



LOGIC STRUCTURE DETERMINATION (LSD) AS A COMPUTER ASSISTED STRUCTURE ELUCIDATION (CASE) FOR MOLECULAR STRUCTURE DETERMINATION OF CYTOTOXIC CEMBRANOIDS FROM SOFT CORAL

Hedi Indra Januar^{1,2}, Neviaty Putri Zamani^{2*}, Dedi Soedharma², and Ekowati Chasanah¹

¹Research and Development Center for Marine and Fisheries Product Competitiveness and Biotechnology, Jalan KS. Tubun Petamburan VI, Slipi, Central Jakarta 10260, Indonesia

²Department of Marine Science, Faculty of Fisheries and Marine Science, Bogor Agricultural University, Indonesia, Jalan Raya Dramaga Kampus IPB, Bogor 16680, Indonesia

Article history:

Received: 18 January 2016; Revised: 20 April 2016; Accepted: 25 April 2016

Abstract

Indonesian tropical soft corals are valuable resources that produce pharmacological cytotoxic cembranoids. However, the manual structure determination in these compounds requires adequate knowledge of organic chemistry. This study presents the application of Logic Structure Determination (LSD) as a freeware Computer Assisted Structure Elucidation (CASE) for automatic molecular structure determination of cembranoid compounds from soft corals species. 12 Nuclear Magnetic Resonance (NMR) datasets of cytotoxic cembranoids were used to evaluate the accuracy of LSD in generating the possible structures. The results of this study shows that LSD generated numerous possible molecular structures as the data input files were only derived from 2D-NMR HMQC (Heteronuclear Multiple-Quantum Correlation), COSY (Homonuclear Correlation Spectroscopy), and HMBC (Heteronuclear Multiple Bond Correlation). The accuracy was significantly enhanced (only 2-4 possible cembranoid structures from each NMR dataset) with the addition of H2BC (Heteronuclear 2 Bond Correlation) experimental data. This may indicate that VLRC (Very Long-Range Correlations) significantly affects LSD capability. Furthermore, LSD with direct 2 bond NMR experimental data is a reliable CASE technique for cembranoid compounds identification. In general, this freeware-CASE has the potential to be applied on other types of small molecule compounds and may serves as a solution for elucidation bottleneck step in studies on Indonesian natural products.

Keywords: cembranoid, computer assisted structure elucidation, logic structure determination, soft coral

1. Introduction

Soft coral is a marine organism that produces metabolites through pharmacological activity. There are 2,628 new chemical compounds from soft corals reported by various researchers between 1990 and 2009 (Leal et al., 2012). A review on soft coral metabolites shows that 61% of all secondary soft corals metabolites is terpenoid class, commonly the cembranoid subclass compound (Rocha et al., 2011). Manual elucidation through this type of natural product needs a careful inspection of spectroscopic Nuclear Magnetic Resonance (NMR) dataset. Numerous combinations of functional groups, such as hydroxyl, carbonyl, epoxy, and lactone, may be attached to the main cembranoid ring system in several positions.

Therefore, manual molecular structure determination of cembranoids requires careful inspections by skillful researchers.

Even though international collaborations have shown increasing new findings on bioactive compounds from marine resources, molecular structure elucidation capacity in Indonesian marine natural products study has not shown significant increase over the last two decades. A bibliographic study shows that only 8 Indonesian scientists contributed as the main principal authors (1st author) in international publications in the theme of the new findings of natural products from Indonesian sponges and soft coral between 1994 and 2015 (Januar, 2016a). This becomes a problem because information on

*Corresponding author.
E-mail: npzamani@gmail.com

metabolites structure is important to evaluate the potency of new drug candidates (Ahlqvist et al., 2015). Therefore, a breakthrough is needed to increase new findings of natural products from Indonesian marine resources, particularly from soft corals species.

Computer Assisted Structure Elucidation (CASE) is a potential technique to solve the complexity and to improve timescales in molecular structure determination process (Blinov et al., 2001; Reynolds & Mazzola, 2015). This program aims to generate all possible structures that are consistent with a particular set of spectroscopic data with minimum amount of human intervention (Jaspars, 1999) as there are evidences of incorrect natural products structures in several publications (Elyashberg et al., 2010). Furthermore, CASE may serves as a solution for limited connectivity of 2D NMR information that may not be able to be solved through manual elucidation (Blinov et al., 2003a). An example of this is quindolinocryptotackieine, a complex $C_{31}H_{20}N_4$ alkaloid. The structure of this compound was elucidated by CASE after a decade effort of manual elucidation by experienced spectroscopist (Blinov et al., 2003b).

One of the most reviewed CASE is StrucEluc, a commercial software from ACD Labs. StrucEluc has proven itself to be a powerful tool in automatic computer assisted structure elucidation (Elyashberg et al., 2002; Jayaseelan & Steinbeck, 2014). However, as a commercial application, the cost of using this software may beyond the average marine natural products laboratory in Indonesia. As an alternative, recent developments by various scientists have produced several freeware CASE techniques under General Public License (GPL). One among these is Logic Structure Determination (LSD) software. So far, previous studies have shown that LSD successfully elucidates flavonoid, alkaloid, and steroid compounds (Mulholland et al., 2006; Elyashberg et al., 2010; Hubert et al., 2015). However, none of those studies had been applied to cembranoid class. This study presents the application of LSD as CASE technique for automatic structure elucidation on cembranoid compounds from soft corals species. The study aims to examine the accuracy of LCD in generating cembranoid structures from 1D and 2D NMR dataset.

2. Material and Methods

The structures of 12 cembranoids from *Nephthea* sp., *Sarcophyton* sp., and *Sinularia* sp. collected from Banda, Bunaken, and Mahengetang Island were used in this study (Figure 1). Figure 1 also shows the main connectivity of each atom in the structure, which was derived from 2D-NMR dataset and was previously used

to manually elucidate the cembranoids structures. All NMR spectra were recorded in our previous study (Chasanah et al., 2014a; 2014b; Januar et al., 2016b) at 3.5-14.3 mg/ml cembranoids (in 0.7 ml deuterated chloroform) on Jeol ECS 400 MHz NMR spectrometer, with spectra referenced to residual 1H and ^{13}C resonances in the deuterated solvents.

The LSD software version 3.4.9 was downloaded from http://eos.univ-reims.fr/LSD/index_ENG.html. The LSD data input file of each cembranoid (based on NMR data) was carefully made according to Nuzillard (2013). The data input file of each cembranoid was made in two versions. The first was a data input file that contains HMQC (Heteronuclear Multiple-Quantum Correlation), COSY (Homonuclear Correlation Spectroscopy), and HMBC (Heteronuclear Multiple Bond Correlation) data NMR experiments. The second was a data input file that contains HMQC, COSY, HMBC, and H2BC (Heteronuclear 2 Bond Correlation) 2D-NMR experiments. This design was aimed to analyze the effects of VLRC (Very Long-Range Correlations), which usually observes COSY and HMBC experiments, on LSD's capabilities to generate the possible structures.

3. Results and Discussion

This study used several types of cembranoids from three soft coral families: hydroxy-cebranoids from *Nephthea* sp., oxy-cebranoids from *Sarcophyton* sp., and lactone-cebranoids from *Sinularia* sp. Genotype analyses by other studies have shown that although soft coral species produce the same type of cembranoid as 14 carbon backbone ring, the variation in functional groups may serve as a chemotype for each soft coral species (Aratake et al., 2012). All of these cembranoids were found to be toxic for tumor cell lines (MCF-7, HeLa, or T47D) in our previous study, with IG_{50} under 50 mg/l (Chasanah et al., 2014a, 2014b; Januar et al., 2016b).

LSD relies on the connectivity data found in 2D-NMR without any reference to chemical shift database (Elyashberg, 2015). This CASE works differently from the commercial ACD StrucEluc, which can automatically import NMR data directly from spectrometer and can count the spectroscopic chemical shift into the calculation of possible structure building. The principle of this software is defining all $^1H^1H$ and $^1H^{13}C$ through-bond chemical shift correlations detected in HMQC, COSY, HMBC, H2BC, or other 2D-NMR spectra of the target compound (Hubert et al., 2015). The user has to define and to input each hybridization state, the number of attached hydrogens, and the electric charge as well as the correlation (by COSY, HMBC, and H2BC, or other

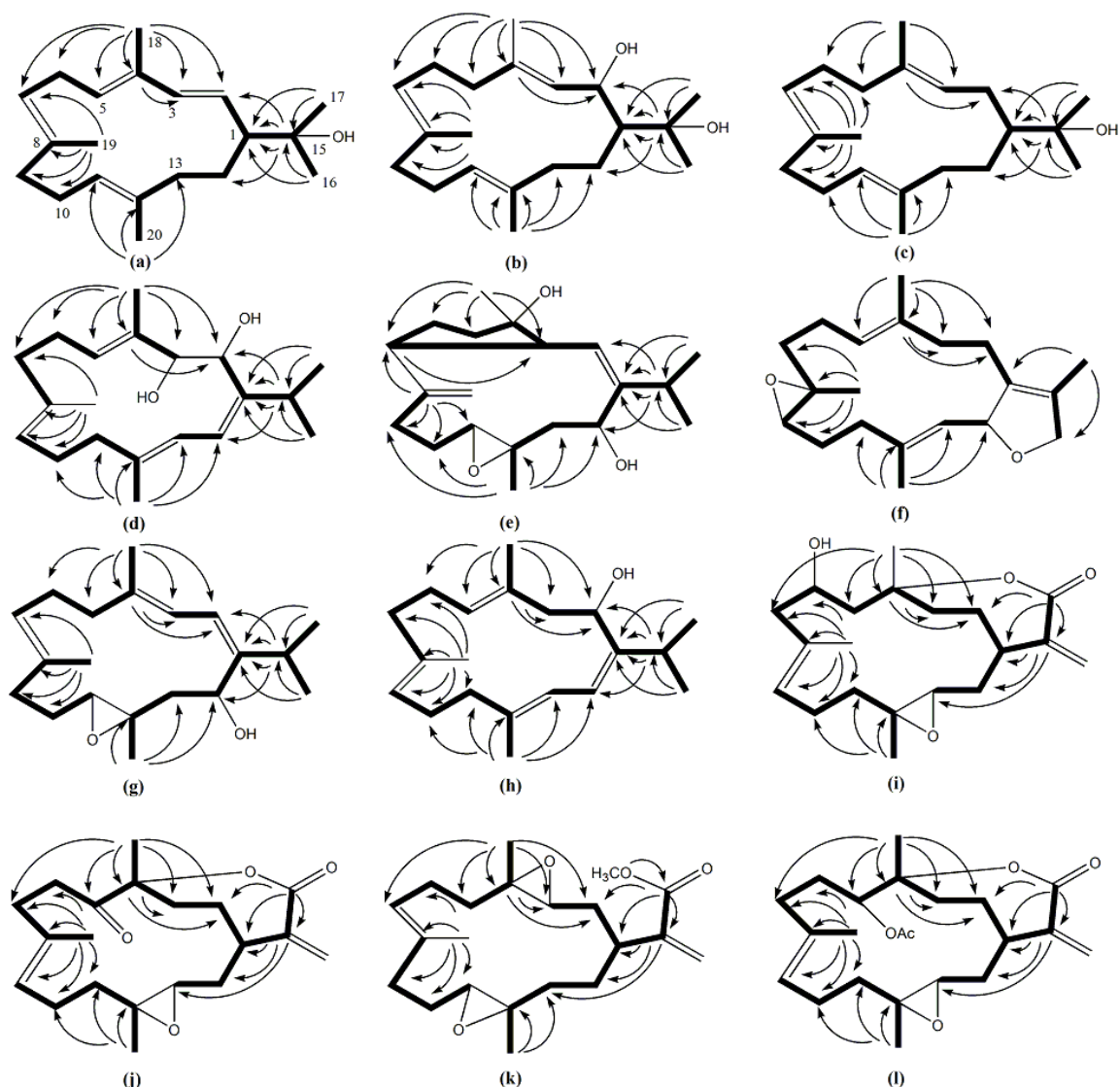


Figure 1. Cytotoxic cembranoid structures from Indonesian *Nephthea* sp. (a-c), *Sarcophyton* sp. (d-h) and *Sinularia* sp. (i-l) with 'arrow' represent COSY (Homonuclear Correlation Spectroscopy) and HMBC (Heteronuclear Multiple Bond Correlation), meanwhile 'bold bond' represent H2BC (Heteronuclear 2 Bond Correlation) (Chasanah et al., 2014a; 2014b; Januar et al., 2016b).

correlation experiment in 2D-NMR) of particular atom in the targeted molecule on this software manually. As a free CASE, this study suggests that the manual input of the dataset is easy to be conducted and only needs basic knowledge of organic chemistry.

The verification into LSD software accuracy showed that this program generated 20-94 possible structures without H2BC experimental data and 2-4 possible structures with H2BC experimental data (Table 1). An important limitation of NMR-based structure elucidation, with manual elucidation or computer assist, is VLCR. VLCR may exist in COSY and HMBC through more than 3 (HMBC) or 4 (COSY) chemical bonds (Mulholland et al., 2006). This study

shows that VLCR significantly affects the LSD calculation to generate the possible structures. In cembranoids, the circular ring without any N or O in the main carbon backbone may affect numerous VLCR in COSY or HMBC spectra. Figure 1 shows the common VLCR on cembranoids from C-18 to C-7, which in manual elucidation may confuse the actual position of C-7. There are solutions to avoid VLCR effects. First, user can decide manually whether each correlation is a VLRC or not based on the qualitative estimation of their intensity (Mulholland et al., 2006). However, as shown in Table 1, the use of HMQC, COSY, and HMBC with the estimation of small intense correlation as VLCR results in a numerous possible

Table 1. Possible structure solutions of 12 cembranoids from soft corals by CASE-LSD (Computer Assisted Structure Elucidation-Logic Structure Determination) calculation from experimental 2D-NMR (2 Dimensions-Nuclear Magnetic Resonance) HMQC (Heteronuclear Multiple-Quantum Correlation), COSY (Homocuclear Correlation Spectroscopy), HMBC (Heteronuclear Multiple Bond Correlation), and H2BC (Heteronuclear 2 Bond Correlation) dataset

Cembranoid Compound	Possible Structure Solution from CASE-LSD	
	By HMQC, COSY, and HMBC data	By HMQC, COSY, HMBC, and H2BC data
a 15-hydroxy-Cembrene	33	2
b 2-Hydroxy-Nephtenol	94	3
c Nephtenol	24	3
d Sarcophytol B	32	4
e Sarcophytolide	45	4
f Sarcophytoxide	20	2
g 11,12-epoxy-Sarcophytol A	35	4
h Sarcophytol A	57	3
i Sinulariolide	20	3
j 5-dehydro-Sinulariolide	22	4
k Flexilarin B	24	3
l Episinularidide Acetate	22	4

structures. It needs careful inspections to manually remove VLRC one by one in COSY or HMBC spectra. An experienced chemist may be able to do this, but it is considered as a time-consuming task for a beginner.

The most potential solution is to use 2 bond correlation NMR data, such as 1.1 ADEQUATE (Abundance Double Quantum Transfer Experiment), INADEQUATE (Incredible Natural Abundance Double Quantum Transfer Experiment), and H2BC (Plainchont et al., 2013). ADEQUATE and INADEQUATE show direct two bond correlation between neighboring carbons. However, both of these experiments need larger amount of samples and longer acquisition time compared to the usual 2D-NMR experiment because the natural viability of ^{13}C is very small. It may also needs high frequency NMR machine. In our previous study, a 400 MHz NMR did not generate sophisticated INADEQUATE spectra (5-10 mg samples). Meanwhile, H2BC is a 2D-NMR method that shows HMBC-type spectrum, almost exclusively only two bond correlations spectra (Nyberg et al., 2005). This experiment relatively has lower demands of NMR

equipment than direct ^{13}C - ^{13}C experiments. Our 400 MHz NMR can generate sophisticated H2BC spectra from 3-5 mg sample.

Figure 2 shows an example of HMBC and H2BC experiments from 11,12-epoxy-sarcophytol A. A VLRC in HMBC appears from H-14 to C-10 and C-11, and also from H-5 to C-1. Signal intensity in HMBC is not regarded to be a reliable parameter to deduce VLRC correlation or 2-3 bond correlation. With H2BC data, deducing the neighboring atom can be done by overlaying the H2BC on HMBC spectra. The interpretation from H2BC significantly enhances the accuracy of LSD in predicting correct cembranoid structures. An example of LSD results with H2BC experimental data (15-hydroxy-Cembrene and 2-hydroxy-Nephtenol) is shown in Figure 3. With H2BC data, the possible structures are detected to be different within isomer, only in the variation at two double bonds assignment. The amount of possible structures increased with the amount and the complexity of functional groups in the cembranoid backbone ring system.

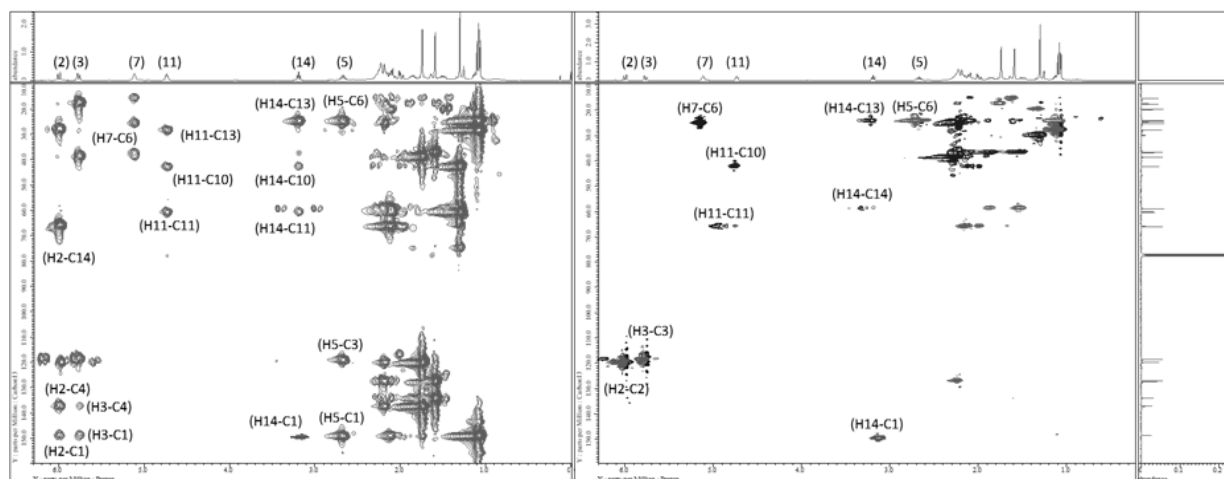


Figure 2. Spectra of HMBC-Heteronuclear Multiple Bond Correlation (a), and H2BC-Heteronuclear 2 Bond Correlation (b) from 11,12-epoxy-Sarcophytol A.

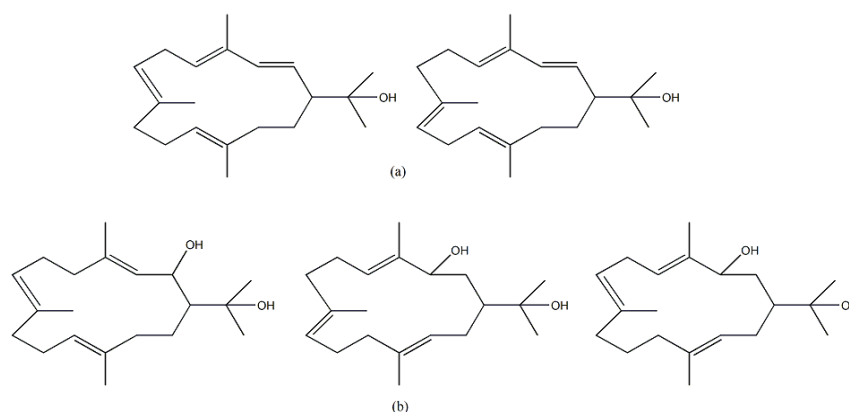


Figure 3. LSD Results from HMQC (Heteronuclear Multiple-Quantum Correlation), COSY (Homocorrelation Spectroscopy), HMBC (Heteronuclear Multiple Bond Correlation), and H2BC (Heteronuclear Bond Correlation) dataset of 15-hydroxy-Cembrane (a) and 2-hydroxy-Nephttenol (b).

In general, this study shows that the application of LSD with H2BC experiment is a reliable CASE for elucidating cembranoids from soft corals. Although it is not as powerful as the commercial CASE and still needs user interferences to assign the properties of each atom, this elucidation method can generate a sophisticated results. The correct structure can be selected among 2-4 possible structures by searching in the open natural products database, such as chemspider (<http://www.chemspider.com>) or NCBI (National Center for Biotechnology Information) database (<https://pubchem.ncbi.nlm.nih.gov>). Other studies have also shown that LSD may be applied on other types of compounds, such as flavonoid, steroids, and alkaloids (Mulholland et al., 2006; Elyashberg et al., 2010; Hubert et al., 2015), and that the application of LSD with direct 2 bond NMR experimental data is a reliable CASE technique to identify small molecule bioprospecting compounds in natural products studies.

4. Conclusion

LSD with direct 2 bond NMR experimental data is a reliable CASE technique to identify cembranoid compounds. This GPL-CASE may also be applied on other types of small molecule bioprospecting compounds and may serve as a potential solution for elucidation bottleneck step in Indonesian marine natural products studies.

References

- Ahlqvist, M., Leandersson, C., Hayes, M.A., Zamora, I. & Thompson, R.A. (2015). Software aided structural elucidation in drug discovery. *Rapid Communications in Mass Spectrometry*, 29(21), 2083-2089.
- Aratake, S., Tomura, T., Saitoh, S., Yokokura, R., Kawanishi, Y., Shinjo, R., Reimer, J.D., Tanaka, J. & Maekawa, H. (2012). Soft coral *Sarcophyton* (Cnidaria: Anthozoa: Octocorallia) species diversity and chemotypes. *PLoS One*, 7(1), p.e30410.

- Blinov, K.A., Elyashberg, M.E., Molodtsov, S.G., Williams, A.J. & Martirosian, E.R. (2001). An expert system for automated structure elucidation utilizing ^1H - ^1H , ^{13}C - ^1H and ^{15}N - ^1H 2D NMR correlations. *Fresenius' Journal of Analytical Chemistry*, 369(7-8), 709-714.
- Blinov, K. A., Carlson, D., Elyashberg, M. E., Martin, G. E., Martirosian, E. R., Molodtsov, S., & Williams, A. J. (2003a). Computer assisted structure elucidation of natural products with limited 2D NMR data: application of the StrucEluc system. *Magnetic Resonance in Chemistry*, 41(5), 359-372.
- Blinov, K. A., Elyashberg, M. E., Martirosian, E. R., Molodtsov, S. G., Williams, A. J., Sharaf, M. M. H., Schiff, P. L., Jr., Crouch, R. C., Martin, G. E., Hadden, C. E., Guido, J. E., & Mills, K. A. (2003b). Quindolinocryptotackeine: the elucidation of a novel indoloquinoline alkaloid structure through the use of computer assisted structure elucidation and 2D NMR. *Magnetic Resonance in Chemistry*, 41(8), 577-584.
- Elyashberg, M.E., Blinov, K.A., Williams, A.J., Martirosian, E.R. & Molodtsov, S.G. (2002). Application of a new expert system for the structure elucidation of natural products from their 1D and 2D NMR data. *Journal of Natural Products*, 65(5), 693-703.
- Elyashberg, M., Williams, A.J. & Blinov, K. (2010). Structural revisions of natural products by Computer-Assisted Structure Elucidation (CASE) systems. *Natural Product Reports*, 27(9), 1296-1328.
- Elyashberg, M. (2015). Identification and structure elucidation by NMR spectroscopy. *Trends in Analytical Chemistry*, 69, 88-97.
- Hubert, J., Chollet, S., Purson, S., Reynaud, R., Harakat, D., Martinez, A., Nuzillard, J.M. & Renault, J.H. (2015). Exploiting the Complementarity between dereplication and Computer-Assisted Structure Elucidation for the Chemical Profiling of Natural Cosmetic Ingredients: Tephrosia purpurea as a Case Study. *Journal of Natural Products*, 78(7), 1609-1617.
- Chasanah, E., Kawung, N. J., Rumampjk, N.D., Mangindaan, R., Mantiri, D., & Januar, H.I. (2014a). Characterization of bioactive substances from Manado waters soft coral. *Technical Research Report*. Indonesian Research and Development Center for Marine and Fisheries Product Competitiveness and Biotechnology.
- Chasanah, E., Januar, H.I., Marraskuranto, E., Susilowati, R., & Pratitis E. (2014b). Effects of climate change into bioactive substances in marine invertebrate from CTI regions. *Technical Research Report*. Indonesian Research and Development Center for Marine and Fisheries Product Competitiveness and Biotechnology.
- Januar, H.I. (2016a). Overview of Authorship Trends at International Biodiscovery Collaboration Research to Indonesian Sponge and Soft Coral Over the Last Two Decades. *Journal of Scientometrics Research*. *In press*.
- Januar, H.I., Zamani, N.P., Soedharma, D., & Chasanah, E. (2016b). Changes in Allelochemical Cembranoid Composition of Soft Coral *Sarcophyton* sp. at Volcanic CO_2 Seeps in Mahengetang Island Indonesia. *Chemistry and Ecology*. *In press*.
- Jaspars, M. (1999). Computer assisted structure elucidation of natural products using two-dimensional NMR spectroscopy. *Natural Product Reports*, 16(2), 241-248.
- Jayaseelan, K.V. & Steinbeck, C. (2014). Building blocks for automated elucidation of metabolites: natural product-likeness for candidate ranking. *BMC Bioinformatics*, 15(1), 234-237.
- Leal, M. C., Puga, J., Serôdio, J., Gomes, N.C.M., & Calado, R. (2012). Trends in the discovery of new marine natural products from invertebrates over the last two decades—where and what are we bioprospecting? *PloS One*, 7(1), p.e30580.
- Mulholland, D.A., Langlois, A., Randrianarivoelosia, M., Derat, E. & Nuzillard, J.M. (2006). The structural elucidation of a novel iridoid derivative from *Tachiadenus longiflorus* (Gentianaceae) using the LSD programme and quantum chemical computations. *Phytochemical Analysis*, 17(2), pp.87-90.
- Nuzillard, J. M. (2013) LSD 4.3.8. http://eos.univ-reims.fr/LSD/MANUAL_ENG.html. Access: 16 January 2016.
- Nyberg, N.T., Duus, J.Ø. & Sørensen, O.W. (2005). Heteronuclear two-bond correlation: suppressing heteronuclear three-bond or higher NMR correlations while enhancing two-bond correlations even for vanishing 2 J CH. *Journal of the American Chemical Society*, 127(17), 6154-6155.
- Plainchont, B., Emerenciano, V. & Nuzillard, J.M. (2013). Recent advances in the structure elucidation of small organic molecules by the LSD software. *Magnetic Resonance in Chemistry*, 51(8), 447-453.
- Reynolds, W.F. & Mazzola, E.P. (2015). Nuclear magnetic resonance in the structural elucidation of natural products. *In Progress in the Chemistry of Organic Natural Products*, 223-309. Springer International Publishing.
- Rocha, J., Peixe, L., Gomes, N. & Calado, R. (2011). Cnidarians as a source of new marine bioactive compounds—An overview of the last decade and future steps for bioprospecting. *Marine Drugs*, 9(10), 1860-1886.