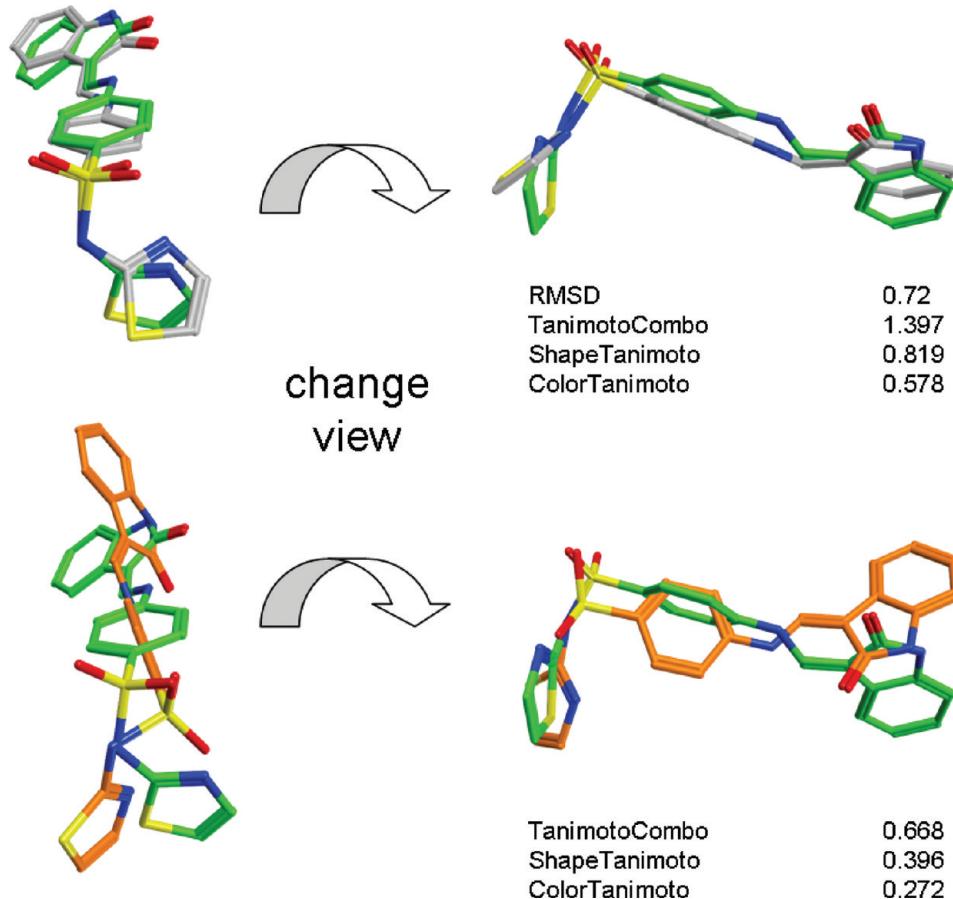


Figure 14. An example from PDB entry 1ke8 to demonstrate that ROCS-optimize alignments do not always achieve maximal Shape or Color match. Top - rmsd-based alignment. Bottom - ROCS-based alignment. The bioactive conformation is shown in green. In this case the rmsd-optimized alignment has the higher Tanimoto scores.

lets TFD easily exceed the usually applied cutoff. However, if one of the central carbons in Figure 8 is replaced for example by an interacting amide and a torsion deviation leads to a totally different interaction direction of this amide a high TFD would be correct. To decide which atoms or functional groups of a small molecule are important is not possible without the context of a protein or a set of known active molecules. Since we developed the TFD to be an independent measure, we therefore did not include functional groups into the calculations.

The TFD does not need an alignment step in the form of rigid body superposition like the relative rmsd, but it needs some kind of atom mapping to ensure comparison of the correct torsion angles. The mapping in our implementation is done by calculating unique SMILES for each conformation. As this results in only one mapping per conformation this could lead to the problem described here. If one looks at the two molecules in Figure 12 one would expect a TFD of 0. Actually the TFD is near 1 because the mapping would assign atom c1 to c5, c2 to c6, c3 to c7, and c4 to c8 resulting in a maximal deviation of 180 degrees for both torsion angles. A solution to this problem could be to generate all possible atom mappings, calculate the corresponding TFDs, and then use the minimal TFD. For the example in Figure 12 a second mapping would assign atom c1 to c7, c2 to c8, c3 to c5, and c4 to c6 resulting in a TFD of 0.

Influence on Data Set Composition on Comparison Measures.

The top left part of Figure 13 illustrates the

composition of the data set with respect to the number of rotatable bonds, e.g. 80% of compounds in the data set have fewer than 8 rotatable bonds. The other two plots show the best average TFD and rmsd, respectively, for every rotatable bond bin. It can clearly be seen that average best rmsd values tend to be better for compounds with fewer rotatable bonds whereas there is no clear dependency of the number of rotatable bonds on the average best TFD. This still holds when the data are binned by the number of heavy atoms instead of rotatable bonds (correlation between number of heavy atoms and number of rotatable bonds $R^2 = 0.6$). The graphic also demonstrates that a cutoff of 2 Å, which is sometimes used²¹ for judging the quality of conformation generators, is not really challenging and that the average best rmsd can easily be tuned by adding molecules with few or removing molecules with many rotatable bonds. Observed differences for different conformation generators do thus also not inform about the quality of the conformation generator unless the same data set was used which is however hardly the case in literature.

TFD Does Not Need an Alignment Step. Prior to calculation of relative rmsd or ROCS scores an alignment step is required. An optimal rigid body superposition should end in a best global match and result in a minimal relative rmsd or maximal ROCS score, respectively. Although ROCS is using a smooth Gaussian function to represent the molecular volume,²⁸ an optimal alignment cannot always be guaranteed. In our test set, in about 5.9% of the test cases, the rmsd-based alignment shows a better TanimotoCombo score. An example for which

this is the case is shown in Figure 14. Tanimoto scores for the ROCS-based alignment (bottom) significantly drop compared to the rmsd-based alignment (top). This demonstrates the complexity of the calculation of similarity based on volumetric data representations.

SUMMARY AND CONCLUSIONS

In the present work we showed multiple examples demonstrating the limited usage of rmsd for structure comparison, especially when rmsd is averaged and applied on data sets to assess the quality of conformational models. We have introduced a new measure for comparing conformations of small molecules called Torsion Fingerprint Deviation and showed that it overcomes major disadvantages of rmsd. ROCS Tanimoto scores are a further established measure to compare conformations of small molecules. We also compared TFD to the ROCS Tanimoto scores. However, since ROCS is commercial software and cannot easily be reimplemented, we do not classify it as generally applicable. In contrast to ROCS, extraction, weighting, and comparison of Torsion Fingerprints between a query molecule and one of its conformations is easy to reimplement and thus as generally applicable as rmsd. The calculation of TFD values is, like the calculation of rmsd values, fully automated and thus also ensures full objectivity. TFD values for single molecules are often highly correlated with rmsd values which makes interpretation of TFD straightforward for rmsd accustomed users. In addition, since TFD is normalized to 1 it is easier to define a single cutoff for conformations of diverse molecules to classify whether they are close to the bioactive conformation or not. Due to its normalization averaged TFD values for evaluation data sets of conformational model generators can still be interpreted intuitively. It is also important to mention that TFD does not depend on an alignment step to identify the best global match. rmsd as well as ROCS scores require an optimal rigid body superposition (alignment step) in order to derive the best global match. We showed for ROCS that this best global match is not always guaranteed by the underlying methods, and the alignment independence of TFD is a clear advantage. To even better mimic the way an experienced modeler would classify conformations the TFD calculation could be extended by an adjustable weighting scheme to account for differences like size or symmetry in terminal rotors. Another improvement would be to implement some kind of awareness for deviations in neighboring torsion angles to account for the problems we showed in the 1acl example. We are convinced that the usage of TFD to compare different conformational model generators even with different evaluation data sets will give more meaningful and comparable results than with averaged relative rmsd values.

AUTHOR INFORMATION

Corresponding Author

*E-mail: rarey@zbh.uni-hamburg.de.

Notes

The authors declare no competing financial interest.

ACKNOWLEDGMENTS

T.S.G. thanks Jérôme Hert for supporting the data analysis.

REFERENCES

- (1) Lengauer, T.; Lemmen, C.; Rarey, M.; Zimmermann, M. Novel technologies for virtual screening. *Drug Discovery Today* **2004**, *9*, 27–34.
- (2) Schneider, G.; Böhm, H.-J. Virtual screening and fast automated docking methods. *Drug Discovery Today* **2002**, *7*, 64–70.
- (3) Kitchen, D. B.; Decornez, H.; Furr, J. R.; Bajorath, J. Docking and scoring in virtual screening for drug discovery: methods and applications. *Nat. Rev. Drug Discovery* **2004**, *3*, 935–949.
- (4) Tuccinardi, T. Docking-based virtual screening: recent developments. *Comb. Chem. High Throughput Screening* **2009**, *12*, 303–14.
- (5) Breda, A.; Basso, L. A.; Santos, D. S.; de Azevedo, W. F., Jr. Virtual Screening of Drugs: Score Functions, Docking, and Drug Design. *Curr. Comput.-Aided Drug Des.* **2008**, *4*, 265–272.
- (6) OpenEye Scientific Software, ROCS; Santa Fe, NM. <http://www.eyesopen.com/rocs> (accessed Mar 08, 2012).
- (7) Chemical Computing Group, *Molecular Operating Environment MOE*, Version 2010.10; Montreal, Quebec, Canada. <http://www.chemcomp.com> (accessed Mar 08, 2012).
- (8) McGann, M. FRED Pose prediction and virtual screening accuracy. *J. Chem. Inf. Model.* **2011**, *51*, 578–596.
- (9) OpenEye Scientific Software, OMEGA; Santa Fe, NM. <http://www.eyesopen.com/omega> (accessed Mar 08, 2012).
- (10) Rarey, M.; Kramer, B.; Lengauer, T.; Klebe, G. A fast flexible docking method using an incremental construction algorithm. *J. Mol. Biol.* **1996**, *261*, 470–489.
- (11) Friesner, R. A.; Banks, J. L.; Murphy, R. B.; Halgren, T. A.; Klicic, J. J.; Mainz, D. T.; Repasky, M. P.; Knoll, E. H.; Shelley, M.; Perry, J. K.; Shaw, D. E.; Francis, P.; Shenkin, P. S. Glide: a new approach for rapid, accurate docking and scoring. 1. Method and assessment of docking accuracy. *J. Med. Chem.* **2004**, *47*, 1739–1749.
- (12) Halgren, T. A.; Murphy, R. B.; Friesner, R. A.; Beard, H. S.; Frye, L. L.; Pollard, W. T.; Banks, J. L. Glide: a new approach for rapid, accurate docking and scoring. 2. Enrichment factors in database screening. *J. Med. Chem.* **2004**, *47*, 1750–1759.
- (13) Kirchmair, J.; Markt, P.; Distinto, S.; Wolber, G.; Langer, T. Evaluation of the performance of 3D virtual screening protocols: RMSD comparisons, enrichment assessments, and decoy selection—what can we learn from earlier mistakes? *J. Comput.-Aided Mol. Des.* **2008**, *22*, 213–228.
- (14) Kirchmair, J.; Laggner, C.; Wolber, G.; Langer, T. Comparative analysis of protein-bound ligand conformations with respect to catalyst's conformational space subsampling algorithms. *J. Chem. Inf. Model.* **2005**, *45*, 422–430.
- (15) Kirchmair, J.; Wolber, G.; Laggner, C.; Langer, T. Comparative performance assessment of the conformational model generators omega and catalyst: a large-scale survey on the retrieval of protein-bound ligand conformations. *J. Chem. Inf. Model.* **2006**, *46*, 1848–1861.
- (16) Abagyan, R. A.; Totrov, M. M. Contact Area Difference (CAD): A Robust Measure to Evaluate Accuracy of Protein Models. *J. Mol. Biol.* **1997**, *268*, 678–685.
- (17) Ai, R.; Fatmi, M. Q.; Chang, C.-E. A. T-Analyst: a program for efficient analysis of protein conformational changes by torsion angles. *J. Comput.-Aided Mol. Des.* **2010**, *24*, 819–827.
- (18) Kroemer, R. T.; Vulpetti, A.; McDonald, J. J.; Rohrer, D. C.; Giordanetto, J.-Y. T. F.; Cotesta, S.; McMartin, C.; Kihlen, M.; Stouten, P. F. W. Assessment of docking poses: interactions-based accuracy classification (IBAC) versus crystal structure deviations. *J. Chem. Inf. Comput. Sci.* **2004**, *44*, 871–881.
- (19) Yusuf, D.; Davis, A. M.; Kleywegt, G. J.; Schmitt, S. An alternative method for the evaluation of docking performance: RSR vs RMSD. *J. Chem. Inf. Model.* **2008**, *48*, 1411–1422.
- (20) Baber, J. C.; Thompson, D. C.; Cross, J. B.; Humblet, C. GARD: A Generally Applicable Replacement for RMSD. *J. Chem. Inf. Model.* **2009**, *49*, 1889–1900.
- (21) Watts, K. S.; Dalal, P.; Murphy, R. B.; Sherman, W.; Friesner, R. A.; Shelley, J. C. ConfGen: a conformational search method for

- efficient generation of bioactive conformers. *J. Chem. Inf. Model.* **2010**, *50*, 534–546.
- (22) Vismara, P. Union of all the minimum cycle bases of a graph. *Electron. J. Combin.* **1997**, *4*, 1–15.
- (23) Floyd, R. Algorithm 97: Shortest path. *Commun. ACM* **1962**, *5*
- (24) Boström, J.; Greenwood, J. R.; Gottfries, J. Assessing the performance of OMEGA with respect to retrieving bioactive conformations. *J. Mol. Graphics Modell.* **2003**, *21*, 449–462.
- (25) Perola, E.; Charifson, P. Conformational analysis of drug-like molecules bound to proteins: an extensive study of ligand reorganization upon binding. *J. Med. Chem.* **2004**, *47*, 2499–510.
- (26) Berman, H. M.; Westbrook, J.; Feng, Z.; Gilliland, G.; Bhat, T. N.; Weissig, H.; Shindyalov, I. N.; Bourne, P. E. The Protein Data Bank. *Nucleic Acids Res.* **2000**, *28*, 235–242.
- (27) Griewel, A.; Kayser, O.; Schlosser, J.; Rarey, M. Conformational sampling for large-scale virtual screening: accuracy versus ensemble size. *J. Chem. Inf. Model.* **2009**, *49*, 2303–2311.
- (28) Grant, J. A.; Gallardo, M. A.; Pickup, B. T. A fast method of molecular shape comparison: A simple application of a Gaussian description of molecular shape. *J. Comput. Chem.* **1996**, *17*, 1653–1666.