Rapid conformation searching. Part 1. Diastereomeric compounds

Jonathan M. Goodman* and Andres Bueno Saz

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

A procedure for rapidly searching the conformation space of diastereomers is presented and evaluated. In the examples considered, it finds the low energy conformations much more quickly than Monte-Carlo conformation searching. The computational cost of this procedure is substantially less than that of a Monte-Carlo, and the conformations that are found are biased towards low energy structures.

Introduction

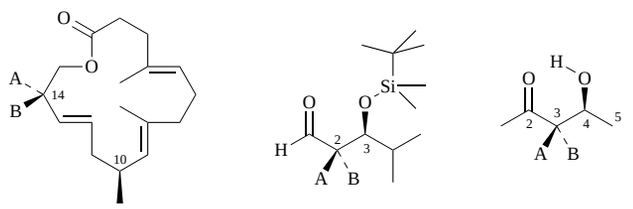
Whenever a new compound is made or contemplated, it is useful to calculate its shape in order to analyse or predict its spectral properties and reactivity. The energy of a molecule may be calculated and minimised using molecular mechanics or molecular orbital methods, creating a structure which is either the global minimum or a local minimum. The latter is much more likely for the majority of compounds, and so some form of conformation searching must be applied to the system to ensure that the global minimum has been found, as well as the low-energy local minima. Conformation searching is often the time limiting step in this process, and so a wide variety of methods has been developed to make it as rapid as possible.[1]

The usual procedure employed to perform a conformation search is to take a single molecule at a time, and subject it to some searching protocol. However, this does not fit well with the sort of questions that are often asked of molecular modellers. A reaction may form diastereomers, which must be differentiated. A library of compounds may need to be compared. Competing transition states may need to be analysed. The effect of a small change in a protecting group may need to be predicted. These questions are very suitable for computational analysis, because the key result will be a difference in energies, rather than an absolute energy, and this can be obtained with greater precision. All of these problems have required multiple conformation searches of all the species that are involved, and so the time required increases linearly with the number of species. In this paper we describe a method which substantially reduces the time required for conformation searches of this sort.

One of the ideas of a Monte-Carlo search[2] is that reasonable conformations are used to generate new reasonable conformations. A partial knowledge of the potential energy surface of a particular molecule is used to generate good structures which will profitably investigate unknown areas of the surface. If a conformation search is required for each of two similar molecules, then the searches should be able to help each other, because the two molecules' potential energy surfaces must have much in common. If a suitable way can be found to allow the searches to communicate, then the combined search should take much less time than performing the two searches in isolation.

One simple way in which information may be shared between the searches is to use the result of one search as the starting point for another. With this in mind, a program FLIP, was written which takes the output from a conformation search and inverts selected stereocentres.† This program was then tested on several diastereomeric systems (**1–4**, Figs. 1 and 2).

† The executable for the program, compiled for Silicon Graphics workstations, is available on the Cambridge Chemistry Department World-Wide-Web server (http://www.ch.cam.ac.uk/MMRG/software/).

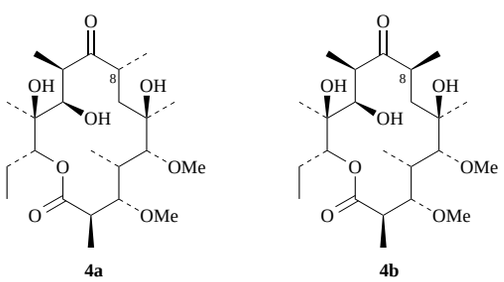


1a A = H; B = Me
1b A = Me; B = H

2a A = H; B = Me
2b A = Me; B = H

3a A = H; B = Me
3b A = Me; B = H

Fig. 1 Trial diastereometric systems



4a **4b**

Fig. 2 Erythromycin analogues

Results and discussion

FLIP† must be given five atom numbers as its input. The first is the atom number of the centre which will be inverted, the second and third are attached atoms which do not move, and the fourth and fifth are atoms which do move. FLIP first finds which atoms need to be moved by exploring all the atoms attached to the moving ones, and then checks that there are no rings which prevent the inversion process. A vector is found which goes through the inversion centre and bisects the direction of the two fixed bonds. The atoms which need to be moved are then rotated by 180° around this vector. This has the effect of inverting the chosen chiral centre, whilst conserving any chiral centres on the side chains which are being moved.

This program allows us to investigate a number of questions: are the potential energy surfaces of diastereomers sufficiently similar for this process to be useful? If a molecule has several chiral centres, which is the best one to invert? How much time can the procedure save? The following examples provide data to answer these questions.

Example 1

The macrocyclic lactone **1a** (Fig. 1) was made by Still and Romero as a substrate for polyepoxidation towards the synthesis of monensin.[3] A full conformation search of this system was not feasible in 1986, but is now possible, although requiring a great deal of computer time. Full conformation searches were carried out on **1a** and **1b**, using MACROMODEL 4.5,[4] and the

Table 1 Conformation searches on **1a** and **1b**

	Time required/ min	Global minimum/ kJ mol^{-1}	Number of conformations up to:		
			4 kJ mol^{-1}	10 kJ mol^{-1}	50 kJ mol^{-1}
Monte-Carlo search					
1a (5000 steps)	1301	108.5	4	57	1312
1b (5000 steps)	1341	109.4	11	66	1234
FLIP **1b** to **1a**					
Invert C-14 of **1b**	148	108.5	4	51	1002
Invert C-10 of **1b**	242	108.5	4	55	966
Monte Carlo search and **1b** (C-14) combined[a]	84	108.5	4	57	1551
Combine searches **1b** (C-10) and **1b** (C-14)[a]	75	108.5	4	58	1353
FLIP **1a** to **1b**					
Invert C-14 of **1a**	220	109.4	11	59	1023
Invert C-10 of **1a**	259	109.4	11	59	962
Monte-Carlo search and **1a** (C-14) combined[a]	50	109.4	11	67	1502
Combine searches **1a** (C-10) and **1a** (C-14)[a]	42	109.4	11	65	1381

[a] The results of these searches were concatenated and reminimised to remove all the duplicate structures.

MM2 forcefield.[5] Each of these found over a thousand local minima as well as the global minimum for each diastereomer. The searches found the lowest energy structures several times each, and this is a normal criterion for stopping the search. They are unlikely to have found all the local minima within 50 kJ mol^{-1} of the global minimum.

The output file of the conformation search of **1a** was subjected to the stereoinverting program, FLIP. Every minimum energy conformation of **1a** was turned into a conformation of **1b**. The structures were then reminimised. The results are given in Table 1. The global minimum was found in every case, and the table lists the number of structures found within 4, 10 and 50 kJ mol^{-1} of the global minimum. FLIP processed the 1312 conformers in a few seconds. The reminimisation of the new conformations was quicker, per structure, than the Monte-Carlo minimisations, because the inversion of a stereocentre does not affect bond lengths, bond angles and the ring conformation and so all the structures generated by FLIP are likely to be close to a local minimum. The total time for the inversion and reminimisation was a sixth of the time required for the Monte-Carlo search of **1b**.

The efficiency of this procedure was assessed by comparing the conformations (Table 1). All structures within 4 kJ mol^{-1} above the global minimum were found. 90% of structures within 10 kJ mol^{-1} were found. More than 80% as many structures within 50 kJ mol^{-1} of the global minimum were found by the stereoinversion program as the Monte-Carlo search. The conformations found by FLIP were compared with the structures found by the Monte-Carlo search. It was discovered that the Monte-Carlo search had not found all structures within 50 kJ mol^{-1} of the global minimum, and that the flipping procedure had generated some new structures. However, all the low energy structures were found by both methods.

We also tried inverting the other stereocentre, to get the enantiomeric compound. This gave rather similar results. We then combined the results of inverting each stereocentre. This produces more structures than were found by the initial Monte-Carlo search, in rather less total time.

We also tried turning **1b** into **1a**. This was quicker than the Monte-Carlo search by a factor of nine, but a slightly lower proportion of the conformations was found. However, all structures within 4 kJ mol^{-1} of the global minimum were found.

The searches for both diastereomers were run for 5000 steps. In the final 500 steps of each search 65 (**1a**) and 58 (**1b**) new structures were found. However, all of these were high in energy (more than 12 kJ mol^{-1} above the global minimum). All of the low energy structures were found several times.

The results suggest that once a conformation search of one diastereomer has been performed, the conformation space of the other diastereomer can be investigated very much more rapidly. In this case, the results suggest that the diastereomer of the macrolide prepared by Still and Romero would not give the correct configuration for monensin at the new chiral centres formed by epoxidation.

Example 2

The aldehydes **2a** and **2b** were made by Paterson *et al*. in the synthesis of swinholide,[6] and their conformational properties are interesting because of the stereocontrol exerted by the β-chiral centre. Complete conformation searches were done on each diastereomer, and then the stereoinversion procedure was applied to each (Table 2). Once again, the potential energy surface of each diastereomer was effectively searched by modifying the other. For both searches, all conformations within 4 kJ mol^{-1} were found, more than 80% of conformations within 10 kJ mol^{-1} of the global minimum, and over 60% of the conformations within 50 kJ mol^{-1}. The time saving for inverting then minimising the conformations rather than running a new conformation search was over 90% in both cases.

The procedure was repeated, inverting the hydroxy centre by swapping the hydrogen and the isopropyl group. This was somewhat less effective, as it took slightly longer and found fewer structures. It might be expected that the inversion making the smallest change to the molecule will be the most effective, and so swapping a hydrogen and a methyl should be preferred to swapping a hydrogen and an isopropyl.

Example 3

The hydroxy ketones, **3**, were chosen because they contain a hydrogen bond, and because their small size allows exhaustive searching. Both of the chiral centres were inverted by swapping all six possible pairs of substituents. In all cases, the global minimum and all structures within 4 kJ mol^{-1} of the global minimum were found. The results are given in Table 3. All the reminimisations took *ca*. 1 min.

The results in Table 3 are given in pairs, because if each centre were a perfect tetrahedron there would only be half as many entries. For example, swapping the hydrogen and ethyl on carbon four would be identical to swapping the hydroxy group and the main chain of the molecule. The carbon atoms are not

Table 2 Conformation searches on **2a** and **2b**

	Time required/ min	Global minimum/ kJ mol^{-1}	Number of conformations up to:		
			4 kJ mol^{-1}	10 kJ mol^{-1}	50 kJ mol^{-1}
Monte-Carlo search					
2a (3000 steps)	543	53.59	14	55	251
2b (3000 steps)	550	49.21	2	17	265
FLIP **2b** to **2a**					
Flip H and Me (C-2) of all **2b** conformations	43	53.59	14	46	164
Flip H and Pr^{i}(C-3) of all **2b** conformations	50	53.95	13	35	125
FLIP **2a** to **2b**					
Flip H and Me (C-2) of all **2a** conformations	38	49.21	2	14	167
Flip H and Pr^{i}(C-3) of all **2a** conformations	44	49.21	2	16	129

Table 3 Conformation searches on **3a** and **3b**

	Time required/ min	Global minimum/ kJ mol^{-1}	Number of conformations up to:		
			4 kJ mol^{-1}	10 kJ mol^{-1}	50 kJ mol^{-1}
Monte-Carlo search					
3a (1000 steps)	36	3.665	2	3	25
3b (1000 steps)	39	5.083	2	2	20

	Conformations up to:			Conformations up to:	
FLIP **3b** to **3a**	10 kJ mol^{-1}	50 kJ mol^{-1}	FLIP **3a** to **3b**	10 kJ mol^{-1}	50 kJ mol^{-1}
FLIP C-3			FLIP C-3		
H and Me	3	17	H and Me	2	17
C-2 and C-4	3	17	C-2 and C-4	2	17
H and C-2	2	13	H and C-2	2	15
Me and C-4	3	14	Me and C-4	2	15
H and C-4	2	15	H and C-4	2	14
Me and C-2	2	15	Me and C-2	2	14
FLIP C-4			FLIP C-4		
H and Me		16	H and Me	2	19
OH and C-3	3	18	OH and C-3	2	19
H and OH	3	15	H and OH	2	14
Me and C-3	3	15	Me and C-3	2	13
H and C-3	3	18	H and C-3	2	15
Me and OH	3	18	Me and OH	2	15

perfectly tetrahedral and this accounts for the small differences in some of the results.

Table 3 confirms that the best results are obtained by making the smallest possible changes to the molecules. It is better to swap a hydrogen for a methyl than a hydrogen for an acetyl.

Example 4

Erythromycin is an important antibiotic, with a very complicated potential energy surface. A conformation search was performed on erythromycin analogue **4a** (Fig. 2), using the AMBER[7] forcefield and the GB/SA solvent model for water,[8] as implemented in MACROMODEL.[4] Six days searching (9999 Monte-Carlo steps) found 2393 structures within 50 kJ mol^{-1} of the lowest energy structure.The lowest energy structure was found only twice, so there can be no great confidence that it corresponded to the global minimum.

The hydrogen and the methyl group at C-8 were inverted, to form **4b**, and the 2393 new structures were minimised. This produced 686 unique conformations and took one day. In order to establish how representative these structures are, a new Monte-Carlo search was run, starting from the lowest energy conformation for **4b**. This took six days, and found 1962 structures. The lowest energy structure was the starting structure, produced by the stereoinversion procedure. The stereoinversion procedure found seven structures within 4 kJ mol^{-1} of the lowest energy structure. The Monte-Carlo search found only six of these structures. Within 10 kJ mol^{-1} the stereoinversion procedure found 31 structures and the Monte-Carlo search found 40. Within 50 kJ mol^{-1} the steroinversion procedure found only 686 structures and the Monte-Carlo search found 1962. The low energy structures are the most important ones, and this result demonstrates that the stereoinversion procedure is biased towards these. The Monte-Carlo search was helped by being given the lowest energy structures found by the FLIP procedure. If the search had begun at a higher energy structure, as it would have done if it had used a modified version of the global minimum of **4a** or a random conformation of **4b**, then it

would probably have been less effective. The lowest energy structure of **4b** was generated by inverting and minimising the 25th lowest energy structure of **4a**.

Conclusions

A method for greatly increasing the speed of conformation searching for diastereomers has been developed. It has been tested on cyclic, and acyclic systems, and on systems with hydrogen-bonding interactions. The best choice of centres to invert are those which produce the smallest change (*i.e.* swapping hydrogen and methyl, rather than hydrogen and isopropyl). If there are several chiral centres, then changing more than one and combining the results gives a more complete conformation search.

The inversion procedure does not find as many structures as a Monte-Carlo search, but it is much faster, and the structures that are missed tend to be high in energy. The distribution of reminimised stereoinverted structures seems to favour low energy conformations, and these are the conformations which are most interesting.

Studies on complementary ways of relating the conformation surfaces of similar molecules are underway.

Acknowledgements

We thank the Royal Society, the Cambridge Centre for Molecular Recognition and NATO for support; Professor Mark Lipton (Purdue) for helpful discussions.

References

1 (*a*) M. Saunders, K. N. Houk, Y. D. Wu, W. C. Still, M. Lipton, G. Chang and W. Guida, *J. Am. Chem. Soc.*, 1990, **112**, 1419; (*b*) J. M. Goodman and W. C. Still, *J. Comp. Chem.*, 1991, **12**, 1110; (*c*) A. R. Leach, *Reviews in Computational Chemistry*, VCH, Weinheim, 1991, vol. 2, pp. 1–55.
2 G. Chang, W. C. Guida and W. C. Still, *J. Am. Chem. Soc.*, 1989, **111**, 4379.
3 W. C. Still and A. G. Romero, *J. Am. Chem. Soc.*, 1986, **108**, 2105.
4 MACROMODEL version 4.5: F. Mohamedi, N. G. J. Richards, W. C. Guida, R. Liskamp, M. Lipton, C. Caulfield, G. Chang, T. Hendrickson and W. C. Still, *J. Comp. Chem.*, 1990, **11**, 440.
5 N. L. Allinger, *J. Am. Chem. Soc.*, 1977, **99**, 8127.
6 I. Paterson, J. G. Cumming, J. D. Smith and R. A. Ward, *Tetrahedron Lett.*, 1994, **35**, 441.
7 (*a*) S. J. Weiner, P. A. Kollman, D. A. Case, U. C. Singh, C. Ghio, G. Alagona, S. Profeta and P. Weiner, *J. Am. Chem. Soc.*, 1984, **106**, 765; (*b*) S. J. Weiner, P. A. Kollman, D. T. Nguyen and D. A. Case, *J. Comp. Chem.*, 1986, **7**, 230.
8 W. C. Still, A. Tempczyk, R. C. Hawley and T. Hendrickson, *J. Am. Chem. Soc.*, 1990, **112**, 6127.

Paper 6/06711E
Received 30*th September* 1996
Accepted 18*th February* 1997