

THE MINIMAL SET OF INTERMOLECULAR INTERACTIONS IN THE STRUCTURES OF SUBSTITUTED PROLINES

ALEXANDRU BANARU^{a,*}, ALEXANDER KOCHNEV^a

ABSTRACT. In this paper the calculation of intermolecular interaction energy based on Gavezzotti–Filippini semi-empirical method was performed for the series of substituted prolines. The initial structural data were mined from Cambridge Structural Database. The topology of each structure was analyzed with respect to a minimal generator set of the space group. It was shown that the least number of strong symmetrically independent interactions required to generate a crystal structure is likely to exceed the cardinality of a minimal generator set.

Keywords: *intermolecular interaction, lattice energy, space group generator, minimal generator set*

INTRODUCTION

Up to the moment the most powerful retrieval tool for structural data of organic compounds is Cambridge Structural Database, CSD [1]. Statistics based on CSD has following applications: crystal structure prediction, prediction of polymorphs, crystal engineering, a quantitative estimation of intermolecular interactions, justification of phase transitions, an insight into evaluable material properties etc. The database was established in 1965 and initially contained published results on structure determination by X-ray and neutron diffraction for compounds having at list one “organic” carbon atom. Thereafter this ambiguous criterion was moderated, and nowadays CSD contains, for instance, carbonyl complexes $M(\text{CO})_n$ of transition metals. Since the late 1990th, an unpublished data is also permitted if it satisfies certain requirements to the reliability. According to some estimates [2], the capacity

^a *Moscow State University, Chemistry Faculty, 1/3 Lenin Hills, 119991, Moscow, Russia*

* *Corresponding author: banaru@phys.chem.msu.ru*

of CSD may reach a million of crystal structures in the nearest future. Each structure is supplemented by a refcode of 6 letters, polymorphs having the same refcodes with slightly different numeric endings. A compound may have several refcodes attributed to, since the structure may be studied by different workgroups and under different conditions.

In the beginning of the XXth century Kitaigorodsky derived crystallographic symmetries favorable for a molecular packing [3]. Preferential symmetries lead to non-uniform distribution of crystal over the space groups [4, 5]. In particular, about a half of organic racemates crystallize in the space group $P2_1/c$ [5]. However, a molecular packing is hardly governed solely by the stereometric preferences that are more substantial for comparatively poor chemical interactions among molecules. More generally, interacting molecules do not recognize space groups but adapt to each other to minimize lattice energy. A space groups just restricts the number of symmetrically independent interactions formed by a given molecule in the crystal structure. In this work we gain an insight into correlation between the energy of intermolecular interactions and the space group.

SPACE GROUP GENERATORS

Space group generators are symmetry operations, whose various products (compositions) generate the whole space group [6]. A minimal generator set is that of the minimal cardinality N_G for a given space group, and $2 \leq N_G \leq 6$ [7]. Such sets are not presented in International crystallographic tables [6] but their cardinalities were calculated [8]. If a crystal consists of symmetrically equal molecules occupying a so-called general position (with the site symmetry of the point group C_1), then N_G equals the minimal number of symmetrically independent intermolecular interactions required for the formation of the corresponding crystal structure [7, 8]. Since such interactions are necessary, one could expect that they are the strongest by the energy of intermolecular interaction. In order to test this hypothesis, we are performing analysis of intermolecular interactions in the series of similar compounds. Do interaction energies satisfy the “broken stick model” [9]? According to this model, widely applicable in social sciences, the distribution of some discrete variable over sequence numbers should exhibit an extremum at the number setting bound to the most principal part of that sequence. The greatest intermolecular interaction energies U_1, U_2, \dots, U_n might exhibit a “brake of the stick” at $n = N_G$.

METHOD

Crystal structures were selected among proline derivatives because this amino acidic frame entails significant variety of intermolecular interactions. The search was performed in CSD ver.5.36 [1]. The following restrictions were applied to the molecular structure: 1) each is a proline ester, 2) each has less than 20 carbon atoms, 3) the chemical unit is sole (no hydrates, solvates etc.). Additionally, structures with no atomic coordinates were rejected. Totally 10 structures were found (Figure 1) and 2 pairs of polymorphs among them (3, 4 and 7, 8).

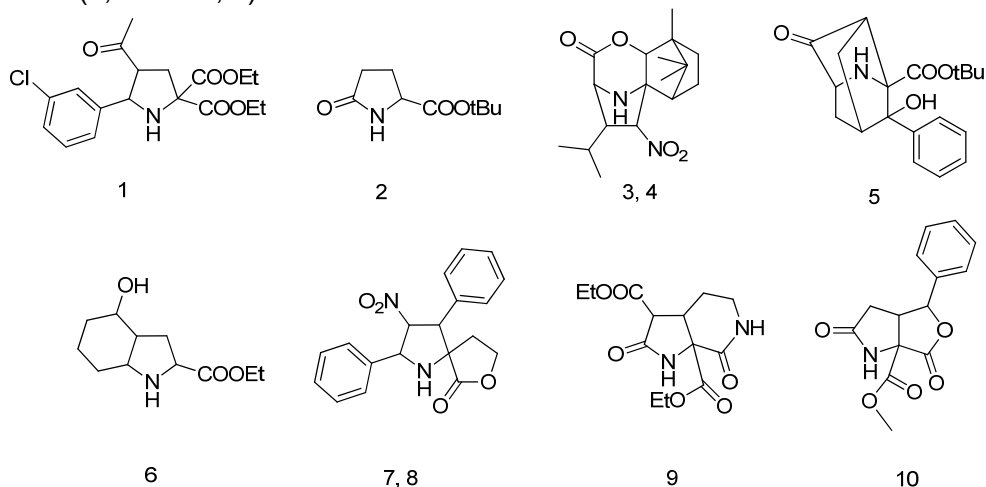


Figure 1. The structural formulas of 1-9

The computation of interaction energy was performed in Mercury [10] with the use of Gavezzotti–Filippini potential 6–exp [11]. In this approach the potential of intermolecular interaction U is the sum of all interatomic potentials between two molecules. Intramolecular interactions were being neglected. For the chart the 10 strongest contacts with the energies U_1, U_2, \dots, U_{10} were considered, because less strong contacts commonly share less than 1% of the lattice energy U_Σ . Denote $U_1 + U_2 + \dots + U_{10} = U_{\text{tot}} \approx U_\Sigma$. The “broken stick” implies that at some n the value of $\Delta U_n/U_{\text{tot}}$, where $\Delta U_n = U_{n+1} - U_n$, should change dramatically.

Each intermolecular interaction corresponds to some symmetry operation. Combining these operations, one can obtain a minimal generator set of the space group. In fact, this is a minimal set of interactions to form the crystal structure. Although for any space group there can be various minimal

sets, we select the only one in the following manner. Assume that such minimal set of interactions has an advantage if it contains interactions with the highest energy. Thus, we select a minimal set of interactions such that its 1st strong interaction is the strongest among all sets. If this point does not lead to a definite set, we select a set such that its 2nd strong interaction is the strongest among those in other sets, etc.

RESULTS AND DISCUSSION

In 6 structures there is an intermolecular H-bond forming either a molecular chain along a screw axis 2_1 (**2**, **5**, **8-9**) and along a translation (**6**), or a center-symmetric dimer (**10**). In each case, assuming that the strongest interaction is necessary for the observed molecular arrangement, the symmetry operator of the H-bond has to be also included in the minimal generator set of the space group (Table 1). If this operator was missing, the generator set would not be minimal. It is surprising that among polymorphs **7-8** H-bond is present just in the latter structure despite they have much more in common.

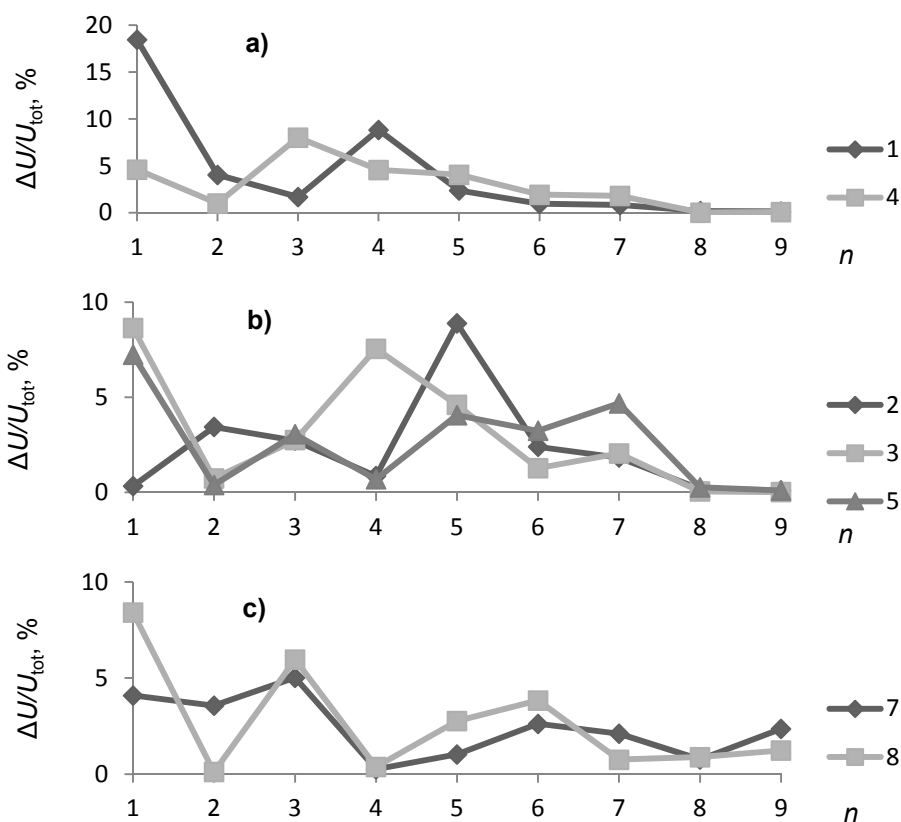
Table 1. Space group, Z , N_G and essential interactions* for the structures **1-10**

No	CSD refcode	Space group, Z	N_G	Essential intermolecular interactions along**
1	AZEFER	$P2_1$, $Z = 2$	3	I. Y II. 2_1 (1 y $\frac{1}{2}$) III. 2_1 (1 y 0) IV. 2_1 ($\frac{1}{2}$ y $\frac{1}{2}$)
2	DICXET	$P2_12_12_1$, $Z = 4$	2	I. X II. 2_1 ($\frac{1}{2}$ y $\frac{3}{4}$) III. 2_1 (x $\frac{1}{4}$ 1)
3	FONYUD	$P2_12_12_1$, $Z = 4$	2	I. X II. 2_1 (0 y $\frac{1}{4}$) III. 2_1 (x $\frac{1}{4}$ 0)
4	FOXCUR	$P2_1$, $Z = 2$	3	I. Y II. 2_1 (0 y 0) III. 2_1 ($\frac{1}{2}$ y $\frac{1}{2}$) IV. 2_1 ($\frac{1}{2}$ y 0)
5	GUDBEM	$P2_12_12_1$, $Z = 4$	2	I. 2_1 (x $\frac{1}{4}$ 1) II. 2_1 (1 y $1\frac{1}{4}$)
6	HEQRIG	$P-1$, $Z = 2$	4	I. X+Y II. Y III. i ($\frac{1}{2}$ $\frac{1}{2}$ 0) IV. X V. i (1 1 0) VI. i (1 $\frac{1}{2}$ 0) VII. i ($\frac{1}{2}$ 1 0) VIII. i ($\frac{1}{2}$ $\frac{1}{2}$ $\frac{1}{2}$)
7	LOGHEV	$P2_1/c$, $Z = 4$	3	I. 2_1 ($\frac{1}{2}$ y $\frac{1}{4}$) II. i ($\frac{1}{2}$ $\frac{1}{2}$ 0) III. X

8	LOJNOO	$P2_1/c$, $Z = 4$	3	<u>I. 2_1 ($\frac{1}{2} y \frac{1}{4}$)</u> II. i ($\frac{1}{2} \frac{1}{2} \frac{1}{2}$) III. X
9	MALREX	$Pbca$, $Z = 8$	3	<u>I. 2_1 ($x \frac{3}{4} 1$)</u> II. b ($\frac{3}{4} y z$) III. a ($x y \frac{3}{4}$)
10	POSMIV	$Pbca$, $Z = 8$	3	<u>I. i ($\frac{1}{2} \frac{1}{2} \frac{1}{2}$)</u> II. X III. i ($0 \frac{1}{2} \frac{1}{2}$) <u>IV. a ($x y \frac{1}{4}$)</u> V. b ($-\frac{1}{4} y z$)

*Assuming that stronger interactions are more advantageous as generators (see explanation in **METHOD**), the interactions are listed as energy decrease (I, II, etc.) till the space group is generated by the symmetry operations (generators) accounting for those interactions. The minimal set is shown in bold.

**X, Y, X+Y denote interactions of a molecule with those related by a corresponding translation; 2_1 – related by a screw axis; i – related by a center of inversion; a , b – related by a glide plane. Intermolecular H-bonds are shown underlined.



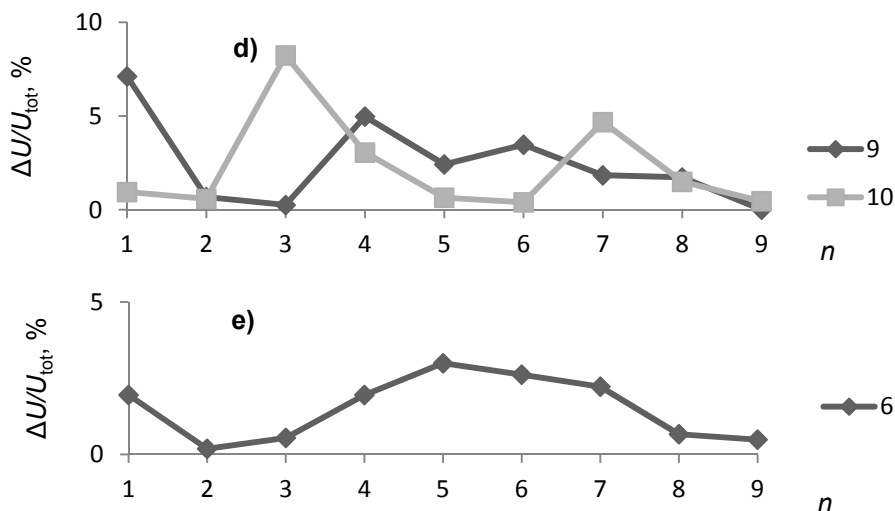


Figure 2. $\Delta U/U_{tot}$ (see explanation in **METHOD**) for the structures of space groups $P2_1$ (a), $P2_12_12_1$ (b), $P2_1/c$ (c), $Pbca$ (d), and $P-1$ (e)

In 4 structures (**1-4**) the 1st strong interaction cannot be included in the minimal generator set because this interaction corresponds to a translation, but not to a screw axis 2_1 in the same direction. While the screw axis can generate the translation: $2_{1(Y)}^2 = Y$, the screw axis cannot be generated by the translation itself. The mentioned result disproves the hypothesis that there should be a “brake of the cane” at the $(N_G + 1)$ th strong contact on the energy chart. On the other hand, in 4 structures (**5, 7-9**) the minimal set includes the 1st, 2nd, etc. strong contacts in a row. In other cases essential interactions are formed in addition to the minimal set, either one interaction (**1-4**), or more than one (**6, 10**). This means that actually the value of the minimal coordination number is commonly more than the number of generators of the space group.

There are intermolecular energy charts shown in Figure 2. Sometimes the maximal value of $\Delta U_n/U_{tot}$ without consideration of the 1st one indicates $n = N_G$ (**1, 7, 8**), but more commonly doesn't. The chart usually has a wavy shape with the floating “wavelength”. As a rule, falling from $n = 1$ to 2, the “wave” rises again at $n = 3$ or 4, and the next rise is at $n = 6$ or 7. As the average energy drops, the “wave” fades out. Since the maximal $\Delta U_n/U_{tot}$ in the majority of structures is located farther than $n = N_G$, distant interactions are to play a very important role in the formation of the crystal structure.

The structure **6** has the crucial intermolecular interaction with the extremely distant location ($n = 8$). It is the interaction over the center of inversion i ($\frac{1}{2} \frac{1}{2} \frac{1}{2}$). There is no possibility to exclude the 8th interaction from the minimal generator set. If it were not for this interaction, **6** would be a layered 2D crystal (XY). Actually, $P-1$ is itself a rather complicated group ($N_G = 4$), and a complicated system of interactions formed by its molecules is not so surprising. Nevertheless, this space group is commonly thought of as one of the simplest, because it belongs to the least symmetric crystal system (triclinic). These opposite points of view should be reconciled.

CONCLUSION

As it follows from the above section, the minimal number of intermolecular interactions is rarely equal to the N_G . It would be more definitive to count the least number of strong contacts sufficient to form a given structure. This entails enduring interest in the manual analysis of intermolecular interactions in a crystal structure.

ACKNOWLEDGEMENTS

The authors are thankful to L. Vasilieva for her help in preparing of this manuscript.

REFERENCES

1. F.H. Allen, *Acta Crystallographica*, **2002**, B58, 380.
2. C.R. Groom, F.H. Allen, *Angewante Chemie International Edition*, **2014**, 53, 662.
3. A.I. Kitaigorodsky, "Molecular Crystals and Molecules", Academic Press, New York, **1973**.
4. L.N. Kuleshova, M. Yu. Antipin, *Russian Chemical Reviews*, **1999**, 68, 1.
5. P.M. Zorkii, P.N. Oleinikov, *Journal of Structural Chemistry*, **2001**, 42, 24.
6. Th. Hahn (ed.), "International Tables for Crystallography, Volume A: Space-group symmetry, 1st online ed.", International Union of Crystallography, **2006**, part 2.2.10.
7. A.M. Banaru, *Moscow University Chemistry Bulletin*, **2009**, 64, 80.
8. E.A. Lord, A.M. Banaru, *Moscow University Chemistry Bulletin*, **2012**, 67, 50.

9. R.H. MacArthur, *Proceedings of National Academy of Sciences USA*, **1957**, *43*, 293.
10. C.F. Macrae, I.J. Bruno, J.A. Chisholm, P.R. Edgington, P. McCabe, E. Pidcock, L. Rodriguez-Monge, R. Taylor, J. van de Streek, P.A. Wood, *Journal of Applied Crystallography*, **2008**, *41*, 466.
11. A. Gavezzotti, G. Filippini, *Journal of Physical Chemistry*, **1994**, *98*, 4831.