Journal of Molecular Graphics and Modelling 29 (2010) 443–449

Contents lists available at ScienceDirect

Journal of Molecular Graphics and Modelling

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

Complementing ultrafast shape recognition with an optical isomerism descriptor

Ting Zhou, Karine Lafleur, Amedeo Caflisch*

Department of Biochemistry, University of Zurich, Winterthurerstrasse 190, CH-8057 Zurich, Switzerland

ARTICLE INFO

Article history:
Received 20 May 2010
Received in revised form 18 August 2010
Accepted 24 August 2010
Available online 29 September 2010

Keywords:
Chirality descriptor
3D database search
Conformational isomerism
Conformer clustering

ABSTRACT

We introduce the mixed product of three vectors spanning four molecular locations as a descriptor of optical isomerism. This descriptor is very efficient as it does not require molecular superposition, and is very robust in discriminating between a given isomer and its mirror image. In particular, conformational isomers that are mirror images of each other, as well as optical isomers have opposite sign of the descriptor value. For efficient database searches, the optical isomerism descriptor can be used to complement an available ultrafast shape recognition (USR) method based solely on distances, which is not able to distinguish enantiomers. By an extensive comparison of the USR-based similarity score with an approach based on Gaussian molecular volume overlap, the accuracy and completeness of the former are discussed.

1. Introduction

Shape complementarity is essential in macromolecular recognition and binding of small molecules to proteins because of the sensitivity of the van der Waals energy at separations close to the optimal distance. There is abundant experimental evidence that small molecules with shape similar to known active compounds are likely to have similar biological activities [1]. Therefore, screening of databases of three-dimensional (3D) molecular structures can be performed by comparison of molecular shapes [2–4]. Several methods have been developed and applied in the past few decades to identify compounds similar to a query molecule [5–10]. They are useful whenever one or more inhibitors of a target protein are known particularly when the 3D structure of the protein is not available.

Recently, a method termed Ultrafast Shape Recognition (USR) has been developed for searching very large databases of molecular structures [11]. Despite its recent publication, USR has already been used in several drug design projects [3,12–15] because of its simplicity and efficiency. Importantly, the molecules do not need to be superposed. Only, the distances between each atom of the molecule and four molecular locations are calculated for USR: the molecular centroid (ctd), the atom closest to ctd (cst), the atom farthest to ctd (fct), and the atom farthest to fct (ftf). The shape of a molecule is then encoded by 12 descriptors independent of the number of atoms. The first descriptor is the mean of atomic distances from ctd $\mu_1^{\mathrm{ctd}} \equiv (1/N)\sum_{j=1}^{N} d_j^{\mathrm{ctd}}$, where d_j is the distance of the jth atom from ctd, and N is the number of atoms in the molecule. The second descriptor is the square root of the second central moment of the distribution of the same atomic distances $\mu_2^{\mathrm{ctd}} \equiv \left[(1/N)\sum_{j=1}^{N}(d_j^{\mathrm{ctd}} - \mu_1^{\mathrm{ctd}})^2\right]^{1/2}$. The third descriptor is the cubic root of the third central moment of the same distribution $\mu_3^{\mathrm{ctd}} \equiv \left[(1/N)\sum_{j=1}^{N}(d_j^{\mathrm{ctd}} - \mu_1^{\mathrm{ctd}})^3\right]^{1/3}$ which is a measure of asymmetry. The remaining nine descriptors are calculated analogously using cst, fct, and ftf. Since only intramolecular distances are used in the 12 descriptors, the USR is not able to distinguish mirror images.

Here, we supplement the original USR method [11] with an optical isomerism descriptor that is able to discriminate a molecule from its mirror image, and is therefore particularly useful for clustering conformers and searching 3D databases. Our extension of USR (called USR:OptIso) is first tested on three pairs of conformations of kinase inhibitors and 15 pairs of different types of isomers. Then similarity scores based on USR and USR:OptIso for 1.6×10^{10} pairs of conformers of 2.7 millions small molecules are compared with the ones based on Gaussian molecular volume overlap [16] calculated by ROCS (OpenEye Scientific Software).

2. Methods

2.1. Optical isomerism descriptor

Considerable efforts have been devoted to symmetry detection in chemistry and chemoinformatics. In particular, several methods have been developed to analyze chirality. These include two-dimensional descriptors [17–19] for the prediction of the major product of stereoselective reactions [20–22], and three-dimensional descriptors (chiral topological indices) as

* Corresponding author. Tel.: +41 44 635 55 21; fax: +41 44 635 68 62.
E-mail address: caflisch@bioc.uzh.ch (A. Caflisch).

1093-3263/$ – see front matter © 2010 Elsevier Inc. All rights reserved.
doi:10.1016/j.jmgm.2010.08.007

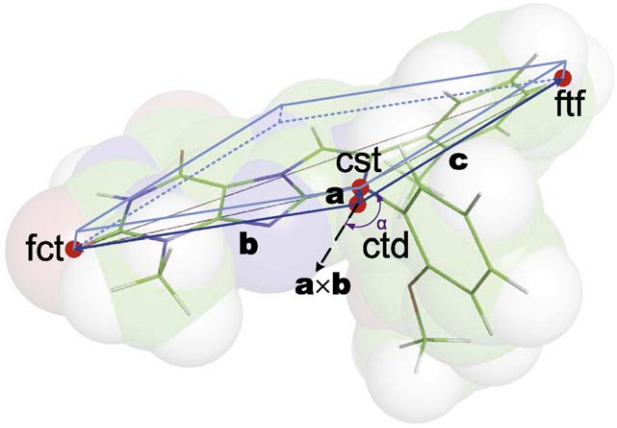

Fig. 1. Optical isomerism descriptor. The four molecular locations of a conformer of compound **1** are denoted with red circles, and the three vectors **a**, **b**, and **c** with blue arrows. The optical isomerism descriptor is the cubic root of the volume of the parallelepiped with blue edges. The sign of the optical isomerism descriptor is negative for this conformer because **c** and **a** × **b** form an obtuse angle (violet angle α). The mirror image of this conformer has a positive value of the optical isomerism descriptor, and is shown in Fig. 3. (For interpretation of the references to color in this figure legend, the reader is referred to the web version of the article.)

complement to distance matrices in quantitative stereochemical structure–activity relationship models [23–29].

Here, the following vectors are introduced for the efficient evaluation of the optical isomerism descriptor: $\mathbf{a} \equiv \mathbf{cst} - \mathbf{ctd}$, $\mathbf{b} \equiv \mathbf{fct} - \mathbf{ctd}$, and $\mathbf{c} \equiv \mathbf{ftf} - \mathbf{ctd}$, where **ctd**, **cst**, **fct**, and **ftf** are the vectors connecting the origin of the coordinates to each of the four molecular locations (Fig. 1). The optical isomerism descriptor is defined as the cubic root of the scalar triple product (or mixed product) of **a**, **b**, and **c**, i.e., optical isomerism descriptor $\equiv [\mathbf{c}\,(\mathbf{a} \times \mathbf{b})]^{1/3}$. The cubic root is used to obtain a unit of length (Å) as for the other 12 descriptors. The computational cost for evaluating the optical isomerism descriptor is neglectable, since the coordinates of the four molecular locations have already been calculated for the other 12 descriptors.

During the writing of this manuscript, Armstrong et al. reported a modification of USR that is able to distinguish enantiomers [30]. They use the *cross* product of two vectors spanning *three* of the four USR molecular locations (ctd, fct, and ftf) to define a fourth location which is different from cst. In contrast, the crucial component of our descriptor is the *triple* product of three vectors spanning all of the *four* locations. Moreover, Armstrong and collaborators replace three of the 12 USR descriptors (those involving cst) whereas we supplement the USR with the optical isomerism descriptor. Note that cst might contain useful information as Ballester et al. reported that the first moment of the distribution of distances from cst had a similar value across active molecules (Fig. 3 in Ref. [3]). In general, the atomic distances to ctd are different from that to cst as the separation between ctd and cst is usually between 0.4 and 2.0 Å for small molecules (Supplementary Material Fig. S-1). Since our optical isomerism descriptor contains more geometrical information than the one of Armstrong et al., it is expected to perform at least equivalently.

2.2. Similarity score

The inverse Manhattan distance is used as similarity score [11]:

$$S_{pq} = \frac{1 \text{ length unit}}{1 \text{ length unit} + \frac{1}{n}\sum_{i=1}^{n} \left| M_i^p - M_i^q \right|},$$

where M_i^p is the ith descriptor of the conformation p. Note that all the USR descriptors have a unit of length, which is Å here, therefore the S_{pq} is dimensionless. The addition of 1 length unit at the denominator prevents a division by zero in the case of identical 3D structures, and yields a similarity score of 1 for them. The similarity scores with the optical isomerism descriptor (S_{pq}^{13}) and without (S_{pq}^{12}) are compared in the next section.

3. Results and discussion

3.1. Discriminatory power of USR:OptIso and usefulness for clustering

The ability of the optical isomerism descriptor to discriminate isomers is presented in Fig. 2. For a molecule and its mirror image the 12 descriptors of the original USR method are identical because they only depend on distance distributions. However, their optical isomerism descriptors are opposite because their four molecular locations are mirror images as well. In contrast, structural isomers, diastereoisomers, and other types of conformers that are not mirror image of each other have different distance distributions. Therefore, the first 12 descriptors are enough to discriminate them.

A recently published inhibitor of the receptor tyrosine kinase Ephrin type-B receptor 4 (EphB4) [13] is used to illustrate the usefulness of the optical isomerism descriptor (Fig. 3). The similarity score S_{pq}^{13} is able to distinguish the two mirror image conformers of compound **1** because of the opposite sign of their optical isomerism descriptors. In contrast, the two conformers of **1** have identical 12 descriptors based on the original USR method [11].

The optical isomerism descriptor is useful for clustering as it can distinguish between different conformers/isomers that would be clustered together by the original USR. In our previous study [13], multiple conformers of **1** were generated by systematic bond-rotation and optimized to their nearest local minima using density functional theory. Two local minima were then considered identical if their similarity score S_{pq} was higher than 0.999. Interestingly, the opposite sign of the optical isomerism descriptor contributes significantly to the identification of mirror images (or pairs of conformers very close to mirror images), in particular when S_{pq}^{12} is close to 1 (Fig. 4). Furthermore, the USR:OptIso was tested on two pairs of isomers of recently published kinase inhibitors (**2** and **3** in Fig. 5) [31,32]. The first 12 descriptors of USR have identical values, while the optical isomerism descriptor reduces the similarity score from 1 to 0.705 for compound **2** and from 1 to 0.650 for compound **3**.

The optical isomerism descriptor can be used for searching (multi-)conformational libraries. As an example, using the similarity score S_{pq}^{13} yields only the conformer similar to the query whereas the conformations that are similar to its mirror image might be retrieved erroneously if one neglects the optical isomerism descriptor. This is a clear advantage of USR:OptIso with respect to the original USR.

Finally, it is necessary to verify that similar conformers of a given molecule yield very similar values of S_{pq}^{12} and S_{pq}^{13}. A set of 100 similar structures of the protein kinase inhibitor PP2 [33] was used for assessing the robustness upon minor structural change of the original USR and USR:OptIso. A scatter plot is presented in Supplementary Mater Fig. S-2. This test indicates that slight changes in the coordinates yield minor changes in both S_{pq}^{12} and S_{pq}^{13} when they are close to 1. Note that the high similarity range (i.e., values close to 1) is the most relevant case for virtual screening as only a small fraction of hits can be tested in practice.

Isomer1	Isomer2	Type of Isomerism	S_{pq}^{12}	S_{pq}^{13}	OID1	OID2	ROCS
		Structural isomerism: functional group isomerism	0.891	0.887	0.153	-0.032	0.999
		Structural isomerism: position isomerism	0.802	0.789	0.396	-0.106	0.825
		Structural isomerism: skeletal isomerism	0.826	0.833	0.838	0.911	0.892
		Conformational isomerism	1.000	0.889	0.815	-0.815	0.667
		Diastereoisomerism: cis/trans	0.979	0.972	0.456	0.334	0.685
		Diastereoisomerism: E/Z	0.922	0.928	0.000	0.000	0.302
		Chirality with one stereogenic centers: R/S	1.000	0.910	0.642	-0.642	0.997
		Chirality without stereogenic centers: allenes	1.000	0.884	0.849	-0.849	0.997
		Chirality without stereogenic centers: alkylidenecycloalkanes	1.000	0.860	1.054	-1.054	0.953
		Chirality without stereogenic centers: spiranes	1.000	0.987	-0.087	0.087	0.963
		Chirality without stereogenic centers: byphenyls - atroposomerism	1.000	0.964	-0.238	0.238	0.993
		Chirality without stereogenic centers: helicenes	1.000	0.790	1.725	-1.725	0.703
		Chirality without stereogenic centers - planar chirality: cyclophanes	1.000	0.759	2.053	-2.053	0.322
		Chirality without stereogenic centers - planar chirality: annulenes	1.000	0.863	1.035	-1.035	0.947
		Chirality without stereogenic centers - planar chirality: cyclalkenes	1.000	0.809	-1.514	1.514	0.957

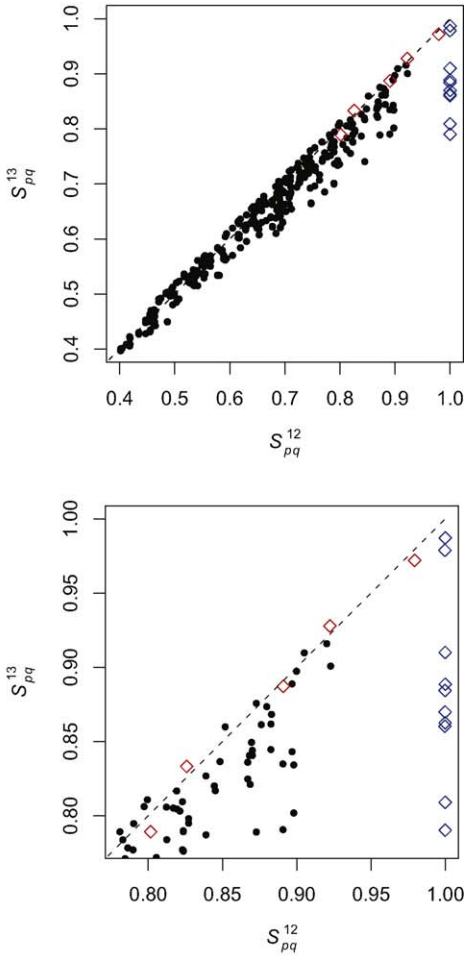

Fig. 2. Application of the optical isomerism descriptors to isomers. All pairs of isomers can be distinguished by the USR:OptIso (13 descriptors) because the optical isomerism descriptors (OID1 and OID2) of mirror image isomers have opposite signs. The scatter plots show the pairwise USR comparisons of these 30 compounds (top right, full data set; bottom right, zoom-in on values close to 1). The original USR (S_{pq}^{12}) and USR:OptIso (S_{pq}^{13}) methods assign close similarities except for 10 pairs (blue diamonds) out of 435. The five pairs of isomers that can be distinguished by both S_{pq}^{12} and S_{pq}^{13} are denoted with red diamonds. The dashed line represents $S_{pq}^{12} = S_{pq}^{13}$. The last column of the table is the ROCS shape Tanimoto [16].

3.2. Potential limitations of the optical isomerism descriptor

The optical isomerism descriptor is the cubic root of the (signed) volume of the parallelepiped defined by three vectors connecting four molecular locations. It is therefore equal to zero whenever the four molecular locations are coplanar. To estimate the frequency of the coplanarity of these four locations, we calculated the optical isomerism descriptor of 100,812,356 poses of about 2.7 million molecules (downloaded from the 2007 version of the ZINC library [34]) generated by high-throughput docking into EphB4 [35]. Strikingly, only 28 poses (of 24 molecules) have an optical isomerism descriptor smaller than 0.01 in absolute value, which indicates that the optical isomerism descriptor is able in the vast majority of cases to clearly discriminate a given isomer from its mirror image. Some of these 28 poses are (close to) coplanar, while in others of them, the majority of the atomic nuclei are in a plane, and the nuclei out of the plane happen to be excluded (they are neither closest nor farthest to a centroid) from the four centroids. Note that a value of the optical isomerism descriptor different from zero does not necessarily imply that a molecule is chiral. For this reason, we prefer to use the term "optical isomerism" rather than "chirality" descriptor.

3.3. Comparison between similarity scores calculated by USR and ROCS

It is interesting to compare the USR-based similarity scores with a metric based on superimposed volume. The correlation coefficient between the similarity score based on Gaussian molecular shape overlap (ROCS shape Tanimoto [16]) and either S_{pq}^{12} or S_{pq}^{13} is 0.64 for the 435 possible pairs of the 30 compounds in Fig. 2. This relatively low correlation is due to the fact that the similarity score evaluations are based on two different procedures. In the former, the similarity is calculated by volume-overlap percentage after structural superposition, whereas in USR which does not require overlap, the similarity is evaluated using distributions of nuclei distances from the molecular centroids. Moreover, neither the iterative maximizing of the overlapped molecular volume in ROCS nor the maximum/minimum function for determining centroids in USR is continuous with respect to the Cartesian coordinates of the atomic nuclei, which are used as input.

For an in-depth comparison of similarity scores calculated by USR or USR:OptIso and ROCS shape Tanimoto the aforementioned

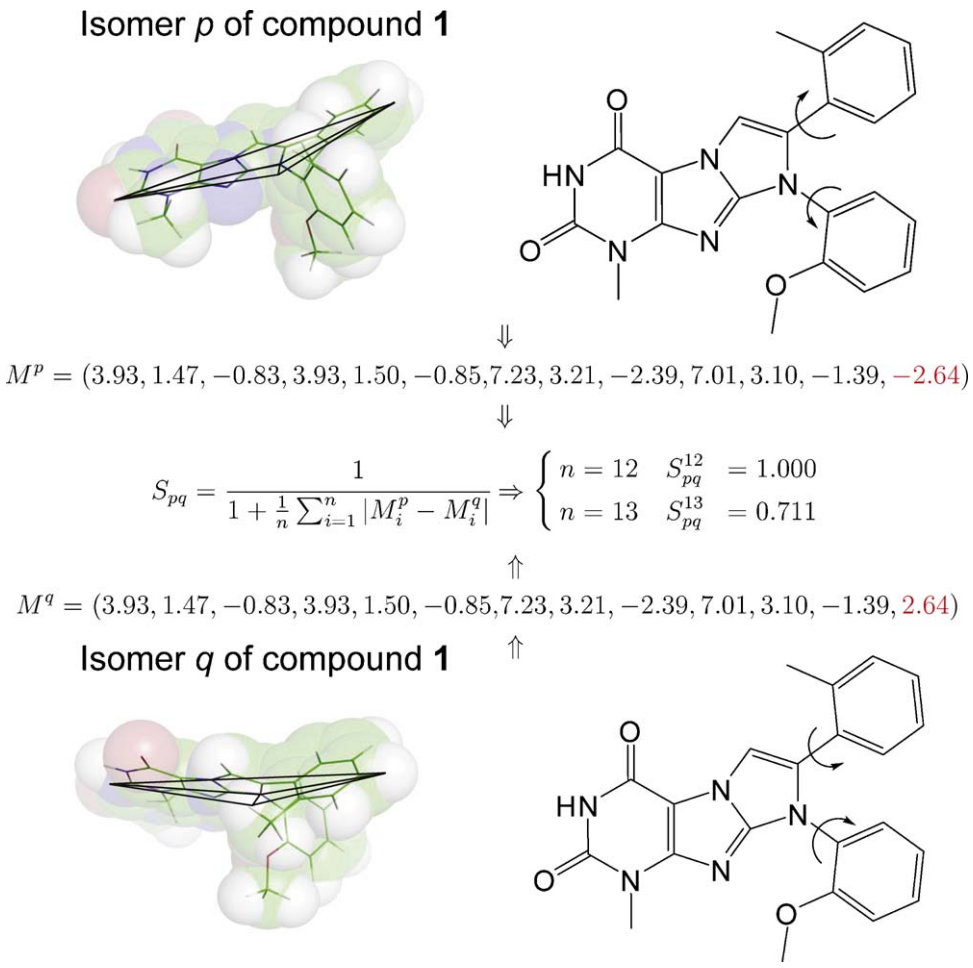

Fig. 3. Two mirror images of the tyrosine kinase inhibitor **1** [13] generated by systematic bond-rotation are distinguished by the optical isomerism descriptor. The four molecular locations are connected with black lines on the 3D structures which are shown with sticks and transparent CPK models. This example shows the high discriminating power of the optical isomerism descriptor whose usage results in a low similarity score $S_{pq}^{13} = 0.711$ for the two mirror images of **1** whereas they are not distinguished by the original USR method ($S_{pq}^{12} = 1.000$). (For interpretation of the references to color in this figure legend, the reader is referred to the web version of the article.)

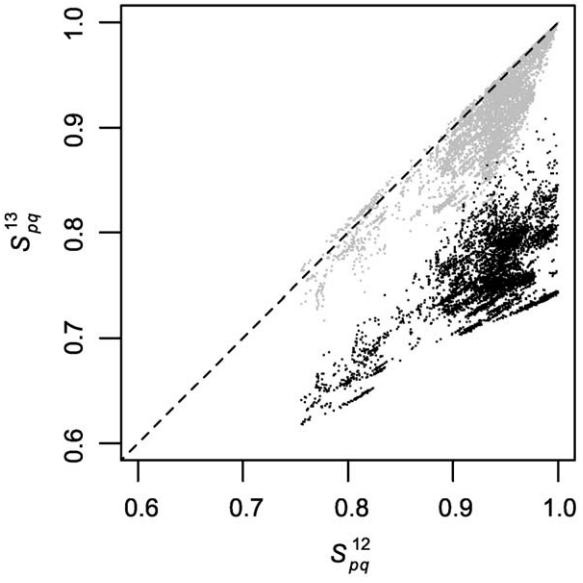

Fig. 4. Scatter plot of the similarity scores with the optical isomerism descriptor (S_{pq}^{13}) and without (S_{pq}^{12}) for the 10,585 pairs of 146 local minima [13] of **1**. The color of each data point illustrates the sign of the optical isomerism descriptor (same and opposite signs are in gray and black, respectively). The dashed line represents $S_{pq}^{12} = S_{pq}^{13}$.

100 million poses of about 2.7 million molecules from ZINC were used to generate 1.6×10^{10} pairs of conformations. Notably, these two methods have very different algorithms to evaluate molecular shape similarity. Cartesian coordinates are the only input required to calculate USR scores, whereas ROCS also needs atomic radii to evaluate the Gaussian molecular volume. To compare the conformation pairs filtered by different similarity cutoffs, the "accuracy" was defined as the number of pairs for which both scores exceeded the cutoff divided by the pairs for which only the USR similarity score exceeded the cutoff (i.e., C/(A + C) in Fig. 6). In the same way, the "completeness" was defined as the ratio of the number of pairs for which both scores exceeded the cutoff to the pairs for which only ROCS shape Tanimotos exceeded the cutoff (i.e., C/(B + C) in Fig. 6).

Both accuracy and completeness increase monotonously with the similarity cutoff (Fig. 6). The accuracy of USR:OptIso has improved compared with the original USR, because the mirror images of the query conformation have been eliminated by the opposite optical isomerism descriptors. For instance, 91.25% of the conformation pairs that have USR:OptIso scores ≥0.968 also have ROCS shape Tanimotos ≥0.968, whereas the percentage decreases to 57.17% for the original USR. If the similarity score cutoff is set to 0.98, the accuracy of the USR:OptIso increases to 99.02% compared to 86.30% for the original one. Panel (a) of Fig. 7 shows "the most inaccurate" example whose USR score is higher than the ROCS shape Tanimoto.

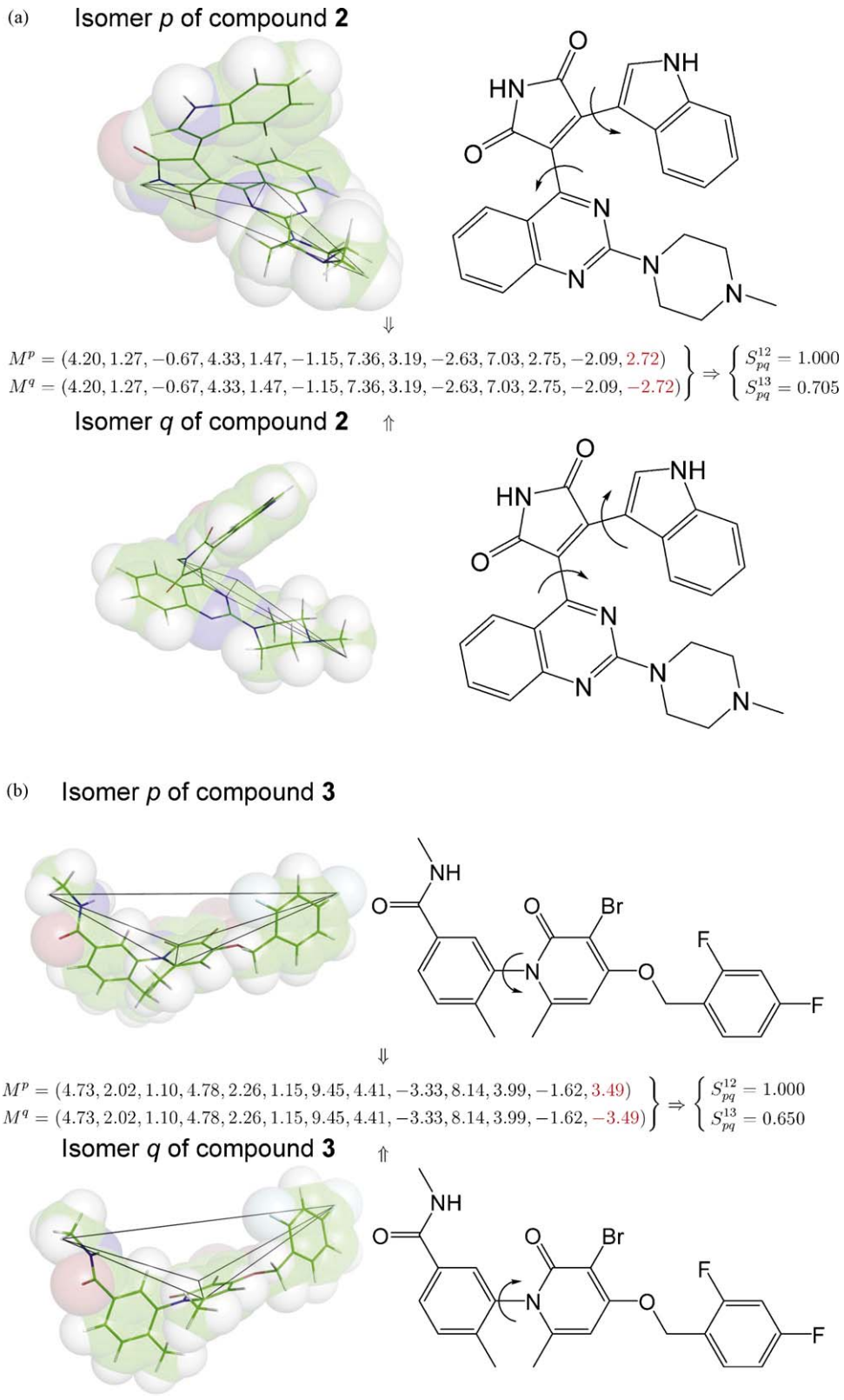

Fig. 5. Discriminatory power of the optical isomerism descriptor. (a) The conformer of the protein kinase C inhibitor **2** observed in the X-ray structure [31] is shown in the top panel, while its mirror image is shown in the bottom panel. The similarity score decreases from 1 to 0.705 by taking into account the optical isomerism descriptor. (b) The active conformer of the p38 MAP kinase inhibitor **3**, shown in the top panel, is >100-fold more potent than its atropisomer shown in the bottom panel [32]. The similarity score decreases from 1 to 0.650 by taking into account the optical isomerism descriptor. (For interpretation of the references to color in this figure legend, the reader is referred to the web version of the article.)

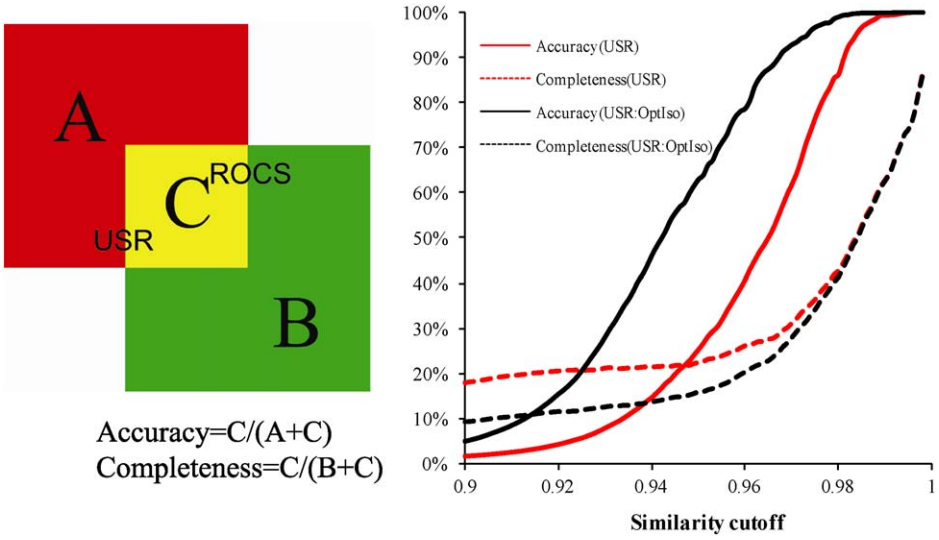

Fig. 6. Accuracy and completeness of USR similarity score in reproducing ROCS shape Tanimoto. In the left panel, the red and the green squares are the sets of conformation pairs for which USR similarity scores and ROCS shape Tanimoto satisfy a particular cutoff, respectively. Their overlap (yellow part) represents the conformation pairs for which both similarity scores satisfy the cutoff. In the right panel, the cutoff varies from 0.9 to 1.0, which covers the range interesting for virtual screening applications. The solid and dashed lines represent accuracy and completeness, respectively, while the red and black color denote the USR and USR:OptIso results, respectively.

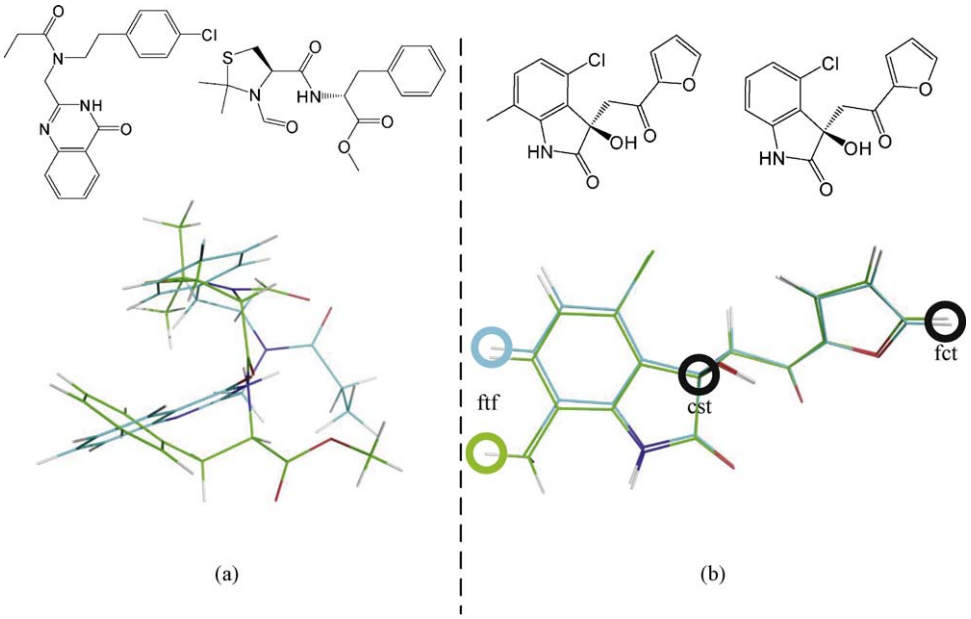

Fig. 7. Examples of large discrepancies between USR and ROCS shape Tanimoto. (a) High USR similarity score but low ROCS shape Tanimoto. The overlaid conformations were optimized by ROCS. The USR and USR:OptIso similarity scores are 0.9807 and 0.9810. The ROCS shape Tanimoto is 0.676. This is the only case out of 1.6×10^{10} conformation pairs whose USR is larger than 0.98 and ROCS shape Tanimoto smaller than 0.7. (b) Low USR similarity score but high ROCS shape Tanimoto. The USR and USR:OptIso similarity scores are 0.6988 and 0.6981. The ROCS shape Tanimoto is 0.999. These two molecules are different in only one substituent. The additional methyl group in the green conformation changes the location of ftf from the phenyl hydrogen (of the cyan conformation) to the methyl hydrogen. (For interpretation of the references to color in this figure legend, the reader is referred to the web version of the article.)

The completeness of USR methods is low. One of the main reasons is that USR overestimates the difference of conformation pairs that are different at the extremity of the molecule where fct and/or ftf are defined. Fig. 7(b) shows an example of a conformation pair whose ROCS shape Tanimoto is higher than the USR similarity score. In this example, USR overestimates the conformational difference because of a change of the ftf location. The similarity calculated by USR:OptIso is often lower than the one calculated by USR (Fig. 4) due to the influence of the optical isomerism descriptor. Therefore, USR:OptIso has lower completeness than USR. This difference in completeness becomes smaller when the similarity cutoff approaches 1. The low completeness is not surprising as ROCS shape Tanimoto distinguishes atomic elements by different van der Waals radii whereas both USR and USR:OptIso treat all atoms equally. Moreover, the optimal volume overlap has to be calculated in ROCS shape Tanimoto for every pair of conformations while the USR methods are significantly more efficient (more than 5000 times faster [11]) as they use only interatomic distances.

4. Conclusions

The optical isomerism descriptor (defined as the mixed product of the three vectors from the molecular centroid to the three molecular locations cst, fct, ftf) is an efficient and robust tool for

shape comparison. It can be used as a supplement of the original 12 USR descriptors, which are based solely on distance distributions, while the optical isomerism descriptor is able to distinguish mirror images. It is therefore helpful for analyzing molecules with stereogenic centers, atropisomerism, and in the clustering of conformers generated by systematic bond-rotation. Moreover, it can be used for the efficient search of molecular conformations that are superposable on the query structure. Finally, a comparison of the USR similarity score with the ROCS shape Tanimoto shows that both accuracy and completeness increase monotonously with the similarity score cutoff. The accuracy of the USR:OptIso similarity score is always higher than the one based on the original USR, and the completeness of USR:OptIso is close to the one of USR in high similarity ranges, which are relevant for virtual screening.

Acknowledgement

We thank Dr. Ballester for providing a detailed description of USR. We thank OpenEye Scientific Software for providing an academic license of ROCS.

Source code: The source code for calculating the optical isomerism descriptor is available at http://code.google.com/p/usrchirality/.

Appendix A. Supplementary data

Supplementary data associated with this article can be found, in the online version, at doi:10.1016/j.jmgm.2010.08.007.

References

[1] G. McGaughey, R. Sheridan, C. Bayly, J. Culberson, C. Kreatsoulas, S. Lindsley, V. Maiorov, J. Truchon, W. Cornell, J. Chem. Inf. Model. 47 (2007) 1504–1519.
[2] T. Rush, J. Grant, L. Mosyak, A. Nicholls, J. Med. Chem. 48 (2005) 1489–1495.
[3] P.J. Ballester, P.W. Finn, W.G. Richards, J. Mol. Graph. Model. 27 (2009) 836–845.
[4] S. Kinnings, R. Jackson, J. Chem. Inf. Model. 49 (2009) 2056–2066.
[5] P. Artymiuk, P. Bath, H. Grindley, C. Pepperrell, A. Poirrette, D. Rice, D. Thorner, D. Wild, P. Willett, F. Allen, J. Chem. Inf. Comput. Sci. 32 (1992) 617–630.
[6] R. Sheridan, S. Kearsley, Drug Discov. Today 7 (2002) 903–911.
[7] I.L. Ruiz, G.C. García, M.A. Gómez-Nieto, J. Chem. Inf. Model. 45 (2005) 1178–1194.
[8] J. Haigh, B. Pickup, J. Grant, A. Nicholls, J. Chem. Inf. Model. 45 (2005) 673–684.
[9] P. Hawkins, A. Skillman, A. Nicholls, J. Med. Chem. 50 (2007) 74–82.
[10] Y. Cao, T. Jiang, T. Girke, Bioinformatics 24 (2008) i366–i374.
[11] P.J. Ballester, W.G. Richards, J. Comput. Chem. 28 (2007) 1711–1723.
[12] E. Cannon, F. Nigsch, J. Mitchell, Chem. Cent. J. 2 (2008) 3.
[13] K. Lafleur, D. Huang, T. Zhou, A. Caflisch, C. Nevado, J. Med. Chem. 52 (2009) 6433–6446.
[14] H. Li, J. Huang, L. Chen, X. Liu, T. Chen, J. Zhu, W. Lu, X. Shen, J. Li, R. Hilgenfeld, H. Jiang, J. Med. Chem. 52 (2009) 4936–4940.
[15] P.J. Ballester, I. Westwood, N. Laurieri, E. Sim, W.G. Richards, J. R. Soc. Interface 7 (2010) 335–342.
[16] J.A. Grant, M.A. Gallardo, B.T. Pickup, J. Comput. Chem. 17 (1996) 1653–1666.
[17] M. Randic, M. Razinger, J. Chem. Inf. Comput. Sci. 36 (1996) 429–441.
[18] M. Randic, J. Chem. Inf. Comput. Sci. 41 (2001) 639–649.
[19] Q.Y. Zhang, J.A. de Sousa, J. Chem. Inf. Model. 46 (2006) 2278–2287.
[20] J.A. de Sousa, J. Gasteiger, J. Chem. Inf. Comput. Sci. 41 (2001) 369–375.
[21] J.A. de Sousa, J. Gasteiger, J. Mol. Graph. Model. 20 (2002) 373–388.
[22] J.A. de Sousa, J. Gasteiger, I. Gutman, D. Vidović, J. Chem. Inf. Comput. Sci. 44 (2004) 831–836.
[23] H.P. Schultz, E.B. Schultz, T.P. Schultz, J. Chem. Inf. Comput. Sci. 35 (1995) 864–870.
[24] A. Golbraikh, D. Bonchev, A. Tropsha, J. Chem. Inf. Comput. Sci. 41 (2001) 147–158.
[25] S. Capozziello, A. Latranzi, Chirality 15 (2003) 227–230.
[26] S. Fujita, Theor. Chem. Acc. 113 (2005) 80–86.
[27] C.S. Yang, C.L. Zhong, QSAR Comb. Sci. 24 (2005) 1047–1055.
[28] M. Dervarics, F. Otvos, T.A. Martinek, J. Chem. Inf. Model. 46 (2006) 1431–1438.
[29] R. Natarajan, S. Basak, T. Neumann, J. Chem. Inf. Model. 47 (2007) 771–775.
[30] M.S. Armstrong, G.M. Morris, P.W. Finn, R. Sharma, W.G. Richards, J. Mol. Graph. Model. 28 (2009) 368–370.
[31] J. Wagner, P.V. Matt, R. Sedrani, R. Albert, N. Cooke, C. Ehrhardt, M. Geiser, G. Rummel, W. Stark, A. Strauss, S. Cowan-Jacob, C. Beerli, G. Weckbecker, J.-P. Evenou, G. Zenke, S. Cottens, J. Med. Chem. 52 (2009) 6193–6196.
[32] L. Xing, H. Shieh, S. Selness, R. Devraj, J. Walker, B. Devadas, H. Hope, R. Compton, J. Schindler, J. Hirsch, A. Benson, R. Kurumbail, R. Stegeman, J. Williams, R. Broadus, Z. Walden, Monahan, J. Biochem. (Mosc.) 48 (2009) 6402–6411.
[33] A. Sturz, B. Bader, K. Thierauch, J. Glienke, Biochem. Biophys. Res. Commun. 313 (2004) 80–88.
[34] J. Irwin, B. Shoichet, J. Chem. Inf. Model. 45 (2005) 177–182.
[35] T. Zhou, A. Caflisch, J. Chem. Inf. Model. 49 (2009) 145–152.