Research Article

Space Group Approximation of a Molecular Crystal by Classifying Molecules for Their Electric Potentials and Roughness on Their Inertial Ellipsoid Surface

Jose Fayos

Rocasolano Institute, CSIC, Rodriguez Ayuso 6, 28022 Madrid, Spain

Correspondence should be addressed to Jose Fayos; jose.fayos@yahoo.es

Received 29 June 2014; Accepted 26 September 2014; Published 16 October 2014

Academic Editor: Sailaja Krishnamurty

Copyright © 2014 Jose Fayos. This is an open access article distributed under the Creative Commons Attribution License, which permits unrestricted use, distribution, and reproduction in any medium, provided the original work is properly cited.

In order to predict the most probable space group where a molecule crystallizes, it is assumed that molecular shape and electric potential distribution on the molecular surface are the main factors or predictors. However, to compare and classify molecules by these two factors seems to be very difficult for in general such different objects. Thus, in order to compare molecules, they are reduced to their inertial ellipsoid in which surface 26 equally spaced points were chosen where a roughness factor and an electric potential due to all atomic charges of the whole molecule are calculated. By this procedure, different molecules encoded by these two predictor vectors can be compared and classified, showing that molecules that crystallize in the same space group have more similar predictor vectors. This result opens the possibility to predict the more probable spatial group associated with a molecule.

1. Introduction

The first hypothesis considered for crystal packing prediction CPP of organic molecules is that the isolated molecule contains the information of its future crystal [1]. So it is for the so-called “blind tests,” the last published in 2011 [2], where several laboratories compete to find the crystal structure of a molecule by diverse calculations, where it is assumed that 95% of the molecules present no polymorph and prefer a cell with a given space group SG. Some approaches were previously done to get molecular crystal structure information by data mining [3, 4]. In the present study it is further assumed that the molecular crystal space group is mainly predetermined by the molecular form or roughness and by the electric potential distribution on the molecular surface, both factors being the best predictors for crystal packing, including the formation of hydrogen bonds. In fact electrostatic forces determine molecular reactions as can be observed by X-ray in the electron density distributions of crystalline molecules [5, 6], were interactions by electrostatic forces (including H-bonds) between equal molecules in a crystal, would determine its crystal packing. The purpose of this work is to classify the molecules into groups by similarity

of the above predictors and to check if these assumed types of packing are correlated with their space group SG, or conversely, to verify that molecules crystallizing in the same SG have similar aggregation predictors. However, comparing these two predictors between usually dissimilar molecules does not seem so simple, unless the molecules could first be reduced to more comparable objects. In a previous work [7], some molecular crystal descriptors like cell axes and the presence of some symmetry elements, but not the SG, were predicted by reducing each molecule to its inertial ellipsoid (adding to each axis the hydrogen VDW radius of 1.17 Å). The same molecular reduction to its inertial ellipsoid is taken here, defined by its three axes: Large, Medium, and Small (L, M, S), where 26 points are added to its surface: 2×3 on the ends of the ellipsoid axes, 3×4 in the edges centers, and 8 in the face centers: all points approximately equidistant from each other. Figure 1(a) shows the numbered 26 points with their sequence of coordinates on the ellipsoid surface, where for clarity the ellipsoid has been deformed to a cube in this figure.

The classical charges q_j for the atoms in every molecule were calculated by Chem3D Pro [8]; then the electrostatic potential V_i on those 26 points of every ellipsoid surface, due

FIGURE 1: (a) Inertial ellipsoid IE reduced to a cube, only for clarity in this figure, with the 26 points on its surface. The 26 (L, M, S) coordinates: (1-1-1 1-1-10 1-1-11 10-1 100 101 11-1 110 111 0-1-1 0-10 0-11 00-1 001 01-1 010 011 -1-1-1 -1-10 -1-11 -10-1 -100 -101 -11-1 -110 -111). (b) HEPJER molecule into its inertial ellipsoid with the 8 points for S = 0. (c) Distribution of 26 potentials V and 26 concavities C of HEPJER on his inertial ellipsoid, from (1-1-1) to (-111).

to the charges q_j of all j atoms in the molecule, was calculated by $V_i = \sum_j (q_j / rij)$, where rij are the distances from point i to the j atoms.

Besides, a roughness factor in those 26 points of every molecular ellipsoid surface, from now concavity factor, was calculated as the average of the distances from each point to their four closest atoms in the molecule. In total two vectors per molecule: 26V for potentials and 26C for concavities, each of 26 components, are both the space group predictors in this work. In fact, the main differences with the previous work [7] are the addition of this concavity vector, not scaling the molecular potential vectors and a more simple space group classification procedure.

Table 1 shows molecular global form parameters such as the average molecular concavity $\langle \text{conc} \rangle$ and its standard deviation $\langle \text{desv} \rangle$, the relationship between the principal inertial axes (M/L, S/L, S/M), and other parameters of internal

symmetry of the molecule, some of them used in the previous work [7].

Table 2 shows the averages of those parameters per space group SG. Although all these form parameters are somehow conditioning the SG of the crystal, it is not easy to find in these tables any similarities between those global form parameters for molecules of the same SG, which justifies extending that global form information among 26P over the inertial ellipsoid.

2. Molecular Space Group Predictors

Due to the relation between the present and previous work [7], the same 31 molecules are also chosen here. Table 3 shows those molecules all having azol group with three different substitutes: a group of 17 molecules 17P21/c with space group

TABLE 1: Global form descriptors for each molecule.

	$\langle \text{conc} \rangle$	desv	M/L	S/L	S/M	PL	mS	mM	mL
FAQROE P21/c	2.076	0.32	0.574	0.371	0.647	2	2	0	1
FAQSAR P21/n	2.152	0.29	0.542	0.363	0.669	2	2	0	1
FAQSEV P21/c	2.041	0.19	0.688	0.372	0.541	2	2	0	1
FAQSOF P21/n	2.334	0.42	0.987	0.516	0.523	2	1	0	0
FAQSUL P21/c	2.175	0.32	0.664	0.380	0.572	2	2	0	1
FAQTAS P21/n	2.111	0.35	0.589	0.370	0.627	1	1	0	0
HDMPYZ P21/c	2.316	0.42	0.795	0.619	0.778	0	0	0	0
HEPHUF P21/n	2.372	0.28	0.592	0.579	0.978	0	0	0	0
HEPJER P21/n	1.774	0.29	0.636	0.245	0.385	2	2	1	1
KOSFUT P21/n	2.164	0.41	0.761	0.697	0.915	0	0	0	0
LEVVAJ P21/n	1.996	0.24	0.544	0.284	0.521	2	2	0	1
RIVBAZ P21/c	2.083	0.53	0.649	0.581	0.895	0	2	0	1
RUPSEA P21/n	1.866	0.21	0.511	0.230	0.449	2	2	1	0
TEHQAY P21/c	2.181	0.46	0.883	0.353	0.399	2	2	1	0
VAXLAH P21/a	2.370	0.43	0.763	0.408	0.535	1	1	0	0
WILBAU P21/c	2.031	0.51	0.644	0.576	0.894	0	1	0	1
YAXZOM P21/n	2.073	0.23	0.439	0.282	0.643	2	2	0	1
HEPJAN Pbca	2.341	0.47	0.916	0.595	0.649	0	0	0	0
MBCPAZ Pbca	2.243	0.22	0.729	0.416	0.570	2	2	0	1
POYXUW Pbca	2.044	0.38	0.456	0.210	0.460	2	2	0	1
BIWWEJ P212121	2.453	0.47	0.361	0.311	0.861	1	1	0	1
PAZDPY P212121	1.913	0.36	0.755	0.271	0.359	2	2	1	0
RUPRID P212121	1.905	0.25	0.584	0.255	0.437	2	2	1	0
BEWLEU Pna21	2.461	0.35	0.651	0.367	0.564	2	1	0	0
HIWJIG01 Pna21	1.944	0.25	0.870	0.401	0.461	1	1	0	0
HIWJIG Pca21	2.141	0.22	0.909	0.521	0.573	1	1	0	0
TAXLOT P-1	2.721	0.58	0.683	0.496	0.727	2	2	0	1
VEHCOA P-1	2.277	0.40	0.776	0.764	0.984	1	2	0	0
BENSES P2/c	2.851	0.50	0.789	0.609	0.772	2	2	1	0
GISZIR C2/c	2.245	0.32	0.581	0.326	0.561	2	2	0	1
PYRZAL10 P21	2.176	0.22	0.606	0.529	0.873	0	0	0	0

$\langle \text{conc} \rangle$ = average (distance(P-4nearest.atoms)) in the 26P, desv = stand.deviation of $\langle \text{conc} \rangle$.

M/L, S/L, S/M for the inertial ellipsoid.

PL = 2, planar molec H's no considered; PL = 1, pseudo-PL; PL = 0, no PL.

mS, mM, mL = 2: molecular symmetry plane m perpendicular to S, M, L.

mS, mM, mL = 1: pseudo m perpendicular to S, M, L.

mS, mM, mL = 0: no m.

TABLE 2: Mean form factors $\langle \text{M/L} \rangle$, $\langle \text{S/L} \rangle$ of the inertial ellipsoid; mean planarity of the molecules $\langle \text{PL} \rangle$ (between max = 2, 1 and min = 0); and mean molecular planes of symmetry $\langle \text{mS} \rangle$, $\langle \text{mM} \rangle$, or $\langle \text{mL} \rangle$ (between max = 2, 1 and min = 0), with their standard deviations, for each SG.

	$\langle \text{M/L} \rangle$	des	$\langle \text{S/L} \rangle$	des	$\langle \text{PL} \rangle$	des	$\langle \text{mS} \rangle$	des	$\langle \text{mM} \rangle$	des	$\langle \text{mL} \rangle$	des
14NoP21/c	0.69	0.16	0.43	0.16	1.43	0.76	1.43	0.76	0.21	0.43	0.36	0.50
3laz	0.68	0.15	0.43	0.15	1.35	0.84	1.42	0.76	0.19	0.40	0.45	0.51
17P21/c	0.66	0.14	0.43	0.14	1.29	0.92	1.41	0.80	0.18	0.39	0.53	0.51
6FAQP21c	0.67	0.16	0.40	0.06	1.83	0.41	1.67	0.52	0	0	0.67	0.52
3Pbca	0.70	0.23	0.41	0.19	1.33	1.15	1.33	1.15	0	0	0.67	0.58
3P212121	0.57	0.20	0.28	0.03	1.67	0.58	1.67	0.58	0.67	0.58	0.33	0.58
3Pna21	0.81	0.14	0.43	0.08	1.33	0.58	1	0	0	0	0	0
2P-1	0.73	0.07	0.63	0.19	1.5	0.71	2	0	0	0	0.50	0.71

TABLE 3: The 31 molecules selected for the present work, with the CSD codes.

Compound	R3	R4	R5
(1) BENSES	(a)	CN	H
(2) BEWLEU	(b)	COOMe	COOMe
(3) BIWWEJ	Ph	H	diMe-NH-CH ₂ -Ph
(4) FAQROE	COOEt	H	H
(5) FAQSAR	COOEt	H	Me
(6) FAQSEV	COOEt	Me	H
(7) FAQSOF	COOEt	Ph	H
(8) FAQSUL	COOEt	Br	H
(9) FAQTAS	COOEt	Br	Me
(10) GISZIR	Me	CN	NH-Ph-4-CF ₃
(11) HDMPYZ	H	(c)	Me
(12) HEPHUF	H	CH ₂ -4-pz	H
(13) HEPJAN	Me	CH ₂ -3,5-diMe-4-pz	Me
(14) HEPJER	3-pz	H	H
(15) HIWJIG	H	NH ₂	H
(16) HIWJIG01	H	NH ₂	H
(17) KOSFUT	H	O-Si(2tBuOH)	C(tBuO)
(18) LEVVAJ	(d)	H	H
(19) MBCPAZ	Me	Br	C(O)(NH ₂)
(20) PAZDPY	N=N ⁺ =N ⁻	Ph	H
(21) POYXUW	NH-COPh	H	Br
(22) PYRZAL10	CH ₂ -C(COO ⁻)(NH ₃ ⁺)	H	H
(23) RIVBAZ	tBu	N=O	tBu
(24) RUPRID	NH ₂	H	Ph
(25) RUPSEA	Ph-4-NO ₂	H	NH ₂
(26) TAXLOT	NH-CH-C(e)(COOEt)	H	H
(27) TEHQAY	H	H	3,5-dimethoxyph
(28) VAXLAH	(f)	COOMe	COOMe
(29) VEHCOA	H	NO ₂	SiMe ₃
(30) WILBAU	tBu	NO ₂	tBu
(31) YAXZOM	(g)	Me	(g)

(a)

(b)

(c)

(d)

(e)

(f)

(g)

TABLE 4: Total averages $\langle 26\langle NV \rangle \rangle$ and $\langle 26\langle NC \rangle \rangle$ for potentials V and concavities C on the inertial ellipsoid IE, and the average of the 26 standard deviations of $26\langle NV \rangle$ and $26\langle NC \rangle$ for each space group SG. $\langle 26\langle NV \rangle \rangle$ is the average of the 26 V averaged (vertical sense in Supplementary Table) between the N molecules of the SG. $\langle 26\langle des\langle NV \rangle \rangle \rangle$ is the average of the 26 standard deviations for the $26\langle NV \rangle$. $\langle 26\langle NC \rangle \rangle$ is the average of the 26 C averaged (vertical sense in Supplementary Table) between the N molecules of the SG. $\langle 26\langle des\langle NC \rangle \rangle \rangle$ is the average of the 26 standard deviations for the $26\langle NC \rangle$. $des\langle 26des \rangle$ is the standard deviation of the last average $\langle 26\langle des\langle NC \rangle \rangle \rangle$.

N molecules	$\langle 26\langle NV \rangle \rangle$	$\langle 26\langle des\langle NV \rangle \rangle \rangle$	$\langle 26\langle NC \rangle \rangle$	$\langle 26\langle des\langle NC \rangle \rangle \rangle$	$des\langle 26des \rangle$
14NoP21/c	0.09	0.10	2.26	0.47	0.12
31azols	0.005	0.09	2.18	0.42	0.11
17P21/c	0.01	0.07	2.11	0.34	0.10
6FAQP21/c	-0.05	0.05	2.15	0.25	0.08
2P-1	-0.003	0.05	2.50	0.52	0.13
3P212121	0.071	0.13	2.02	0.38	0.11
3Pna21	-0.003	0.05	2.18	0.35	0.07
3Pbca	-0.01	0.06	2.21	0.34	0.18

$\langle 26\langle des\langle NV \rangle \rangle \rangle$ shows the V similarity for the N molecules in the same point of IE.

$\langle 26\langle des\langle NC \rangle \rangle \rangle$ shows the C similarity for the N molecules in the same point of IE.

$des\langle 26des \rangle$ shows the similarity between the $26\langle des\langle NC \rangle \rangle$.

SG P21/c (also P21/n or P21/a), six of them 6FAQP21/c sharing also a COOEt group in the same substituent, eleven molecules distributed among several SG: 3 in Pbca, 3 in P212121, 3 in Pna21, 2 in P-1, and other three molecules crystallizing in different SG.

Finally, in order to compare properly both molecular vectors, $26V$ for potentials and $26C$ for concavities, between different molecules, each molecule was reoriented within its inertial ellipsoid IE, by using the symmetry planes $L = 0$, $M = 0$, or $S = 0$, in order to have the molecular largest potential V -weighted octant, among the eight octants, in the IE octant (+L+M+S), where V -weighted is calculated with the potentials of the seven points of the octant in the proportion: $[3V(111) + 2V(110 + 101 + 011) + V(100 + 010 + 001)]$. The Supplementary Table (available online at <http://dx.doi.org/10.1155/2014/737480>) shows the final comparable values of $26V$ and $26C$ for the reoriented 31 molecules. It is important to see in Tables 1 and 3 that the presence of two molecules in the same space group SG occurs not only for little changes in their structure like between FAQROE and FAQSAR but also for big changes like for FAQROE and KOSFUT, suggesting that the 26 potentials and 26 concavities sequences around the inertial ellipsoid will determine better the space group.

It is interesting to consider here that all the molecules could also be reoriented to have the lowest, instead of the largest, potential V -weighted octant in (+L+M+S). Although the relative orientation between the largest and lowest V -weighted octants is not the same for all molecules, however the average distribution on the inertial ellipsoid IE surface of the $26V$ and $-26V$, respectively, is similar for both orientations of the 31 molecules, which reinforces this analysis. Finally the largest V -weighted octant was taken for molecular reorientation. Figure 2(a) shows first the distribution along the IE of the $26V$ averaged for the 31 azole molecules $26\langle 31V \rangle$, together with the 26 standard deviations of these averages $26std\langle 31V \rangle$, showing two saw-tooth peaks, around (110)-(111) and (010)-(011), as expected (see Figure 1(a)) for molecules

oriented with the largest potential V -weighted octant in (+L+M+S).

The centrosymmetric molecule HEPJER almost planar in (L, M, 0) is shown in Figure 1(b) into a schematic inertial ellipsoid IE, because it is more suitable for its simplicity to understand the sequence of concavities and potentials along the 26 points on its IE shown in Figure 1(c). The concavities C are maximal in the 8 molecular LM0 contour points, the remaining concavities on either side of plane $S = 0$ being lower. The symmetrical potentials V have lower values on 0MS points away from the molecule and higher peaks especially on LM0 points.

3. Analysis of Averaged Potentials and Concavities by Space Group

Figure 2(a) shows the distribution from (1-1-1) to (-111) along the inertial ellipsoid of the 26 potentials averaged for the N molecules $26\langle NV \rangle$ belonging to different groups: the total of 31az, the 14NotP21/c, the 17P21/c, and the 6FAQP21/c. It also shows the distribution of the 26 standard deviations of those averages $des(26\langle NV \rangle)$ per group, which indicate the similarity between the potentials of the N molecules in each one of the 26 points of their inertial ellipsoid. In general $des(26\langle NV \rangle)$ are maximum in singular points with maximum or minimum values of $\langle NV \rangle$ and in particular are clearly superior for the first two groups of mixed space groups compared to the last two P21/c groups, as also shown in Table 4 with the total averages under the column $\langle des(26\langle NV \rangle) \rangle$. Figure 2(a) also shows that the distribution of the $26\langle NV \rangle$ for 31az and 14azNotP21/c is similar with two peaks in saw-tooth form, although being less pronounced the second peak for 31az. The distributions of the $26\langle NV \rangle$ for the equal space groups 17P21/c and 6FAQP21/c differ from the previous with the loss of the second saw-tooth peak and an increased negative potential V on (0-10).

FIGURE 2: (a) Distribution of $26\langle NV \rangle$ for N molecules: 26 components of averages $\langle NV \rangle$ with the standard deviations for groups 31az, 14NoP21/c, 17P21/c, and 6FAQP21/c. Comparison between the $26\langle NV \rangle$ for 17P21/c and 14NoP21/c. (b) Distribution of the $26\langle NC \rangle$ concavities and the corresponding standard deviations: 26 components of $\langle NC \rangle$ averages for the groups 31az, 14NoP21/c, 17P21/c, and 6FAQP21/c.

FIGURE 3: Continued.

FIGURE 3: Distribution of the 26V potentials and 26C concavities for molecules of the groups 3Pbca, 3P212121, 3Pna21, and 2P-1.

Therefore, although the overall molecular shape factors of Tables 1 and 2 and the contents of the molecules in Table 3 do not seem to indicate similarity between molecules within each space group SG, however the distributions of the $26\langle NV \rangle$ potentials along the 26 points of the inertial ellipsoid appear to have some similarity between the molecules of the same SG. In fact, the analysis of Figure 2(a) and the values of $\langle 26(\text{des}\langle NV \rangle) \rangle$ in Table 4 suggest that there exists an association of the potential 26V predictor vector of a molecule with its molecular crystal or space group SG.

Figure 2(b) shows the average distribution of the 26 concavities $26\langle NC \rangle$ and of the corresponding standard deviations $\text{des}(26\langle NC \rangle)$ for the four groups of N molecules: 3laz, 14NotP21/c, 17P21/c, and 6FAQP21/c. While for 14azNotP21/c, the distribution of $26\langle NC \rangle$ is pseudosymmetrical around the center (00 ± 1), that symmetry tends to break for 17azP21/c and more for 6FAQP21/c molecules, especially in the area where their $26\langle NV \rangle$ potentials lose the second saw-tooth peak. The distribution of $26\langle NC \rangle$ for 3laz involves molecules of equal and different space group being intermediate as might be expected. Table 4 also shows the quantitative differences between the average distributions of $26\langle NC \rangle$ concavities for different groups of molecules. The average values under $\langle 26(\text{des}\langle NC \rangle) \rangle$ showing more similar distributions for the 26 concavities between the molecules of the groups 6FAQP21/c and 17P21/c than between the molecules of the group 14NotP21/c, while an intermediate similarity for the total of 3laz. This shows the association of the predictor vector 26C of molecular concavities with the space group SG of the molecular crystal, parallel to the previous association observed of the predictor vector 26V of potentials with the space group SG.

Figure 3 shows the molecular distributions of 26V potentials and 26C concavities for the four minority space groups: Pbca, P212121, Pna21, and P-1. Although this analysis is less significant with only 3, 3, 3, and 2 molecules each, it also notes some similarity between the distributions of its 26V and of

its 26C for molecules with the same space group SG, except for P-1 and BIWWEJ (unique molecule in Table 2 with its 26V “calculated” positive). Furthermore, Table 4 shows that the average standard deviation values $\langle 26(\text{des}\langle NV \rangle) \rangle$ and $\langle 26(\text{des}\langle NC \rangle) \rangle$ for these four space groups do not deviate too much from those of groups 6FAQP21/c and 17P21/c, with the exceptions described above.

Finally, Figure 2(b) also shows the differences between the average potentials $\langle V \rangle$ and average concavities $\langle C \rangle$ distributions along the inertial ellipsoid surface for the 17P21/c and 14NonP21/c molecular groups. While for $L = 1$ there is quite similitude between the V 's and between the C 's of both groups, for $L = 0$ the second saw-tooth disappearance for the V 's is accompanied with some C 's distribution variations between both groups, and for $L = -1$ besides the notable V 's distribution differences there are drastic differences between the C 's distributions for both molecular groups.

4. Conclusion

Assuming that the molecular form and the potential distribution on its surface were the major predictors for crystal packing, it is not easy to compare these properties between different molecules, in order to classify them by their space group. To enable this comparison, molecules are reduced to their inertial ellipsoids IE with 26 singular points equally spaced on the surface, in which 26 potentials and 26 concavity factors are calculated. These two molecular vectors 26V and 26C are taken as molecular packing or space group predictors for the molecular crystal (assuming no polymorphism). Comparing both predictors between 31 molecules, there is more similarity between them for molecules crystallizing in the same space group SG than between molecules with different SG. This suggests that each space group would have its own mean distribution of their 26V potentials and 26C concavities on a virtual inertial ellipsoid, which would

enable predicting the probable space group of a molecular crystal by calculating the 26V and 26C distributions on its molecular inertial ellipsoid. Foreknowledge of the probable space group associated with a molecule would facilitate the total crystal prediction CPP to perform other crystal engineering calculation. For example, if the above predicted space group SG (associated with a crystalline form) were not convenient for pharmaceutical processes [1], a molecular modification simulation could be tried to change the SG avoiding that molecular aggregation.

Summary of Symbols

IE:	Inertial ellipsoid of the molecule
SG:	Crystal space group
L, M, S:	Large, medium, small axes of IE
P:	One of 26 points on the IE surface
V:	Electric potential at one P
C:	Concavity at one P
N:	Number of molecules in a group
des:	Standard deviation of an average
26V:	Vector with the 26V on IE surface
26C:	Vector with the 26C on IE surface
$\langle NV \rangle$:	NV average on a P of a molecular group
$\langle NC \rangle$:	NC average on a P of a molecular group
des $\langle NV \rangle$:	Standard deviation of average $\langle NV \rangle$
des $\langle NC \rangle$:	Standard deviation of average $\langle NC \rangle$
$\langle 26\langle NV \rangle \rangle$:	Average of the total 26 $\langle NV \rangle$ of a group
$\langle 26\langle NC \rangle \rangle$:	Average of the total 26 $\langle NC \rangle$ of a group
$\langle 26(\text{des}\langle NV \rangle) \rangle$:	Average of the 26(des $\langle NV \rangle$)
$\langle 26(\text{des}\langle NC \rangle) \rangle$:	Average of the 26(des $\langle NC \rangle$).

Conflict of Interests

The author declares that there is no conflict of interests regarding the publication of this paper.

References

- [1] S. L. Price, "Predicting crystal structures of organic compounds," *Chemical Society Reviews*, vol. 43, pp. 2098–2111, 2014.
- [2] D. A. Bardwell, C. S. Adjiman, Y. A. Arnautova et al., "Towards crystal structure prediction of complex organic compounds—a report on the fifth blind test," *Acta Crystallographica Section B: Structural Science*, vol. 67, part 6, pp. 535–551, 2011.
- [3] J. Fayos and F. H. Cano, "Crystal-packing prediction by neural networks," *Crystal Growth and Design*, vol. 2, no. 6, pp. 591–599, 2002.
- [4] J. Fayos, L. Infantes, and F. H. Cano, "Neural network prediction of secondary structure in crystals: hydrogen-bond systems in pyrazole derivatives," *Crystal Growth and Design*, vol. 5, no. 1, pp. 191–200, 2005.
- [5] H. Nakatsuji, S. Kanayama, S. Harada, and T. Yonezawa, "Electrostatic force theory for a molecule and interacting molecules. 7. Ab initio verification of the force concepts based on the floating wave functions of ammonia, methyl(1+) ion, and ammonia(1+) ion," *Journal of the American Chemical Society*, vol. 100, no. 24, pp. 7528–7534, 1978.
- [6] Y. Honda and H. Nakatsuji, "Force concept for predicting the geometries of molecules in an external electric field," *Chemical Physics Letters*, vol. 293, no. 3–4, pp. 230–238, 1998.
- [7] J. Fayos, "Molecular crystal prediction approach by molecular similarity: data mining on molecular aggregation predictors and crystal descriptors," *Crystal Growth and Design*, vol. 9, no. 7, pp. 3142–3153, 2009.
- [8] F. H. Allen, "The Cambridge structural database: a quarter of a million crystal structures and rising," *Acta Crystallographica B*, vol. 58, no. 1, part 3, pp. 380–388, 2002.