



Regular Article

Chemical curation to improve data accuracy: recent development of the 3DMET database

Miki H. Maeda, Tomoki Yonezawa and Tomomi Komaba

Advanced Analysis Center, National Agriculture and Food Research Organization, Tsukuba, Ibaraki 305-8602, Japan

Received October 12, 2017; accepted February 6, 2018

We have developed a three-dimensional structure database of natural metabolites (3DMET). Early development of the 3DMET database relied on content auto-generated from 2D-structures of other chemical databases. From 2009, we began manual curation, obtaining new compounds from published works. In the process of curation, problems of digitizing 3D-structures from structure drawings of documents were accumulated. As the same as auto-generation, structure drawings should be also paid attention about stereochemistry. Our experiences in manual curation of 3DMET, as described herein, may be useful to others in this field of research and for the development of supporting systems of a chemical structure database. Manual curation is still necessary for proper database entry of the 3D-configurations of chiral atoms, a problem encountered frequently among natural products.

Key words: chemical curation, natural products, absolute configuration, chirality, molecular docking

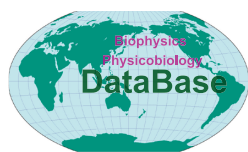
Today, there are some well-known databases of chemicals. COMPOUND, one of the earliest databases of chemicals with 2D-structures, is a part of KEGG LIGAND [1] as a

metabolic database. We tried to extract natural products by links to the metabolic map. However, we could not automatically distinguish natural products from the others, due to some natural products not linked to metabolic map. That was the reason why we began to develop 3DMET.

Chembridge Structural Database (CSD) [2], Ligand Box [3], KNApSAcK-3D [4], ZINC [5], PubChem-3D [6], NCI DIS 3D database [7] are well-known chemical database including 3D-structure. All of them stock huge number of records and they are still increasing. CSD is the world's repository of small-molecule organic and metal-organic crystal structures and is distinguished from the other chemical databases because they collect modeled structures automatically generated by their employed programs. The 3D-structures of other 3D-structure databases are automatically generated. The 3DMET database is novel as a manually curated database.

Our ultimate goal for the 3DMET database is to be a basic resource for functional prediction of proteins with unknown function. To achieve the goal, molecular docking technique would be applied to estimate natural ligands. Structure quality enough to perform molecular docking is inquired as a virtual library. How accurate is the quality? In programs of molecular docking, conformation sampling is the most essential part and mainly categorized into two kinds of algorithms as described in the next paragraph. For the detailed information about molecular docking, reviews by Meng,

Corresponding author: Miki H. Maeda, Advanced Analysis Center, National Agriculture and Food Research Organization, 2-1-2 Kannondai, Tsukuba, Ibaraki 305-8602, Japan.
e-mail: mmaeda@naro.affrc.go.jp



◀ Significance ▶

Several chemical structure databases currently exist. The purpose of each is guided by the research needs of the database developer. The 3DMET database has been created to be a resource of manually curated 3D-structures of natural products. In this work we describe our experiences of the manual curation process which may be useful to curators of other chemical databases or developers of supporting systems.



et al. [8] and Pagadala, *et al.* [9] could be referred.

Surflex-Dock [10] detects matching conformation against pharmacophore-like points, consisting of hydrogen-bond donor/acceptor points and hydrophobic points in the pocket. Many conformations from the input structure are generated and evaluated. FlexX [11] performs molecular docking by incremental construction. The input structure is divided into fragments. At first, a fragment is set near an interaction point. Then adjacent fragment are connected to the first fragment. After repeating this process, the final conformation is obtained. In each algorithm, stereochemistry of input structure is important because the resulted conformation would reflect the initial stereochemistry. Therefore, quality of structure in dataset is a key point for accurate molecular docking. However, the structures may not need to be critical minimized conformations because resulted conformations are usually changed from the initial ones given as their input structures. Newly developed programs were not tested yet, but combinations of 3D-generators and minimizers we previously checked were not accurate for atom chirality when we started manual curation [12,13].

Thus, our dataset should be accurate about stereochemistry of all atoms and bonds in a molecule. Due to inaccurate structures by automatic conversion, manual curation was introduced in 2009. The earliest curators checked the correctness of the 3D-structures automatically converted from the publicly available 2D-structure databases. Entries were checked by the correspondence of canonical SMILES [14] and/or InChI [15], and they were corrected if the structures were incorrect. After 3D-generation from the set of public databases, we began to build the database through collection of diverse natural compounds from literature. Because it was difficult to employ persons with enough skill to do chemical curation, computer tools were also applied on the possible steps. The 3D-structure generation process were divided into the following four steps; (1) Scanning document pages to PDF files, (2) generation of coordinates by optical chemical structure reader, (3) 2D-3D conversion, and (4) energy minimization. Steps (2) through (4) could be automated. The resulting energy-minimized 3D-structures were confirmed by chemical curators. Our experience has shown that this process did not always result in a 3D-structure suitable to meet the objectives of 3DMET [13].

In this report, we introduce fully manual curation to collect high-quality 3D-structures of natural products for continued development of the 3DMET database. This is a traditional approach for the development of similar databases. Although this process is less automated and therefore take more human time for development, we believe the quality of the database is substantially better and will allow for accurate functional prediction for proteins of unknown function. From our accumulated knowledge of curation, problems peculiar to manual curation were observed. For each such problem, we established appropriate measures to ensure structure correctness. A description of how we

addressed these challenges for accurate 3D-structure creation maybe useful for database developers and for database users. Therefore, we present a case report of our manual curation efforts in development of the 3D-structure chemical database 3DMET in order to share our experiences.

Materials and Methods

Detection of duplicated structures in a dataset

As a dataset of general natural products, compound structures described in a book, ROMPP Encyclopedia Natural Products (Thieme Medical Publishers, Inc., Germany), were generated. For the discernment of chemical structures, molecular specification strings were compared. The specifications were calculated by InChI (International Chemical Identifier) 1.04 [15] and SMILES Tool Kit version 4.95 by Daylight Inc. A set of compounds with completely the same strings were considered to be one compound. The redundant entries were confirmed by curators because different compounds rarely have the same name or InChI/SMILES in the cases as described in the section of Results and Discussion.

Investigation of suspicious description

Articles as data resource were based on the database of Natural Product Updates (NPU) 2010 and 2011 published by Royal Society of Chemistry (RSC). From the listed literatures of NPU 2010 and 2011, the articles as reports about newly isolated or structurally updated natural products were individually investigated. The compounds to verify about the descriptions were picked up. The problems were classified to several categories: shortage of information, correspondence error within the articles, discrepancy to experimental data, probable mistake, etc.

Investigation of methods to determine molecular structures

Target resources were articles reporting newly isolated natural products in the database of NPU 2013 by RSC. From the articles, newly reported compounds were counted and classified by with and without chiral center in the structures of the compounds. About the compounds that absolute configurations of the atoms were completely defined, the methods applied to define stereochemistry were also investigated.

Results and Discussion

Detection of duplicated entries

Detection of redundancy is necessary to prevent duplication of structure entries. Due to its better accuracy compared to SMILES [12], InChI is employed as our standard program to detect correspondence between two structures. However, as the result about the ROMPP dataset, in some cases, InChI is not able to identify an appropriate correspondence for our policy. Identical InChI strings generated between two compounds were typically caused by one of three reasons; tautomers, atropisomers, and another stereochemical

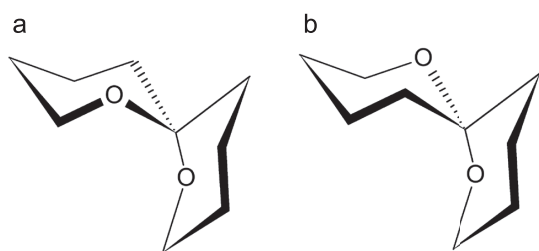


Figure 1 Stereoisomers of oleans; a, (R)-olean and b, (S)-olean.

isomers such as spiro compounds, for example, (R)-olean (Fig. 1a) and (S)-oleans (Fig. 1b).

Tautomers are not distinguished by InChI because it was designed to regard tautomeric conformers as the same. In 3DMET, tautomers are defined as individual entries. Therefore, InChI is not suitable to distinguish tautomers. For this purpose, canonical SMILES is used because it does not consider molecular tautomerism. We added a check process by using SMILES after using InChI to maximize automatic discrimination.

Atropisomers are stereoisomers derived from prevented rotation about a single bond. Atropisomers are the same connection among atoms. However, individual conformers can be isolated with separated physicochemical properties due to constraints in rotation. Because 3DMET is also a web-based 3D-structure dictionary, we allow atropisomers as separated

entries. This redundancy cannot be detected by InChI and SMILES and should be confirmed by manual curation.

Generally, R or S in stereochemistry indicates atom chirality attached with four substituted residues. However, in spiro compounds, a pair of connected side chains makes stereochemistry result from the different position of two rings, as shown in Figure 1. Although rarely occurring in natural compounds, the two resulting isomers are impossible to distinguish by InChI or SMILES and are therefore difficult to detect. For this special case of stereochemical compounds our experience extends only to (R)- and (S)-oleans.

Accuracy of text in literature

To accelerate data accuracy, good curators are needed to verify chemical description from literature. Misspellings are commonly observed in publication text. In addition, other questionable descriptions derived from common reasons were also observed as shown in Table 1. Of the reported compounds, 2,057 (18.3%) required confirmation. Below, we describe how these entries were processed for the 3DMET database.

The most frequent problem is insufficient structural information. Although not technically a mistake, it creates a serious difficulty for curators in database development. The percentage of entries requiring confirmation was 80.4%. Most of these were due to unclear drawing of defined chiral

Table 1 Curation issues that required detailed confirmation from the source publication

Issue	Compounds	Previously reported	Newly reported
Compounds requiring confirmation	2,507 (100%)		
<i>Shortage of information</i>			
unclear drawing of defined chiral atoms	1,579 (63.02%)	778	801
lacked compound name	436 (17.39%)	269	167
<i>Correspondence error within the article</i>			
correct name but wrong structure	93 (3.71%)	24	69
inverted drawing of sugar	77 (3.07%)	33	44
wrong name but correct structure	43 (1.72%)	13	30
unresponding chiral definition	52 (2.07%)	1	51
<i>Discrepancy to experimental data</i>			
to NMR spectrum	10 (0.40%)	1	9
to Mass spectrum	3 (0.12%)	0	3
to Circular dichroism	2 (0.08%)	0	2
to X-ray analysis	1 (0.04%)	0	1
<i>Probable mistake</i>			
misspelling of name	43 (1.72%)	21	22
different structure from previous report	139 (5.54%)	135	4
inconsistent names within article	11 (0.44%)	2	9
unresponding name and drawing	15 (0.60%)	15	0
<i>Miscellaneous</i>			
distortion in generated 3D-structure	2 (0.08%)	0	2

This data were determined from 13,881 newly structure-defined natural compounds reported in the articles listed in Natural Product Updates 2010 to 2011. The value provided in the columns indicates the number of compounds with the defined inaccurate. Numbers in the parentheses are the rate against all compounds requiring confirmation.

atoms (63.0%). For example, hydrogens were often left out from the structure drawing of chemically well-known structures, such as steroids. Therefore, omitted hydrogens should be supplemented by curators in order to restore the structure including clearly understandable stereochemistry from its backbone name. Another difficulty occurs when the published compound lacks a name, as above this also creates a difficulty for the chemical database curators. Of the compounds we have processed, 17.4% were classified in this category (Table 1). In most cases, newly reported compounds were not named but rather describe as a compound, numbered and a picture drawing provided. This type of curation problem occurred frequently in the previously reported compounds. When curation of such compounds is performed, an IUPAC name is given for database information by using Marvin Beans 15.11.9 of ChemAxon Ltd in our 3DMET team.

The remaining 20% of the target compounds to confirm (3.5% of compounds overall) derived from mistakes. The errors were attributed to discrepancy between name and structure, discrepancy between experimental data and picture drawing of compound structures, and so on (Table 1).

Correspondence errors were often observed between structure and name. The numbers of wrong structure, wrong name, or inverted sugar were 93, 43, and 77, respectively. As detailed in Table 1, simple mistakes to draw structures and special wrong-structures consisting of sugars drawn inverted were separately counted. The inverted structures of sugars were drawn in glycosides. The errors in the sugar structures may be caused by any program bugs to draw chemical structures because the similar mistakes were seen in the other works. Sometimes, it was observed that atomic chiral description (R or S) of the name did not correspond to the structure drawing. This is more often detected in newly reported compounds than in previously reported. One of the reasons seemed to be that authors use an IUPAC name of a previously reported compound as its partial structure name. That is possible for the newly reported structure without change of chirality at all position. However, when substitution is occurred at any position of the chiral atoms, atomic chirality may be inversed. The inversed stereochemistry of new compounds similar to previously reported ones should be carefully checked.

Not only correspondence between compound name and structure but also that of 2D-picture drawing and estimated structure from experimental data also sometimes did not correlate. Wrong descriptions were observed among nuclear magnetic resonance spectrum (NMR), mass spectrum (MS), circular dichroism (CD) and X-ray analysis. However, overall such miss-drawings were much less than the other errors.

Mistakes were also observed in the text. Not considering the omission of common information, the most frequent mistake was the miss-copied structure from its former report (5.5% of all compounds to require confirmation). These compound structures were revised in consideration of the pre-

vious reports. Misspellings were also revised in a similar manner. In addition, it was observed as rare examples that reported structures could not become chemically appropriate three-dimensional structures (“distortion in generated 3D-structure”). Therefore, skillful chemical curators who can identify errors and determine the best course of treatment are very important to streamline this process.

Accuracy of structures in literature

To increase data reliability, expert curators are required to discern literature contents. To estimate the accuracy of literature, journal articles describing newly found natural compounds were investigated. Table 2 shows investigated journals and reported compounds including those which are newly isolated and those which were being updated (806 articles, 2671 new or updated compounds, and 2175 compounds with chiral atoms). For literature resources related to natural products, Journal of Natural Products by American Chemical Society published the most articles. A total of 173 articles describing 892 compounds (721 chiral compounds out of them, 84.2%) in the journal. The percentage of chiral compounds out of the newly reported compounds was 82.8%. Highlighting again atom, chirality is an important factor in consideration of structures of natural products.

Because of importance of chirality, determination of absolute configuration for each compound is needed. When a chemical structure is published, the structure is generally accepted as being correct. However, the structure may be drawn as a relative configuration. Stereochemistry of undefined chiral atoms is usually drawn with either a wavy line or a straight line. A wavy line between two atoms clearly means undefined chirality and can be easily recognized by readers. An explicitly represented molecular structure with stereochemistry will be recognized as its absolute configuration. However, the structure sometimes means one of the relative configurations. For curation, we need to estimate such kinds of the structures from the text and the data of the articles to avoid excluding any possible configuration.

From the above background, the rate to detect absolute configuration was investigated from the articles described in NPU issues of 2013. Number and percentage of defined configuration are shown in Figure 2. Not all compounds were completely defined as absolute or relative, and partially defined compounds were also observed. Compounds consisting of absolute configuration part, relative configuration part and undefined part were not observed in this sample article set. They were rarely found in the other years. In this investigated article set, 1,694 compounds completely defined absolute configuration were reported (77.9% of all newly reported compounds). In the remaining compounds, 166 compounds of which all atoms were defined as relative configurations could be shown as one of the relative configurations. In addition, the stereochemistry of 304 compounds (totally 13.9%, shown as “atoms undefined configuration”) were partially defined or not defined at all.

Table 2 Journals investigated as literature resources about structure definition of natural products

Journal	Articles	New compounds	
		Total	With chiral center
J. Nat. Prod.	173	892	721
Fitoterapia	96	288	228
Nat. Prod. Commun.	83	150	98
Tetrahedron Lett.	71	162	147
Helv. Chim. Acta	70	206	175
Org. Lett.	62	126	112
Bioorg. Med. Chem. Lett.	59	169	130
Tetrahedron	38	162	141
Chem. Pharm. Bull.	34	146	105
Chem. Biodiversity	28	92	75
Heterocycles	22	52	39
Biosci. Biotechnol. Biochem.	18	35	25
Eur. J. Org. Chem.	17	101	97
Org. Biomol. Chem.	6	38	35
RSC Adv.	5	14	14
Chem. Eur. J.	4	6	6
Nat. Prod. Res.	4	5	4
J. Med. Chem.	3	3	3
Nat. Prod. Sci., Korea	3	3	3
Angew. Chem., Int. Ed.	2	9	9
Bull. Chem. Soc. Jpn.	2	4	4
Phytother. Res.	2	3	1
Chem. Commun.	1	1	1
Chem. Lett.	1	1	1
J. Am. Chem. Soc.	1	1	1
Z. Naturforsch., B: Chem. Sci.	1	2	0
Total	806	2671	2175

The articles were selected based on the database, Natural Product Updates on 2013.

Next, we determined relationship between the methods to define stereochemistry and the success rate of it, in order to estimate possible application to curation. All reported compounds were analyzed by 1D-NMR and MS. In addition, the techniques of X-ray, NMR, CD, and optical rotatory dispersion (ORD) as shown in Table 3 were used to define stereochemistry of the compounds. X-ray and NMR are essential

to define connection of atoms. The methods of CD, ORD and Mosher's method (usually modified Mosher's method [16]) are used to obtain additional information about stereochemistry. For example, after connectivity of atoms and/or relative configuration is defined by NMR, the configuration of chiral atoms is generally evaluated by CD or ORD. The modified Mosher's method will be also utilized if applicable

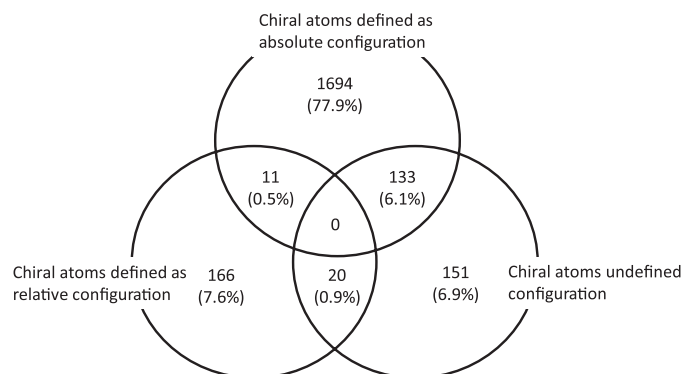


Figure 2 Defined chiral atom configurations of a molecule shown by Venn's diagram. Overlap of two configurations means that a molecule includes two parts of configurations.

Table 3 Success rate of defining absolute configuration by each analytical method

Methods	Compound number		Rate (%)	
	Completely defined	All		
X-ray	158	171	92.4	
NMR	NOESY/ROESY	1298	1625	79.9
	Mosher's method	110	126	87.3
Circular dichroism (CD)	486	523	92.9	
Optical rotatory dispersion (ORD)	193	201	96.0	

NMR, nucleic magnetic resonance; NOESY, nuclear Overhauser enhancement and exchange spectroscopy; and ROESY, rotating frame nuclear Overhauser effect spectroscopy. The number of compounds completely defined by each method and the number of all compounds analyzed by the method are listed. Multiple methods could be applied to one compound (refer to Table 4).

to the molecule. Thus, multiple methods listed in Table 3 could be applied to one compound. Detailed combinations of techniques are indicated in Table 4.

As shown in Table 3, X-ray analysis permitted 92.4% of absolute structures. It is noteworthy that about 8% of structures detected by X-ray analysis did not define the stereochemistry. This is also true for NMR. Even if NOESY and/or ROESY analysis are added, about 20% of absolute configuration cannot be determined. The CD and ORD methods cannot define configurations by themselves. However, by combining with X-ray and NMR, over 90% of molecules can be determined as absolute configurations. X-ray analysis is the most certain method to determine absolute configuration. However, the frequency by NMR (NOESY and ROESY) is 7.5 fold of that by X-ray analysis (Table 4). It may be caused by the difficulty of crystallization. As the results, classification of description about methods in articles could allow to determine whether absolute configuration of chemical structure was defined. It may be developed to a support system of chemical curation.

Information about next release of 3DMET

As described above, 3DMET is currently being developed as a fully curated database. The total number of in-house curated chemical compounds is more than 30,000, for which duplication has not yet been accessed. We are currently working to identify duplicates among them. InChI is being used with verification by curators. Following this, we will update the 3DMET database with 13,000 new entries including the ROMPP set as a new release. The remaining structures will be added in several times as soon as it is organized.

To distinguish the earlier not manual curated data of the 3DMET database, the new curated data entries are given IDs beginning with an L and followed by a continuous five digit number similar to the entries of B series of version 2. The new entries will be published as L series, meaning natural compounds fully curated from literature.

Table 4 Combination of analytical methods to determine absolute configuration

Analytical methods					Number
X-ray	NMR		CD	ORD	
	NOESY or ROESY	Mosher's method			
✓					40
✓	✓				79
✓	✓	✓			1
✓	✓		✓		20
✓	✓			✓	5
✓		✓			1
✓			✓		6
✓				✓	3
✓			✓	✓	3
	✓				708
	✓	✓			57
	✓		✓		301
	✓			✓	71
	✓	✓	✓		31
	✓		✓	✓	24
	✓	✓		✓	1
		✓			19
			✓		81
				✓	66
			✓	✓	20
Total					1537

All reported compounds were analyzed by 1D-NMR and MS. Number indicates the number of compounds defined by one or multiple checked methods on the line. An abbreviation of each method is the same as Table 3. Determination by minor techniques is not listed here and the total number of each method is different from the number of "defined as absolute configuration" in Figure 2.

Conclusion

In development of an accurate chemical compound database, the greatest difficulty is to obtain good chemical curators. Because corrections are often needed to the structure or to its name provided in the original published text, curators should be chemically educated. This is a particular problem for the development of databases containing chemical structure information. Frequently, structure data include incorrect descriptions involving: their names, such as misspelling; their structures, such as drawing errors; or stereochemically undefined structures presented as absolute structures in the figures. To avoid errors from such compound structure reporting inconsistencies, good chemical curators are most essential for development of accurate database.

Acknowledgment

We thank Ms. Noriko Noguchi and Ms. Mariko Watanabe for contribution as chemical curators. Special thanks go to Ms. Mikiko Takahashi for critical reading of this manuscript

and her advice from the view of a natural products chemist. I sincerely acknowledge to Dr. Franz J. St. John of USDA Forest Service for his detailed English editing.

This research was supported by MEXT/JSPS KAKENHI (Grant Number 218065, 238062, 248057, and 257008) and Research Grant from NARO Gender Equality Program. The computer environment including MOE 2011 of Agriculture, Forestry and Fisheries Research Information Technology Center (AFFRIT), Ministry of Agriculture and Forestry and Fisheries (MAFF) was used. The 3DMET database is opened to access from the virtual laboratory service of AFFRIT.

Conflicts of Interest

The authors declare no competing financial interest.

Author Contributions

MM directed the entire project. MM and TY co-wrote this manuscript. MM analyzed factors about data duplication. TY investigated description about structure analysis in literature. TK investigated suspicious description in literature.

References

- [1] Goto, S., Okuno, Y., Hattori, M., Nishioka, T. & Kanehisa, M. LIGAND: database of chemical compounds and reactions in biological pathways. *Nucleic Acids Res.* **30**, 402–404 (2002).
- [2] Groom, C. R., Bruno, I. J., Lightfoot, M. P. & Ward, S. C. The Cambridge Structural Database. *Acta Crystallogr B Struct. Sci. Cryst. Eng. Mater.* **72**, 171–179 (2016).
- [3] Kawabata, T., Sugihara, Y., Fukunishi, Y. & Nakamura, H. LigandBox: A database for 3D structures of chemical compounds. *Biophysics* **9**, 113–121 (2013).
- [4] Nakamura, K., Shimura, N., Otabe, Y., Hirai-Morita, A., Nakamura, Y., Ono, N., *et al.* KNApSACk-3D: a three-dimensional structure database of plant metabolites. *Plant Cell Physiol.* **54**, e4 (2013).
- [5] Irwin, J. J., Sterling, T., Mysinger, M. M., Bolstad, E. S. & Coleman, R. G. ZINC: A Free Tool to Discover Chemistry for Biology. *J. Chem. Inf. Model.* **52**, 1757–1768 (2012).
- [6] Bolton, E. E., Chen, J., Kim, S., Han, L., He, S., Shi, W., *et al.* PubChem3D: a new resource for scientists. *J. Cheminform.* **3**, 32 (2011).
- [7] Milne, G. W., Nicklaus, M. C., Driscoll, J. S., Wang, S. & Zaharevitz, D. National Cancer Institute Drug Information System 3D database. *J. Chem. Inf. Comput. Sci.* **34**, 1219–1224, (1994).
- [8] Meng, X.-Y., Zhang, H.-X., Mezei, M. & Cui, M. Molecular Docking: A powerful approach for structure-based drug discovery. *Curr. Comput. Aided Drug Des.* **7**, 146–157 (2011).
- [9] Pagadala, N. S., Syed, K. & Tuszynski, J. Software for molecular docking: a review. *Biophys. Rev.* **9**, 91–102 (2017).
- [10] Jain, A. N. Surflex-Dock 2.1: robust performance from ligand energetic modeling, ring flexibility, and knowledge-based search. *J. Comput. Aided Mol. Des.* **21**, 281–306 (2007).
- [11] Rarey, M., Kramer, B., Lengauer, T. & Klebe, G. A fast flexible docking method using an incremental construction algorithm. *J. Mol. Biol.* **261**, 470–489 (1996).
- [12] Maeda, M. & Kondo, K. Three-Dimensional Structure Database of Natural Metabolites (3DMET): A Novel Database of Curated 3D Structures. *J. Chem. Inf. Model.* **53**, 527–533 (2013).
- [13] Maeda, M. Current Challenges in Development of a Database of Three-Dimensional Chemical Structures. *Front. Bioeng. Biotechnol.* **3**, 66 (2015).
- [14] Weininger, D. SMILES, a chemical language and information system. 1. Introduction to methodology and encoding rules. *J. Chem. Inf. Comput. Sci.* **28**, 31–36 (1988).
- [15] Heller, S., McNaught, A., Stein, S., Tchekhovskoi, D. & Pletnev, I. InChI—the worldwide chemical structure identifier standard. *J. Cheminform.* **5**, 7 (2013).
- [16] Ohtani, I., Kusumi, T., Kashman, Y. & Kakisawa, H. High-field FT NMR application of Mosher's method. The absolute configuration of marine terpenoids. *J. Am. Chem. Soc.* **113**, 4092–4096 (1991).

This article is licensed under the Creative Commons Attribution-NonCommercial-ShareAlike 4.0 International License. To view a copy of this license, visit <https://creativecommons.org/licenses/by-nc-sa/4.0/>.

