ORGANIC LETTERS

2008
Vol. 10, No. 5
789−792

Total Structure Determination of Grassypeptolide, a New Marine Cyanobacterial Cytotoxin

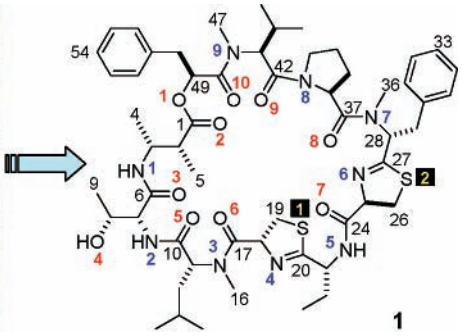
Jason C. Kwan,[†] James R. Rocca,[‡] Khalil A. Abboud,[§] Valerie J. Paul,[||] and Hendrik Luesch*,[†]

Department of Medicinal Chemistry and McKnight Brain Institute, University of Florida, Gainesville, Florida 32610, Department of Chemistry, University of Florida, Gainesville, Florida, 32611, and Smithsonian Marine Station, Fort Pierce, Florida 34949

luesch@cop.ufl.edu

Received December 5, 2007

ABSTRACT



A collection of the cyanobacterium *Lyngbya confervoides* off Grassy Key in Florida yielded grassypeptolide (1), a 31-membered macrocyclic depsipeptide with unusually high D-amino acid content, two thiazolines, and one *β*-amino acid. We report the rigorous 3D structure determination and conformational analysis in solution and solid state by NMR, MS, X-ray crystallography, chemical degradation, and molecular modeling involving distance geometry and restrained molecular dynamics. Grassypeptolide (1) inhibited cancer cell growth with IC_{50} values from 1.0 to 4.2 *μ*M.

Marine cyanobacteria continue to be a rich source of novel bioactive metabolites, including many cytotoxins predominantly originating from *Lyngbya majuscula*.[1] Other *Lyngbya* species are less explored. Recently, we have focused our efforts on marine cyanobacteria in Florida waters and reported several potent elastase inhibitors from *Lyngbya confervoides*.[2] Here, we describe the cytotoxicity-guided isolation of a new cytotoxic depsipeptide, grassypeptolide (**1**), from an extract of another *L. confervoides* collected in the Florida Keys. This cyanobacterium was of interest because it inhibited settlement of coral larvae (*Porites asteroides*) and reduced survival of coral recruits.[3] Grassypeptolide (**1**) contains some unusual residues, such as the *β*-amino acid 2-methyl-3-aminobutyric acid (Maba, C1−5) and 2-aminobutyric acid (Aba, C20−23). Until now, the Aba unit had precedence only in sponge metabolites,[4] whereas the Maba unit was found in one other cyanobacterial compound, guineamide B.[5] Additionally, compound **1** consists of an unusually high number of D-amino acid units.

[†] Department of Medicinal Chemistry, University of Florida.
[‡] McKnight Brain Institute, University of Florida.
[§] Department of Chemistry, University of Florida.
[||] Smithsonian Marine Station.

(1) Gerwick, W. H.; Tan, L. T.; Sitachitta, N. *Alkaloids Chem. Biol.* **2001**, *57*, 75−184.

(2) (a) Matthew, S.; Ross, C.; Rocca, J. R.; Paul, V. J.; Luesch, H. *J. Nat. Prod.* **2007**, *70*, 124−127. (b) Taori, K.; Matthew, S.; Rocca, J. R.; Paul, V. J.; Luesch, H. *J. Nat. Prod.* **2007**, *70*, 1593−1600.

(3) Kuffner, I. B.; Walters, L. J.; Becerro, M. A.; Paul, V. J.; Ritson-Williams, R.; Beach, K. S. *Mar. Ecol. Prog. Ser.* **2006**, *323*, 107−117.

10.1021/ol702946d CCC: $40.75
Published on Web 01/26/2008

The tandem thiazoline rings flanking the D-Aba derived moiety are reminiscent of the lissoclinamides and the patellamides (Figure 1), which are cyclic peptides containing

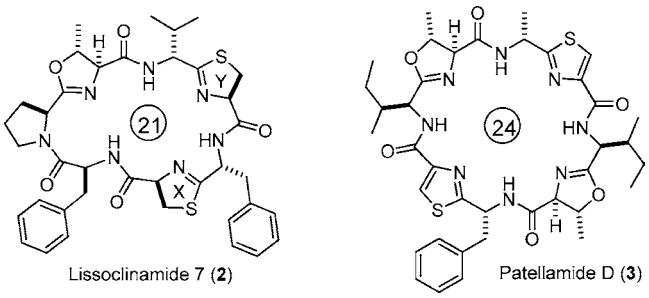


Figure 1. Macrocyclic marine metabolites closely related to **1**.

up to four cysteine- and serine-derived cyclocondensation products and which tend to contain D-amino acids.[6,7] Lissoclinamide 7 (**2**),[6] closest related to **1** and the most cytotoxic of the series, has two thiazoline rings with the same arrangement and stereoconfiguration as **1**, yet the macrocycle is only 21-membered in **2** as opposed to 31-membered in **1**. Although both the lissoclinamides and the patellamides were originally isolated from the ascidian *Lissoclinum patella*, the biosynthetic gene clusters were recently found in the obligate symbiotic cyanobacterium *Prochloron didemni*.[7] Remarkably, these compounds are synthesized ribosomally (followed by post-translational modification) rather than by nonribosomal peptide synthetases (NRPS).[7] Grassypeptolide (**1**) is the first reported compound with tandem thiazoline rings in the depicted arrangement to be produced by an independently living cyanobacterium. Considering the number of nonribosomal peptide residues, it is probably made via an NRPS-like pathway.

Samples of *L. confervoides* were collected off Grassy Key.[8] The nonpolar extract (EtOAc/MeOH 1:1) was fractionated over HP-20 resin followed by silica chromatography and reversed-phase HPLC to afford **1** {$[\alpha]^{20}_{D}$ +76 (*c* 0.1, CH_2Cl_2)}. NMR data combined with a $[M + H]^+$ peak at *m*/*z* 1102.5438 in the HRESIMS of **1** suggested a molecular formula of $C_{56}H_{79}N_9O_{10}S_2$ (calcd for $C_{56}H_{80}N_9O_{10}S_2$, 1102.5464). The ^{1}H NMR spectrum of **1** in $CDCl_3$ was indicative of a peptide by displaying three secondary amide doublets (δ_H 7.12, 7.40, 7.53), three putative *N*-Me tertiary amide singlets (δ_H 2.78, 3.11, 3.15), and several resonances characteristic for α-protons of amino acids (δ_H ~4 to ~5). Considering the IR spectrum, which exhibited bands due to ester (1733 cm^{-1}) and amide (1640 cm^{-1}) carbonyl stretch vibrations, **1** appeared to be a depsipeptide.

Analysis of the ^{1}H NMR, ^{13}C NMR, APT, COSY, HMQC, HMBC, and ROESY spectra recorded in $CDCl_3$ revealed the presence of two regular α-amino acid units (threonine, C6−9; proline; C37−41), two N-methylated α-amino acids (*N*-methylleucine, C10−16; *N*-methylvaline, C42−47), one *β*-amino acid (Maba, C1−5), phenyllactic acid (Pla, C48−56), a *N*-methylphenylalanine-derived thiazoline carboxylic acid unit (*N*-Me-Phe-thn-ca; C24−36), and a thiazoline carboxylic acid moiety derived from Aba (Aba-thn-ca; C17−23) (Table 1). The presence of the two thiazoline rings was deduced from the chemical shifts of vicinally coupled H-18 (δ_H 5.32) and H_2-19ab (δ_H 3.58/3.27) as well as H-25 (δ_H 5.30) and H_2-26ab (δ_H 3.70) combined with HMBC correlations of these spin systems to putative carbonyl-derived carbons from Aba [C-20 (δ_C 178.5)] and *N*-Me-Phe [C-27 (δ_C 177.2)], respectively. In addition, 1D selective TOCSY experiments revealed homoallylic coupling in both thiazoline rings between H-18/H-21 and H-25/H-28. HMBC analysis (Table 1) readily established the connectivity of the units as shown for **1**, which was further confirmed by interresidue ROESY correlations. Notably, there was an unusual four-bond correlation between H-2 and C49, which could have arisen because of a planar "W" conformation.[9]

Compound **1** was hydrolyzed with 6 N HCl (110 °C, 18 h) and the hydrolyzate subjected to chiral HPLC, revealing the presence of D-Aba, *N*-Me-D-Phe, L-Pro, *N*-Me-L-Val, L-Pla, and D-*allo*-Thr in the molecule, but the correct assignment for *N*-Me-Leu remained unclear. A sample of **1** was also subjected to ozonolysis prior to hydrolysis in an attempt to detect cysteic acid (Cya) and hence deduce the configuration of the thiazoline rings. However, peaks for both L- and D-Cya were detected by chiral HPLC, preventing unambiguous configurational assignment. Marfey's analysis[10] of the hydrolyzed ozonolysis product was carried out to ascertain the configuration of the Maba,[11,12] *N*-Me-Leu and Cya units, using 1-fluoro-2,4-dinitrophenyl-5-L-leucinamide (L-FDLA) as the derivatizing agent. Reversed-phase HPLC of **1** derivatized with L-FDLA allowed the assignment of (2*R*,3*R*)-Maba.[13] In addition, the presence of D-*allo*-Thr, *N*-Me-D-Phe,[14] L-Pro, and *N*-Me-L-Val was confirmed, and *N*-Me-D-Leu could be unambiguously assigned. L-FDLA adducts for L- or D-Cya were quantified by LC−MS and

(4) (a) Fusetani, N.; Sugawara, T.; Matsunaga, S.; Hirota, H. *J. Am. Chem. Soc.* **1991**, *113*, 7811−7812. (b) Kobayashi, J.; Itagaki, F.; Shigemori, H.; Ishibashi, M.; Takahashi, K.; Ogura, M.; Nagasawa, S.; Nakamura, T.; Hirota, H.; Ohta, T.; Nozoe, S. *J. Am. Chem. Soc.* **1991**, *113*, 7812−7813. (c) Nakao, Y.; Fujita, M.; Warabi, K.; Matsunaga, S.; Fusetani, N. *J. Am. Chem. Soc.* **2000**, *122*, 10462−10463.

(5) Tan, L. T.; Sitachitta, N.; Gerwick, W. H. *J. Nat. Prod.* **2003**, *66*, 764−771.

(6) Wipf, P.; Fritch, P. C.; Geib, S. J.; Sefler, A. M. *J. Am. Chem. Soc.* **1998**, *120*, 4105−4112.

(7) (a) Schmidt, E. W.; Nelson, J. T.; Rasko, D. A.; Sudek, S.; Eisen, J. A.; Haygood, M. G.; Ravel, J. *Proc. Natl. Acad. Sci. U.S.A.* **2005**, *102*, 7315−7320. (b) Long, P. F.; Dunlap, W. C.; Battershill, C. N.; Jaspars, M. *ChemBioChem* **2005**, *6*, 1760−1765. (c) Donia, M. S.; Hathaway, B. J.; Sudek, S.; Haygood, M. G.; Rosovitz, M. J.; Ravel, J.; Schmidt, E. W. *Nat. Chem. Biol.* **2006**, *2*, 729−735.

(8) Collection site position: 24°43.381′N, 80°51.696′W.

(9) Claridge, T. D. W. *High-Resolution NMR Techniques in Organic Chemistry*; Elsevier: San Diego, CA, 1999.

(10) Marfey, P. *Carlsberg Res. Commun.* **1984**, *49*, 591−596.

(11) Only the 2*R*,3*R* and 2*R*,3*S* standards were used. The elution times of the 2*S*,3*S* and 2*S*,3*R* isomers were deduced by derivatizing these standards with DL-FDLA.

(12) (a) Fujii, K.; Ikai, Y.; Mayumi, T.; Oka, H.; Suzuki, M.; Harada, K.-I. *Anal. Chem.* **1997**, *69*, 3346−3352. (b) Fujii, K.; Ikai, Y.; Oka, H.; Suzuki, M.; Harada, K.-I. *Anal. Chem.* **1997**, *69*, 5146−5151.

(13) Two peaks were observed corresponding to (2*R*,3*R*)- and (2*S*,3*R*)-Maba in the approximate ratio of 2.5:1. This is consistent with chromatograms of the standards, which were obtained from the corresponding *N*-benzoylated-*O*-methyl esters. The latter also showed some epimerization at the 2-position during hydrolysis.

Table 1. ^{1}H and ^{13}C NMR Spectral Data for Compound **1** in $CDCl_3$ (δ in ppm, J in Hz) at 500 MHz (^{1}H) and 100 MHz (^{13}C)

C/H	δ_H (J)	δ_C[a]	HMBC[b]
1		172.5, s	
2	2.51, qd (6.9, 6.2)	45.5, d	1, 3, 4, 5, 49[d]
3	4.18, dqd (6.8, 6.7, 6.2)	48.6, d	1, 2, 4, 5, 6
4	1.16, d (6.7)	19.7,[c] q	2, 3
5	1.10, d (6.9)	14.6, q	1, 2, 3
NH	7.40, br d (6.8)		
6		169.8, s	
7	4.45, dd (7.8, 6.4)	59.2, d	6, 8, 9, 10
8	4.02, dq (6.4, 6.2)	68.8, d	
9	1.23, d (6.2)	19.7,[c] q	7, 8
OH	3.96, br		
NH	7.12, d (7.8)		7, 8, 10
10		170.3, s	
11	4.92, br	56.7, d	13, 17
12a	1.85, m	36.9, t	10, 11, 13, 14, 15
12b	1.72, ddd (−14, 8.1, 6.2)		10, 11, 13, 14, 15
13	1.55, m	25.1, d	11, 14, 15
14	0.95, d (6.6)	23.2, q	12, 13, 15
15	0.90, d (6.5)	22.1, q	12, 13, 14
16	3.15, s	32.3, q	11, 17
17		170.4, s	
18	5.32, ddd (9.5, 9.1, 1.8)	77.8, d	17, 19, 20
19a	3.58, dd (−9.9, 9.1)	33.4, t	17, 18, 20
19b	3.27, dd (−9.9, 9.5)		17, 18, 20
20		178.5, s	
21	4.64, m	54.4, d	20
22a	2.18, m	25.2, t	23
22b	1.97, m		20, 21, 23
23	0.96, t (7.2)	11.0, q	21, 22
NH	7.53, d (7.9)		21, 22, 24
24		171.0, s	
25	5.30, m	79.3, d	24, 27
26a/b	3.70, m (2H)	37.7, t	24, 25, 27
27		177.2, s	
28	3.83, dd (9, 3.5)	69.0, d	27, 37
29a	3.57, dd (−13.9, 9)	35.3, t	27, 28, 30, 31/35
29b	3.44, dd (−13.9, 3.5)		28, 30, 31/35
30		138.2, s	
31/35	7.35, m	129.8, d	29, 33
32/34	7.34, m	128.7, d	30
33	7.25, m	126.7, d	
36	2.78, s	39.6, q	28, 37
37		173.0, s	
38	4.77, dd (7.4, 5.5)	57.0, d	37, 39, 40, 41, 42
39a/b	2.04, m (2H)	27.5, t	37, 38, 40, 41
40a	2.12, m	24.8, t	38, 39, 41
40b	1.86, m		38, 39, 41
41a	3.69, m	47.6, t	38, 39, 40
41b	3.60, m		39, 40
42		167.8, s	
43	4.93, d (10.9)	60.3, d	42, 44, 45, 46, 47, 48
44	2.42, dqq (10.9, 6.7, 6.4)	27.3, d	42, 43, 45, 46
45	0.97, d (6.4)	19.5, q	43, 44
46	0.87, d (6.7)	18.2, q	43, 44, 45
47	3.11, s	30.3, q	43, 48
48		171.1, s	
49	5.40, dd (9.9, 3.5)	72.0, d	1, 50, 51
50a	3.12, dd (−14.5, 9.9)	37.2, t	49, 51, 52/56
50b	3.00, dd (−14.5, 3.5)		48, 51, 52/56
51		135.6, s	
52/56	7.21, m	129.2, d	50, 54
53/55	7.30, m	128.6, d	51
54	7.26, m	127.3, d	52/56

[a] Multiplicity deduced from APT and HMQC spectra. [b] Protons showing long-range correlation to indicated carbon. [c] These carbons have the same chemical shift. [d] An unusual four-bond HMBC[9] (see text).

found to be present in the ratio of 1.64:1, indicating that either the thiazolines were of opposite configuration producing cysteic acids in different yields or that epimerization of one or both units had occurred. However, presumably at least one thiazoline had to have *R* configuration because of the excess L-Cya produced.

Attempts were then made to crystallize the compound. Eventually, a small yield of crystals was produced using a mixture of dichloromethane and methanol.[15] The resulting X-ray structure (Figure 2) confirmed the gross 2D arrangement and all of the previously assigned stereocenters. Additionally, both thiazolines could be assigned as *R*, confirming that significant epimerization had occurred under the reaction conditions.

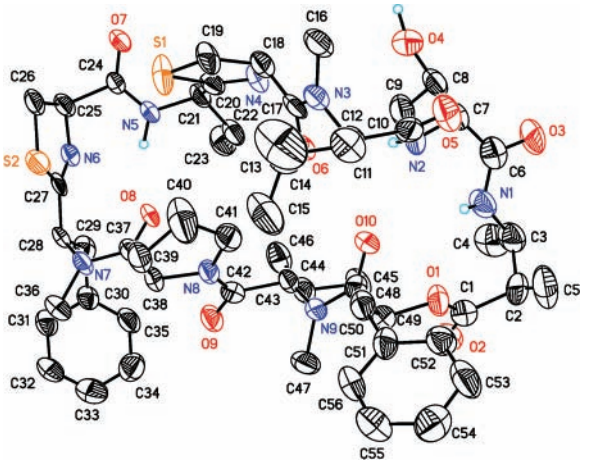


Figure 2. Displacement ellipsoids (50% probability) for the X-ray crystal structure of grassypeptolide (**1**).

The crystal structure shows hydrogen bonds between the NH (at N1) of Maba to the Thr N (N2; 2.35 Å) and the Pla ester O (O1; 2.52 Å). Another hydrogen bond occurs between the Thr NH and the carbonyl of Aba-thn-ca (O6; 2.37 Å). At the opposite site of the macrocycle, a tight turn at *N*-Me-Phe-thn-ca is stabilized by a hydrogen bond (2.04 Å) between the Pro carbonyl (O8) and the NH of Aba-thn-ca (at N5), with the angle between the planes of the thiazoline rings at almost 90°. Analogous turns occur in patellamide D (**3**) at the oxazoline rings, while the thiazoles are planar.[16] In **2**, a turn is centered around the other thiazoline (ring X), and there is a hydrogen bond between the NH of Phe and the N of the adjacent thiazoline, rather than across the turn. The angle between thiazoline planes is still close to 90°, but one is twisted so that its plane is parallel to that of the macrocycle.

(14) Marfey's adducts of *N*-Me-D-Phe and *N*-Me-L-Leu co-eluted; however, the relative intensity of the corresponding peak was reduced and *N*-Me-D-Glu was generated when stringent ozonolysis conditions (25 °C) were employed. Thus, a *N*-Me-D-Phe residue was present in **1**.

(15) One methanol molecule was seen incorporated into the lattice. Upon drying, the crystals would quickly degrade, presumably due to methanol loss.

(16) Schmitz, F. J.; Ksebati, M. B.; Chang, J. S.; Wang, J. L.; Hossain, M. B.; van der Helm, D.; Engel, M. H.; Serban, A.; Silfer, J. A. *J. Org. Chem.* **1989**, *54*, 3463−3472.

Several aspects of the NMR data for **1** suggested that the solution structure was similar to the X-ray stucture.[17] First, ROESY data suggested that all amide bonds were trans in solution, as they are in the solid state. Second, three calculated Φ angles from $^3J_{NH-\alpha H}$ values[18] were similar to those observed in the X-ray structure. Third, the planar "W" suggested by the four-bond HMBC between H-2 and C49 was present in the X-ray stucture.

To investigate the solution structure, 46 distance constraints were derived from ROESY spectra[17] and three dihedral angle constraints from coupling constants (NH-αH). Using a previously established molecular modeling protocol suitable for cyclodepsipeptides,[19] 10 randomly drawn structures of **1** were subjected to distance geometry,[20] followed by simulated annealing and finally restrained molecular dynamics simulation for 1 ns. The modeled structures could be divided into two distinct conformational families. Six structures (Figure 3) bore striking similarity to the X-ray

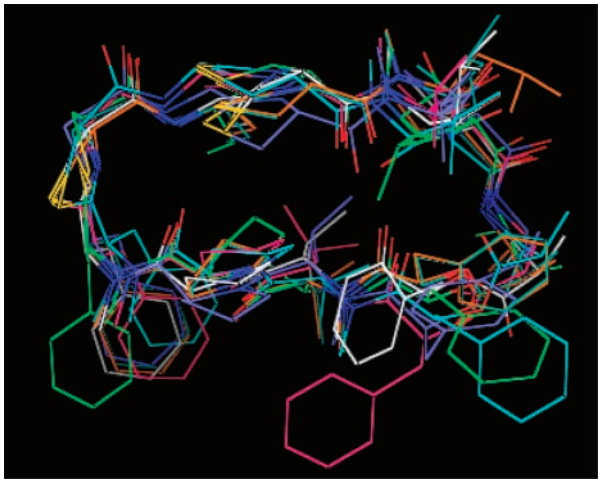


Figure 3. Lowest-energy conformational family most consistent with the ROESY data (X-ray structure overlaid in green).

structure. The other four structures (Figure S1, Supporting Information) had altered macrocyclic ring conformation due to a differing orientation of the Pla-Maba-Thr region but consistently violated the same constraint between one Maba methyl (H_3-5) and H-11 (*N*-Me-Leu). Additionally, this second conformational family exhibited more constraint violations in general and had higher energies.[21] Thus, structures in the conformational family similar to the X-ray structure are in better agreement with the ROESY data, although there were not enough constraints in the Pla-Maba-Thr region (due to signal overlap) to ensure convergence of all 10 random structures to the same conformation.

The antiproliferative activity of **1** was evaluated in four cell lines derived from human osteosarcoma (U2OS), cervical carcinoma (HeLa), colorectal adenocarcinoma (HT29), and neuroblastoma (IMR-32). Compound **1** showed moderate broad-spectrum activity with IC_{50} values of 2.2, 1.0, 1.5, and 4.2 μM, respectively. This data is within the range of IC_{50} values reported for **2** (53.7 nM to 21.5 μM), but in different cell lines which were not tested here.[6,22] Previously, it has been shown that the thiazolines of **2** are important to its cytotoxic activity.[6] It is tempting to speculate that this motif might also be responsible for the activity of **1**, and that it might indicate a shared mechanism of action with the lissoclinamides and patellamides.

Acknowledgment. This article was developed under the auspices of the Florida Sea Grant College Program with support from NOAA, Office of Sea Grant, U.S. Department of Commerce, Grant No. NA06OAR4170014. Field work at Keys Marine Lab was supported by a U.S. Geological Survey State Partnership Program grant with matching funds from the University of Central Florida to Drs. I. Kuffner, L. Walters, V. Paul, and K. Beach. We acknowledge NSF for funding through the External User Program of the National High Magnetic Field Laboratory (NHMFL), which supported our NMR studies at the Advanced Magnetic Resonance Imaging and Spectroscopy (AMRIS) facility in the McKnight Brain Institute of the University of Florida (UF). The 600 MHz 1-mm triple-resonance HTS cryogenic probe used for 1D TOCSY and NMR in DMSO-d_6 was developed through collaboration between UF, NHMFL, and Bruker Biospin.[23] We thank the staff of the Keys Marine Lab and M. Becerro and R. Ritson-Williams (both Smithsonian Marine Station) for assistance with collections. K.A.A. acknowledges the NSF and UF for funding of the purchase of the X-ray equipment. We thank Prof. E. Juaristi (Universidad Autonoma del Estado de Morelos, Mexico) for kindly providing standards for the Maba unit. We thank Drs. A. Wilczynski and S. Sivendran for technical assistance and J. Johnson for LC−MS analysis (all UF). Credit for the photograph of the cyanobacterium at its collection site belongs to R. Ritson-Williams. This is contribution no. 716 from the Smithsonian Marine Station at Fort Pierce.

Supporting Information Available: Experimental methods, NMR data, Figure S1, X-ray diffraction data, and CIF structure file for compound **1**. This material is available free of charge via the Internet at http://pubs.acs.org.

OL702946D

(17) For conformational analysis in solution, NMR data for **1** in DMSO-d_6 was used, as the differing overlap of peaks allowed the unambiguous assignment of more correlations across units. The four-bond HMBC referred to was only observed in $CDCl_3$.

(18) Using the modified Karplus equation, $^3J_{NH-\alpha H} = 8.40 \cos^2 \Phi - 1.36 \cos \Phi + 0.33$; see: Vögeli, B.; Ying, J.; Grishaev, A.; Bax, A. *J. Am. Chem. Soc.* **2007**, *129*, 9377−9385.

(19) Luesch, H.; Yoshida, W. Y.; Moore, R. E.; Paul, V. J.; Corbett, T. H. *J. Am. Chem. Soc.* **2001**, *123*, 5418−5423.

(20) Kuntz, I. D.; Thomason, J. F.; Oshiro, C. M. *Methods Enzymol.* **1989**, *177*, 159−204 and references therein.

(21) In the X-ray structure-like family, the average number of distance constraints over 1 Å was 1.3, and the average energy was 17.507 kcal/mol. The other family had an average number of 7 distance constraint violations over 1 Å and the average energy of the structures was 30.052 kcal/mol.

(22) Hawkins, C. J.; Lavin, M. F.; Marshall, K. A.; van den Brenk, A. L; Watters, D. J. *J. Med. Chem.* **1990**, *33*, 1634−1638.

(23) Brey, W.; Edison, A. S.; Nast, R. E.; Rocca, J. R.; Saha, S.; Withers, R. S. *J. Magn. Reson.* **2006**, *179*, 290−293.