

---

# Organic Crystal Structure Predictions

and their application to X-ray PD structure solution

*Alexander Dzyabchenko*

Karpov Institute of Physical Chemistry  
10 Vorontsovo pole, 103064 Moscow, RUSSIA  
E-mail [adz@cc.nifhi.ac.ru](mailto:adz@cc.nifhi.ac.ru)



---

# Why do we need crystal structure predictions of small organic molecules?

- To understand solid state phenomena: phase transitions, solid-state reactions, crystal disorder, crystal growth, *etc.*
  - To help X-ray structure determinations with powder and electron diffraction methods
  - To assist solution of polymorphism problems in drug design and food industry
  - To elaborate and verify the force fields for biomolecules
-

# Structure seeker : the first crystal structure prediction tool (1946)

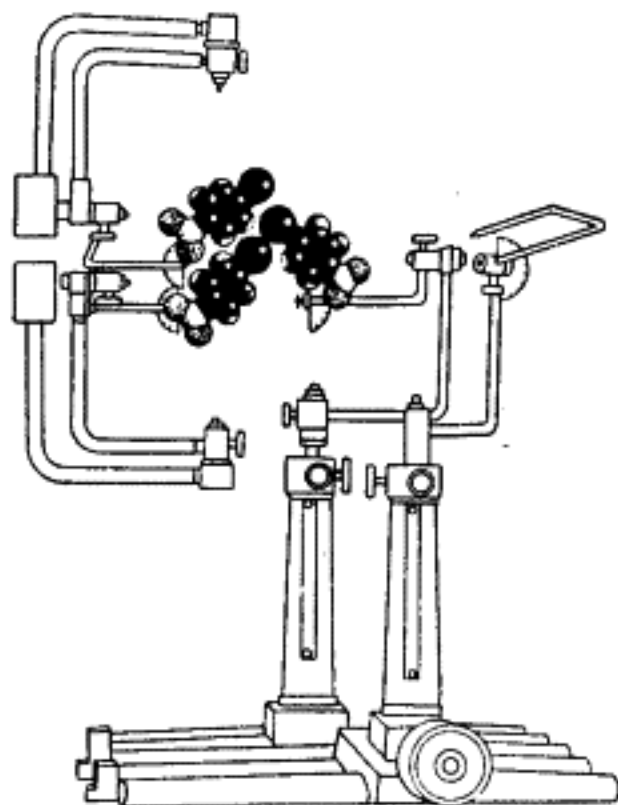


Рис. 1.1. Упаковка молекул на  
структуронскателе.

From the book: А.И.Китайгородский. Молекулярные кристаллы, Наука, 1971. (с. 11)

# Historical overview: use of packing optimization in (single-crystal) structure determination

---

Author(s), year	Method or function minimized	Compound(s)
Burke, <i>et al.</i> 1963	Manual search	Dibenzotetraazapenthalene
Williams, 1966	DF, <i>R</i>	Dibenzoylmethane
Zugenmaier, Sarco, 1972	(DF+ <i>R</i> )	Carbohydrates
Hass, <i>et al.</i> , 1975	Visual search, (DF+ <i>R</i> )	Spirodienone
Dzyabchenko <i>et al.</i> , 1980	DF	Aromatic quinones

---

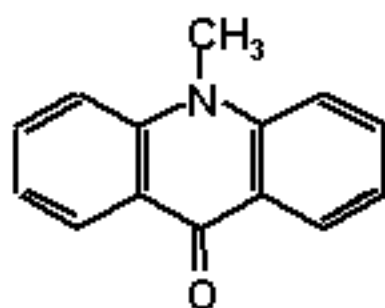
**DF: distance function of quadratic type;  
R: R-factor function**

---

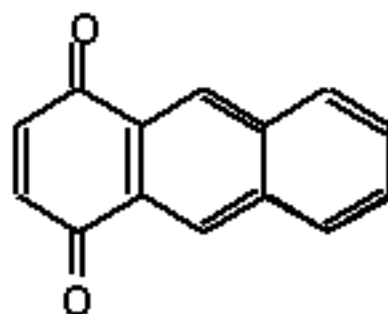
## Progress in theoretical methods and tools for X-ray energy minimization

Autor(s)	Year	Method or tool
Kitaigorodskii	1946	Structure seeker
Kitaigorodskii	1957	The AAP method
Williams	1964	PACK2
Rabinovich & Hirshfeld	1966	SEARCH
Gel'fand <i>et al.</i>	1966	'Method of valleys'
Hirshfeld	1968	Cheshire groups
Williams	1969	PCK2
Williams	1971	Acc. convergence
Williams	1972	PCK6
Huler & Warshel	1974	MCA
Scheraga et al.	1975	EPEN/2
Dzyabchenko	1979	PMC
Busing	1981	WMIN
Dzyabchenko	1994	CRYCOM
van Eijck & Kroon	1997	Structure clustering

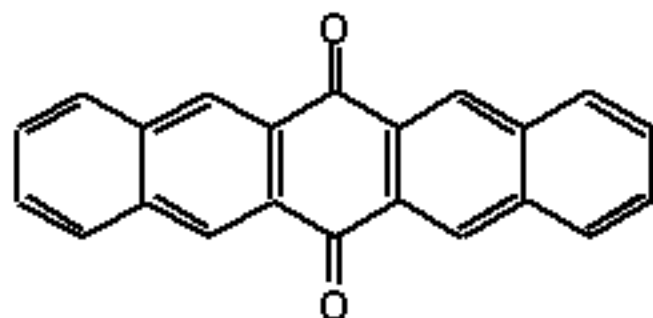
# Early examples of crystal structure solution by global energy minimization (Dzyabchenko *et al.*, 1979-1984)



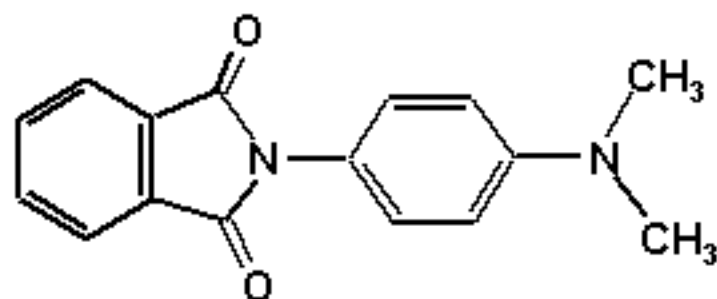
N-Methylacridone,  
 $P2_12_12_1$ ,  $Z=4$



1,4-Anthraquinone,  
 $P2_1/c$ ,  $Z=4$



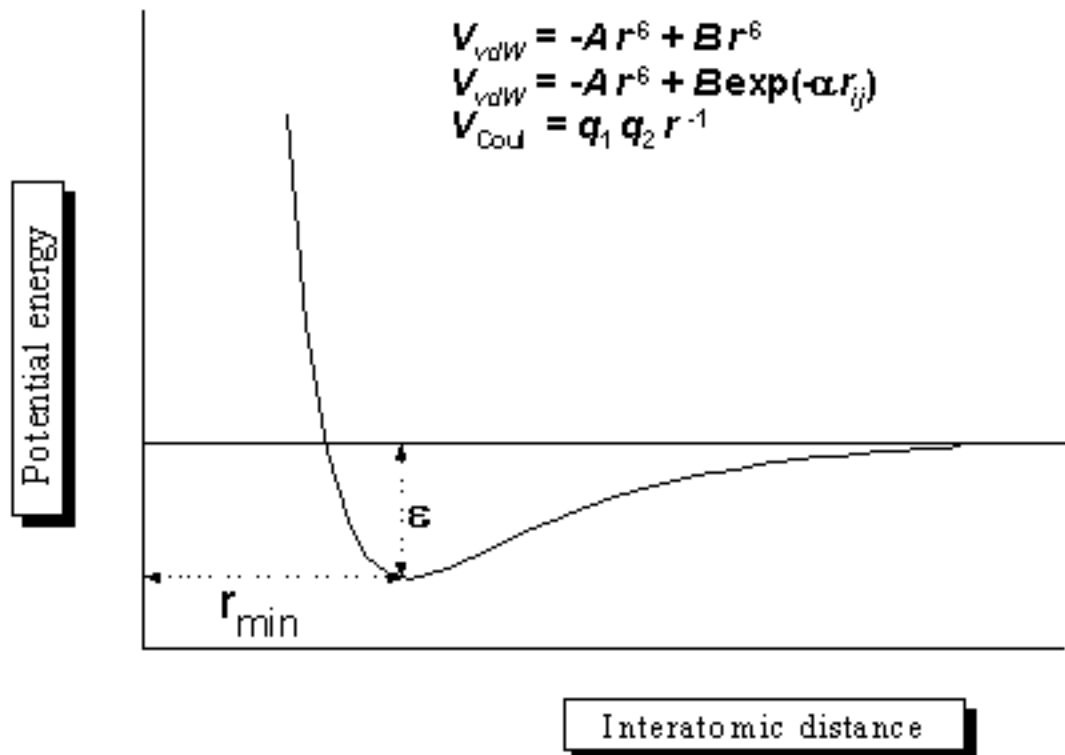
Pentacene-6,13-quinone,  
 $P2_1/c$ ,  $Z=2(-1)$



N,N-(Dimethylaminophenyl)-  
phthalimide,  $Pbcn$ ,  $Z=4(2)$

# The Atom-Atom Potential Method

## 1. Empirical atom-atom potentials:



## 2. Rigid-body approximation

$$3. F_{\text{eqv}} = \min \{E(a, b, \dots, \gamma, x_0, y_0, z_0, \varphi, \theta, \psi) + pV(a, b, \dots, \gamma)\}$$

---

# Where do the crystal force fields come from?

- Empirical FF:

Calibration by modeling of known crystal structures and their physical properties (Mirskaya and Kitaigorodskii; Williams; Scheraga *et al.*)

- *Ab initio* quantum chemical FF

- calculation of dimers, trimers, *etc.* (Mooij, van Eijck, Kroon)

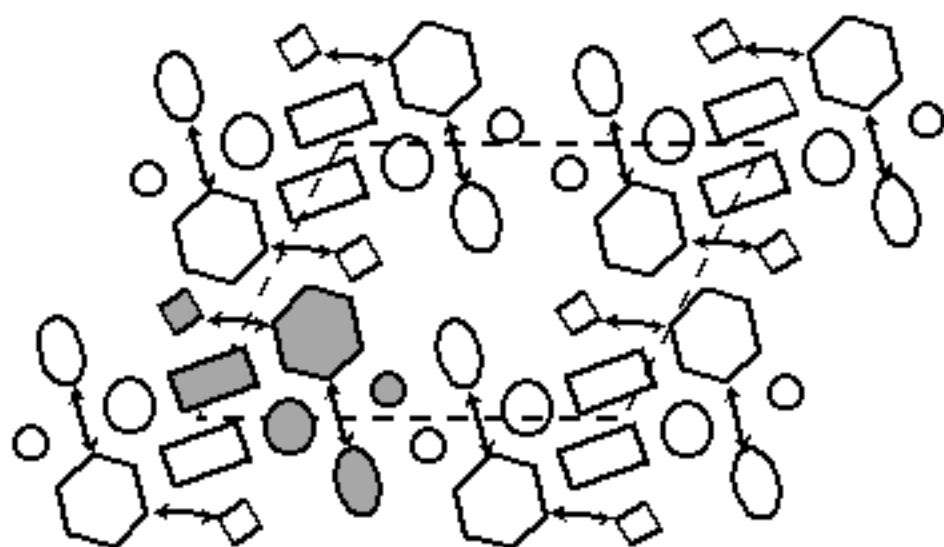
---



# Typical organic crystal structural classes (Belsky, Zorky, 1977, 1995)

Space group	<i>Z</i>	Site symmetry	% within named type	% of total no. of structures	
				1977	1995
<b>A: Chiral</b>					
<i>P2<sub>1</sub>2<sub>1</sub>2<sub>1</sub></i>	4	1	59.9	16.0	14.2
<i>P2<sub>1</sub></i>	2	1	28.5	7.0	6.8
<b>B: Racemic</b>					
<i>P2<sub>1</sub>/c</i>	4	1	54.9	28.6	30.3
<i>P1</i>	2	1	16.6	5.8	9.1
<i>Pbca</i>	8	1	7.6	4.1	4.2
<i>C2/c</i>	8	1	5.4	2.7	3.0
<i>C2/c</i>	4	2	3.4	2.6	1.9
<i>Pna2<sub>1</sub></i>	4	1	3.4	2.1	1.9
<b>C: Achiral</b>					
<i>P2<sub>1</sub>/c</i>	2	-1	42.0	6.2	4.2
<i>P1</i>	1	-1	13.6	1.9	1.4
<i>Pnma</i>	4	<i>m</i>	12.2	1.9	1.2
<b>Bisystemic</b>					
<i>P2<sub>1</sub>/c</i>	8	1,1	-	2.5	3.0
<i>P1</i>	4	1,1	-	-	2.0
<i>P2<sub>1</sub></i>	4	1,1	-	1.3	1.5

# A supramolecular solid in terms of rigid molecular fragments



- A supramolecular (organic or inorganic) solid is described as a system of  $N$  rigid fragments of various shapes and dimensions.
- Fragments involved into a flexible molecule are connected by virtue of link (bond variation) potentials
- Atoms of different fragments interact *via* van der Waals and electrostatic forces (unless they are not involved into a link connection). Link potentials are scaled to preserve standard geometry about flexible bonds. (They work as soft geometric constraints rather than as true potentials.)
- Torsional potentials can optionally be defined in addition to link potentials.

---

# Lattice-energy landscapes: Symmetry

*Acta Cryst.* (1983). **A39**, 941–946

## Symmetry of the Lattice-Energy Functional of a Molecular Crystal

BY A. V. DZYABCHENKO

*Karpov Institute of Physical Chemistry, ul. Obukha 10, Moscow 107120, USSR*

(Received 7 April 1983; accepted 4 July 1983)

### Abstract

In calculating the lattice-energy hypersurface by the systematic variation of the molecular rigid-body parameters and the lattice constants, the ranges to be scanned depend on the molecular symmetry and on the space group. A generalization of Hirshfeld's approach [Hirshfeld (1968). *Acta Cryst.* **A24**, 301–311] applicable to the case of variable lattice constants is suggested. The symmetry of the multidimensional parameter space is defined by the direct product of the molecular point group and a normalizer  $N_A(F)$  of the space group  $F$ . The normalizer  $N_A(F)$  is a group of affine transformations of the crystal axes that leave invariant the coordinates of equivalent positions. An asymmetric unit of the parameter space is obtained through keeping the lattice constants within such ranges that satisfy the Niggli reduced-cell conditions.

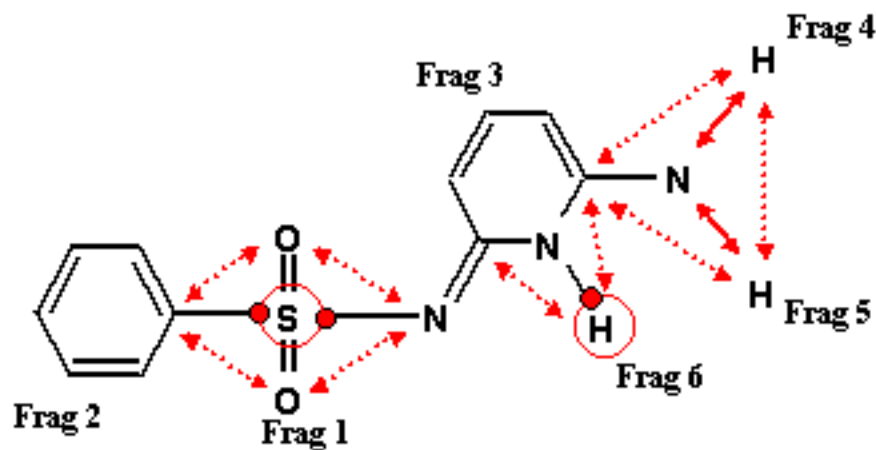
solutions of the phase problem. They were also employed in structure determinations by molecular-packing analysis when the cell dimensions were available from experiment (Williams, 1969; Dzyabchenko, Zavodnik & Belsky, 1979).

In lattice-energy calculations with variable lattice constants a more general approach is called for. This is the subject of the present work. It concerns only triclinic, monoclinic and orthorhombic crystals while those of higher symmetry are treated in terms of Hirshfeld's original theory. The results can be directly applied in calculations of the global lattice-energy minimum, and may also be useful whenever a comparison of crystal structures is made, say, when comparing isomorphous structures.

**Invariant transformations of the unit-cell axes**

---

# Flexible molecule as a system of rigid fragments



Red filled circles are auxiliary force centers.

Dotted red lines show the next-neighbor interactions under the control of link potentials to hold the bond angles at standard values

---

# Program PMC

(*'Packing of Molecules in Crystals'*)

Minimizes lattice energy  $E$  or a penalty function:

$$F = w_1 E + w_2 R$$

of a system of molecules composed of rigid fragments which interact with non-bonded potentials of the '6-exp-1' and '6-12-1' type and connected with each other *via* quadratic *link* (bond variation) and *torsional* potentials

Fast calculation due to:

- *convergence acceleration* for the lattice sums
- *contact symmetry* taken into account
- *analytical derivatives*
- supporting *pair lists* of atoms

Global minimization strategies:

- *systematic grid* search
- method of *valleys*

Supplementary program PMCDAT:

- preparation, testing and transformation of input data to PMC (and CRYCOM)

Links to other programs (such as for structure visualization):

- *via* lists of atomic coordinates in SYBYL or PDB format

---

# Structure prediction and PD study of a metastable polymorph of Piracetam

*Acta Cryst.* (1995). **B51**, 182–187

## **Structure of a Metastable Phase of Piracetam From X-ray Powder Diffraction Using the Atom–Atom Potential Method**

BY D. LOUËR AND M. LOUËR

*Laboratoire de Cristallogénie, CSIM (URA CNRS 1495), Université de Rennes I, Avenue du Général Leclerc,  
35042 Rennes CEDEX, France*

V. A. DZYABCHENKO

*Karpov Institute of Physical Chemistry, ul. Obukha 10, Moscow 107120, Russia*

AND V. AGAPONOV AND R. CEOLIN

*Laboratoire de Chimie Physique, Faculté de Pharmacie, 31 avenue de Monge, 37200 Tours CEDEX, France*

*(Received 19 April 1993; accepted 29 July 1994)*

---

# Prediction of benzene structures

Dzyabchenko (1984)	Global search for benzene structures in four space groups. Assessment of force fields for hydrocarbons
Dzyabchenko, Bazilevskii (1985)	Prediction of high pressure structures by constant-volume and enthalpy minimization
Dzyabchenko (1985)	Transformation paths and potential barriers
Dzyabchenko (1986)	Structural changes at negative pressures
Dzyabchenko (1987 and 1989)	Prediction of structures and space-group symmetry for $Z'=2$ and 4, resp. <i>Emergence of symmetry</i>
Shoda et al. (1994, 1995)	MD simulation, then enthalpy minimization for 20 structures from ( $N=4$ )
Tajima et al. (1995)	Prediction of structure and symmetry ( $N=8$ )
Gibson, Scheraga (1995)	Prediction of structure and symmetry ( $N=8$ )
Chaka et al. (1996)	Search in 14 space groups
Williams (1996)	Monte Carlo search ( $N=2$ and 4)
Leusen (1996)	Monte Carlo search
Thiéry, Rérat (1996)	Search for high-pressure polymorphs. Assigned to benzene II a monoclinic structure predicted by Dzyabchenko (1984)
van Eijck et al. (1998)	Search in all groups lower than tetragonal. Suggested a $P4_32_12$ , $Z=4(2)$ model to the PD data of Thiéry and Leger (1988) for benzene II

---

---

# CRYCOM

*Acta Cryst.* (1994). B50, 414–425

## Method of Crystal-Structure Similarity Searching

By A. V. DZYABCHENKO

*Karpov Institute of Physical Chemistry, ul.Obukha 10, Moscow 103064, Russia*

*(Received 8 July 1993; accepted 2 December 1993)*

### Abstract

An algorithm is presented which is used to recognize the similarity of crystal structures by matching a description of one structure against the equivalent descriptions of a second. Equivalent descriptions are considered since there are multiple possible ways of choosing the crystal axes and their origin, the asymmetric unit and atom numbering. Another idea used in the algorithm is to use positional and rotational parameters of the molecular fragments to describe the principal crystal structure. The two structures are

considered to be similar if in such descriptions the corresponding cell parameters and the parameters of corresponding fragments differ from each other within the limits specified. As a result of such similarity searching, the type of structural similarity, including the transformation matrix of crystallographic axes and the atom correspondences, is determined. The algorithm has been put to practical use in the *CRYCOM* program. This program is compatible with the Cambridge Structural Database (CSD) through interaction with the *GEOM78* program of CSD. The latter program is also used in

---



---

## Why do we need to compare crystal structures?

- To recognize a similarity of *different* structures
  - To determine the space-group symmetry
  - To find phase transition paths
-

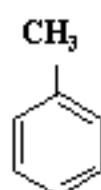
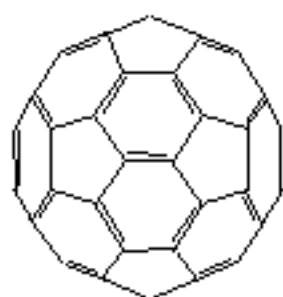
# Fullerene C<sub>60</sub>: Experimental discovery and theoretical structure predictions

Author(s)	Year	Event
Kroto <i>et al.</i>	1985	Discovery of fullerene molecule
Kroto, Krätschmer <i>et al.</i>	1988 1990	Fullerene C <sub>60</sub> obtained as a bulk material
Heiney, <i>et al.</i>	1991	Room-temperature phase of C <sub>60</sub> (fcc, space group <i>Fm3m</i> , Z=4)
Guo, Karasawa & Goddard III	1991	The LT structure predicted to be orthorhombic, <i>Cmca</i> , Z=4
David <i>et al.</i>	1991	Experimental LT structure of C <sub>60</sub> is in fact cubic, <i>Pa3</i> , Z=4 ('P')
David <i>et al.</i>	1992	LT phase contains P and H packings
Lu, Li & Martin	1992	A 'bond-bond' charge model predicts the structure to be <i>Pa3</i> (P)
Cheng, Klein	1992	A 'bond-atom' model predicts the structure to be <i>Pa3</i> (P)
Burgos, Halac, & Bonadeo	1994	A 'bond-pentagon' model predicts both P and H in correct order
Pintschovius & Chaplot	1995	A 'split-bond charge' model offered
Dzyabchenko	1995	The bond-bond model verified with the list of predicted structures.

# Predicted crystal structures of C<sub>60</sub>

## Bond - bond charge model

Structure	Lattice	$E_{\text{tot}}$ kcal/mol	$E_{\text{Coul}}$ kcal/mol	Density g/cm <sup>3</sup>
<b><i>Pa3-P</i></b>	<b>cubic</b>	<b>-44.4</b>	<b>-4.2</b>	<b>1.79</b>
<b><i>Pa3-H</i></b>	<b>cubic</b>	<b>-43.7</b>	<b>-1.8</b>	<b>1.81</b>
<i>P2<sub>1</sub>/c, Z'=2</i>	monocl	-43.2	-2.9	1.79
<i>P2<sub>1</sub>/n, Z=1</i>	monocl	-42.9	-3.2	1.78
<i>Pbca, Z=8</i>	orthorh	-42.8	-1.8	1.80
<i>R-3</i>	rhomb	-42.7	-2.7	1.79
<b><i>Cmca</i></b>	<b>orthorh</b>	<b>-42.5</b>	<b>0.2</b>	<b>1.82</b>
<i>P-1</i>	triclinic	-42.5	-2.3	1.79



# Structure prediction of the 1:1 C<sub>60</sub> toluene solvate confirmed by X-ray powder diffraction

## ■ Computational details:

Searched space group	C1c1
No. of independent molecules	two (one C <sub>60</sub> and one toluene)
Type of atom-atom potential	LJ + Coulomb (6-12-1)
No. of starting structures	~250
Program	PMC-2000

## Comparison of the predicted and final PD structures

	Predicted	Refined
Space group	Cmc2 <sub>1</sub>	~Cmc2 <sub>1</sub>
Z	4	4
Cell parameters (Å)		
a	10.33	10.37
b	31.02	31.67
c	9.88	10.13
C <sub>60</sub> parameters:		
u	0.000	0.000
v	0.141	0.142
w	0.250	0.250
φ (deg)	22.7	22.8
Toluene parameters		
u	0.500	0.500
v	0.064	0.058
w	-0.195	-0.206
φ (deg)	-22.7	-14.7

# Topochemical Polymerization of $C_{70}$ Controlled by Monomer Crystal Packing

Alexander V. Soldatov,<sup>1\*†</sup> Georg Roth,<sup>3</sup>  
 Alexander Dzyabchenko,<sup>4</sup> Dan Johnels,<sup>5</sup> Sergei Lebedkin,<sup>2</sup>  
 Christoph Meingast,<sup>1</sup> Bertil Sundqvist,<sup>6</sup> Miro Haluska,<sup>7</sup>  
 Hans Kuzmany<sup>7</sup>

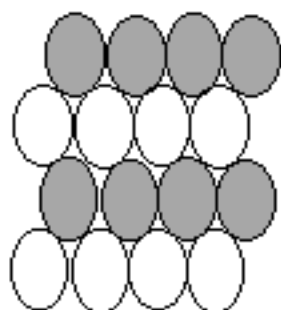
Polymeric forms of  $C_{60}$  are now well known, but numerous attempts to obtain  $C_{70}$  in a polymeric state have yielded only dimers. Polymeric  $C_{70}$  has now been synthesized by treatment of hexagonally packed  $C_{70}$  single crystals under moderate hydrostatic pressure (2 gigapascals) at elevated temperature (300°C), which confirms predictions from our modeling of polymeric structures of  $C_{70}$ . Single-crystal x-ray diffraction shows that the molecules are bridged into polymeric zigzag chains that extend along the *c* axis of the parent structure. Solid-state nuclear magnetic resonance and Raman data provide evidence for covalent chemical bonding between the  $C_{70}$  cages.

**Table 1.** Predicted and observed crystal packings of  $C_{70}$  polymer and dimer (26). The observed monomer structure is from (21), *Z* is the number of fullerenes per unit cell, and the van der Waals lattice energy (per one  $C_{70}$ ) was minimized with the Lennard-Jones 6-12 atom-atom potential with parameters  $r_0 = 3.7$  Å,  $\epsilon = -0.072$  kcal/mol for carbons using the program PMC (22). ND, not determined.

	Monoclinic dimer	Trigonal dimer	Orthorhombic polymer	Hexagonal polymer	Observed polymer	Observed monomer
Space group	<i>C2/m</i>	<i>P3<sub>1</sub></i>	<i>Ccmm</i>	<i>P6<sub>3</sub>/mcm</i>	<i>Ccmm</i>	<i>Pnam</i>
<i>Z</i>	4	6	4	6	4	4
Crystal packing	HP	CP	HP	HP	HP	HP
Lattice energy (kcal/mol)	-43.3	-42.8	-60.90	-51.64	ND	ND
Density (g/cm <sup>3</sup> )	1.844	1.835	1.916	1.736	1.804	1.733
Lattice parameters						
<i>a</i> (Å)	16.93	9.97	16.95	18.13	17.30	17.35
<i>b</i> (Å)	10.06	9.97	10.13	18.13	9.99	9.84
<i>c</i> (Å)	17.86	53.01	16.97	16.97	17.92	18.34
$\alpha$ (degrees)	90	90	90	90	90	90
$\beta$ (degrees)	84.3	90	90	90	90	90
$\gamma$ (degrees)	90	120	90	120	90	90

# Pressure-polymerized $C_{70}$

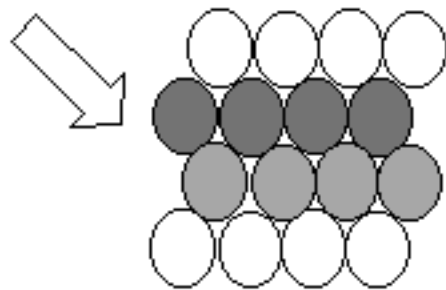
*Quasi-hcp (ABAB...)*



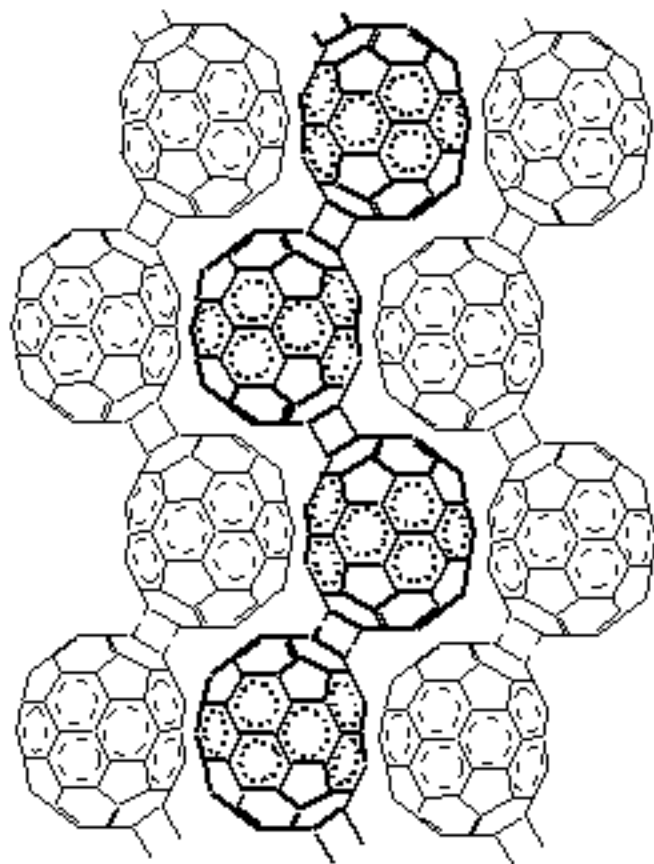
Polymerizes



*Quasi-fcc (ABCABC...)*



Does not  
polymerize!

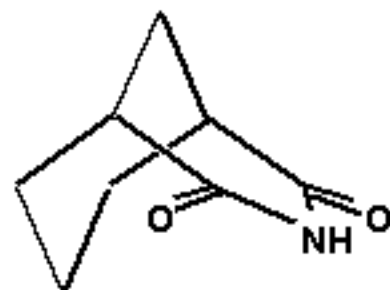


This polymer has been actually synthesized (Soldatov et al. 2001) soon after its structure and the formation conditions were predicted theoretically

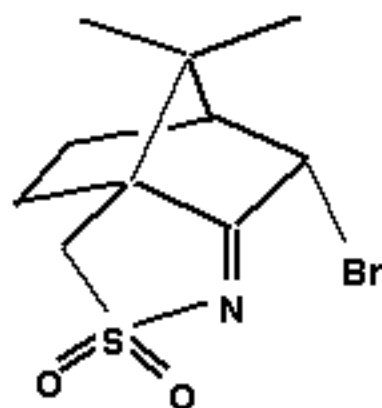
# Crystal structure prediction of small organic molecules: a second blind test (CSP-2001)

<b>Participant</b>	<b>Affiliation</b>	<b>Country</b>
Ammon	Univ. Maryland	USA
Dunitz	ETH-Zurich	Switzerland
Dzyabchenko	Karpov Institute	Russia
Erk	BASF AG	Germany
Gavezzotti	Univ. Milano	Italy
Hofmann	GMD-SCAI	Germany
Leusen	Accelrys Ltd.	UK
Lommerse		The Netherlands
Mooij	Astex Techn. Ltd.	UK
Motherwell	CCDC	UK
Price	Univ. College London	UK
Scheraga	Cornell Univ.	USA
Schmidt	Clariant GmbH	Germany
Schweizer	ETH-Zurich	Switzerland
van Eijck	Univ. Utrecht	The Netherlands
Verwer	Univ. Nijmegen	The Netherlands
Williams	Univ. Louisville	USA

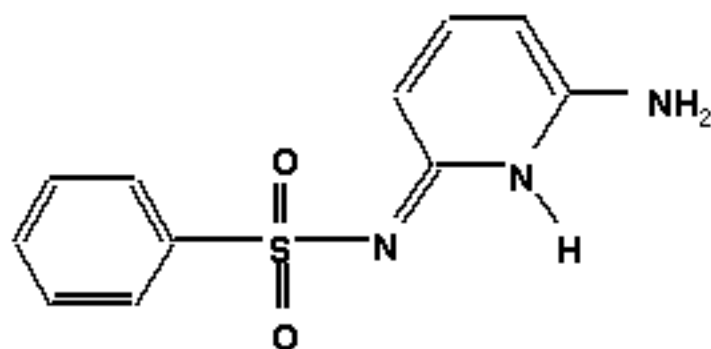
# CSP2001: The test molecules



3-Azabicyclo(3.3.1)nona-2,4-dione (I)



3-Bromocamphoryl sulfonimide (II)



*iso*-Sulfapyridine (III)



# Structural classes explored

---

Molecule	I and III	II
Space groups	<i>P2<sub>1</sub>/c</i> , ( <i>P2<sub>1</sub>/n</i> ), <i>Pbca</i> , <i>C2/c</i> , ( <i>I2/c</i> , <i>C2/n</i> ), <i>P-1</i> , <i>Pna2<sub>1</sub></i> , <i>P2<sub>1</sub>2<sub>1</sub>2<sub>1</sub></i> , <i>Cc</i> , ( <i>Cn</i> , <i>Ic</i> ), <i>C2</i> , and <i>P2<sub>1</sub></i>	<i>P2<sub>1</sub></i> and <i>P2<sub>1</sub>2<sub>1</sub>2<sub>1</sub></i>

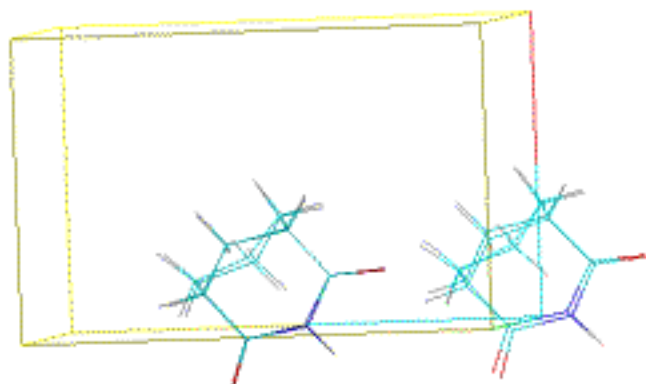
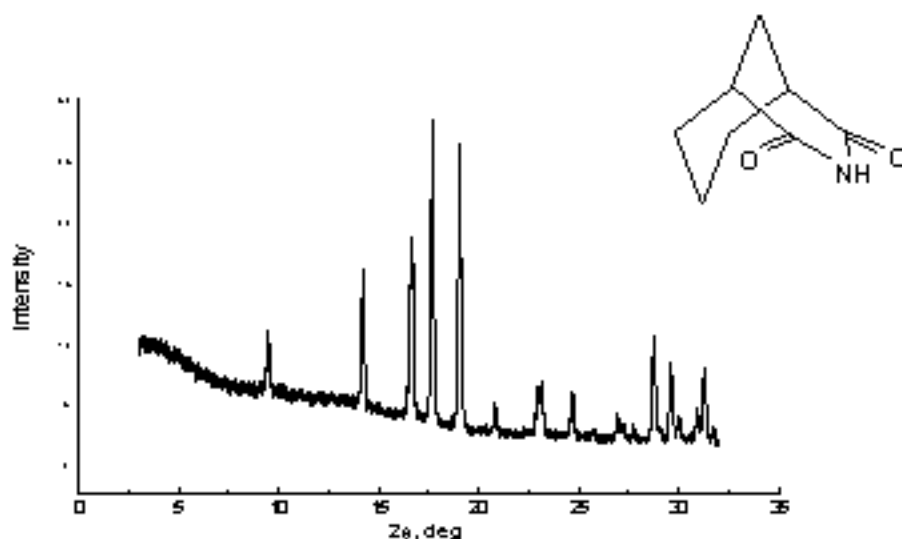
---

# CSP01: *Ab initio* results

Structure	Space group	Energy, kcal/mol	Density, g/cm <sup>3</sup>
I/1	<i>P2<sub>1</sub>/c</i>	0.00	1.371
I/2	<i>P2<sub>1</sub>/c</i>	0.19	1.340
I/3	<i>P-1</i>	0.29	1.312
<b>I/31</b>	<b><i>P2<sub>1</sub>/c</i></b>	<b>1.45</b>	<b>1.368</b>
<b>I/exp</b>	<b><i>P2<sub>1</sub>/c</i></b>		<b>1.338</b>
II/1	<i>P2<sub>1</sub>2<sub>1</sub>2<sub>1</sub></i>	0.00	1.716
II/2	<i>P2<sub>1</sub>2<sub>1</sub>2<sub>1</sub></i>	0.09	1.722
II/3	<i>P2<sub>1</sub>2<sub>1</sub>2<sub>1</sub></i>	0.43	1.691
<b>II/5</b>	<b><i>P2<sub>1</sub>2<sub>1</sub>2<sub>1</sub></i></b>	<b>0.58</b>	<b>1.702</b>
<b>II/exp</b>	<b><i>P2<sub>1</sub>2<sub>1</sub>2<sub>1</sub></i></b>		<b>1.608</b>
III/1	<i>Pbca</i>	0.00	1.526
III/2	<i>Pbca</i>	0.85	1.461
III/3	<i>Pbca</i>	1.48	1.494
<b>III/45*</b>	<b><i>P2<sub>1</sub>/c</i></b>	<b>4.00</b>	<b>1.431</b>
<b>III/exp</b>	<b><i>P2<sub>1</sub>/c</i></b>		<b>1.492</b>

\* Ranked among the minima found in *P2<sub>1</sub>/c*

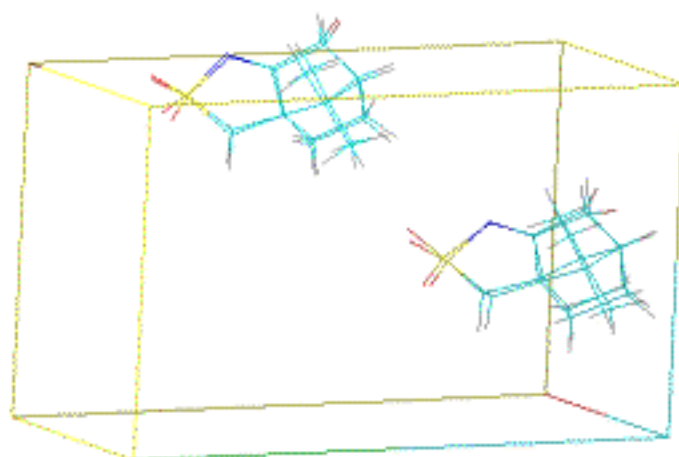
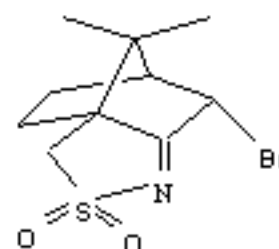
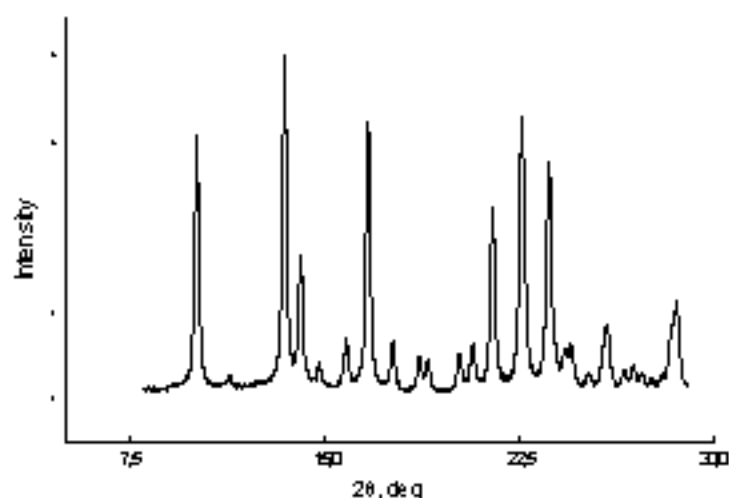
# CSP01 Powder Step: Molecule I



**Top:** powder pattern of compound I.

**Bottom:** superimpositions of the experimental arrangement of molecule with the energy-minimized one (right pair) and of the experimental with the molecule from the model refined with PD data (left pair).

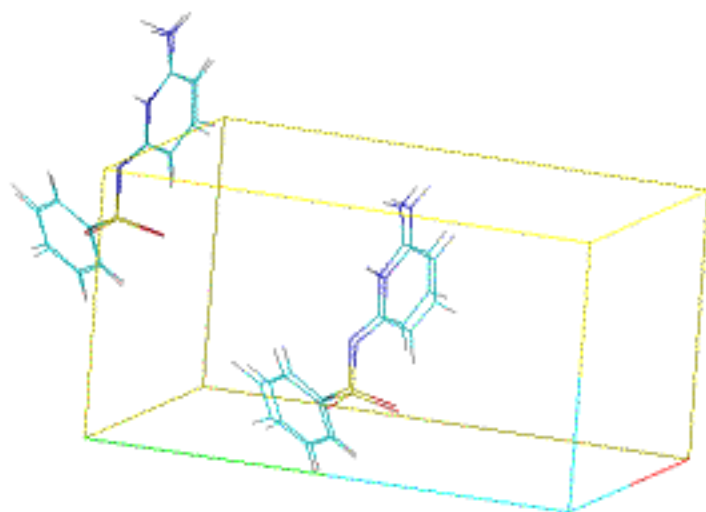
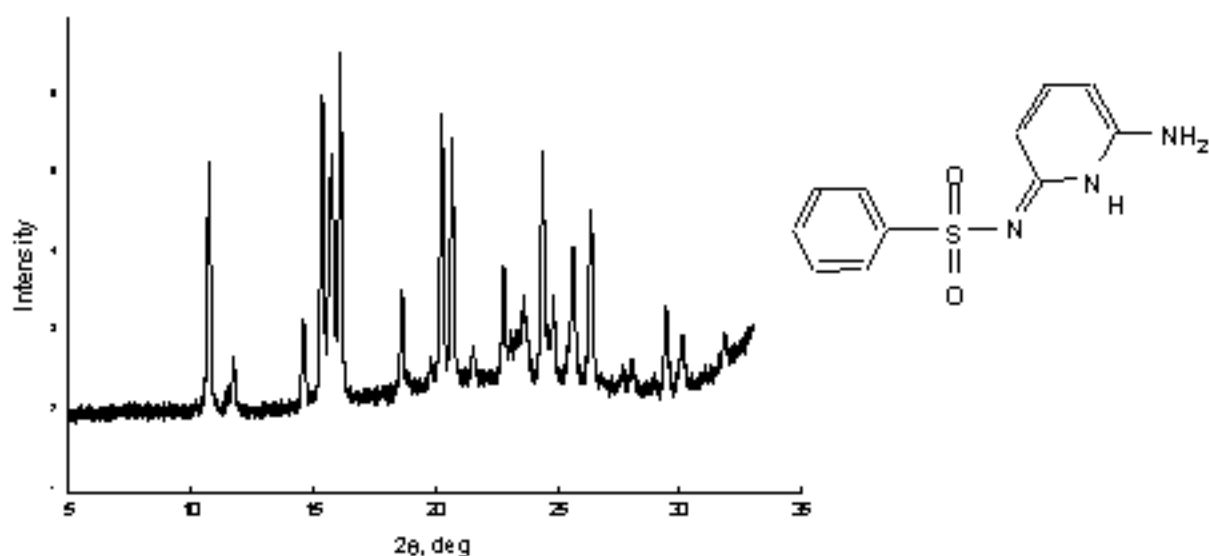
# CSP01 Powder Step: Molecule II



**Top:** Powder pattern of compound II.

**Bottom:** superimpositions of the experimental arrangement of molecule II in the unit cell with the energy-minimized one (lower right pair) and of the experimental with that after refinement with a combined 'Energy + X-ray PD pattern' function (upper left).

# CSP01 Powder Step: Molecule III



**Top:** Powder pattern of compound II.

**Bottom:** superimpositions of the experimental arrangement of molecule III in the unit cell with the energy-minimized one (center of the unit cell) and that after refinement with PD data (upper left).

---

**Weissbuch I., Lahav M., Leiserowitz L.**  
**(2003):**

## **Stereochemical control and Monitoring of Crystal Nucleation**

- Supersaturated solutions contain molecular clusters of various arrangements and shapes, some of which resemble the crystals into which they develop.
  - Control of crystal nucleation may be achieved employing tailor-made additives which are either nucleation inhibitors or promoters in achieving kinetic resolution of enantiomers and induced precipitation of particular crystal polymorphs.
-

# Structure predictions confirmed experimentally

Substance or polymorph	When predicted	Experimental confirmation
A high-pressure monoclinic phase of benzene	1985	Thiéry, Rérat, (1996) <sup>a</sup>
<i>Metastable polymorph of piracetam<sup>b</sup></i>	1995	<i>Louer et al. (1996)</i>
A high-pressure orthorhombic ( <i>Cmca</i> ) phase of $C_{60}$	1998	
Packing of the chain polymer in the O phase of $C_{60}$	1996	Moret et al. (1997)
Packing of the layer polymer in the T phase of $C_{60}$	1997	Moret et al. (2000); Narimbetov et al. (2003)
Packing of the layer polymer in the R phase of $C_{60}$	1997	Chen (2002)
Polymerized $C_{70}$	2001	Soldatov et al. (2001)

---

Acknowledgements to:

V. Agafonov and R. Ceolin (Tours University, France)  
M. Descamps and F. Affouard (Lille-1 Univ, France)  
H. Scheraga (Cornell University, USA)

Financial support from:

- Russian Foundation for Basic Research (1995-1997, 1999-2001)
  - University of Tours, France  
CNRS (1999);
  - Lille university (2000)  
Conseil Regional de Nord pas de Calais (2001)
  - NSF (2001, 2002)
-