Tetrahedron 66 (2010) 6437–6444

Contents lists available at ScienceDirect

Tetrahedron

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

The stereochemical assignment of acyclic polyols: a computational study of the NMR data of a library of stereopentad sequences from polyketide natural products

Steven G. Smith, Julia A. Channon, Ian Paterson, Jonathan M. Goodman *

Department of Chemistry, University of Cambridge, Lensfield Road, Cambridge, CB2 1EW, UK

ARTICLE INFO

Article history:
Received 18 March 2010
Received in revised form 1 June 2010
Accepted 1 June 2010
Available online 17 June 2010

Dedicated to Professor Steve Ley on the occasion of his receipt of the Tetrahedron Prize

Keywords:
Polypropionate
Stereochemistry assignment
NMR shift library
GIAO calculations

ABSTRACT

The reliable stereochemical assignment of flexible molecules, such as acyclic polypropionates is an enormously challenging task. This is illustrated by the NMR chemical shifts for a complete set of sixteen diastereomeric stereopentads whose experimental data is reported here for the first time. Although the experimental spectra are very similar to each other, analysis of the similarity between the shifts of different diastereoisomers reveals that some diastereoisomers are much more distinctive than others. In addition, the NMR shifts of the sixteen compounds have also been calculated using DFT GIAO calculations, and the use of our recently developed CP3 parameter for structure assignment is illustrated for these molecules. Even in cases where the experimental spectra are very similar, our CP3 parameter makes possible the correct assignment of pairs of diastereoisomers with high confidence.

1. Introduction

The polyketide family of natural products represents a large and diverse array of structurally complex compounds that display a wide range of biological activities including antibiotic, antifungal and anti-cancer properties.[1] Polypropionates are an important class of polyketides and many polyketides contain substantial polypropionate portions in a variety of stereochemical configurations. The stereopentad unit, consisting of five contiguous stereocentres bearing alternating methyl and hydroxyl groups as in 1 and 2, is a common polypropionate motif, with over 180 low molecular weight (<465 amu) molecules containing this unit being described in the literature since 2001 alone. Stereopentad building blocks, such as **1** and **2** have been used in the stereocontrolled synthesis of many polyketide natural products including discodermolide, oleandolide, reidispongiolide and spirangien (Fig. 2).[2]

We have previously outlined the synthesis of all sixteen diastereoisomers of the protected stereopentad represented by the structures in Figure 1.[3] Thirteen of these were synthesised with a benzyl protecting group (**1**) and the remaining three as their TBDPS-protected derivatives **2**. The synthetic approach (illustrated in Fig. 3) was based on a stereoselective boron aldol reaction,

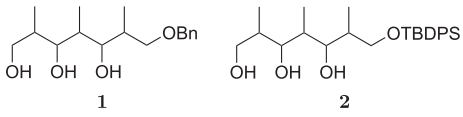


Figure 1. Benzyl and TBDPS-protected stereopentads.

followed by hydroxyl-directed *syn* or *anti* reduction, and then a stereoselective hydration.

The stereopentads **1c**, **1d**, **1e**, **1f**, **1g**, **1h**, **1k**, **1l**, **1m**, **1n**, **1o**, **1p** (see Fig. 4) could be obtained by this route. TBDPS-protected isomers **2a**, **2i** and **2j** could be obtained by protecting group manipulation from the enantiomers of **1f**, **1k** and **1o**, respectively, and isomer **1b** (benzyl protected) was accessed from the diol en-route to stereopentads **1f** and **1n** by a selective oxidation and reduction sequence, followed by hydration. The stereochemistry of the synthesised stereopentads was confirmed by acetal formation from the diols, and also by debenzylation of the triols to give either symmetric or asymmetric tetraols.

This set of stereopentads provide an intriguing library of NMR spectra that could be useful for assigning the stereochemistry of related polypropionates, following the database approach of Kishi,[4] which has recently been extended by Ardisson et al.[5] A table of the experimental data for these molecules is therefore presented.

Also presented is an analysis of how similar the spectra of these molecules are to each other. This is of interest because, if one is

* Corresponding author. Tel.: +44 1223 336434; fax: +44 1223 763076; e-mail address: j.m.goodman@ch.cam.ac.uk (J.M. Goodman).

0040-4020/$ – see front matter © 2010 Elsevier Ltd. All rights reserved.
doi:10.1016/j.tet.2010.06.022

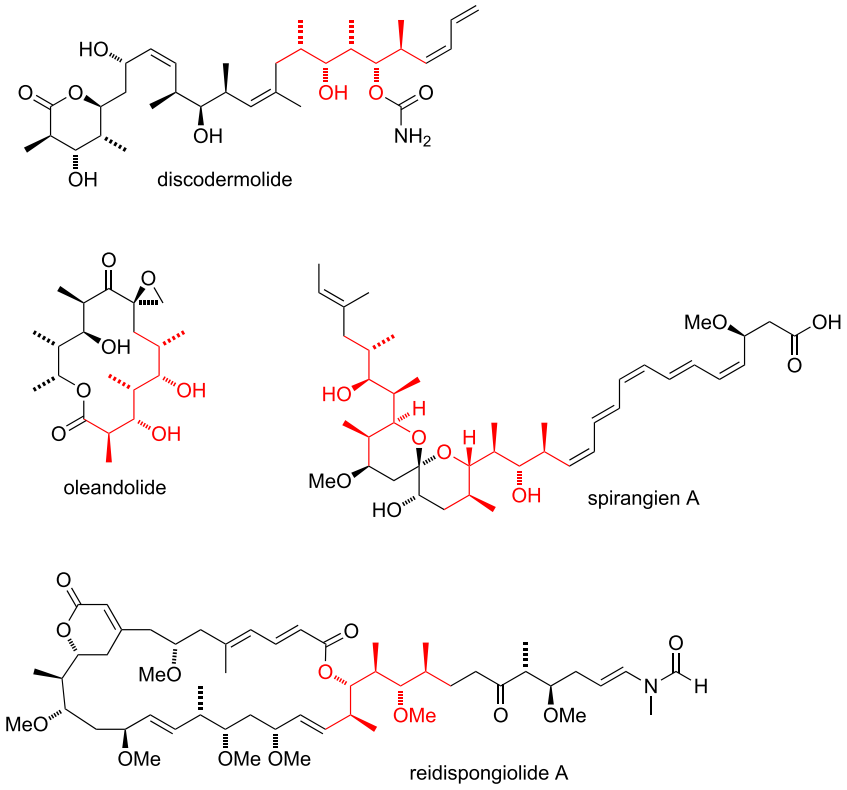


Figure 2. Some natural products containing the stereopentad motif.

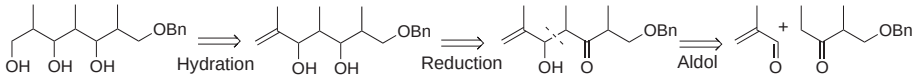


Figure 3. Route to the stereoisomers of **1**.

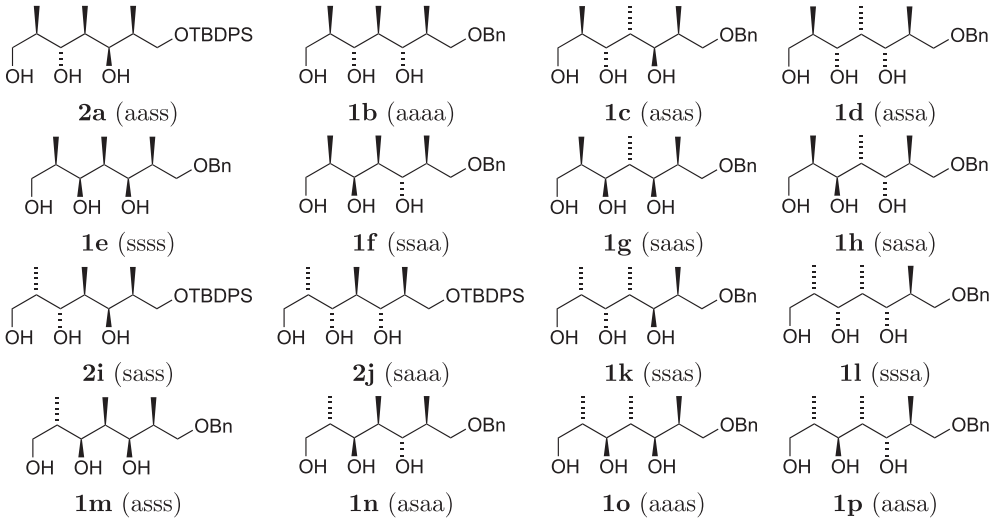
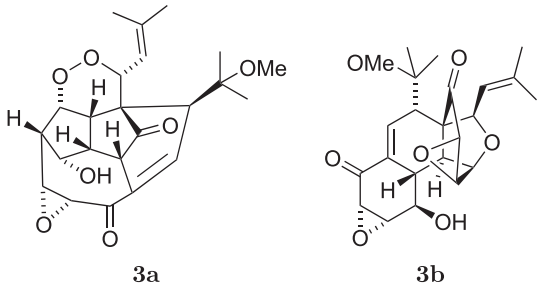


Figure 4. Stereopentads **1** and **2**. The aass notation refers to the *syn* or *anti* relationships between adjacent stereocentres.

synthesising these types of structure or assigning the stereochemistry of a natural product containing the stereopentad motif, it is vital to have confidence that the stereochemistry has been assigned correctly. Our syntheses were designed to ensure that there was no ambiguity. Several diastereoisomers have very similar spectra, however, and it is difficult to distinguish them reliably in the absence of strong corroborative information from a synthetic sequence.

The issue of whether two different molecules can have spectra too similar to be readily distinguished has recently been investigated for the proposed and revised structures of hexacyclinol (**3a** and **3b**, Fig. 5) by Bagno and Saielli.[6] DFT calculation of ^{13}C and ^{1}H NMR shifts and $^{1}H-^{1}H$ coupling constants was used to indicate that, in this case, the two structures are expected to have significantly different NMR spectra. Interestingly, however, protons



3a 3b

Figure 5. Hexacyclinol: Originally proposed structure **3a** and revised structure **3b**.

closest to the sites of major structural difference (for example, next to the ether and peroxide groups) did not always show the largest differences in calculated shift, while significant differences were observed for many protons and carbons that appear from the structures to be in rather similar environments. This suggests that the relationship between structural difference and chemical shift difference is complex, and it may well be that ^{1}H and ^{13}C spectra alone are insufficient to distinguish some structures.

Stereopentads **1** and **2** show great structural similarity to each other, so their NMR spectra can be expected to be very similar. In addition, in such flexible systems it is not obvious how the different stereochemical configurations will manifest themselves as differences in chemical shift. It is, therefore, of interest to investigate how different the spectra of the stereopentad diastereoisomers are to each other, to identify whether particular groups of isomers are easier to distinguish than others, and to investigate whether a greater degree of difference in stereochemistry (for example, having more stereocentres with a different configuration) gives rise to more easily distinguishable spectra.

If two spectra are sufficiently different to be distinguishable, one is then faced with the question of which belongs to which diastereoisomer. With flexible systems, such as **1** and **2** assigning, which is which is not a trivial issue. In such cases, one of the tools available for stereochemical assignment is GIAO NMR shift calculation. There has recently been much interest in this approach and the technique, pioneered by Bifulco,[7,8] has since played key roles in the stereostructure assignment or reassignment of several natural products including hexacyclinol,[9] maitotoxin,[10] applidinones A–C,[11] gloriosaols A and B,[12] kadlongilactones D and F,[13] artarborol,[14] obtusallenes V–VII[15], elatenyne,[16] spiroleucettadine,[17] samoquasine A,[18] mururin C,[19] hassananes,[20] ketopelenolides C and D,[21] 6β-hydroxyhyoscyamine,[22] dolichodial,[23] hypurticin,[24] santalol derivatives[25] and fusapyrones.[26] The effect of using different levels of theory at various stages in the NMR shift calculation has also been extensively investigated,[27–36] and the area has been reviewed.[37]

We have recently shown that GIAO NMR shift calculation, used in conjunction with our new CP3 parameter,[38] can be a powerful tool for assigning a pair of diastereoisomers when one has experimental data for both. The CP3 parameter works by comparing differences in calculated shift to differences in experimental shift in a way that simultaneously reduces systematic errors in the calculated shifts and focuses on those experimental shifts, which are most useful for structure assignment. A full description of this parameter, including how to calculate it either 'by hand' or using our web applet, has been published elsewhere.[38] An estimate of the confidence in the conclusion can be obtained by using Bayes' theorem together with a knowledge of the values of CP3 expected for both correct and incorrect assignments, and the standard deviations of these values. This approach was also described in detail in our previous publication.[38] The set of diastereomeric stereopentads in Figure 4 provides an excellent opportunity to validate our methodology for stereochemical assignment.

This paper has three sections. In the first section, the experimental ^{13}C and ^{1}H shifts for the sixteen stereopentads shown in Figure 4 are presented. Second, the similarity of the diastereoisomers to each other is analysed. Finally, it is shown how GIAO NMR shift calculation can be used to reliably assign pairs of diastereoisomers to structures, even in cases where the previous analysis has indicated that the experimental spectra are very similar and so their assignment based on NMR alone may be uncertain.

2. Results and discussion

2.1. Experimental NMR shifts

The experimental data for the stereopentads are shown in Table 1 (^{13}C data) and Table 2 (^{1}H data). Assignment of the data is based on multiplicities, coupling constant analysis and intensities, and alternative assignments are indicated.

The numbering system used is indicated in Figure 6.

Table 1
Experimental ^{13}C data for the stereopentads in Figure 4

2a[a] aass	**1b** aaaa	**1c** asas	**1d** assa	**1e** ssss	**1f** ssaa	**1g** saas	**1h** sasa	**2i**[a] sass	**2j**[a] saaa	**1k** ssas	**1l** sssa	**1m** asss	**1n** asaa	**1o** aaas	**1p** aasa
11.5	14.5	10.2	4.1	7.1	13.5	8.7	9.6	9.2	8.6	10.3	6.1	5.5	11.1	9.4	10.4
13.6	15.3	10.9	12.7	12.4	13.6	9.4	9.7	11.5	13.1	10.8	12.1	13.2	13.4	13.9	12.7
13.8	15.6	13.2	13.6	13.6	14.0	12.7	13.0	13.4	15.5	13.5	13.0	13.4	13.6	15.4	13.6
19.2	34.8	35.5	35.4	36.8	35.7	35.0	35.7	19.2	19.0	35.2	35.9	36.3	35.1	34.9	34.4
26.9	35.9	36.9	35.8	37.3	36.0	36.0	36.6	26.6	26.7	37.7	36.7	36.8	35.9	35.7	35.6
35.6	39.2	37.1	37.3	38.2	38.9	37.9	37.2	36.4	35.3	37.9	38.1	37.2	37.5	38.8	37.8
37.5	65.1	69.5	69.2	66.4	66.3	68.1	67.7	37.2	35.9	65.9	66.8	69.1	69.6	64.6	69.4
38.8	72.7	73.5	73.6	73.4	73.2	73.5	73.6	38.4	39.0	73.2	73.6	73.3	73.9	73.7	73.7
66.7	73.2	75.4	76.9	73.9	73.9	75.7	77.2	67.9	66.2	73.3	77.7	74.1	77.2	76.1	76.7
69.1	82.9	76.8	83.4	77.7	76.9	79.8	81.5	68.3	68.4	75.2	78.1	79.4	77.7	80.7	77.0
73.7	83.2	77.2	83.6	78.0	83.8	80.3	81.6	75.3	80.0	76.5	81.2	82.6	83.7	82.8	83.9
84.1	127.8	127.6	127.8	127.6	128.0	127.6	127.7	76.7	83.5	127.6	127.7	127.5	128.0	127.6	127.6
127.7	128.0	127.8	128.0	127.6	128.3	127.8	127.9	127.7	127.8	127.7	127.9	127.6	128.3	127.9	127.7
127.7	128.6	128.4	128.6	128.4	128.8	128.5	128.5	127.7	127.9	128.4	128.5	128.4	128.8	128.5	128.5
129.7	135.2	137.8	137.2	138.2	137.5	137.7	137.4	129.7	130.0	137.9	137.4	139.4	137.4	137.7	137.2
129.8								129.8	130.1						
132.9								133.1	132.0						
133.2								133.3	132.2						
135.5								135.5	135.5						
135.7								135.6	135.6						

[a] Data for the TBDPS-protected stereopentad (see Fig. 4).

Table 2
Experimental ^{1}H data for the stereopentads in Figure 4

Atom	**2a**[e] aass	**1b** aaaa	**1c** asas	**1d** assa	**1e** ssss	**1f** ssaa	**1g** saas	**1h** sasa	**2i**[e] sass	**2j**[e] saaa	**1k** ssas	**1l** sssa	**1m** asss	**1n** asaa	**1o** aaas	**1p** aasa
H8, H9, H0	0.75	0.78	0.70	0.73	0.95	0.79	0.69	0.79	0.80	0.68	0.83	0.79	0.64	0.72	0.78	0.75
	1.04	1.16	0.90	0.78	0.97	1.06	0.94	0.89	0.97	1.07	0.96	0.97	0.88	0.78	1.10	0.75
	1.04	1.16	1.05	0.95	1.05	1.09	0.95	1.01	1.12	1.22	1.09	1.01	1.02	1.11	1.20	1.10
H17	1.04								1.07	1.08						
H2, H4, H6	1.83	1.75–1.68	1.78–1.71	1.73–1.68	1.82–1.72	1.81–1.71	1.73	1.72[b]	1.72–1.64	1.84–1.78	1.79–1.71	1.76[b]	1.72[b]	1.82–1.74	1.89–1.80	1.85–1.78
	2.07–1.92	1.97–1.84	1.98–1.81	1.94–1.85	1.82–1.72	1.89–1.82	1.83–1.76	1.84–1.79[a]	1.82–1.73	1.90–1.85	1.88–1.80	1.94–1.85	1.82[a]	1.97–1.86	1.89–1.80	2.06–2.01
	2.07–1.92	2.13–2.06	1.98–1.81	2.09–1.98	1.93[c]	2.17–2.04	1.98–1.92	2.02[c]	1.96–1.82	1.99–1.91	1.94[c]	2.08–1.95	1.94[c]	2.29–2.18	1.98–1.90	2.06–2.01
H7	3.57–3.52	3.48[d]	3.54	3.48	3.40	3.43	3.58	3.51	3.60	3.59	3.57–3.46	3.48	3.48	3.58	3.59–3.51	3.70–3.45
	3.57–3.52	3.58–3.50	3.58	3.70–3.58	3.40	3.51	3.61	3.65	3.62	3.71	3.57–3.46	3.61		3.68	3.59–3.51	3.70–3.45
H1	3.62	3.58–3.50	3.65	3.70–3.58	3.55	3.45	3.65	3.69	3.68	3.86	3.57–3.46	3.67–3.61	3.58	3.63–3.59	3.65–3.60	3.70–3.45
	3.76	3.78	3.65	3.70–3.58	3.56	3.53	3.79	3.71	3.75	3.90	3.57–3.46	3.67–3.61	3.62	3.63–3.59	3.65–3.60	3.70–3.45
H3,H5	3.57–3.52	3.62	3.74	3.72	3.66	3.59	3.81	3.82	3.88	3.64	3.75	3.67–3.61	3.63	3.80	3.84	3.70–3.45
	3.98	3.93	3.96	3.85	3.71	3.86	3.82	3.90	4.01	3.86	3.91	3.82	3.72	3.88	3.95	3.94
H11		4.52	4.49	4.52	4.46	4.50	4.51	4.52			4.49	4.51	4.46	4.52	4.50	4.52
		4.56	4.49	4.52	4.46	4.50	4.51	4.56			4.49	4.51	4.48	4.55	4.54	4.52
H13, H14, H15		7.38–7.28	7.35–7.25	7.39–7.26	7.32–7.25	7.32–7.25	7.38–7.28	7.36–7.32			7.38–7.25	7.35–7.25	7.32–7.22	7.38–7.25	7.39–7.31	7.36–7.32
H19–25	7.47–7.30								7.46–7.32	7.50–7.36						
	7.68–7.60								7.68–7.62	7.75–7.60						

[a] This signal is H2.
[b] This signal is H4.
[c] This signal is H6.
[d] This signal is H7.
[e] Data for the TBDPS-protected stereopentad (see Fig. 4).

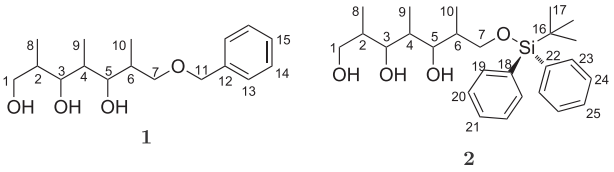


Figure 6. Numbering system used for the stereopentads.

2.2. How similar are the spectra?

Which isomers have the most similar NMR spectra to each other? This is important, since in the synthesis of these molecules it is vital to know whether the isomer in question has a sufficiently similar spectrum to other isomers that there is the potential for confusion.

To answer this question, the mean absolute difference (MAD = $\frac{1}{N}\sum|\delta^i_B - \delta^i_A|$) in shift between all possible pairs of diastereoisomers has been calculated. In carrying out this calculation, all shifts reported as ranges in Table 2 have been averaged (for the purposes of calculating differences to other isomers). To allow comparison between the benzyl and TBDPS-protected isomers, carbon and hydrogen atoms in the protecting groups were ignored, and also the carbon directly attached to the protecting group oxygen (C7) since the shift of this carbon is expected to be significantly affected by the nature of the protecting group. Where the experimental data were unassigned, the shifts to be removed were selected by lining up the experimental shifts with those calculated using the ChemDraw program. The MAD results for the carbon and proton data were then combined using the geometric mean: $\text{MAD} = \sqrt{\text{MAD}_C \times \text{MAD}_H}$. The MAD values for all the structures are illustrated in Figure 7. In order to make the patterns in the 256 MAD values clearer, the sixteen structures have been sorted to minimise the sum of the squares of the differences between adjacent MAD values, so that regions of high and low similarity should be grouped. The pairs of similar structures (yellow) are all close to the major diagonal, so adjacent structures in this illustration are rather similar. Few of the yellow squares are away from the major diagonal, so structures are only similar if they are adjacent or close to being adjacent. The order of the isomers is cyclic, so the last isomers are similar to the first, and each isomer has been included twice in Figure 7 to show this more clearly.

To generate Figure 8, all structures, which differed by more than one standard deviation from the mean (0.31 ppm) were assumed to be distinguishable, and so the diagram focuses on similar structures for which the differences are less than 0.31 ppm. The distances between the spots representing the structures correspond to the MAD between pairs of structures, discounting the larger MAD values, but including both the MADs to nearest neighbours and all small (<0.31) values. The figure shows how some of the structures are quite distinctive (**1n**, **1p**, **2a**, **1f**, **1b**, **1o**, **2j**) but there is a group of structures, which are all quite similar to each other. For example, 1l is close to quite a number of other structures, and will, therefore, be hard to positively identify from ^{1}H and ^{13}C NMR data alone.

The relationship between stereochemistry and similarity is far from simple. It is not necessarily the case that isomers with similar stereochemistry will have similar spectra and vice versa. For example, isomers **1b** and **1d** have quite different spectra (they are far apart in Fig. 8 and give a dark green square in Fig. 7; see also the raw experimental data in Table 1 and Table 2), and yet they only differ at a single stereocentre. Another example is **2a** and **2i**, which differ only at the methyl bearing stereocentre at the end of the pentad (C2) and yet are far apart in Figure 8 and give a dark green square in Figure 7. Table 1 reveals that the largest difference in their ^{13}C spectra is the shift of the most deshielded sp^3 carbon (84.1 vs 76.7 ppm). There are also many other pairs that differ by a single

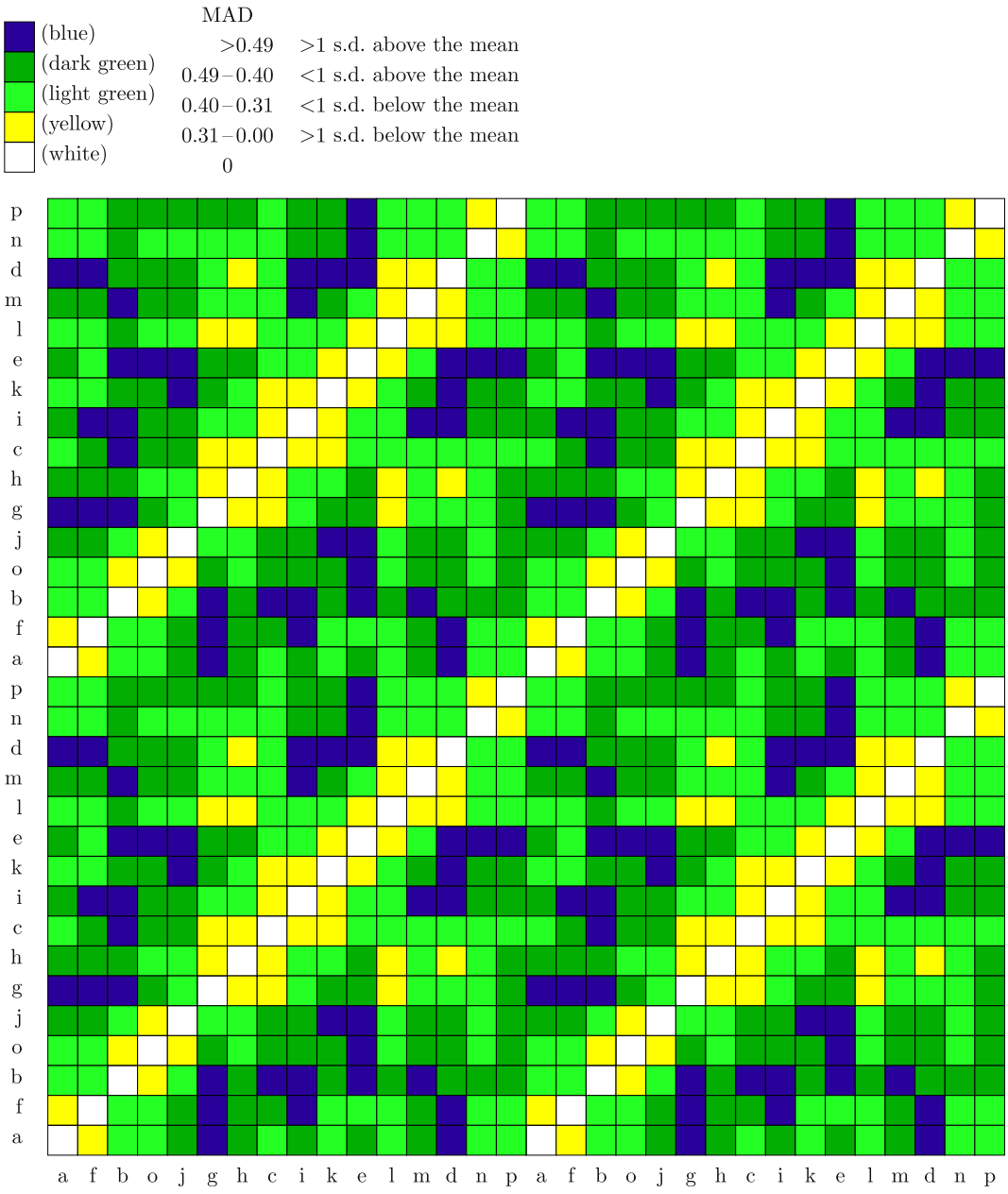


Figure 7. Matrix representation of how similar the stereopentads in Figure 4 are to each other.

stereocentre and yet give large differences in Figure 7; examples include **f** and **h**, **i** and **m**, **k** and **o**.

On the other hand, some isomers differ in configuration at two stereocentres and yet have similar spectra. The maximum difference in relative stereochemistry is two stereocentres since if two isomers differ at, for example, four stereocentres this is the same as the enantiomer of one differing from the other at only one stereocentre. For example, isomers **1g** and **1l** give similar spectra (yellow square in Fig. 7) despite having a very different pattern of stereocentres (Fig. 4). Other pairs of isomers different at two stereocentres and yet giving similar spectra (yellow squares in Fig. 7) include **h** and **l**, **e** and **k**, **c** and **i**.

Some patterns can be identified. For example, because the stereopentads are nearly symmetrical (having a protecting group at one end and a free OH at the other) changes in several stereocentres can sometimes produce a very similar molecule. Compounds **1l** and **1m** (close in Fig. 8 and yellow square in Fig. 7) are an example of this: although they differ in relative configuration at two stereocentres, **1l** can be converted to **1m** (or at least the enantiomer of **1m**) by swapping the OBn and terminal **OH** groups. In fact, there are six pairs that are related in this way (**a**+**f**, **c**+**h**, **i**+**k**, **j**+**o**, **l**+**m** and **n**+**p**) and all of them gave similar spectra (yellow squares in Fig. 7). All but one of the blue squares (most different spectra) in Figure 7 and over half of the dark green squares (next most different spectra) are from pairs of isomers differing at more than one stereocentre, so differing at several centres does seem to increase the chances of the spectra being significantly different but by no means guarantees it.

Finally, in the context of Hoffmann's rule that methyl groups at the centre of a *syn–syn* stereotriad gave unusually low shifts,[39] it may be noted that the four isomers containing this arrangement (**1d**, **1e**, **1l** and **1m**) display a particularly low shift in their ^{13}C spectra (Table 1) and are adjacent to each other in Figure 7 and 8. Nevertheless, it is not true that all isomers containing this arrangement have similar spectra to each other (**1d** and **1e** give a blue square, i.e., most different spectra, in Fig. 7), or that they necessarily

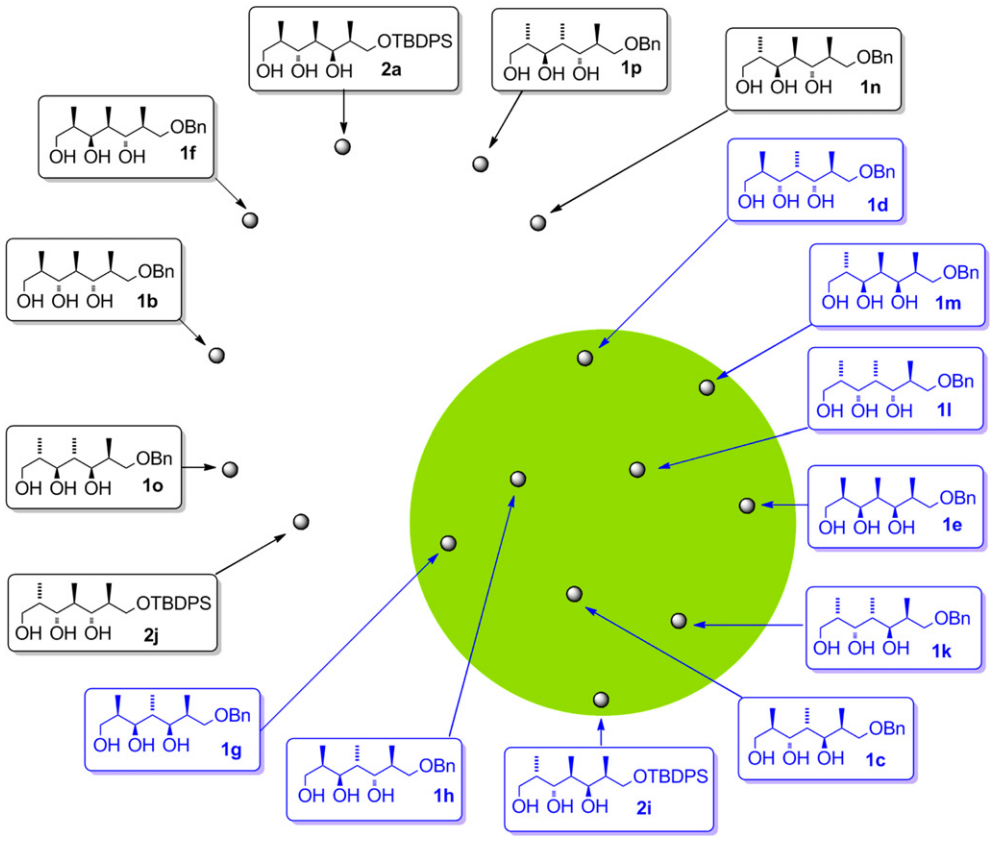


Figure 8. Cyclic representation of how similar the stereopentads are to each other.

have very different spectra from other isomers. For example **1e** has a similar spectrum to **1k** (yellow in Fig. 7) and **1l** has a similar spectrum to **1h**.

2.3. Using NMR shift calculation to distinguish between diastereoisomers

Some isomers have very similar NMR shifts and so are difficult to distinguish from each other. Our recently developed methodology for distinguishing between isomers using GIAO NMR shift calculation and the CP3 parameter[38] could be used to distinguish these molecules.

In order to test the ability of our approach to distinguish between diastereoisomers of **1** and **2**, the assignment of each of the 120 ($^{16}C_2$) possible pairs of the sixteen isomers using our CP3 parameter was investigated. In the same way as for generating Figure 7, the protecting group atoms and C7 were ignored in order to allow comparison between the benzyl and TBDPS-protected stereopentads. However the GIAO calculated shifts were now used in place of the ChemDraw ones to identify the shifts to be removed. The ChemDraw and GIAO calculations were in agreement about, which shifts to remove in almost all cases, and in the few cases where there was a disagreement the two alternative shifts to be removed were very similar. The confidence in each structural assignment made was also calculated, using Bayes theorem with the method we have previously described and with the necessary expectation values and standard deviations taken from our previous publication.[38]

Our CP3 parameter assigned 99 of the 120 pairs both correctly with high confidence (over 85% certain) and 82 of these pairs were assigned with over 99% certainty. Nineteen pairs did not give a clear assignment, and only two were assigned incorrectly with some confidence (more than 85% but less than 95%). Full details of the results of assigning each of the 120 pairs using CP3 and also using the correlation coefficient and mean absolute error may be found in the Supplementary data, Table S3.

Comparing the results with Figure 7, most of the pairs that CP3 (using both ^{13}C and 1H data) did not assign with high confidence correspond to pairs that have similar spectra (yellow and light green squares in Fig. 7). However, it is not necessarily the case that pairs of isomers with similar spectra according to Figure 7 are difficult to assign, since many such pairs are correctly assigned by CP3 with good confidence. Indeed, almost two-thirds (11/16) of the distinct pairs represented by yellow squares (most similar spectra) in Figure 7 are correctly assigned with over 99% confidence by CP3. Presumably, this is because CP3 focuses on the particular shifts in any two spectra that show the greatest differences. The two pairs, which were assigned incorrectly (**1c** and **1f**; **1c** and **1h**) both include **1c**, which is in the middle of the region of similar structures (green circle, Fig. 8). The confusion of **1c** and **1h** arises from the great similarity of these two structures. The reason for the confusion of **1c** and **1f** is less clear.

In the great majority of cases, CP3 can correctly distinguish isomers with high confidence. In about 16% of the remaining comparisons, confidence is low. In less than 2% of cases, the incorrect assignment is made with medium confidence.

3. Computational details

All molecular mechanics calculations were performed using Macromodel[40] (Version 9.5) interfaced to the Maestro[41] (Version 8.0) program. All conformational searches used the Monte Carlo Multiple Minimum[42] (MCMM) or Systematic Pseudo Monte Carlo[43] (SPMC) method and the MMFF force field.[44] The searches were done in the gas phase, with a 50 kJ mol^{-1} upper energy limit and using 250,000 steps. This number of steps turned out to be sufficiently large that, even for these highly flexible molecules, all low energy conformers were found

by an average of well over 20 times, so giving a very high degree of certainty that all low energy conformers had been found.

Quantum mechanical calculations were carried out using Jaguar[45] (Version 7.0). As in our previous studies,[16,38] the widely used B3LYP functional[46] and 6-31G(d,p) basis set[47] were employed for all calculations. NMR shielding constant calculation used the GIAO method.[48]

Our previous studies have shown that single point ab initio gas phase calculations on MMFF geometries (i.e., with no computationally expensive ab initio geometry optimisation or solvent models) gave good results for shift calculation, and the same approach was employed in the current work. The following procedure was therefore used for NMR shift calculation. First, a molecular mechanics conformational search was carried out using the MMFF force field (gas phase). Secondly, all identified conformers within 10 kJ mol^{-1} of the global minimum were subjected to single point ab initio calculations of energy and GIAO shielding constants at the B3LYP/6-31G(d,p) level (again in the gas phase). The choice of energy cutoff is necessarily a compromise between minimising computer time and the risk of missing important conformers (as judged by the subsequent ab initio energies) due to inaccurate ordering of the conformer energies by the MMFF force field. We have shown previously[16,38] that an energy cutoff of 10 kJ mol^{-1} is generally sufficient to give good results with a minimum of computation effort. This assumption turns out to be justified for the stereopentad molecules investigated here by the large number of successful assignments made by CP3. However, in order to investigate whether including more high energy molecular mechanics conformers might have given even better results (though at much greater computational cost) all the shifts were recalculated using an MMFF energy cutoff of 25 kJ mol^{-1} (i.e., including many more conformers for each isomer). Although the accuracy of the calculated shifts for a few of the isomers did show some improvement, there was little change in accuracy for most isomers and some even showed a decrease in accuracy. Further, the number of successful structure assignments by CP3 showed no improvement on that obtained using the lower energy cutoff. We therefore consider the higher energy cutoff to be an unnecessary computational expense for these particular molecules. Full details of this investigation may be found in the Supplementary data.

To calculate NMR shifts for a particular species, the shielding constants were first averaged over symmetry-related positions in each conformer and then subjected to Boltzmann averaging over the conformers i according to

$$\sigma^x = \frac{\sum_i \sigma_i^x \exp(-E_i/RT)}{\sum_i \exp(-E_i/RT)} \quad (1)$$

where σ^x is the Boltzmann averaged shielding constant for nucleus x, σ_i^x is the shielding constant for nucleus x in conformer i, and E_i is the potential energy of conformer i (relative to the global minimum), obtained from the single point ab initio calculation. The temperature T was taken as 298 K.

Chemical shifts were then calculated according to

$$\delta^x_{\text{calcd}} = \frac{\sigma^o - \sigma^x}{1 - \sigma^o/10^6}, \quad (2)$$

where δ^x_{calc} calcd is the calculated shift for nucleus x (in ppm), σ^x is the shielding constant for nucleus x from Eq. 1 and σ^o is the shielding constant for the carbon or proton nuclei in tetramethylsilane (TMS), which was obtained from a B3LYP/6-31G(d,p) calculation on TMS.

4. Conclusions

The experimental NMR data for a set of sixteen diastereomeric stereopentads are reported, and it is shown that these diastereoisomers vary in how easy they are to distinguish by NMR. The relationship between structural similarity and spectral similarity is not always straightforward, although there are some patterns. The graphical representations of similar and distinct diastereoisomers (Fig. 7 and 8) are therefore an important guide to the level of confidence that can be placed in the stereostructure assignment of a new stereopentad-containing polyketide. If, for example, the proposed structure contains one of the distinctive stereopentads (**1n**, **1p**, **2a**, **1f**, **1b**, **1o**, **2j** in Fig. 8) one can have good confidence that the structure has not been confused with another isomer. On the other hand, if the proposed structure represents one of those in the cluster of similar diastereoisomers, there is much more potential for confusion and more caution is required about the stereochemistry.

When it comes to assigning pairs of diastereoisomers (for example, as the major and minor products of a partially stereoselective reaction), Figure 7 and 8 are again an important guide as to whether the two isomers in question are likely to have sufficiently different spectra to be distinguished with good confidence. However, even if the spectra are very similar, a clear assignment is often still possible using GIAO calculated shifts combined with our CP3 parameter. This methodology for stereostructure assignment has been shown here to give excellent results in terms of assigning stereochemistry with high and quantifiable confidence, even for these very challenging stereopentad molecules.

Acknowledgements

We thank the University of Cambridge (S.G.S.) and Unilever for financial support. We thank Dr. Charlotte Bolton for the IT support.

Supplementary data

Complete experimental data for the stereopentads in Figure 4, calculated shifts for these molecules, coordinates and energies of the calculated structures, full details of assigning the structures using CP3 and other parameters, and the results of increasing the number of conformers of the stereopentads included in the shift calculations. This material is available free of charge via the Internet at http://www.elsevier.com. Supplementary data associated with this article can be found in online version at doi:10.1016/j.tet.2010.06.022. These data include MOL files and InChIKeys of the most important compounds described in this article.

References and notes

1. Rohr, J. *Angew. Chem., Int. Ed.* **2000**, *39*, 2847–2849.
2. Discodermolide: (a) Paterson, I.; Florence, G. J.; Gerlach, K.; Scott, J. P. *Angew. Chem.* **2000**, *112*, 385–388; (b) Paterson, I.; Florence, G. J.; Gerlach, K.; Scott, J. P.; Sereinig, N. *J. Am. Chem. Soc.* **2001**, *123*, 9535–9544; Oleandolide: Paterson, I.; Norcross, R. D.; Ward, R. A.; Romea, P.; Lister, M. A. *J. Am. Chem. Soc.* **1994**, *116*, 11287–11314; Reidispongiolide: Paterson, I.; Ashton, K.; Britton, R.; Cecere, G.; Chouraqui, G.; Florence, G. J.; Stafford, J. *Angew. Chem., Int. Ed.* **2007**, *46*, 6167–6171; Spirangien: Paterson, I.; Findlay, A. D.; Noti, C. *Chem.—Asian J.* **2009**, *4*, 594–611.
3. Channon, J. A.; Paterson, I. *Tetrahedron Lett.* **1992**, *33*, 797–800.
4. (a) Kobayashi, Y.; Lee, J.; Tezuka, K.; Kishi, Y. *Org. Lett.* **1999**, *1*, 2177–2180; (b) Lee, J.; Kobayashi, Y.; Tezuka, K.; Kishi, Y. *Org. Lett.* **1999**, *1*, 2181–2184; (c) Tan, Y. K. C.-H.; Kishi, Y. *J. Am. Chem. Soc.* **2001**, *123*, 2076–2078; (d) Higashibayashi, S.; Czechtizky, W.; Kobayashi, Y.; Kishi, Y. *J. Am. Chem. Soc.* **2003**, *125*, 14379–14393.
5. Fleury, E.; Lannou, M.-I.; Bistri, O.; Sautel, F.; Massiot, G.; Pancrazi, A.; Ardisson, J. *Eur. J. Org. Chem.* **2009**, *2009*, 4992–5001.
6. Saielli, G.; Bagno, A. *Org. Lett.* **2009**, *11*, 1409–1412.
7. Barone, G.; Gomez-Paloma, L.; Duca, D.; Silvestri, A.; Riccio, R.; Bifulco, G. *Chem.—Eur. J.* **2002**, *8*, 3233–3239.
8. Barone, G.; Duca, D.; Silvestri, A.; Gomez-Paloma, L.; Riccio, R.; Bifulco, G. *Chem.—Eur. J.* **2002**, *8*, 3240–3245.
9. Rychnovsky, S. D. *Org. Lett.* **2006**, *8*, 2895–2898.
10. Nicolaou, K. C.; Frederick, M. O. *Angew. Chem., Int. Ed.* **2007**, *46*, 5278–5282.
11. Aiello, A.; Fattorusso, E.; Luciano, P.; Mangoni, A.; Menna, M. *Eur. J. Org. Chem.* **2005**, *2005*, 5024–5030.

12. Bassarello, C.; Bifulco, G.; Montoro, P.; Skhirtladze, A.; Kemertelidze, E.; Pizza, C.; Piacente, S. *Tetrahedron* **2007**, *63*, 148–154.
13. Pu, J.-X.; Huang, S.-X.; Ren, J.; Xiao, W.-L.; Li, L.-M.; Li, R.-T.; Li, L.-B.; Liao, T.-G.; Lou, L.-G.; Zhu, H.-J.; Sun, H.-D. *J. Nat. Prod.* **2007**, *70*, 1707–1711.
14. Fattorusso, C.; Stendardo, E.; Appendino, G.; Fattorusso, E.; Luciano, P.; Romano, A.; Taglialatela-Scafati, O. *Org. Lett.* **2007**, *9*, 2377–2380.
15. Braddock, D. C.; Rzepa, H. S. *J. Nat. Prod.* **2008**, *71*, 728–730.
16. Smith, S. G.; Paton, R. S.; Burton, J. W.; Goodman, J. M. *J. Org. Chem.* **2008**, *73*, 4053–4062.
17. White, K. N.; Amagata, T.; Oliver, A. G.; Tenney, K.; Wenzel, P. J.; Crews, P. *J. Org. Chem.* **2008**, *73*, 8719–8722.
18. Timmons, C.; Wipf, P. *J. Org. Chem.* **2008**, *73*, 9168–9170.
19. Hu, G.; Liu, K.; Williams, L. *J. Org. Lett.* **2008**, *10*, 5493–5496.
20. Yang, J.; Huang, S.-X.; Zhao, Q.-S. *J. Phys. Chem. A* **2008**, *112*, 12132–12139.
21. Fattorusso, E.; Luciano, P.; Romano, A.; Taglialatela-Scafati, O.; Appendino, G.; Borriello, M.; Fattorusso, C. *J. Nat. Prod.* **2008**, *71*, 1988–1992.
22. Muñoz, M. A.; Joseph-Nathan, P. *Magn. Reson. Chem.* **2009**, *47*, 578–584.
23. Wang, B.; Dossey, A. T.; Walse, S. S.; Edison, A. S.; Merz, K. M., Jr. *J. Nat. Prod.* **2009**, *72*, 709–713.
24. Mendoza-Espinoza, J. A.; López-Vallejo, F.; Fragoso-Serrano, M.; Pereda-Miranda, R.; Cerda-García-Rojas, C. M. *J. Nat. Prod.* **2009**, *72*, 700–708.
25. Stappen, I.; Buchbauer, G.; Robien, W.; Wolchann, P. *Magn. Reson. Chem.* **2009**, *47*, 720–726.
26. Honma, M.; Kudo, S.; Takada, N.; Tanaka, K.; Miura, T.; Hashimoto, M. *Bioorg. Med. Chem. Lett.* **2010**, *20*, 709–712.
27. Cheeseman, J. R.; Trucks, G. W.; Keith, T. A.; Frisch, M. J. *J. Chem. Phys.* **2006**, *104*, 5497–5509.
28. Forsyth, D. A.; Sebag, A. B. *J. Am. Chem. Soc.* **1997**, *119*, 9483–9494.
29. Giesen, D. J.; Zumbulyadis, N. *Phys. Chem. Chem. Phys.* **2002**, *4*, 5498–5507.
30. Cimino, P.; Gomez-Paloma, L.; Duca, D.; Riccio, R.; Bifulco, G. *Magn. Reson. Chem.* **2004**, *42*, S26–S33.
31. Tormena, C. F.; da Silva, G. V. *J. Chem. Phys. Lett.* **2004**, *398*, 466–470.
32. Bagno, A.; Rastrelli, F.; Saielli, G. *Chem.—Eur. J.* **2006**, *12*, 5514–5525.
33. Wu, A.; Zhang, Y.; Xu, X.; Yan, Y. *J. Comput. Chem.* **2007**, *28*, 2431–2442.
34. Wiitala, K. W.; Hoyle, T. R.; Cramer, C. J. *J. Chem. Theory Comput.* **2006**, *2*, 1085–1092.
35. Zhao, Y.; Truhlar, D. G. *J. Phys. Chem. A* **2008**, *112*, 6794–6799.
36. Sarotti, A. M.; Pellegrinet, S. C. *J. Org. Chem.* **2009**, *74*, 7254–7260.
37. Bifulco, G.; Dambruoso, P.; Gomez-Paloma, L.; Riccio, R. *Chem. Rev.* **2007**, *107*, 3744–3779.
38. Smith, S. G.; Goodman, J. M. *J. Org. Chem.* **2009**, *74*, 4597–4607.
39. Hoffmann, R. W.; Weidmann, U. *Chem. Ber.* **1985**, *118*, 3980–3992.
40. Mohamadi, F.; Richards, N. G. J.; Guida, W. C.; Liskamp, R.; Lipton, M.; Caufield, C.; Chang, G.; Hendrickson, T.; Still, W. C. *J. Comput. Chem.* **1990**, *11*, 440–467.
41. *Maestro, Version 8.0, Schriödinger*; LLC: New York, NY, 2007.
42. Chang, G.; Guida, W. C.; Still, W. C. *J. Am. Chem. Soc.* **1989**, *111*, 4379–4386.
43. Goodman, J. M.; Still, W. C. *J. Comput. Chem.* **1991**, *12*, 1110–1117.
44. (a) Halgren, T. A. *J. Comput. Chem.* **1996**, *17*, 490–519; (b) Halgren, T. A. *J. Comput. Chem.* **1996**, *17*, 520–552; (c) Halgren, T. A. *J. Comput. Chem.* **1996**, *17*, 553–586; (d) Halgren, T. A.; Nachbar, R. B. *J. Comput. Chem.* **1996**, *17*, 587–615; (e) Halgren, T. A.; Nachbar, R. B. *J. Comput. Chem.* **1996**, *17*, 587–615; (f) Halgren, T. A. *J. Comput. Chem.* **1996**, *17*, 616–641; (g) Halgren, T. A. *J. Comput. Chem.* **1999**, *20*, 720–729; (h) Halgren, T. A. *J. Comput. Chem.* **1999**, *20*, 730–748.
45. *Jaguar, Version 7.0, Schrödinger*; LLC: New York, NY, 2007.
46. (a) Becke, A. D. *Phys. Rev. A* **1988**, *38*, 3098–3100; (b) Lee, C.; Yang, W.; Parr, R. G. *Phys. Rev. B* **1988**, *37*, 785–789; (c) Becke, A. D. *J. Chem. Phys.* **1993**, *98*, 5648–5652; (d) Stephens, P. J.; Devlin, F. J.; Chabalowski, C. F.; Frisch, M. J. *J. Phys. Chem.* **1994**, *98*, 11623–11627.
47. Hehre, W. J.; Radom, L.; Schleyer, P. v. R.; Pople, J. A. *Ab Initio Molecular Orbital Theory*; Wiley: New York, NY, 1986.
48. (a) London, F. *J. Phys. Radium* **1937**, *8*, 397–409; (b) Ditchfield, R. *J. Chem. Phys.* **1972**, *56*, 5688–5691; (c) Wolinski, K.; Hinton, J. F.; Pulay, P. *J. Am. Chem. Soc.* **1990**, *112*, 8251–8260.