

# Comparing hypothetical structures generated in the third Cambridge blind test of crystal structure prediction

**Bouke P. van Eijck**Utrecht University (Retired), Department of  
Crystal and Structural Chemistry, Padualaan 8,  
3584 CH Utrecht, The NetherlandsCorrespondence e-mail:  
b.p.vaneijck@chem.uu.nlReceived 12 May 2005  
Accepted 16 August 2005

In the third Cambridge blind test of crystal structure prediction, participants submitted extended lists of up to 100 hypothetical structures. In this paper these lists are analyzed for the two small semi-rigid molecules, hydantoin and azetidine, by performing a new energy minimization using an accurate force field, and grouping these newly minimized structures into clusters of equivalent structures. Many participants found the same low-energy structures, but no list appeared to be complete even for the structures with one independent molecule in the asymmetric unit. This may well be due to the fact that a cutoff at even 100 structures cannot ensure the presence of a structure that has a relatively high ranking in another force field. Moreover, some structures should have possibly been discarded because they correspond to transition states rather than true energy minima. The r.m.s. deviation between energies in corresponding clusters was calculated to compare the reported relative crystal energies for each pair of participants. Some groups of force fields show a reasonably good correspondence, yet the order of magnitude of their discrepancies is comparable to the energy differences between, say, the first ten structures of lowest energy. Therefore, even if we assume that energy is a sufficient criterion, it is not surprising that crystal structure predictions are still inconsistent and unreliable.

## 1. Introduction

The prediction of crystal structures is a field of research that, after interesting early attempts, could only be developed more systematically with the advent in the early 1990s of sufficiently fast computers. Various methods have been developed to generate possible crystal structures without any experimental information other than the chemical structural formula and the empirical knowledge incorporated in most force fields. One common result of these efforts is that there is generally a multitude of possible structures within a relatively small energy window of a few  $\text{kJ mol}^{-1}$ . It is surprising that there are so many ways to pack molecules with nearly equal efficiency into a space-filling pattern, but the evidence is unmistakable. The consequence is that crystal structure prediction is still very difficult: one has to find a ranking for the hypothetical crystal structures to predict the most probable polymorphs. Usually this ranking is based on the static crystal energy, but in fact one should search for the global minimum of Gibbs free energy. Moreover, kinetic effects in crystal growth and the mechanical stability of the crystal should be considered. Unfortunately, all these effects as well as the inherent uncertainty of any force field combine to make the choice of possible observable polymorphs out of the multitude of hypothetical crystal structures a very difficult task.

**Table 1**  
Acceptable predictions in the three Cambridge blind tests.

Molecule	Category	Participants	Rankings	Year
(Ia)	(i), form 1	11	1, 1, 1, 3	1999
(Ib)	(i), form 2	11	–	1999
(II)	(ii)	8	2	1999
(III)	(iii)	11	1	1999
(IV)	(i)	15	2, 3	2001
(V)	(ii)	15	1, 1, 1, 2	2001
(VI)	(iii)	11	–	2001
(VII)	(i)	6	1	1999
(VIII)	(i)	15	1, 1, 1, 2	2004
(IX)	(ii)	15	1	2004
(X)	(iii)	15	–	2004
(XI)	(i), $Z' = 2$	18	–	2004

Thus, it is not surprising that methodology is the main subject in many publications, which implies that substances with known structures are studied. Even for the most scrupulous investigator there is the temptation to continue the work until the experimental structure is at least found, albeit possibly with a disappointingly high energy ranking. Therefore, it was an excellent initiative of the Cambridge Crystallographic Data Centre (CCDC) to organize blind tests in 1999 (Lommerse *et al.*, 2000), 2001 (Motherwell *et al.*, 2002) and in

2004 (Day *et al.*, 2005). In each test the chemical diagrams of three types of molecules were given:

(i) small rigid molecules (fewer than 25 atoms) with only C, H, N and O atoms;

(ii) possibly larger rigid molecules containing less common elements;

(iii) molecules with several torsional degrees of flexibility.

Each participant was allowed to present three predictions per molecule. The number of participants was 11 in the first test, 15 in the second one and 18 in the last one – although not every participant attempted every molecule. The molecules studied are collected in Fig. 1 and a summary of the results is given in Table 1.

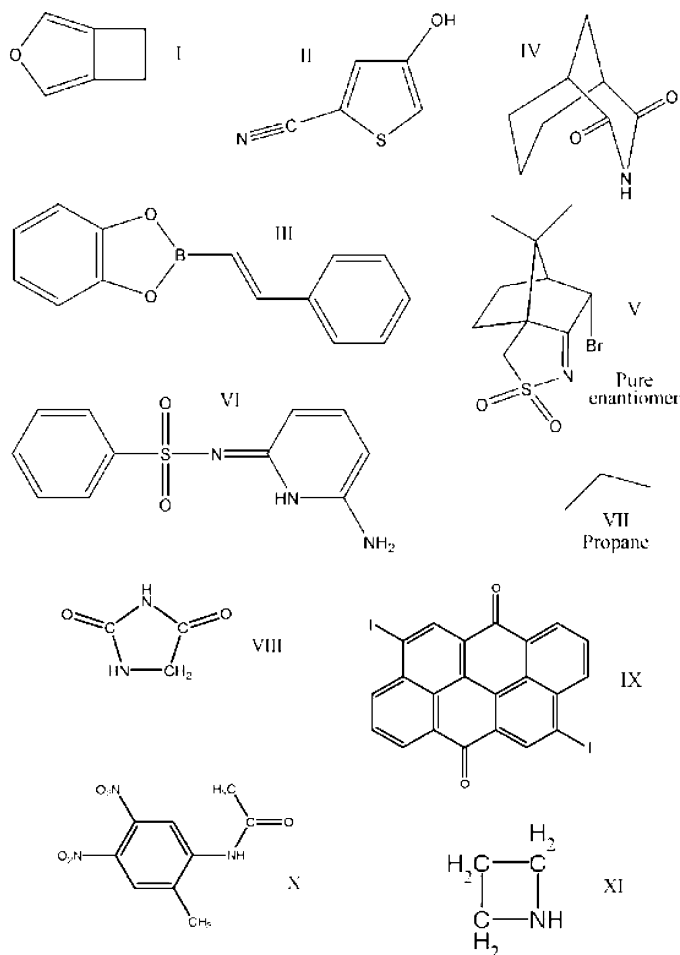
Not surprisingly, it was found that category (iii) is the most difficult: in the three tests only one correct prediction was made. The results for the first two categories are better, but still the overall chance of a correct prediction is below 10%. A failure can be caused in two ways: the structure generation may have failed to produce an observed polymorph at all or its ranking may be too high to be recognized as a real solution. In the latter case there is a magical excuse for the failure: the observed structure is metastable and a careful experimental search for polymorphs would find the predicted one. Interestingly, such a situation has recently been reported (Hulme *et al.*, 2005).

In the 2004 blind test the participants were invited to submit not only their three predictions, but also their best 100 (or so) structures. These extended lists were deposited as supplementary material. This presents an opportunity to perform a large-scale comparison, which is the object of the present paper. Firstly, the overlap between equivalent structures in the lists should indicate the degree of completeness in the structure generations. Secondly, comparing the relative energies of equivalent structures should give a measure of the correspondence between the different force fields. This form of analysis was not explicitly envisaged at the time, so it is understandable that quite a few participants were content to submit significantly fewer than 100 structures. As will be shown here, even that number may be too low for a completely satisfactory analysis.

For this comparison we studied the two molecules of the most favorable category: hydantoin (VIII) and azetidine (XI). Their molecular structures are shown in Fig. 2. Table 2 lists the blind test participants whose extended results were available, with a short summary of their methods. For more details and full literature references the original paper and its supplementary material (Day *et al.*, 2005) should be consulted. In the blind test it was given that the compounds crystallized with at most two independent molecules in the unit cell. Therefore, in this work we disregarded all the reported structures with  $Z' > 2$  which could not be reduced to lower  $Z'$ . This left 746 structures for hydantoin and 900 for azetidine.

## 2. Clustering of equivalent structures

The various lists contain many equivalent structures; indeed, in the ideal case there should be a one-to-one correspondence



**Figure 1**  
The molecules in the three Cambridge blind tests.

**Table 2**

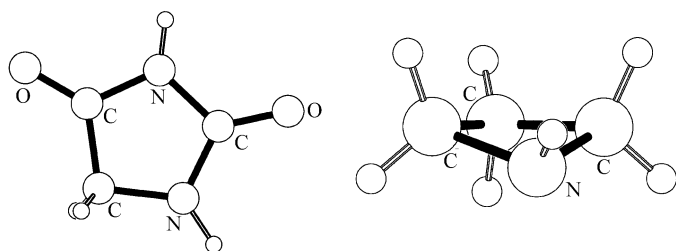
Summary of methodologies.

$N_h$  and  $N_a$  are the numbers of structures for hydantoin and azetidine, respectively. Search: MC: Monte Carlo; GS: grid-based systematic; GA: genetic algorithm; PR: Sobol' pseudo-random; R: random; S: stepwise construction *via* dimers and layers. No. of space groups investigated:  $S_1$ :  $Z' \leq 1$ ,  $S_2$ :  $Z' > 1$ . EM: energy minimization with flexible (F) or rigid (R) molecules. Electrostatics: PC: *ab initio* fitted point charges; PCS: same + satellite charges; DM: *ab initio* distributed multipoles; FF: charges from force field definition. vdW: van der Waals: Various exp-6 models unless noted otherwise: LJ: Lennard-Jones 12-6.

Code	Participant	$N_h$	$N_a$	Search	$S_1$	$S_2$	EM	Electrostatics	vdW
A	Scheraga & Arnautova	9	–	MC	1	1	R	PC	
B	Boerrigter	100	100	MC	1	1	F	PC	
C	Dzyabchenko	–	100	GS	12	0	R	PC, FF	LJ
D	Della Valle & Venuti	100	100	PR	1	1	R	PC	
E	Erk	99	48	GS	4	1	F	PC	
F	Facelli <i>et al.</i>	104	102	GA	14	14	F	PC	LJ
H	Hofmann	100	100	R	10	0	R	From database	
K	Pantelides & Karamertzanis	71	93	PR	59	59	R	PCS	
L	Liang	15	20	MC	12	0	F	PC, FF	
M	Motherwell	–	103	GA	10	0	R	None	
P	Price <i>et al.</i>	45	24	GS	11	0	R	DM	
S	Schweizer & Dunitz	–	22	S	6	0	R	PC	
V	van Eijck	100	100	R	13	5	F	DM	
W	Verwer	10	10	MC	12	0	F	PC, FF	
Y	Day	69	100	MC	17	10	R	DM	

between all the structures in all the lists. The actual situation is dramatically different. The problem of determining which structures are equivalent is far less trivial than it might seem. Several publications (Dzyabchenko, 1994; Chisholm & Motherwell, 2005; Willighagen *et al.*, 2005) discuss working algorithms. Our procedure is based on a comparison of the intermolecular atomic distances (Mooij *et al.*, 1998). Each structure is expanded to  $P1$  and for each combination of two atom types a list is made of distances up to 5 Å. Comparing these lists for one structure gives its number of independent molecules and comparing two structures reveals their similarity. A distance criterion of 0.05 Å was used; the procedure works reliably only when the energies were previously minimized in the same force field.

The clustering problem is illustrated in Fig. 3, which shows schematically the energy as a function of one of the many geometrical degrees of freedom as calculated in two possible force fields. The question is whether or not the reported structures A, B, C or D should be considered as equivalent. The geometrical difference is intuitively the most simple criterion: in Fig. 3, only B and C might be clustered together. Obviously this decision depends on the rather arbitrary choice of a distance cutoff. Moreover, in this study structures have been created with quite different methods which would make it impossible to order the resulting clusters by energy.

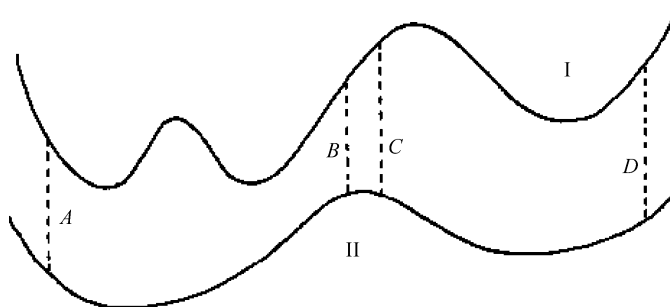


**Figure 2**

Molecular structures of hydantoin (left) and azetidine (right), showing the H(N) atom *trans* to the CCC plane.

Therefore, it was decided to compare the structures after energy minimization. The comparison of the geometries becomes much easier and the energy ranking follows automatically. Force field I will cluster only B and C, whereas force field II will produce the clusters (A, B) and (C, D). So the problem is replaced by the choice of force field, but at least there is in principle a best choice.

For hydantoin and azetidine we have carried out energy minimizations in two force fields, considering the molecules as flexible throughout. The simplest force field had point charges (PC) from 6-31G\*\* wavefunctions, whereas in the more elaborate one a distributed multipole model (DM) was used with atomic multipoles obtained by fitting to the electrostatic potential of an SCF/DZ(2d<sup>ON</sup>) wavefunction (Mooij *et al.*, 1999). All the *ab initio* calculations were carried out with the program GAMESS-UK (Guest *et al.*, 2005), with charges and multipoles being obtained using MOLDEN (Schaftenaar & Noordik, 2000). In both force fields intermolecular dispersion and repulsion were modeled by a Buckingham potential taken from Coombes *et al.* (1996). As the DM force field had produced acceptably low relative energies for the observed structures, it was considered adequate for the clustering purpose. For the other two compounds of the blind test no



**Figure 3**

Illustrating the possible correspondences between energy minima in two force fields.

force field of acceptable accuracy was available, so these compounds were not included in this work.

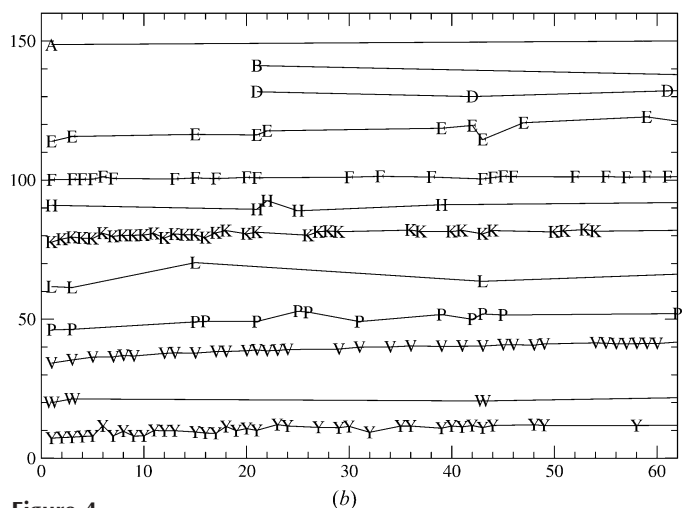
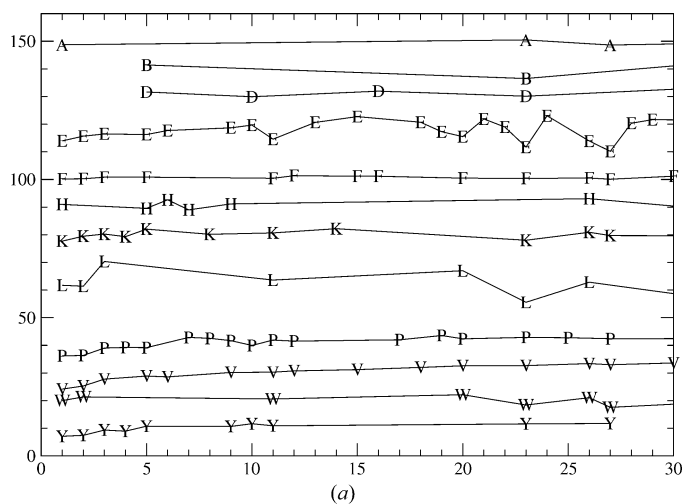
Intramolecular parameters were adjusted to reproduce 6-31G\*-optimized models of the free molecules. Hydantoin was treated as an essentially rigid molecule, and charges and multipoles were calculated only once and used for every structure. For azetidine the situation is different, because the nitrogen-bonded hydrogen atom may point towards the CCC plane (*trans*), away from it (*cis*) or somewhere in between. For this reason the charges and multipoles were calculated separately for every structure and updated when a torsional angle changed more than  $5^\circ$ . The intramolecular energy should now also be taken into account. The best procedure is to calculate this contribution directly and iteratively (van Eijck *et al.*, 2001), but this would cost an excessive amount of computer time. As *ab initio* calculations showed a significant preference for the *trans* form (Day *et al.*, 2005), it was considered sufficient to set an additional intramolecular H–N–C–C torsion potential so as to avoid *cis* structures. As it turned out, the

experimental structure contained two independent *trans* molecules (Fig. 2).

After the initial selection of structures with  $Z' \leq 2$ , no symmetry analysis was performed either before or after the energy minimizations. The space groups were kept as reported by the participants, except where limitations in the *UPACK* program (which cannot handle space groups with high symmetry or molecules on special positions) necessitated the use of a space group with lower symmetry. Thus, one aspect that was little studied is possible symmetry breaking. This may cause more structures to become equivalent and it may also produce  $Z' > 2$ . In fact, three force fields suggested symmetry breaking for the experimental structure of azetidine (Day *et al.*, 2005), leading to  $Z' = 4$  with the fatal implication that this structure should have been rejected according to the rules of the blind test!

Upon clustering of the energy-minimized structures it was found that the situation illustrated in Fig. 3 occurs occasionally. Nevertheless, the two force fields gave nearly the same results, at least for the low-energy structures. As expected, the clusters obtained by direct comparison of the geometries were dependent on the size of the molecular coordination sphere and on the choice of the distance criterion. Using the settings described in the blind test report (Day *et al.*, 2005), the three methods produced identical results when looking for the experimental structures: they all indicated the same clusters of structures as those published in that report.

An exception to this generally favourable situation occurred for planar ring conformations of azetidine in structures submitted by three participants. Upon energy minimization these structures were forced into the *trans* form, but the geometry changes were significant enough to cause some degree of arbitrariness in the clustering. Rather than excluding these entries altogether, as might have been done with a purely geometric criterion, these somewhat doubtful entries were retained in the procedure.



**Figure 4**  
Clusters for hydantoin with  $Z' \leq 1$  (upper) and all  $Z'$  (lower). Each entry on the horizontal axis labels a cluster of structures. Code letters (Table 2) indicate the participants who reported that cluster. The vertical axis gives the reported energies ( $\text{kJ mol}^{-1}$ ). Horizontal lines connect the clusters for each participant, each with an individual and arbitrary zero energy level. The experimental structure is cluster 1.

### 3. Completeness of structure generation<sup>1</sup>

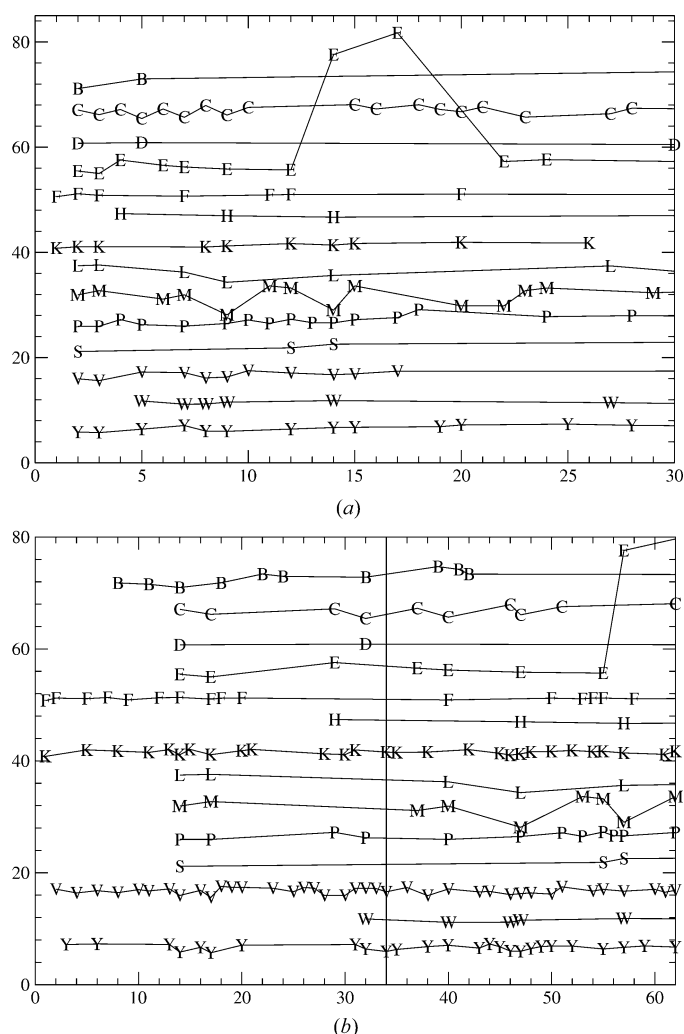
This section discusses clustering together all the structures that could conceivably be equivalent. When in either of the two force fields (DM and PC) structures became identical, they were collected in the same cluster. For the situation given in Fig. 3 all four structures would be clustered together (but note that this would not be the case if structure C was absent). The idea is that a comparable convergence could also occur during the actual crystallization process, where thermal motion can overcome barriers of a few  $\text{kJ mol}^{-1}$ .

In Figs. 4 and 5 each number on the horizontal axis labels a cluster containing structures from at least one participant. These clusters are ordered according to the energy calculated in the DM force field, which was considered to be the more reliable of the two. For each participant the structures are

<sup>1</sup> Supplementary data for this paper are available from the IUCr electronic archives (Reference: NS5004). Services for accessing these data are described at the back of the journal.

indicated by the one-letter code defined in Table 2 and linked by lines just for clarity. Each line starts at the cluster with the lowest energy and leaves at the right-hand side in the direction of the first cluster outside the figure. The vertical axis gives the energy values reported by each participant, converted into  $\text{kJ mol}^{-1}$  if necessary. All the energies reported have their individual zero levels, which have no physical significance. Therefore, each curve was shifted with an arbitrary value, chosen to keep the figure distinguishable: energy differences are only relevant within each line. Frequently, structures from one participant converged to the same cluster, in which case the lowest reported energy was retained.

As we wish to assess the completeness of a set of possible structures, it is important to distinguish between the case where  $Z' \leq 1$  and the others. Structures with  $Z' > 1$  are notoriously numerous and not all participants endeavoured to generate such structures (Table 2). The results for  $Z' > 1$  have been presented in the figures up to cluster 62, which corresponds to the  $Z' \leq 1$  cluster 17 for hydantoin and 15 for azetidine.



**Figure 5** Clusters for azetidine with  $Z' \leq 1$  (upper) and all  $Z'$  (lower). For an explanation see Fig. 4. The vertical line indicates the experimental structure.

The experimental structure for hydantoin (Fig. 4) corresponds to cluster 1. It was present in the lists of ten participants. This is a favourable compound for crystal structure prediction, as most participants agree that there is an energy gap between the global minimum and the next best structures. The blind test for this molecule was flawed by the fact that the experimental structure was disclosed before the date the predictions were to be submitted. So the obvious objection is that a high number of correct ‘predictions’ is suspect. This work shows that there is also a reasonable correspondence between hypothetical structures, to which no suspicion can be attached. We believe, therefore, in the integrity of the submissions.

For azetidine (Fig. 5) the situation is different. The experimental structure ( $Z' = 2$ ) corresponds to cluster 34, as indicated by the vertical line in Fig. 5. There are distressingly many structures within a very small energy range. Against this background it is already gratifying that the experimental structure was encountered by three participants (*K*, *V* and *Y*), albeit with ranks  $>3$ . The remarkable discontinuities in curve *E* will be explained in §4. The global energy minimum in the DM force field has been found by only two participants (*F* and *K*). This structure has the unusual space group  $I4_1/a$  ( $Z' = 1$ ), so it could only be found by studying this space group explicitly or by considering the space group  $C2/c$  with  $Z' = 2$ , which could also have produced it.

As a general conclusion, it can be seen that no participant found all the structures, not even for  $Z' \leq 1$ . There may be a simple explanation for this apparent failure: structures absent from the figures may have actually been found by a certain participant, but not reported because its original ranking was larger than 100. This will occur especially for those who included  $Z' = 2$  structures in their lists: for instance, the curves *K* and *Y* for hydantoin are remarkably well filled for all  $Z'$ , leaving fewer structures for  $Z' \leq 1$ . Thus, the questions of search completeness and energy ranking cannot be separated as easily as was hoped initially.

There are few structures that have been encountered only once, which suggests that the combined efforts of the participants create a fairly comprehensive picture of the possible low-energy structures. Apart from the tetragonal space group discussed above, there were no obvious common features (except, of course,  $Z'$ ) of the structures that were found less consistently.

Three participants (*A*, *B* and *D*) only generated structures in the space group  $P1$ , with several values of  $Z'$ . In principle, this approach could generate structures in all the possible space groups in one run, but in practice the success rate is low (van Eijck & Kroon, 2000).

#### 4. Comparison of energies

An equally important question concerns the correspondence between the energies of equivalent structures in the different force fields. Even without clustering, it is possible to minimize the energies of all the structures in a preferred force field and to compare the energies with the submitted ones. Unfortu-



**Table 3**  
Comparing pairs of participants for hydantoin.

	<i>B</i>	<i>E</i>	<i>H</i>	<i>W</i>	<i>K</i>	<i>F</i>	<i>A</i>	<i>Y</i>	<i>P</i>	<i>V</i>	<i>L</i>	<i>D</i>
<i>B</i>	29	1.9	–	–	–	–	–	–	–	–	–	5.3
<i>E</i>	6	67	3.6	0.7	2.1	3.2	3.0	3.6	4.1	4.3	5.0	3.1
<i>H</i>	2	14	34	–	2.8	1.8	–	–	3.6	3.5	5.2	3.3
<i>W</i>	1	9	3	10	1.7	1.5	2.4	3.6	4.1	4.2	4.8	–
<i>K</i>	2	10	4	9	68	1.1	2.8	1.4	2.6	3.0	6.6	–
<i>F</i>	3	22	10	10	25	92	1.0	1.6	2.6	2.9	6.6	2.2
<i>A</i>	1	5	3	5	6	6	7	–	–	4.2	5.0	–
<i>Y</i>	2	9	3	5	31	18	3	45	1.0	1.4	6.1	–
<i>P</i>	2	12	4	6	9	9	3	9	19	1.1	6.1	2.1
<i>V</i>	2	16	6	8	30	30	5	28	11	99	7.9	2.3
<i>L</i>	1	8	4	8	8	10	4	5	6	8	10	–
<i>D</i>	17	7	4	1	2	5	1	3	4	4	1	56

Entries on the diagonal and lower part of the table indicate the numbers of common structures for each pair of participants, entries above the diagonal give the r.m.s. energy difference ( $\text{kJ mol}^{-1}$ ).

**Table 4**  
Comparing pairs of participants for azetidine.

	<i>B</i>	<i>W</i>	<i>K</i>	<i>F</i>	<i>Y</i>	<i>P</i>	<i>V</i>	<i>C</i>	<i>D</i>	<i>E</i>	<i>H</i>	<i>S</i>	<i>L</i>	<i>M</i>
<i>B</i>	18	–	0.6	2.1	–	–	0.5	–	–	–	–	–	–	–
<i>W</i>	1	8	–	–	0.6	0.2	0.5	1.0	–	–	–	–	1.3	–
<i>K</i>	5	3	52	0.4	0.3	0.3	0.3	0.9	–	0.6	–	–	1.7	2.2
<i>F</i>	4	1	12	78	0.7	0.6	0.7	0.4	–	0.7	–	–	–	1.4
<i>Y</i>	2	5	23	7	46	0.6	0.4	1.0	–	1.1	–	–	1.6	2.2
<i>P</i>	2	4	6	5	9	17	0.5	0.7	–	1.4	0.9	–	1.6	1.9
<i>V</i>	5	5	24	10	26	10	68	1.3	–	1.6	–	–	1.7	2.0
<i>C</i>	3	6	6	4	10	10	8	34	0.9	4.5	1.4	–	1.1	1.5
<i>D</i>	2	1	1	1	2	2	2	6	51	–	–	–	–	–
<i>E</i>	2	3	5	4	6	7	7	6	1	16	–	1.1	4.2	2.0
<i>H</i>	0	2	2	1	2	4	2	5	3	2	50	1.2	–	1.6
<i>S</i>	1	1	3	2	3	3	3	1	2	4	5	13	–	2.2
<i>L</i>	1	6	4	3	6	5	5	8	1	6	2	2	11	0.9
<i>M</i>	1	3	7	6	8	9	8	11	3	6	4	4	5	28

For an explanation see Table 3.

nately, the choice of that force field will introduce a rather subjective bias. Therefore, once again we resorted to comparing clusters of structures. The idea this time was to select on a more restrictive basis of geometry correspondence, because energy minimization may destroy the effects that are being established. However, the clusters became just too small to obtain useful results and we kept the previous clustering.

One possibility would be to shift all the reported energies to a common level and calculate an r.m.s. discrepancy between the energies of the structures in each cluster. However, this approach suffers from a generally low number of structures in each cluster, as well as from the objection that one outlying set of energies can obscure the correspondence within the entire set. A more satisfactory method is to compare the energies of pairs of participants in clusters where both have entries. For each pair with at least four common structures the energy differences were calculated for these common structures, shifted to produce zero average difference and used to calculate r.m.s. energy differences.

Here a difficulty arose for the structures submitted by Erk, who investigated planar rings as well as *trans* molecules for azetidine. Widely different energies were reported, presumably due to an intramolecular effect. Our energy minimization forced all the molecules into the *trans* form, making such

energy differences nonsensical. This is the cause of the peculiar *E* curves in Fig. 5. Therefore, only the *trans* structures were retained for the energy comparison. Motherwell and Facelli employed only planar molecules and these data were supposed to remain meaningful after the energy minimization. A comparable problem occurred for hydantoin, where Price submitted structures with planar as well as slightly deformed molecules with different energy schemes; only the first ones were retained.

Tables 3 and 4 present the numbers of common structures and the r.m.s. values for hydantoin and azetidine, respectively. The numbers on the diagonal are generally smaller than the corresponding entries in Table 2, because structures with  $Z' > 2$  were eliminated and a few others converged to the same energy minimum. The first comment on the tables must be that they cannot in themselves indicate whether a force field is 'good' or 'bad'. If all force fields would give the same erroneous results except one, all the r.m.s. values would be large for that one result.

It was attempted to arrange the tables so as to keep the small r.m.s. values close to the diagonal. This

should facilitate a search for families of entries where all the pairs of r.m.s. values are small. However, it is necessary to attach importance to rather small differences before we can distinguish the following families: for hydantoin (*W*, *K*, *F*) and (*Y*, *P*, *V*), and for azetidine, both these groups combined. Interestingly, it is even more difficult to find common features in the energy protocols for these families. All that we observe is that (*Y*, *P*, *V*) are the force fields that use distributed multipoles to calculate the Coulomb energy. Within these families the r.m.s. values are of the order of  $1.3 \text{ kJ mol}^{-1}$  for hydantoin and  $0.5 \text{ kJ mol}^{-1}$  for azetidine.

## 5. Discussion and conclusions

As indicated above, any conclusions depend somewhat on our method of clustering, which is necessarily arbitrary. Our method is the result of comparing structures after energy minimization in a chosen force field, whose main feature is the use of distributed multipoles (DM), which is supposed to be acceptably accurate. This approach may cause structures that appear to be geometrically widely different to converge to the same energy minimum, a process that could well be similar to real crystallization. Another method or a different force field would give results which are different in detail, but not in their

general aspects. However, these 'details' may include the apparent merits of various methods and investigators. The present study is unavoidably biased in favour of force fields that use distributed multipoles.

Even for 'simple' molecules in the blind test category (i) it is not a trivial matter to generate a complete set of hypothetically possible structures. No participant has obtained all the structures generated by the collaborative workshop together, not even when we restrict ourselves to the structures with one independent molecule in the unit cell. Many participants, including the present author, felt that at least this part of the problem was essentially solved for rigid molecules in common space groups with  $Z' = 1$ , as few experimental structures were missed. This view may have been too optimistic, but, as discussed above, it is also possible that the actual situation is better because the energy ranking is so uncertain that 100 structures are not enough to ensure comparability between the submissions. This is an important lesson for subsequent blind tests.

Another consideration is that we did not investigate whether the energy minima might be transition states, because we did not study second derivatives or remove imposed symmetry elements. Participants who did take that trouble are at a disadvantage because they may have eliminated, and thus 'missed', some structures that occur incorrectly in the lists used in the present work.

The situation is better for hydantoin (Fig. 4) than for azetidine (Fig. 5), where there are more structures within a small energy range. It is remarkable that the results are apparently not appreciably worse for the  $Z' > 1$  structures, which are much more numerous. However, as only four participants studied more than one space group with  $Z' = 2$ , the total set of structures may be less complete than for  $Z' \leq 1$ .

The brute-force crystal structure generation method of Karamertzanis & Pantelides (2005) has set a new standard, taking into account 59 space groups with  $Z' = 1$  as well as  $Z' = 2$ . This should find unusual space groups that are generally neglected. At present, this approach needs large computer resources, a problem that tends to diminish spontaneously with time. Yet even this method apparently missed structures that were found by others. Moreover, the occurrence of unusual space groups is not a major source of failure in structure prediction and an indirect approach may accomplish the same (van Eijck, 2002). For example, the DM global minimum for azetidine in space group  $I4_1/a$  was indeed found by Karamertzanis and Pantelides, but also by Facelli, presumably by considering the space group  $C2/c$  with  $Z' = 2$ .

The two molecules studied have rather different schemes of hydrogen bonding: hydantoin has four atoms each involved in an  $N-H \cdots O=C$  bond, whereas azetidine has one center acting both as a donor and an acceptor in two  $N-H \cdots N$  bonds. In our calculations the multipole electrostatic interaction energy is about four times smaller in the latter compound. Thus, it is not surprising that azetidine has many more structures of nearly the same energy, as there are fewer

hydrogen bonds to be incorporated into a space-filling network.

The same difference may be responsible for the different accuracies in the energy calculations, as witnessed by the r.m.s. values between pairs of reported energy levels (Tables 3 and 4). These values must be viewed in relation to the energy differences between possible structures. In the DM force field used in this work, cluster 25 is at  $4.5 \text{ kJ mol}^{-1}$  for hydantoin and  $2.0 \text{ kJ mol}^{-1}$  for azetidine. Of course, for any individual structure the discrepancy between two force fields can easily be twice as large as the r.m.s. value. So, if two force fields have an r.m.s. difference over 2 or  $1 \text{ kJ mol}^{-1}$ , respectively, at least one of them is inadequate to distinguish between the first 25 hypothetical structures.

It was seen that groups of more than two participants with mutual r.m.s. values below these limits are  $(Y, P, V)$  and  $(W, K, F)$ . Interestingly, for azetidine these groups can be combined into one, but for hydantoin they definitely cannot be. These six participants were the only ones who could predict the hydantoin structure by energy alone with rankings 4 or better (Day *et al.*, 2005). Participants  $Y, K, V$  all had a ranking of 1; moreover, they were the only ones who had the experimental azetidine structure somewhere in their lists of predicted structures (albeit never within the official first three submissions). Since this structure has two independent molecules in the unit cell, this not a trivial accomplishment.

Regarding the reliability of force fields, Day *et al.* (2005) have recently published a much-needed survey of crystal structure predictions for 50 molecules in the blind test category (i). They confirmed the notion, advocated by Price and coworkers in many publications, that the predictions are significantly improved by replacing point charges by distributed multipoles. Again, this causes a large increase of computer time. The approach of Karamertzanis & Pantelides (2004), who place additional point charges on non-atomic positions, is comparable in principle and may turn out to be a useful alternative. Force fields with these properties are  $Y, P, V, K$ , which indeed performed well in the present study.

Day *et al.* (2005) observed that crystal structures are more difficult to predict when they contain hydrogen bonds. They suggested that such crystals are more easily trapped in metastable polymorphic forms. The two compounds studied here do not conform to this general pattern. Moreover, in our study (van Eijck *et al.*, 2001) of six hexapyranoses, which are hydrogen-bonding substances *par excellence*, we found no evidence that the thermodynamic approach was inadequate. Of course, energy alone cannot be the only criterion for successful crystal structure prediction, as thermal effects and crystal growth conditions must be important. However, any study of such effects will need a fairly complete set of hypothetical structures with reliable energies as a necessary condition for more refined work. It appears that this desirable state of affairs has been only partially reached at the moment.

I am grateful to Panos Karamertzanis, Graeme Day and Sally Price for very helpful comments.

## References

- Chisholm, J. A. & Motherwell, S. (2005). *J. Appl. Cryst.* **38**, 228–231.
- Coombes, D. S., Price, S. L., Willock, D. J. & Leslie, M. (1996). *J. Phys. Chem.* **100**, 7352–7360.
- Day, G. M., Motherwell, W. D. S., Ammon, H. L., Boerrigter, S. X. M., Della Valle, R. G., Venuti, E., Dzyabchenko, A., Dunitz, J. D., Schweizer, B., van Eijck, B. P., Erk, P., Facelli, J. C., Bazterra, V. E., Ferraro, M. B., Hofmann, D. W. M. *et al.* (2005). *Acta Cryst.* **B61**, 511–527.
- Day, G. M., Motherwell, W. D. S. & Jones, W. (2005). *Cryst. Growth Des.* **5**, 1023–1033.
- Dzyabchenko, A. V. (1994). *Acta Cryst.* **B50**, 414–425.
- Eijck, B. P. van (2002). *J. Comput. Chem.* **23**, 456–462.
- Eijck, B. P. van & Kroon, J. (2000). *Acta Cryst.* **B56**, 535–542.
- Eijck, B. P. van, Mooij, W. T. M. & Kroon, J. (2001). *J. Phys. Chem. B*, **105**, 10573–10578.
- Guest, M. F., Bush, I. J., van Dam, H. J. J., Sherwood, P., Thomas, J. M. H., van Lenthe, J. H., Havenith, R. W. A. & Kendrick, J. (2005). *Mol. Phys.* **103**, 719–747.
- Hulme, A. T., Price, S. L. & Tocher, D. A. (2005). *J. Am. Chem. Soc.* **127**, 1116–1117.
- Karamertzanis, P. G. & Pantelides, C. C. (2004). *Mol. Sim.* **30**, 413–436.
- Karamertzanis, P. G. & Pantelides, C. C. (2005). *J. Comput. Chem.* **26**, 304–324.
- Lommerse, J. P. M., Motherwell, W. D. S., Ammon, H. L., Dunitz, J. D., Gavezzotti, A., Hofmann, D. W. M., Leusen, F. J. J., Mooij, W. T. M., Price, S. L., Schweizer, B., Schmidt, M. U., van Eijck, B. P., Verwer, P. & Williams, D. E. (2000). *Acta Cryst.* **B56**, 697–714.
- Mooij, W. T. M., van Duijneveldt, F. B., van Duijneveldt–van de Rijdt, J. G. C. M. & van Eijck, B. P. (1999). *J. Phys. Chem. A*, **103**, 9872–9882.
- Mooij, W. T. M., van Eijck, B. P., Price, S. L., Verwer, P. & Kroon, J. (1998). *J. Comput. Chem.* **19**, 459–474.
- Motherwell, W. D. S., Ammon, H. L., Dunitz, J. D., Dzyabchenko, A., Erk, P., Gavezzotti, A., Hofmann, D. W. M., Leusen, F. J. J., Lommerse, J. P. M., Mooij, W. T. M., Price, S. L., Scheraga, H., Schweizer, B., Schmidt, M. U., van Eijck, B. P. *et al.* (2002). *Acta Cryst.* **B58**, 647–661.
- Schaftenaar, G. & Noordik, J. H. (2000). *J. Comput.-Aided Mol. Des.* **14**, 123–134.
- Willighagen, E. L., Wehrens, R., Verwer, P., de Gelder, R. & Buydens, L. M. C. (2005). *Acta Cryst.* **B61**, 29–36.