Acta Cryst. (1995). B51, 99–103

Attempted Prediction of the Crystal Structures of Six Monosaccharides

BY BOUKE P. VAN EIJCK, WIJNAND T. M. MOOIJ AND JAN KROON

Department of Crystal and Structural Chemistry, Bijvoet Center for Biomolecular Research, Utrecht University, Padualaan 8, 3584 CH Utrecht, The Netherlands

(*Received* 17 *July* 1994; *accepted* 19 *August* 1994)

Abstract

A method is reported to generate possible crystal structures of the six hexopyranoses where comparison with an X-ray determination is possible. In these molecules, internal degrees of freedom are all-important. Using essentially only the information that the space group is $P2_12_12_1$ with one molecule in the asymmetric unit, a systematic search was made for all low-energy crystal structures of these substances. The energies were minimized with respect to nine lattice and rigid-body parameters and six intramolecular dihedral angles. The number of possible structures within the range 10 kcal mol^{-1} is of the order 1000. In all cases, the experimental structure was among them, and in four cases this was either the structure with the lowest energy or only a few tenths of a kcal mol^{-1} higher. However, in the two other cases the relative energy of the experimental structure was over 5 kcal mol^{-1}. Such calculations can provide a sensitive test for force fields.

Introduction

Six years ago Maddox (1988) made the provocative statement: 'One of the continuing scandals in the physical sciences is that it remains in general impossible to predict the structure of even the simplest crystalline solids from a knowledge of their chemical composition'. Since then a few researchers have responded to this challenge.

Gavezzotti (1991) was the first to report significant progress in this field. He started by coupling pairs of molecules by one of a few symmetry operations, and selecting those with a promisingly low interaction energy for further consideration. A search for possible translations then led directly to space groups with two molecules per cell, whereas space groups with four molecules per cell can be obtained after a second clustering of the molecule pairs. In favorable cases, the crystal structures of hydrocarbons could be reproduced in this way. Independently, Gdanitz (1992) developed a different approach where the space group is chosen in advance and the cell dimensions and the molecular positions are determined by Monte Carlo simulated annealing. He was able to reproduce the crystal structures of hexamethylbenzene and ethene; later on (Karfunkel & Gdanitz, 1992; Karfunkel, Leusen & Gdanitz, 1993), the method was extended with considerable success to molecules that also contain O and N atoms. Recently, Holden, Du & Ammon (1993) used a concept of frequently occurring coordination patterns to develop a rather general method for the prediction of crystal structures of organic compounds. They compared four possible space groups and found that the structure with the lowest energy was either the X-ray structure, or was only slightly higher in energy.

All these promising results refer to structures that are constructed from rigid molecules, often with a pronounced shape which might limit the number of acceptable stacking patterns. Some authors announce an extension to molecules with internal degrees of freedom, but such applications have, to our knowledge, not yet been published.

Some of our recent efforts have been directed towards the development of a reliable force field for carbohydrates (Kouwijzer, van Eijck, Kroes & Kroon, 1993). In the present work, we shall attempt to predict the possible low-energy crystal structures of hexopyranoses and compare them with the experimental data. In these molecules, there are six internal degrees of freedom that play an essential role: the dihedral angle ω = O5—C5—C6—O6, which defines the conformation of the exocyclic CH_2OH group, and five C—C—O—H dihedral angles. Any fair attempt to predict such structures must consider these parameters as unknown in advance.

Carbohydrate structures are dominated by hydrogen bonds, and we suspect that Gavezzotti's clustering method might not lead the way to the molecular clusters actually present in the full structure. Perlstein (1992, 1994) has shown that the latter often correspond to local energy minima no more than a few kcal mol^{-1} above the global minimum of the corresponding free cluster, but he excluded hydrogen-bonded structures from the investigation. Taking acetic acid (Jönsson, 1971) as an example, it is hard to see how the clustering method could avoid ending up with a (incorrect) dimer structure. We thus considered the entire crystal structure from the beginning. As Karfunkel & Gdanitz (1992) reported difficulties with hydrogen-bonded structures, we did not use their Monte Carlo method but relied

on a systematic search instead. Such a search is at present only feasible for structures with one independent molecule in the asymmetric unit, where the number of degrees of freedom is limited: six intramolecular dihedral angles together with the cell edges and the positions and orientations of the molecule in the cell give 15 parameters to be determined for each crystal structure. Fortunately, this class covers more than 90% of the entries in the Cambridge Structural Database (Padmaja, Ramakumar & Viswamitra, 1990), although this percentage may be lower for hydrogen-bonded molecules (Gavezzotti & Filippini, 1994). Six out of the seven reported hexopyranose structures fall into this category, the exception being α-D-mannose (Longchambon, Avenel & Neuman, 1976). They all crystallize in space group $P2_12_12_1$ and we based our strategy on this knowledge. For chiral compounds, this is again not as restrictive as it may seem: $P2_12_12_1$ accounts for more than half of such molecules and *ca* 90% would be covered after extension of the search to space groups $P2_1$ and $P1$, which does not seem excessively difficult.

Method

Preliminaries

For each substance, we constructed a standard input molecular structure with the hydroxymethyl group in the TG conformation ($\omega = 180°$) and the five hydroxyl dihedral angles set to zero. The structures of the free molecules were energy-minimized under these restraints, and then transformed to principal axes. Thus, essentially all experimental information was removed, except that the $^4C_1$ ring configuration (Allen & Fortier, 1993) was assumed throughout.

Any position of these basic blocks in a crystal cell can be obtained by rotation about Euler angles φ, θ, ψ (the exact convention was taken from Allen & Tildesley, 1987), followed by translations X, Y, Z parallel to the cell axes a, b, c. The space spanned by these nine variables is huge, but fortunately it can be reduced by symmetry: each structure is invariant by translations over half a cell edge, or by rotation over 180° about the cell axes (Table 1). It is seen that the range of Euler angles can be limited to $0 \le \varphi < 2\pi$, $0 \le \theta < \frac{1}{2}\pi$, $0 \le \psi < \pi$; all in all, the search effort can generally be reduced by a factor of 32. The space-group symmetry may allow further reduction: in $P2_12_12_1$, the three axes are equivalent and can be interchanged, so the condition $a < b < c$ can be applied.

In the calculations, the same structure is often obtained from different starting points, not necessarily with all parameters within the standard ranges indicated above. A comparison with the experimental structure must also be made. To verify whether or not two structures are equivalent, the consequences of an interchange of axes must be carefully studied; it is advantageous to consider these interchanges as 90° rotations about the cell axes. From the equivalent positions (Hahn, 1983), it can be deduced that such an interchange necessitates a shift of one quarter along the cell diagonal. As an example, the relations between the parameters of two equivalent structures connected by a C_z^4 rotation are given in Table 2.

Table 1. *Equivalency relations for all space groups*

$\varphi' = \varphi + \pi$	$\theta' = \pi - \theta$	$\psi' = \pi - \psi$	$X' = X$	$Y' = -Y$	$Z' = -Z$
$\varphi' = \varphi + \pi$	$\theta' = \pi - \theta$	$\psi' = -\psi$	$X' = -X$	$Y' = Y$	$Z' = -Z$
$\varphi' = \varphi$	$\theta' = \theta$	$\psi' = \pi + \psi$	$X' = -X$	$Y' = -Y$	$Z' = Z$
$X' = X + \frac{1}{2}a$					
$X' = Y + \frac{1}{2}b$					
$Z' = Z + \frac{1}{2}c$					

In earlier work (Kouwijzer, van Eijck, Kroes & Kroon, 1993), we concluded that the *GROMOS* force field (van Gunsteren & Berendsen, 1987) performed very well for monosaccharides, so we also used it here. In this force field, all CH and CH_2 groups are treated as united atoms, and chirality around these centers is maintained by harmonic potentials for improper dihedral angles. The hydroxyl H atoms are considered explicitly and are modeled with a point charge only. No special hydrogen-bond potential is used; electrostatic interactions and the O···O Lennard–Jones interaction being sufficient. No Ewald summations are carried out, the concept of unbreakable neutral charge groups being used instead; only the interactions between charge groups within a given cutoff radius were taken into account. The calculation of the energy and its derivatives with respect to the parameters was painstakingly optimized, using pair lists to speed up energy minimizations.

Search algorithm

The search procedure consists of two stages. In the first stage, the molecules are considered as rigid units. Inspired by the idea behind the *CHEAT* force field (Grootenhuis & Haasnoot, 1993), we removed the hydroxyl H atoms and, more generally, even all electrostatic interactions. A rough replacement of the hydrogen bonds thus neglected was obtained by greatly enhancing the O···O Lennard–Jones potential to a minimum energy of 5.2 kcal mol^{-1} at 2.8 Å. This oversimplified force field reproduced fairly well the main features of the experimental structures in a short molecular dynamics simulation. As long as the correct starting structures for the following stage are obtained, the precise properties of the force field are not important. The main requirement is simplicity to obtain great speed for the calculations of the energy and its derivatives. For the bulky exocyclic CH_2OH group, such a united-atom approach is not applicable: all calculations were done separately for the three possibilities $\omega = -60°$ (GG), $\omega = 60°$ (GT), $\omega = 180°$ (TG).

The next scan is over Euler angles. The φ-range was scanned in 18 steps of 20°. The ψ-step should be inversely proportional to $\sin\theta$, and so the ψ-range of

Table 2. *Equivalency relations for rotation C_z^4 in space group $P2_12_12_1$*

$a' = b$	$b' = a$	$c' = c$
$X' = -Y + b/4$	$Y' = X + a/4$	$Z' = Z + c/4$
$A'_{1i} = -A_{2i}$	$A'_{2i} = A_{1i}$	$A'_{3i} = A_{3i}$

A is the Euler matrix (Allen & Tildesley, 1987) and $i = 1, 2, 3$.

180° was divided into 9 steps for $\theta = 80°$, 8 for $\theta = 60°$, 6 for $\theta = 40°$, 4 for $\theta = 20°$ and 1 for $\theta = 0°$, for a total of 504 orientations. The following scan is over cell edges: b and c were varied in 1 Å steps and a was calculated from the experimental density. This information is not critical, as an error of 10% corresponding to less than 1 Å in a can be easily tolerated. With the restrictions $a < b \leq c$, $a > 4.5$ Å and $c < 20$ Å, the number of different cells was 43. The final scan is over X, Y, Z translations, which were varied between zero and half the cell edges in 1 Å steps.

All in all, over 8 million trial structures were examined. The cutoff radius for energy calculations was rather small (7 Å). Depending on several energy thresholds, a structure was rejected or kept for further consideration at various points in the algorithm. This consisted of a preliminary energy minimization with respect to the nine variables φ, θ, ψ, a, b, c, X, Y and Z. We found that the best convergence was obtained by starting with a few steepest-descent steps without line optimization, followed by a regular conjugate gradients procedure (Press, Flannery, Teukolsky & Vetterling, 1986). For every set of Euler angles, a first clustering of equivalent solutions was done. After a further conjugate gradients minimization cycle, the entire set of structures was again reduced by clustering of equivalent solutions; all possibilities exemplified in Tables 1 and 2 were checked. For each of the three ω-values the parameters of the 2000 structures with lowest energy were taken over to the second stage.

In this second stage the hydroxyl H atoms are introduced, with five dihedral angles as additional parameters to be determined from various starting positions. Now, of course, the standard force field was used, except for the cutoff radius which was still 7 Å. The same energy minimization procedure was used; in the steepest-descent phase, the weights of these five internal parameters were greatly enhanced with respect to the nine crystallographic parameters. To our surprise it was sufficient to use two starting positions for each dihedral angle, corresponding to 32 calculations for each structure selected in the first stage. Each calculation is now more time-consuming since Coulomb contributions and intramolecular interactions can no longer be neglected. For the structures that survive all energy thresholds, the second stage is finished with an all-atom energy minimization where only the cell axes are fixed. In the end equivalent solutions are clustered, the results for the three ω-values are merged and the structures with lowest energy, with complete sets of atomic coordinates, are collected. On a Silicon Graphics Challenge computer (150 MHz processor R4400), the first stage took *ca* 20 h per structure and the second stage *ca* 10 h.

Results

At the end of the search the number of remaining structures was of the order 1000, within the energy range *ca* 10 kcal mol^{-1}. This list is now so limited that more effort can be spent on each candidate. In fact, the problem changes its character here: it is no longer a matter of generating possible solutions, but rather of finding out which of them corresponds to the most stable structure.

In all cases, the experimental structure was present. There is never any doubt about which calculated structure corresponds to this experimental one, Euler angles being within 4°, cell edges within 0.6 Å, molecular centers within 0.4 Å, and hydroxyl dihedrals within 15°. The one problematic case is β-D-glucose, where a redetermination of the crystal structure (Kouwijzer, van Eijck, Kooijman & Kroon, 1995) indicated a value of 168° for the dihedral angle C5—C4—O4—H4, in contrast to an earlier value of −143° (Chu & Jeffrey, 1968). Both values were found within 25° as possible solutions.

However, in this stage the experimental structure never corresponded to that calculated with the lowest energy (Table 3, entry *U*7). This should not be surprising as the cutoff radius was unrealistically short. Therefore, the energy calculation was then refined to improve the selection of the correct solution from the list. All atomic coordinates as well as the cell axes were varied to minimize the energy. First the cutoff radius was increased to 10 Å (Table 3, entry *U*10). Then the aliphatic H atoms were added in order to use a recently developed all-atom force field (Kouwijzer, van Eijck, Kooijman & Kroon, 1994). Table 3 shows that this gave quite an improvement: for the first four substances, the correct solution was either first or second in ranking, with a relative energy of at most 0.2 kcal mol^{-1}, which is inside any reasonable estimate of the accuracy of the force field. Increasing the cutoff radius to 30 Å was not really helpful, as might have been expected since the force field was not parameterized for such a value. In any case, it is clear that the ranking of all solutions depends dramatically on such details of the force field.

At present we can offer no explanation for the systematically bad ranking of β-D-glucose and β-D-galactose, in comparison with the other four structures where the results are very satisfactory. A 20 ps molecular dynamics simulation at room temperature did not change the relative energies significantly. We report the atomic coordinates of the structures with lowest energy for these compounds in Table 4, to allow interested readers to test

Table 3. *Results of the crystal structure prediction*

	*U*7		*U*10		*A*10		*A*30	
	R	*E*	*R*	*E*	*R*	*E*	*R*	*E*
α-D-Galactose	32	3.7	8	2.5	2	0.2	1	
α-D-Glucose	21	2.3	8	1.5	2	0.2	3	1.1
α-D-Talose	58	3.6	20	2.4	1		1	
β-D-Allose	3	0.5	6	0.7	2	0.1	2	0.9
β-D-Galactose	29	4.8	105	4.9	231	6.0	88	4.9
β-D-Glucose	356	5.6	213	5.7	374	5.9	384	5.1

E is the energy (kcal mol^{-1}) of the correct structure with respect to the structure of lowest energy, *R* is the corresponding ranking number of the correct structure. *U* refers to the standard *GROMOS* united-atom force field, *A* to the all-atom force field; the number after these codes gives the cutoff radius in Å. Column *U*7 gives the results as they emerge from the search procedure, the other three are obtained from supposedly better force fields.

References for experimental structures: α-D-galactose: Jeffrey & Shiono (1987), Kouwijzer, van Eijck, Kooijman & Kroon (1995); α-D-glucose: Brown & Levy (1979); α-D-talose: Ohanessian, Avenel, Kanters & Smits (1977); β-D-allose: Kroon-Batenburg, van der Sluis & Kanters (1984); β-D-galactose: Longchambon, Ohanessian, Avenel & Neuman (1975); β-D-glucose: Chu & Jeffrey (1968); Kouwijzer, van Eijck, Kooijman & Kroon (1995).

their favorite force field on these hypothetical structures in comparison with the published experimental ones.

Table 4. *Fractional coordinates of the structure with lowest energy for β-D-galactose and β-D-glucose, space group $P2_12_12_1$*

	β-D-Galactose $a = 7.49$ Å $b = 7.78$ Å $c = 11.66$ Å			β-D-Glucose $a = 6.05$ Å $b = 9.35$ Å $c = 11.87$ Å		
	x	*y*	*z*	*x*	*y*	*z*
C1	0.356	0.384	0.280	0.639	0.251	0.145
C2	0.248	0.223	0.310	0.427	0.281	0.211
C3	0.061	0.267	0.353	0.266	0.378	0.147
C4	0.063	0.408	0.446	0.374	0.501	0.083
C5	0.167	0.562	0.397	0.579	0.452	0.018
C6	0.176	0.715	0.479	0.704	0.579	−0.033
O1	0.542	0.345	0.270	0.816	0.214	0.219
O2	0.232	0.115	0.209	0.315	0.147	0.232
O3	−0.027	0.114	0.394	0.110	0.436	0.227
O4	0.135	0.352	0.554	0.215	0.559	0.005
O5	0.348	0.515	0.367	0.732	0.372	0.088
O6	0.224	0.868	0.417	0.800	0.673	0.050
H(O1)	0.575	0.263	0.336	0.819	0.106	0.232
H(O2)	0.340	0.036	0.202	0.250	0.147	0.312
H(O3)	0.068	0.020	0.409	−0.026	0.370	0.224
H(O4)	0.272	0.368	0.555	0.063	0.564	0.043
H(O6)	0.322	0.839	0.358	0.776	0.779	0.027
H(C1)	0.303	0.439	0.200	0.606	0.165	0.083
H(C2)	0.319	0.155	0.379	0.473	0.334	0.290
H(C3)	−0.013	0.319	0.280	0.179	0.311	0.086
H(C4)	−0.076	0.450	0.457	0.423	0.583	0.144
H(C5)	0.099	0.604	0.319	0.525	0.382	−0.051
H(C61)	0.276	0.687	0.545	0.589	0.637	−0.088
H(C62)	0.046	0.731	0.522	0.833	0.535	−0.089

Discussion

Although the *GROMOS* force field was never designed for this kind of work, it performs remarkably well. As pointed out by Karfunkel & Gdanitz (1992), a severe test for a force field is not only that it should reproduce the observed crystal structure well, but also that it should be shown that this structure is global rather than just a local minimum. Considerations of temperature (the calculations apply to 0 K) and crystal growth (kinetic effects and critical cluster size could be important) may cast some doubts on the validity of a simple energy criterion: the experimental structure might be one of several possible polymorphs, not necessarily the one with the lowest energy. Nevertheless, as the discrepancies for β-D-glucose and β-D-galactose are over 5 kcal mol^{-1}, we feel that improvement of the force field should be our first aim now. We have noted that in all calculated structures, the hydrogen bonds tend to be rather short and that, correspondingly, the density is often somewhat high. It is noteworthy that this applies especially to the low-energy structures reported in Table 4, where the density is about 10% over the average experimental value. More fundamentally, Karfunkel & Gdanitz (1992) have found that the atomic partial charges must be optimized for each molecule individually in order to obtain reliable results. For flexible molecules this is difficult to do, and we suspect that in the end it may be necessary to take explicit polarization terms into account. Then, further shortening of the list with possible crystal structures and the subsequent use of more sophisticated methods like molecular dynamics should become possible. Although complete *ab initio* crystal structure determinations are not yet reliable, the possibility of combining such information with X-ray powder data (Holden, Du & Ammon, 1993; Harris, Tremayne, Lightfoot & Bruce, 1994) is also definitely promising.

We are impressed by the astonishingly large number of structures within a few kcal mol^{-1}. This is equivalent to predicting a large tendency towards polymorphism. This tendency appears to be larger than in the other published structure predictions, and one of the referees suggested that it might be related to the notoriously poor crystallization propensity of sugars. Some of the structures differ only in the positions of the hydroxyl H atoms, but most are entirely different from each other. Evidently, the directional properties of the five hydrogen bonds (Steiner & Saenger, 1992; Gavezzotti & Filippini, 1994) can be easily accommodated in many different ways. This is entirely due to the conformational freedom of the hydroxyl groups.

References

Allen, F. H. & Fortier, S. (1993). *Acta Cryst.* B**49**, 1021–1031.

Allen, M. P. & Tildesley, D. J. (1987). *Computer Simulation of Liquids*, p. 86. Oxford: Oxford University Press.

Brown, G. M. & Levy, H. A. (1979). *Acta Cryst.* B**35**, 656–659.

Chu, S. S. C. & Jeffrey, G. A. (1968). *Acta Cryst.* B**24**, 830–838.

Gavezzotti, A. (1991). *J. Am. Chem. Soc.* **113**, 4622–4629.

Gavezzotti, A. & Filippini, G. (1994). *J. Phys. Chem.* **98**, 4831–4837.

Gdanitz, R. J. (1992). *Chem. Phys. Lett.* **190**, 391–396.

GROOTENHUIS, P. D. J. & HAASNOOT, C. A. G. (1993). *Mol. Simul.* **10**, 75–95.
GUNSTEREN, W. F. VAN & BERENDSEN, H. J. C. (1987). *GROMOS. Groningen Molecular Simulation Package*. Univ. of Groningen, The Netherlands.
HAHN, T. (1983). *International Tables for X-ray Crystallography*, Vol. A. Dordrecht: Kluwer Academic Publishers.
HARRIS, K. D. M., TREMAYNE, M., LIGHTFOOT, P. & BRUCE, P. G. (1994). *J. Am. Chem. Soc.* **116**, 3543–3547.
HOLDEN, J. R., DU, Z. & AMMON, H. L. (1993). *J. Comput. Chem.* **14**, 422–437.
JEFFREY, G. A. & SHIONO, R. (1977). *Acta Cryst.* B**33**, 2700–2701.
JÖNSSON, P.-G. (1971). *Acta Cryst.* B**27**, 893–898.
KARFUNKEL, H. R. & GDANITZ, R. J. (1992). *J. Comput. Chem.* **13**, 1171–1183.
KARFUNKEL, H. R., LEUSEN, F. J. J. & GDANITZ, R. J. (1993). *J. Comput.-Aided Mat. Des.* **1**, 177–185.
KOUWIJZER, M. L. C. E., VAN EIJCK, B. P., KOOIJMAN, H. & KROON, J. (1995). *Acta Cryst.* In the press.
KOUWIJZER, M. L. C. E., VAN EIJCK, B. P., KROES, S. J. & KROON, J. (1993). *J. Comp. Chem.* **14**, 1281–1289.
KROON-BATENBURG, L. M. J., VAN DER MSLUIS, P. & KANTERS, J. A. (1984). *Acta Cryst.* C**40**, 1863–1865.
LONGCHAMBON, F., AVENEL, D. & NEUMAN, A. (1976). *Acta Cryst.* B**32**, 1822–1826.
LONGCHAMBON, F., OHANESSIAN, J., AVENEL, D. & NEUMAN, A. (1975). *Acta Cryst.* B**31**, 2623–2627.
MADDOX, J. (1988). *Nature*, **335**, 201.
OHANESSIAN, J., AVENEL, D., KANTERS, J. A. & SMITS, D. (1977). *Acta Cryst.* B**33**, 1063–1066.
PADMAJA, N., RAMAKUMAR, S. & VISWAMITRA, M. A. (1990). *Acta Cryst.* A**46**, 725–730.
PERLSTEIN, J. (1992). *J. Am. Chem. Soc.* **114**, 1955–1963.
PERLSTEIN, J. (1994). *J. Am. Chem. Soc.* **116**, 455–470.
PRESS, W. H., FLANNERY, B. P., TEUKOLSKY, S. A. & VETTERLING, W. T. (1986). *Numerical Recipes, the Art of Scientific Computing*. Cambridge: Cambridge University Press.
STEINER, T. & SAENGER, W. (1992). *Acta Cryst.* B**48**, 819–827.

Acta Cryst. (1995). B**51**, 103–107

A Database Study of Intermolecular NH···O Hydrogen Bonds for Carboxylates, Sulfonates and Monohydrogen Phosphonates

BY B. PIRARD, G. BAUDOUX AND F. DURANT

Laboratoire de Chimie Moléculaire Structurale, Facultés Universitaires Notre-Dame de la Paix, rue de Bruxelles 61, *B*-5000 *Namur, Belgium*

(*Received* 22 *April* 1994; *accepted* 15 *September* 1994)

Abstract

A search of the Cambridge Structural Database (CSD, version 5.05, 1993) was performed in order to compare the geometrical features of the hydrogen bonds involving on the one hand amino groups and on the other hand carboxylates, sulfonates or monohydrogen phosponates. Phosphonates were not considered because only four entries containing amino and phosphonate moieities were located in the CSD. The hydroxylic group of monohydrogen phosphonates primarily acts as a hydrogen-bond donor. The three moieties under study show NH···O hydrogen bonds with similar geometrical features. This statistical analysis has focused on the hydrogen-bond distances and angles and on the distributions of the H atoms around the acceptor O atoms of carboxylates, sulfonates or monohydrogen phosphonates.

Introduction

The medicinal chemistry literature contains numerous examples of bioisosteric modifications (Davies, Williams & Smith, 1987, and references therein; Lipinski, 1986, and references therein). Bioisosterism is one of the strategies used to develop more potent and less toxic analogues of a lead compound. Within this scope, crystallographic databases such as the Cambridge Structural Database (CSD) (Allen, Kennard & Taylor, 1983) provide an invaluable source of information about potential intermolecular interactions between drug molecules and macromolecules (Klebe, 1994). Crystal packing data enable one to suggest similar patterns of interactions for different functional groups with a common partner (Allen, 1992, and references therein) and hence to consider new bioisosteric replacements.

This strategy has been applied to carboxylic acids and to some of their isosteres. Phosphinic (Howson, Mistry, Broekman & Hills, 1993; Baylis, Campbell & Dingwall, 1984), phosphonic (Chiefari, Galanopoulos, Janowski, Kerr & Prager, 1987; Lipinski, 1986, and references therein), sulfonic acids (Abbenante & Prager, 1992; Lipinski, 1986, and references therein), sulfonamides and tetrazoles (Davies, Williams & Smith, 1987, and references therein; Lipinski, 1986, and references therein) are current surrogates for carboxylic groups. For instance, the carboxylic group of γ-aminobutyric acid, an inhibitory neutrotransmitter, has been replaced by phosphinic, phosphonic or sulfonic groups (Bowery, 1993). In this contribution, we have focused on the