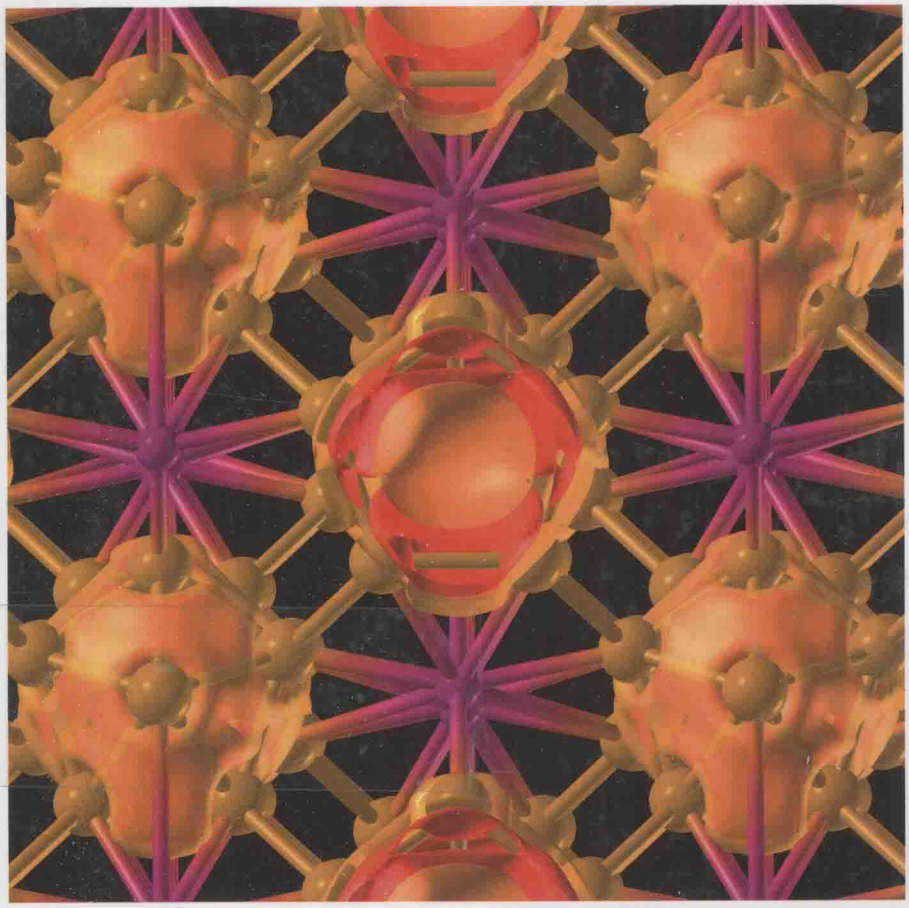
Edited by Artem R. Oganov

WILEY-VCH

Modern Methods of Crystal Structure Prediction



Edited by Artem R. Oganov

Modern Methods of Crystal Structure Prediction

WILEY-VCH Verlag GmbH & Co. KGaA

The Editor

Artem R. Oganov
Department of Geosciences and
Department of Physics and Astronomy
State University of New York
Stony Brook, NY 11794
USA
artem.oganov@sunysb.edu

Library of Congress Card No.: applied for

British Library Cataloguing-in-Publication Data
A catalogue record for this book is available from the British Library.

Bibliographic information published by the Deutsche Nationalbibliothek
The Deutsche Nationalbibliothek lists this publication in the Deutsche Nationalbibliografie; detailed bibliographic data are available on the Internet at <http://dnb.d-nb.de>.

Composition Laserwords Private Ltd., Chennai, India
Printing and Binding Fabulous Printers Pte. Ltd., Singapore
Cover Design Adam Design, Weinheim

Printed in Singapore
Printed on acid-free paper

ISBN: 978-3-527-40939-6

Edited by
Artem R. Oganov

Modern Methods of
Crystal Structure Prediction

Related Titles

Desiraju, G. R. (ed.)

Crystal Design

Structure and Function

2003

ISBN: 978-0-470-84333-8

Westbrook, J. H., Fleischer, R. L. (eds.)

Intermetallic Compounds, Volume 1, Crystal Structures

2000

ISBN: 978-0-471-60880-6

Desiraju, G. R. (ed.)

The Crystal as a Supramolecular Entity

1996

ISBN: 978-0-471-95015-8

List of Contributors

Vladislav A. Blatov
Samara State University
Chemistry Department
Ac. Pavlov St. 1
443011 Samara
Russia

Salah Eddine Boulfelfel
MPI CPfS
Noethnitzerstr. 40
01187 Dresden
Germany

Stefan Goedecker
Basel University
Department of Physics und
Astronomy
Klingelbergstrasse 82
4056 Basel
Switzerland

Richard G. Hennig
Cornell University
Department of Materials
Science and Engineering
214 Bard Hall
Ithaca, NY 14853-1501
USA

Martin Jansen
Max Planck Institute for
Solid State Research
Heisenbergstr. 1
70569 Stuttgart
Germany

Stefano Leoni
MPI CPfS
Noethnitzerstr. 40
01187 Dresden
Germany

Andriy O. Lyakhov
Stony Brook University
Department of Geosciences and
New York Center for
Computational Science
Stony Brook, NY 11794-2100
USA

Roman Martoňák
Comenius University
Department of Experimental
Physics
Mlynská dolina F2
842 48 Bratislava
Slovakia

Modern Methods of Crystal Structure Prediction. Edited by Artem R. Oganov

ISBN: 978-3-527-40939-6

Artem R. Oganov
Stony Brook University
Department of Geosciences,
Department of Physics and
Astronomy and New York Center
for Computational Science
Stony Brook, NY 11794-2100
USA

and

Moscow State University
Geology Department
119992 Moscow
Russia

Davide M. Proserpio
Università degli Studi di Milano
Dipartimento di Chimica
Strutturale e Stereochimica
Inorganica (DCSSI)
Via G. Venezian 21
20133 Milano
Italy

J. Christian Schön
Max Planck Institute for
Solid State Research
Heisenbergstr. 1
70569 Stuttgart
Germany

William W. Tipton
Cornell University
Department of Materials
Science and Engineering
214 Bard Hall
Ithaca, NY 14853-1501
USA

Mario Valle
Data Analysis and
Visualization Services
Swiss National Supercomputing
Centre (CSCS)
Cantonale Galleria 2
Manno, 6928
Switzerland

David J. Wales
University of Cambridge
Department of Chemistry
Lensfield Road
Cambridge, CB2 1EW
UK

Scott M.Woodley
University College London
Department of Chemistry
Gower Street
London WC1E 6 BT
UK

Introduction: Crystal Structure Prediction, a Formidable Problem

Artem R. Oganov

The famous 1988 editorial in *Nature* by John Maddox [1] stated:

> *"One of the continuing scandals in the physical sciences is that it remains in general impossible to predict the structure of even the simplest crystalline solids from a knowledge of their chemical composition".*

The central topic of the present volume is to review the state of the art in resolving this "scandal". Crystal structure is arguably the most important piece of information about a material, as it determines – directly or indirectly – pretty much all properties of a material. Knowing the structure, one can compute a large number of properties of a material, even before it is synthesized – hence the crucial importance of structure prediction for computational materials design. When the structure is unknown and cannot be predicted, very little can be said about the material.

Until recently, it was widely believed that crystal structures are fundamentally unpredictable [1–3] – as human behavior, or earthquakes, or long-term behavior of stock exchange. However, the situation began to change dramatically in 2003–2006, and this avalanche-like development of this important field can be called a scientific revolution that continues to this day. The aim of this book is to present some of the most important modern approaches to the formidable problem of crystal structure prediction.

What do we exactly mean by "crystal structure prediction problem"? For each chemical composition there are an infinite number of possible atomic arrangements that can, in principle, be obtained in the laboratory – these correspond to all possible local minima of the free energy. Among these, at each thermodynamic conditions (pressure, temperature, chemical potential) there are a finite number of special structures, extreme in some sense – the lowest energy (i.e. the most stable structures), the highest/lowest value of some other property (hardness, density, band gap, superconducting Tc, ...), or highest rate of nucleation (corresponding to kinetically preferred phases). Prediction of these structures is a well-defined and crucially important problem. In the simplest and most important case, by crystal structure prediction we mean finding, at given P-T conditions, the stable crystal structure knowing only the chemical formula.

Modern Methods of Crystal Structure Prediction. Edited by Artem R. Oganov

ISBN: 978-3-527-40939-6

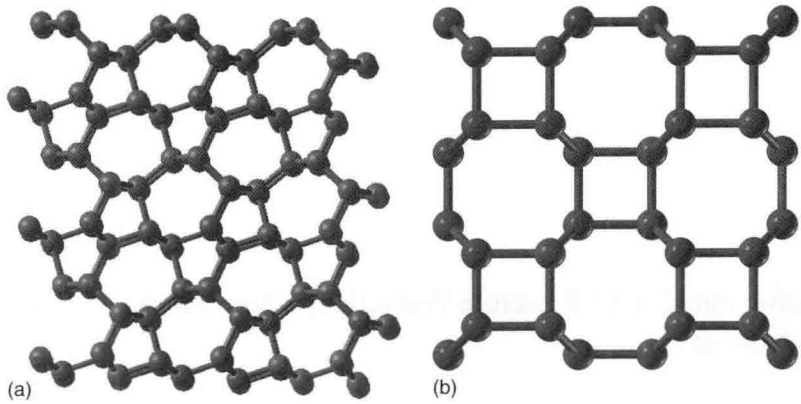


Figure 1 Structures of metastable superhard sp^3-bonded allotropes of carbon: (a) M-carbon [5] and (b) bct4-carbon [6].

Many types of approaches have been proposed to address this problem. Some are topological (as reviewed in the chapter by Blatov and Proserpio [4]) and aim at constructing the simplest topologies consistent with what we know about the chemistry of the system. This way, assuming sp^2-hybridization of carbon atoms one would arrive at 2H-graphite structure, and assuming sp^3-hybridization one would find the diamond and lonsdaleite structures – and a vast array of interesting metastable structures, including clathrates, M-carbon (Figure 1a), bct-carbon (Figure 1b) and other possible allotropes.

Topological approaches often appeal to symmetry, since in the vast majority of cases stable crystal structures do display some symmetry – the asymmetric point group 1 (or corresponding to it space group P1) is very rare (see Table 1). The ubiquity of symmetry may simplify the task of structure prediction, and not only in topological approaches.

Other approaches are based on empirical correlations and involve either *structural diagrams* [8–10] or *data mining approaches* [11, 12]. In either case, a large database of known stable crystal structures is required. While data mining approaches involve advanced machine learning concepts and are capable of predicting not only stable structures, but also the likelihood of compound formation in multinary systems (a formidable task too!), structural diagrams are much more empirical and limited in their scope. In these, one frequently uses ionic radii or the so-called "pseudopotential radii" [9], both of which (especially the latter) lack strict physical meaning and uniqueness. Instead of the "pseudopotential radii" one could use other quantities, such as the chemical scale or the Mendeleev number – the resulting empirical structure diagrams seem to have a good ability to separate structure types (e.g., Figure 2), and thus have predictive power.

The most unbiased, non-empirical and hence most generally applicable approaches are based on *computational optimization* – i.e. explicit calculations of the (free) energy and exploration of its landscape with the aim of finding the most stable arrangement of the atoms. These approaches are the main focus of

Table 1 Distribution of 280 000 chemical compounds over the 32 point groups. Note somewhat different frequencies for inorganic (I) and organic (O) compounds. (data collected by G. Johnson and published in [7]).

	I	O		I	O
1	0.67%	1.24%	422	0.40%	0.48%
$\bar{1}$	13.87	19.18	4mm	0.30	0.09
2	2.21	6.70	$\bar{4}$2m	0.82	0.34
M	1.30	1.46	4/mmm	4.53	0.69
2/m	34.63	44.81	6	0.41	0.22
222	3.56	10.13	$\bar{6}$	0.07	0.01
mm2	3.32	3.31	6/m	0.82	0.17
mmm	12.07	784	622	0.24	0.05
3	0.36	0.32	6mm	0.45	0.03
$\bar{3}$	1.21	0.58	$\bar{6}$m2	0.41	0.02
32	0.54	0.22	6/mmm	2.82	0.05
3m	0.74	0.22	23	0.44	0.09
$\bar{3}$m	3.18	0.25	m3	0.84	0.15
4	0.19	0.25	432	0.13	0.01
$\bar{4}$	0.25	0.18	$\bar{4}$3m	1.42	0.11
4/m	1.17	0.67	m3m	6.66	0.12

this book. Among the advantages are (i) the explicit calculation of the optimized quantity of interest (e.g., the energy), (ii) unbiased search techniques for exploring the energy landscape can – unlike the previously mentioned approaches, assuming knowledge of material's chemistry and likely crystal structures – arrive at completely unexpected results and truly novel structures. For instance, who would guess (based on whatever chemical knowledge) that boron under pressure would assume a NaCl-type structure composed of B_2 and B_{12} clusters with partially ionic bonding between the two? Who would guess that, when compressed to 2 million atmospheres, sodium assumes a structure unknown for any other element and becomes a transparent dielectric? Nevertheless, this is exactly what happens [13, 14], and these phenomena were first predicted using optimization techniques and only then confirmed experimentally.

When considering crystal structure prediction as an optimization problem – i.e. the problem of finding the global minimum of the energy landscape, certain properties of this landscape need to be explored. First, the number of distinct points on the landscape can be estimated as:

$$C = \binom{V/\delta^3}{N} \prod_i \binom{N}{n_i} \tag{1}$$

where N is the number of atoms in the unit cell of volume V, δ is a relevant discretization parameter (for instance, 1 Å) and n_i is the number of atoms of i-th

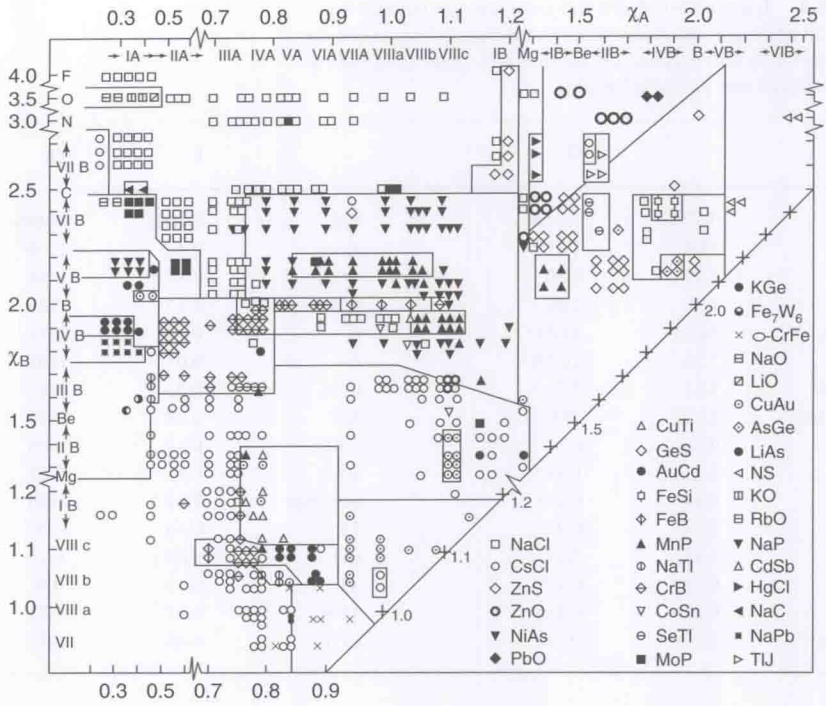


Figure 2 Pettifor's structure diagram for 574 AB compounds (from [10]).

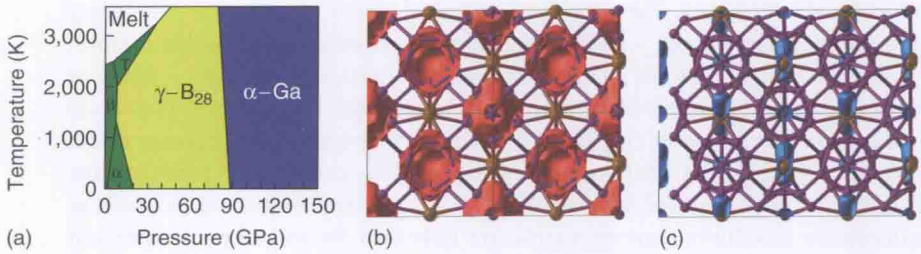


Figure 3 Boron: (a) its schematic phase diagram (from [13]) and distribution of electrons corresponding (b) bottom and (c) top of the valence band in γ-B_{28} [15].

type in the unit cell. C is astronomically large (roughly, $\sim 10^N$ if one uses $\delta = 1$ Å and typical atomic volume of 10 Å^3).

It is useful to consider the dimensionality of the energy landscape:

$$d = 3N + 3 \tag{2}$$

where $3N$-3 degrees of freedom are the atomic positions, and the remaining six dimensions are lattice parameters. For a system with 100 atoms in the unit cell, the landscape is 303-dimensional!

Equation (1) implies that the difficulty of crystal structure prediction increases exponentially with system size (or landscape dimensionality) and it thus poses an NP-hard problem (which is a shorthand of "non-deterministic polynomial-time hard", meaning that the scaling of the problem with the system size is faster than any polynomial). Such high-dimensional problems with astronomically large numbers of possible solutions imply that simple exhaustive search strategies are unfeasible.

Great simplification of the problem can be achieved if structures are relaxed, i.e. brought to the nearest local energy minima. During relaxation, certain correlations between atomic positions set in – interatomic distances adjust to reasonable values, and unfavorable interactions are avoided. The intrinsic dimensionality is thus equal to a reduced value:

$$d^* = 3N + 3 - \kappa \tag{3}$$

where κ is the (non-integer) number of correlated dimensions. Just doing relaxation, great simplifications of the global optimization problem can be achieved – for example, the dimensionality drops from 99 to 11.6 for $Mg_{16}O_{16}$ (a really simple system), while the decrease is less substantial for chemically complex systems – from 39 to 32.5 for $Mg_4N_4H_4$. To appreciate this simplification of the problem, we remind that the number of local minima depends exponentially on the intrinsic dimensionality:

$$C^* = \exp(\beta d^*) \tag{4}$$

This implies that any efficient search method must include structure relaxation (i.e. local optimization). Even simple random sampling, when combined with local optimization, can deliver correct solutions – although only for very small systems, roughly $N < 8$–10 ([17, 18], see chapter by Tipton and Hennig [19] in this volume). Much larger systems can be treated by more advanced methods, such as simulated annealing ([20, 21], see chapter by Schön and Jansen [22]), metadynamics ([23, 24], see chapter by Martoňák [25]), basin hopping ([26], see chapter by Wales in this volume [27]), minima hopping ([28], see chapter by Goedecker [29]), or evolutionary algorithms ([5, 30, 31], see chapter by Lyakhov et al. [32]). Many of the above methods rely on the fact that in usual chemical systems good (i.e. low energy) structures share some similarities, i.e. are located relatively close to each other on the landscape, forming the so-called energy funnels, low-energy regions of configuration space. This gives the landscape a benign overall shape – such as the one shown in Figure 5. Exploiting the fact (assumed for a long time by chemists, but now proven on real systems – see, e.g. [16]) that in real chemical systems there are only a few (or just one) energy funnels, allows further gains of efficiency of structure predictions. Nowadays, systems with a few hundred degrees of freedom can be treated by some of these methods – perfectly adequate for most inorganic and organic systems. Extending this limit to much larger systems may

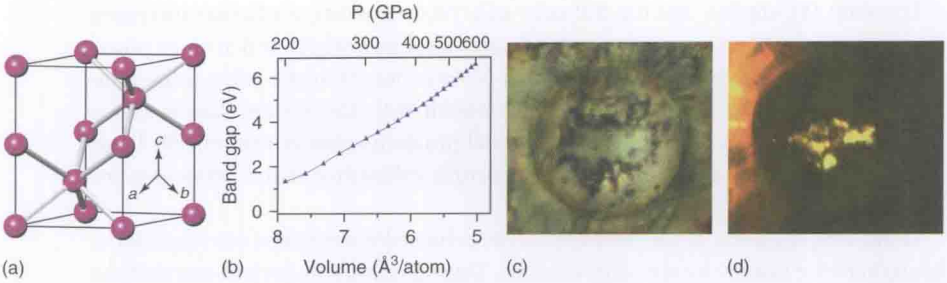


Figure 4 *hP*4 phase of sodium: (a) crystal structure, (b) band gap computed in the GW approximation, and optical photographs of a sodium sample at (c) 110 GPa (where sodium is a white reflecting metal) and (d) at 199 GPa (where it is a red transparent insulator). After [14].

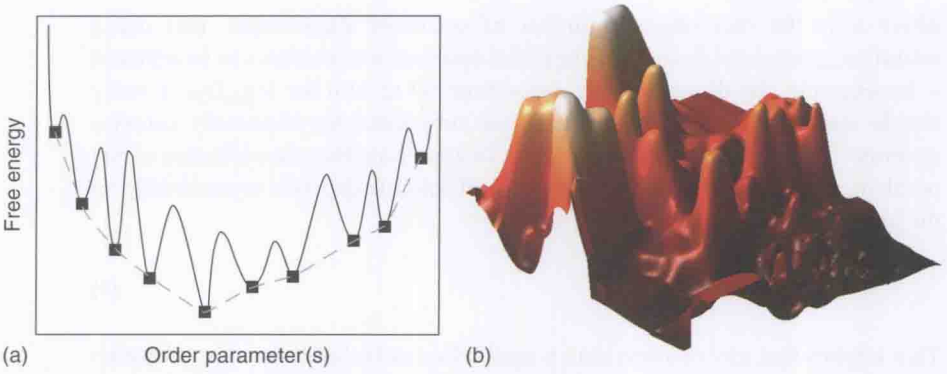


Figure 5 Energy landscape: (a) schematic illustration showing the full landscape (solid line) and reduced landscape (dashed line interpolating local minima points), (b) 2D-projection of the reduced landscape of Au_8Pd_4 (done using the method presented in [16]) showing all low-energy structures clustered in one region of configuration space.

enable us to treat biologically important systems and address such problems as protein folding. There are already some steps in this direction (see, e.g. [29]).

Next level of complexity is to ask if we can predict not just the stable structure, but also the whole set of stable chemical compositions (and the corresponding structures) in multicomponent system. This means that we are dealing with a complex landscape consisting of compositional and structural coordinates, and instead of a single ground state we should have a set of ground states located on the so-called convex hull (Figure 6). There are some encouraging steps in solving this problem [33–35].

We can also consider landscapes of properties other than the (free) energy. In this case, hybrid optimization needs to be performed – combining local optimization

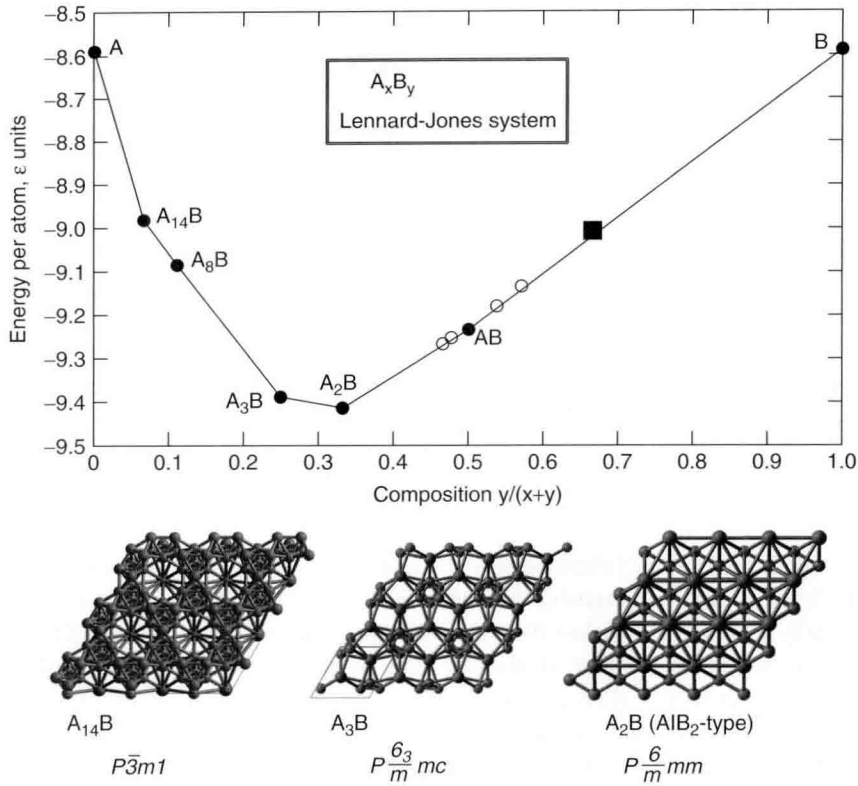


Figure 6 Examples of a simultaneous prediction of stable compositions and corresponding structures in a binary Lennard-Jones system with variable composition. Solid circles denote ground states, open circles – metastable solutions. See [35] for the potential model. From [35]. It is amazing that such a simple system gives such a wealth of complex ground states.

with respect to the energy (to find possible (quasi)equilibrium states) and global optimization with respect to the property of interest. This allows us to find structures (and compositions), corresponding to the desired values of the physical property of interest. Figure 7 gives an example of such optimization, utilizing the evolutionary algorithm [5, 30, 31] to search for the hardest possible structure of SiO_2.

This search employed the model of hardness [37] extended by Lyakhov and Oganov [36, 38], who also questioned whether diamond is the hardest carbon allotrope. The answer was that it indeed is (with the theoretical hardness of 89.4 GPa, within the error bars of to the experimental values [39]), but a number of other allotropes come close to it – for instance, lonsdaleite (89.3 GPa), bct4-carbon (theoretical hardness 84.2 GPa) and M-carbon (theoretical hardness of 83.4 GPa) shown in Figure 1. Both M-carbon [40] and bct4-carbon [6] (and private communication from a talented young researcher X.-F. Zhou, August 2009) structures were proposed as explanations for the experimentally observed new superhard allotrope

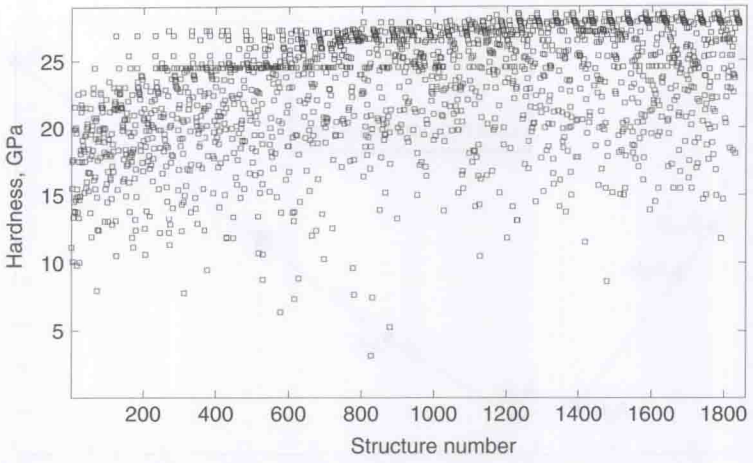


Figure 7 Evolutionary search for the hardest structure of SiO_2. From [36].

("superhard graphite") obtained by cold compression of graphite above 17 GPa [41]. Both structures are metastable and both match experimental observations almost equally well, but nevertheless there is a way of deciding which one is more likely to be the "superhard graphite" of Mao [41]. This brings us to the next major unsolved problem – *prediction of synthesizability of a metastable phase.* Indeed, the structure with optimal properties will frequently be metastable, and will be of interest only if it can be synthesized. This requires that the activation barrier for its formation from an available precursor (in this case, graphite) be lower than the barrier of formation of any other structure. The best approach for computing the absolute activation energies of solid-solid phase transitions is the transition path sampling

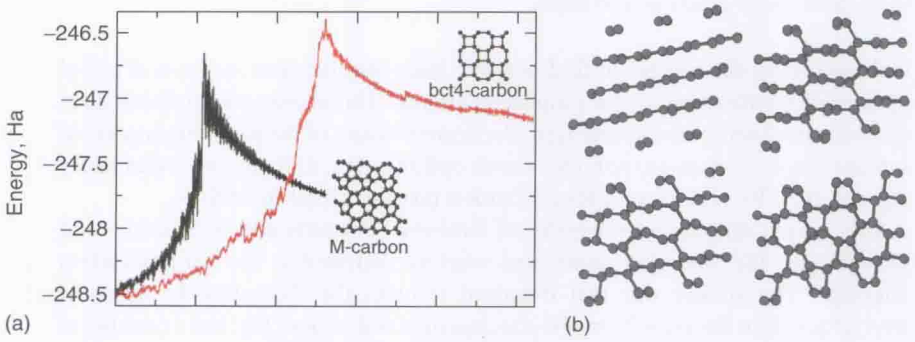


Figure 8 Synethesizability of M-carbon vs bct4-carbon: (a) energy profiles along the graphite-M-carbon (black) and graphite-bct4 (red) transition pathways and (b) the transition pathway graphite-M-carbon. The energies in given per supercell of 144 atoms. From [44]. Courtesy of S.E. Boulfelfel.

method [42], reviewed in this volume by Leoni and Boulfelfel [43]. One of the main advantages is that this method enables studies of nucleation and growth of the new phase, and its absolute activation barriers are meaningful (unlike those obtained by most other methods). As shown in Figure 8, the lower computed energy barrier clearly favors M-carbon over bct4-carbon [44].

It is my hope that this volume, reviewing most of the major methods of crystal structure prediction, all the way from topological approaches [4] to optimization methods [19, 22, 25, 27, 29, 32] and methods to appraise synthesizability of a material [43], will be useful to a wide readership of physicists, chemists, materials scientists and earth scientists. This volume also presents, in the Appendix, the first attempt to systematically compare different optimization strategies for a set of very challenging inorganic structure prediction problems [45]. The methods described in this volume should motivate further research into the structure and properties of materials, and will (probably quite soon) widely enable computational design of new functional materials. We are witnessing the dawn of a new era, where crystal structure prediction will no longer be an intractable problem.

I am grateful to Salah Eddine Boulfelfel, Andriy O. Lyakhov, Mario Valle, Feiwu Zhang, and Qiang Zhu, as well as former postdoc Yanming Ma and graduate student Colin W. Glass. I would also like to thank Wiley-VCH and its editors, in particular Anja Tschoertner, for their professionalism in preparing this book for publication. This work is supported by grants from Intel Corporation, Rosnauka (Russia, Contract No. 02.740.11.5102), Research Foundation of Stony Brook University, the National Natural Science Foundation of China (grant No. 10910263), and DARPA (grant #54751). Finally, I would like to express my gratitude to all the authors of this volume – it has been enormous pleasure to organize this book and edit it.

References

1. Maddox, J. (1988) Crystals from first principles. *Nature*, **335**, 201.
2. Gavezzotti, A. (1994) Are crystal structures predictable? *Acc. Chem. Res.*, **27**, 309–314.
3. Ball, P. (1996) Materials chemistry – scandal of crystal design... *Nature*, **381**, 648–650.
4. Blatov, V.A. and Proserpio, D.M. (2010) Periodic-Graph Approaches in Crystal Structure Prediction. **This volume.**
5. Oganov, A.R. and Glass, C.W. (2006) Crystal structure prediction using *ab initio* evolutionary techniques: principles and applications. *J. Chem. Phys.*, **124**, art. 244704.
6. Umemoto, K., Wentzcovitch, R.M., Saito, S., and Miyake, T. (2010) Body-centered tetragonal C_4: a viable sp^3 carbon allotrope. *Phys. Rev. Lett.*, **104**, 125504.
7. Newnham, R.E. *Properties of materials.* Oxford University Press (2005).
8. Mooser, E. and Pearson, W.B. (1959) On the crystal chemistry of normal valence compounds. *Acta Cryst.*, **12**, 1015–1022.
9. Burdett, J.K., Price, G.D., and Price, S.L. (1981) Factors influencing solid-state structure – an analysis using pseudopotential radii structural maps. *Phys. Rev.*, **B24**, 2903–2912.
10. Pettifor, D.G. (1984) A chemical scale for crystal structure maps. *Solid State Commun.*, **51**, 31–34.
11. Curtarolo, S., Morgan, D., Persson, K., Rodgers, J., and Ceder, G. (2003) Predicting crystal structures with data

mining of quantum calculations. *Phys. Rev. Lett.*, **91**, art. 135503.

12. Fischer, C.C., Tibbets, K.J., and Ceder, G. (2006) Predicting crystal structure by merging data mining with quantum mechanics. *Nature Materials*, **5**, 641–646.
13. Oganov, A.R., Chen, J., Gatti, C., Ma, Y.-Z., Ma, Y.-M., Glass, C.W., Liu, Z., Yu, T., Kurakevych, O.O., and Solozhenko, V.L. (2009) Ionic high-pressure form of elemental boron. *Nature*, **457**, 863–867.
14. Ma, Y., Eremets, M.I., Oganov, A.R., Xie, Y., Trojan, I., Medvedev, S., Lyakhov, A.O., Valle, M., and Prakapenka, V. (2009) Transparent dense sodium. *Nature*, **458**, 182–185.
15. Oganov, A.R. and Solozhenko, V.L. (2009) Boron: a hunt for superhard polymorphs. *J. Superhard Materials*, **31**, 285–291.
16. Oganov, A.R. and Valle, M. (2009) How to quantify energy landscapes of solids. *J. Chem. Phys.*, **130**, 104504.
17. Freeman, C.M., Newsam, J.M., Levine, S.M., and Catlow, C.R.A. (1993) Inorganic crystal structure prediction using simplified potentials and experimental unit cells – application to the polymorphs of titanium dioxide. *J. Mater. Chem.*, **3**, 531–535.
18. Schmidt, M.U. and Englert, U. (1996) Prediction of crystal structures. *J. Chem. Soc. – Dalton Trans.*, **10**, 2077–2082.
19. Tipton, W.W. and Hennig, R.G. (2010) Random search methods. **This volume**.
20. Pannetier, J., Bassasalsina, J., Rodriguez-Carvajal, J., and Caignaert, V. (1990) Prediction of crystal structures from crystal chemistry rules by simulated annealing. *Nature,*, **346**, 343–345.
21. Schön, J.C. and Jansen, M. (1996) First step towards planning of syntheses in solid-state chemistry: determination of promising structure candidates by global optimisation. *Angew. Chem. – Int. Ed.*, **35**, 1287–1304.
22. Schön, J.C. and Jansen, M. (2010) Predicting solid compounds using simulated annealing. **This volume**.
23. Martoňák, R., Laio, A., and Parrinello, M. (2003) Predicting crystal structures: The Parrinello-Rahman method revisited. *Phys. Rev. Lett.*, **90**, 075503.
24. Martoňák, R., Donadio, D., Oganov, A.R., and Parrinello, M. (2006) Crystal structure transformations in SiO_2 from classical and ab initio metadynamics. *Nature Materials*, **5**, 623–626.
25. Martoňák, R. (2010) Simulation of structural phase transitions in crystals: the metadynamics approach. **This volume**.
26. Wales, D.J. and Doye, J.P.K. (1997) Global optimization by basin-hopping and the lowest energy structures of Lennard-Jones clusters containing up to 110 atoms. *J. Phys. Chem.*, **A101**, 5111–5116.
27. Wales, D.J. (2010) Energy landscapes and structure prediction using basin-hopping. **This volume**.
28. Goedecker, S. (2004) Minima hopping: An efficient search method for the global minimum of the potential energy surface of complex molecular systems. *J. Chem. Phys.*, **120**, 9911–9917.
29. Goedecker, S. (2010) Global optimization with the minima hopping method. **This volume**.
30. Glass, C.W., Oganov, A.R., and Hansen, N. (2006) USPEX – evolutionary crystal structure prediction. *Comp. Phys. Comm.*, **175**, 713–720.
31. Lyakhov, A.O., Oganov, A.R., and Valle, M. (2010) How to predict large and complex crystal structures. *Comp. Phys. Comm.*, **181**, 1623–1632.
32. Lyakhov, A.O., Oganov, A.R., and Valle, M. (2010) Crystal structure prediction using evolutionary approach. **This volume**.
33. Wang, Y. and Oganov, A.R. (2008) Research on the evolutionary prediction of very complex crystal structures. IEEE Computational Intelligence Society Walter Karplus. Summer Research Grant 2008 Final Report. *http://www.ieee-cis.org/_files/EAC_Research_2008_Report_WangYanchao.pdf*
34. Trimarchi, G., Freeman, A.J., and Zunger, A. (2009) Predicting stable stoichiometries of compounds via evolutionary global space-group optimization. *Phys Rev*, **B80**, 092101.