Mol Divers (2007) 11:115–118
DOI 10.1007/s11030-007-9065-5

FULL LENGTH PAPER

Natural macrocyclic molecules have a possible limited structural diversity

Aaron T. Frank · Nicola S. Farina · Nahed Sawwan · Orrette R. Wauchope · Mo Qi · Edyta M. Brzostowska · Wang Chan · Frank W. Grasso · Paul Haberfield · Alexander Greer

Received: 2 December 2006 / Accepted: 25 October 2007 / Published online: 20 November 2007

Abstract This paper examines ring size patterns of natural product macrocycles. Evidence is presented that natural macrocycles containing 14-, 16-, and 18-membered rings are of frequent occurrence based on a data mining study. The results raise a question about the limited diversity of macrocycle ring sizes and the nature of the constraints that may cause them. The data suggest that the preference bears no relationship to the odd–even frequency in natural fatty acids. The trends reported here, along with those reported previously (Wessjohann et al. (2005) Mol Divers 9:171), may be generalized to better understand the possible structure preferences of natural macrocycles.

Keywords Natural Product macrocycles · Ring sizes · Structure patterns

Introduction

The molecular diversity of natural product macrocycles is a topic of interest to many synthetic chemists and biochemists. The recent database analysis of Wessjohann et al. in 2005 highlighted natural macrocycles in terms of their ring size, molecular weight, and frequency of sub-structural motifs [1]. We believe that such studies are timely. There is an availability of data on natural macrocycles, which make it possible to examine possible structural preferences [2]. This led us to examine further whether macrocycles of natural product origin are of a limited diversity.

A. T. Frank · N. S. Farina · N. Sawwan · O. R. Wauchope · M. Qi · E. M. Brzostowska · W. Chan · P. Haberfield · A. Greer (✉)
Departments of Chemistry, The Graduate Center and City University of New York (CUNY), Brooklyn College, Brooklyn, NY 11210, USA
e-mail: agreer@brooklyn.cuny.edu

F. W. Grasso
Department of Psychology, The Graduate Center and City University of New York (CUNY), Brooklyn College, Brooklyn, NY 11210, USA

Experimental section

Some 4,000 natural products were surveyed using the database Combined Chemical Dictionary (Chapman & Hall/CRC Press), SciFinder Scholar (Chemical Abstract Service), and primary literature sources from 43 scientific journals. Two compendia were also used: *Biosynthetic Products for Cancer Chemotherapy* (Plenum Press) and *Anticancer Drugs From Animals, Plants, and Microorganisms* (Wiley-Interscience) [3,4]. The molecular weight of the macrocycles ranged from a low of 220 amu to a high of 1,611 amu. Three separate individuals scored the patterns of ring size data without consulting one another. The reliability of using this method ranged from 95% to 98%. There is a significant difference between the frequencies of occurrence of odd and even macrocycles for many macrocycle sizes. Three analyses allowed us to evaluate the hypothesis that even and odd macrocycle ring sizes differed in their contributions to the observed deviation from homogeneity of frequency. For these analyses the critical values were adjusted to maintain a 0.01 significance level. A standardized residuals analysis indicated that macrocycles sizes 11, 12, 13, 14, 16, 18, 19, 20 and 24 occurred with frequencies that deviated significantly from a uniform distribution ($p < 0.001$). A comparison of the frequency of occurrence even and odd macrocycle sizes pooled across all observed lengths indicated even numbered macrocycles occurred with greater frequency ($\chi^2(1) = 23.03\ p \ll 0.001$) than odd numbered rings. A χ^2 test was conducted for independence between macrocycle ring size (short versus long

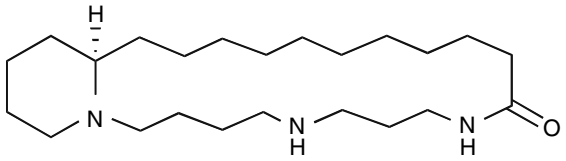

Fig. 1 (−)-Isooncinotine

split at ring size 26) and macrocycle parity (even–odd), where no evidence was found for the association between ring size and parity.

Results and discussion

The recognition of a structure pattern on the natural selection of macrocycles was sought. Thus, a data mining study was applied to explore ring size frequency in natural product macrocycles (macrocyclic secondary metabolites). Consideration of the number of ring atoms in the structure included saturated, unsaturated, and aromatic systems, which are monocyclic, some of which have small rings fused or contain bridges that connect rings. Ring size is counted according to the smallest internal macrocycle ($n \geq 11$). For example, (−)-isooncinotine [5,6], a macrocyclic spermidine alkaloid, is considered as a 22-membered ring (Fig. 1). For this study, two databases were used, the Chemical Combined Dictionary and SciFinder Scholar, along with information collected from 43 primary journal sources. About 4,000 natural products were examined from a collection of approximately 200,000 available compounds for which no systematic route exists for checking ring size. It is not possible to search the Chemical Combined Dictionary or SciFinder Scholar databases for all natural macrocycle candidates because odd or even rings cannot be counted reliably with automated search techniques. Our smaller scale search of 4,000 natural products, on the other hand, is found to be highly precise, and the statistical analysis suggests good accuracy. The counting of ring sizes has been checked and rechecked. We sought ring size information about macrocycle structures regardless of whether any reports existed on their bioactivity.

From our analysis, it is estimated that 92% of the literature sources used antibiotic screening as an impetus to determine the natural macrocycle structure. In contrast, 8% utilized chemical screening where the macrocycle structure was sought neglecting possible biological activity.

Figure 2 shows the number of macrocyclic compounds examined plotted against ring size. A periodicity is observed. Even size 14-, 16-, and 18-membered rings are preferred compared to odd-membered rings and compared to ring systems greater than 19 atoms in size. Natural 14-, 16-, and 18-membered rings predominate in heterocycle, carbocycle, and cyclic peptide structures derived from terrestrial and

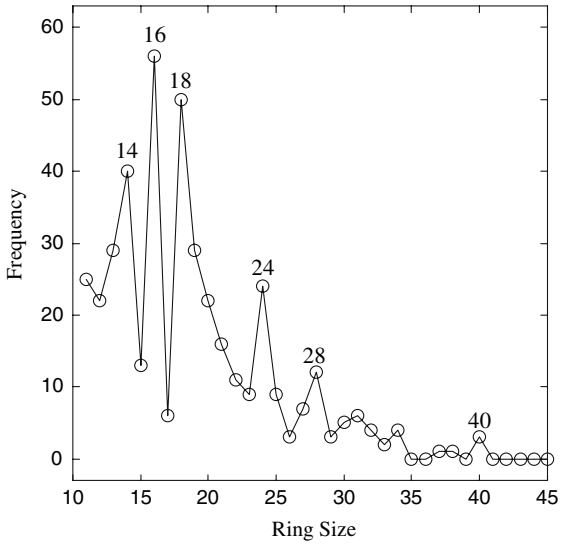

Fig. 2 Survey of the abundance of macrocycle ring sizes in natural products

marine sources. This analysis is used to gauge whether natural macrocycles contain repetitive structures reaching similar conclusions as reported previously [1].

Interestingly, the structural features of many of the compounds show an origin independent of a two-carbon (acetate) bioinsertion pathway and appear to bear no relationship to the odd–even frequency in natural fatty acids. Many of the compounds surveyed contain an odd number of carbons. The distribution of atom types (i.e., O, N, S heterocycles versus carbocycles) does not reveal a pattern to account for the odd–even preference either. Our analysis revealed the following atom type distribution: 50%, oxygen heterocycles (including lactones and cyclic ethers); 32%, nitrogen heterocycles (including cyclic peptides and aza-rings); 17%, carbocycles; 1% sulfur heterocycles.

The odd–even oscillation pattern exists in the ring series $n = 13$–19 representing 54% of the natural product macrocycles surveyed. Macrocycles ranging from 11- to 12-, and 20- to 34-membered rings are less common in nature (11 and 33%, respectively). Ring sizes larger than 35 are rare and their numbers are far less abundant (2%). A statistical analysis allowed us to determine that this even-odd trend is not due to random chance. The frequency of occurrence of the 51 macrocycle sizes analysed indicated a distribution that deviated significantly from uniform ($\chi^2(51) = 1112.34$ $p \ll 0.01$). The ring sizes in the data set proceed from 11 to 60 (total = 50) and along with a separate ring size of 72. Thus, 51 different ring sizes were analyzed. The jump from 60 to 72 in the sequence does not reduce the quality of the analysis since each ring size is treated as a category.

One might ask whether this even-over-odd trend for macrocycles has a direct correlation to a particular biosynthetic

pathway or efficiency in molecule construction. That is, do ring closure reactions favor the even-membered macrocycles? Such a question about possible ring closure preference cannot be readily deduced in living systems due to the high number and complexity of biosynthetic processes. For example, there is intense interest surrounding favored macrocyclization mechanisms in nonribosomal peptide and polyketide synthases [7–9]. However, in laboratory controlled organic synthesis odd–even effects have been noted in the ring formation of macrocyclic esters [10], macrocyclic alpha-nitrile imines [11–13], polyester imides [14], and acyloin condensations [10,15] where even-membered rings are produced more often than the corresponding odd-membered rings [16]. The predominant formation of even-numbered cycles compared to the odd-numbered cycles is often explained by lower strain, better conformational alignment, favored "backbiting", and minimal transannular interactions [13,17–20].

To favor the formation of a small-ring compound it is required that the conformation available to the small chain be compatible with the ring closure reaction (Baldwin's rules) [21–23]. To favor the formation of a larger ring, the process of bringing together the two ends of a chain proved to be more difficult because of the large number of possible conformations and the unfavourable entropy of activation [20,24,25]. It is likely that whatever the actual ring closure reaction, it is necessary to bring the reacting ends of the linear precursor molecule into close proximity [26]. We suggest that the driving force for this step is the hydrophobic force, which causes the straight chain molecule to fold so as to achieve the least surface area. For efficient macrocyclization (gauged by chain length versus percent yield), the two terminal units of a chain may be captured on a surface, such as the surface of the sodium metal in the acyloin reaction, or the surface of the thorium salt in the Ruzicka cyclization [13]. This reaction still relies on the random meeting of the ends, but it can exclude some conformations, compared to those where the two ends are not on the catalytic surface, thereby reducing the entropy.

In conclusion, patterns in which organisms generate secondary metabolites are important to identify. We cannot conceive of a hypothesis to explain why even ring-sized natural macrocycles would be selected by collectors or screeners over odd-ring sized cyclic compounds. Since many natural product molecules possess cyclic structures (∼20%), reference points are needed to better understand cyclization biogenesis patterns from terrestrial and marine sources.

Acknowledgments We thank the National Institutes of Health (GM076168-01) and PSC-CUNY for financial support. O. R. W. and W. C. are the recipients of the David Davidson Memorial Summer Research Award donated by Herbert and Estelle Meislich. The reviewer is thanked for helpful comments.

References

1. Wessjohann LA, Ruijter E, Garcia-Rivera D, Brandt W (2005) What can a chemist learn from nature's macrocycles—A brief, conceptual view. Mol Divers 9:171–186
2. Databases, such as Chemical Combined Dictionary may be used. The primary literature, such as the *Journal of Natural Products* also provides a ready source of structural information on natural macrocycles
3. Pettit GR (1977–1989) Biosynthetic products for cancer chemotherapy. Plenum Press, New York
4. Pettit GR (1994) Anticancer drugs from animals, plants, and microorganisms. Wiley-Interscience, New York
5. Guggisberg A, Badawi MM, Hesse M, Schmid H (1974) Alkaloids. 151. Structure of the macrocyclic spermidine alkaloids, oncinotine, neooncinotine, and isooncinotine. Helv Chim Acta 57:414–434
6. Scheiper B, Glorius F, Leitner A, Fuerstner A (2004) Catalysis-based enantioselective total synthesis of the macrocyclic spermidine alkaloid isooncinotine. Proc Nat Acad Sci 101:11960–11965
7. Kopp F, Marahiel M (2007) Macrocyclization strategies in polyketide and nonribosomal peptide biosynthesis. Nat Prod Rep 24:735–749
8. Walsh CT (2004) Polyketide and nonribosomal peptide antibiotics: modularity and versatility. Science 303:1805–1810
9. Shen B, Liu W, Nonaka K (2003) Enediyne natural products:. biosynthesis and prospect towards engineering novel antitumor agents. Curr Med Chem 10:2317–2325
10. Kricheldorf HR, Richter M, Schwarz G (2002) Macrocycles. 19. Cyclization in the nematic phase? polyesters derived from hydroquinone 4-hydroxybenzoate and aliphatic dicarboxylic acids. Macromolecules 35:5449–5453
11. Fleming FF, Shook BC (2002) Nitrile anion cyclizations. Tetrahedron 58:1–23
12. Curran DP, Liu W (1999) Radical cyclization/fragmentation reactions of dicyano-cyclopropanes to enaminonitriles. A radical alternative to the Thorpe-Ziegler reaction. Synlett 1:117–119
13. Sicher J (1962) Stereochemistry of many-membered rings. In: de la Mare PBD, Klyne W (eds) Progress in stereochemistry. Butterworths, Washinton, pp 202–235
14. Shaikh AA, Schwarz G, Kricheldorf HR (2003) Macrocycles 23. Odd-even effect in the cyclization of poly(ester imide)s derived from catechols. Polymer 44:2221–2230
15. Spanagel EW, Carothers WH (1935) Polymerization and ring formation. XXV. Macrocyclic esters. J Am Chem Soc 57:929–934
16. Limited information is available for laboratory prepared 11-, 13-, and 15- membered rings: Date J (1976) Multistep conformational interconversion mechanisms. Top Stereochem 9:199–270
17. Wandel H, Wiest O (2002) Enediynes in 11-membered rings. Synthesis, structure, and reactivity of highly strained but unusually stable macrocycles. J Org Chem 67:388–393
18. Eliel EL, Allinger NL, Angyal SJ, Morrison GA (1965) Conformational analysis. Wiley Interscience, New York, pp 213–226
19. Peterson PE (1972) Heterocycles containing a d-orbital acceptor atom. Consideration of the dependence of structural and reactivity effects on whether the number of ring atoms is odd or even. J Org Chem 37:4180–4182
20. Nemba RM, Ngouhouo F (1994) On the enumeration of chiral and achiral skeletons of position isomers of homosubstituted monocyclic cycloalkanes with a ring size n (odd or even). Tetrahedron 50:6663–6670

21. Johnson CD (1993) Stereoelectronic effects in the formation of 5- and 6-membered rings: the role of Baldwin's rules. Acc Chem Res 26:476–482
22. Piccirilli JA (1999) Do enzymes obey the Baldwin rules? A mechanistic imperative in enzymic cyclization reactions. Chem Biol 6:59–64
23. Deslongchamps P (1983) Organic chemistry series In: Stereoelectronic effects in organic chemistry, vol 1. Pergamon Press, New York
24. Meng Q, Hesse M (1992) Macrocycles 161. Top Curr Chem 161:107
25. Ercolani G (1998) Physical basis of self-assembly macrocyclizations. J Phys Chem B 102:5699–5703
26. Illuminati G, Mandolini L (1981) Ring closure reactions of bifunctional chain molecules. Acc Chem Res 14:95–102