Table 4. *Fourier coefficients for the positional parameters* ($\times 10^4$) *defined by* (1)

The center of mass is denoted by G, and θ and φ are the polar and azimuth angles (°) of the molecular direction.

		Mean	A_1	B_1	A_2	B_2	A_3	B_3	A_4	B_4
G	x	−983	0	−248	4	0	0	−33	1	0
G	y	658	0	194	12	0	0	3	5	0
G	z	0·0	−1·8	0	0	−0·3	−0·7	0	0	−0·3
	θ	90·0	0·77	0	0	−0·24	0·01	0	0	0·21
	φ	−25·43	0	6·55	−0·02	0	0	0·57	−0·38	0
S	x	1146	0	−132	17	0	0	−12	8	0
S	y	−81	0	370	8	0	0	21	−6	0
C	x	−1687	0	−282	3	0	0	−27	5	0
C	y	894	0	133	13	0	0	1	11	0
N1	x	−2795	26	−362	6	−4	1	−44	−2	17
N1	y	1279	−5	29	22	−2	3	−12	7	9
N2	x	−2795	−26	−362	6	4	−1	−44	−2	−17
N2	y	1279	5	29	22	2	−3	−12	7	−9

the same modulation. If the modulation wavenumber is not just $\frac{1}{9}\mathbf{c}_0^*$, then the modulation is incommensurate and the displacement pattern changes from that of model 1 to model 2 along the c axis with a long period.

If we regard this ninefold superstructure as an 'incommensurate' one, then the local structure around $z=\frac{1}{4}$ and $\frac{3}{4}$ corresponds to 'almost commensurate' regions (microdomains of ferroelectric phase I, the polar sense changing alternately along the b axis), which are separated by 'discommensurations' (or domain walls) around $z=0$ and $\frac{1}{2}$. It should be noted that the discommensuration width is about two basic cell dimensions in thiourea.

The authors wish to thank Mr S. Takagi and Mr Y. Yoshida for their help in computer calculations. This work was partly supported by a Grand-in-aid for Scientific Research from the Ministry of Education, Science and Culture.

References

ELCOMBE, M. M. & TAYLOR, J. C. (1968). *Acta Cryst.* A**24**, 410–420.

GOLDSMITH, G. J. & WHITE, J. G. (1959). *J. Chem. Phys.* **31**, 1175–1187.

International Tables for X-ray Crystallography (1974). Vol. IV. Birmingham: Kynoch Press. (Present distributor Kluwer Academic Publishers, Dordrecht.)

ISHIBASHI, Y. & SHIBA, H. (1978). *J. Phys. Soc. Jpn*, **45**, 409–413.

KOVALEV, O. A. (1965). *Irreducible Representations of the Space Groups*. New York: Gordon & Breach.

MCKENZIE, D. R. (1975). *J. Phys. C*, **8**, 1605–1619.

MASHIYAMA, H., HASEBE, K. & TANISAKI, S. (1988). *J. Phys. Soc. Jpn*, **57**, 166–175.

MOUDDEN, A. H., DENOYER, F., LAMBERT, M. & FITZGERALD, W. (1979). *Solid State Commun.* **32**, 933–936.

SAKURAI, T. (1967). Editor. *Universal Crystallographic Computation Program System*. The Crystallographic Society of Japan, Tokyo, Japan.

SHIOZAKI, Y. (1971). *Ferroelectrics*, **2**, 245–260.

SOLOMON, A. L. (1956). *Phys. Rev.* **104**, 1191.

TANISAKI, S., MASHIYAMA, H. & HASEBE, K. (1984). 39th Annu. Meet. Phys. Soc. Jpn, Fukuoka, Advanced Abstracts Vol. 2, p. 101.

TANISAKI, S. & NAKAMURA, N. (1970). *J. Phys. Soc. Jpn*, **28**, Suppl. 293–294.

YAMAMOTO, A. (1980). *Phys. Rev. B*, **22**, 373–379.

Acta Cryst. (1988). B**44**, 445–448

Can Statistical Analysis of Structural Parameters from Different Crystal Environments Lead to Quantitative Energy Relationships?

BY HANS BEAT BÜRGI

Laboratory for Chemical and Mineralogical Crystallography, University of Berne, CH-3012 Berne, Switzerland

AND JACK D. DUNITZ

Organic Chemistry Laboratory, Swiss Federal Institute of Technology, ETH-Zentrum, CH-8092 Zürich, Switzerland

(*Received* 13 *October* 1987; *accepted* 5 *April* 1988)

Abstract

Two recent attempts to derive quantitative energy relationships from statistical analysis of structural parameters observed in different crystal environments are examined. Both are based on the assumption of a Boltzmann-like distribution for the probability of a structure being observed in a deformed state. This assumption is shown to be untenable. An alternative model taking explicit account of the perturbing *forces* responsible for structural deformation is then considered. Although low-energy regions of molecular potential energy surfaces can certainly be recognized and mapped from distributions of observed structures,

0108-7681/88/040445-04$03.00

the derivation of quantitative energy relationships from statistical analysis alone is not possible without introducing arbitrary and unwarranted assumptions.

Introduction

The structure of a molecule in a crystal environment is not necessarily identical with the equilibrium structure of the isolated molecule, *i.e.*, the forces exerted on a molecule by its environment in a crystal can deform it to a greater or lesser extent. In general, stiff structural parameters (such as bond lengths) will hardly change from one environment to another, whereas soft conformational parameters will show a greater variability. Studies of the structural variability of molecules and molecular fragments in different crystal environments thus lead to the mapping of low-energy molecular deformation paths, which can often serve as useful models for chemical reaction paths (for a review see Bürgi & Dunitz, 1983). In this article we discuss two recent attempts to derive more quantitative energy relationships by statistical analysis of observed structural variability and show that both are untenable. We then consider an alternative approach to this type of problem and draw some conclusions from it.

Models based on Boltzmann-like distributions

In their recent statistical analysis of the flexible furanose ring, Bartenev, Kameneva & Lipanov (1987) introduce quantitative energy considerations by assuming that the probability W_G of a structure being observed in a deformed state is determined by the free-energy difference ΔG between the deformed and energetically optimal conformation of an isolated molecule. They assume a Boltzmann-like distribution $W_G=\exp(-\Delta G/RT_c)$ where RT_c characterizes the 'mean energy of deformation due to intermolecular interactions in crystals, by analogy with the probability $\exp(-\Delta G/RT)$ of the deformed conformation being realized in solution because of thermal motion with the mean energy RT.' From a statistical analysis of the distribution of about 250 observed furanose conformations in a variety of nucleoside and nucleotide crystals and fitting to the above type of expression with $T = 300$ K, Bartenev, Kameneva & Lipanov derive a barrier of 13·8 kJ mol^{-1} for interconversion of the C2′-*endo* and C3′-*endo* conformations along the pseudorotation pathway. This is lower than the barrier estimated by NMR (18·8 kJ mol^{-1}), so these authors then argue that the appropriate value of T_c in the expression for W_G 'must be higher than room temperature to allow the process of crystallization'. From comparison of the two barriers, they suggest that T_c should be taken as about $1{\cdot}4 \times 300$ K.

Two questions immediately arise. Has the assumption of a Boltzmann-like distribution any serious foundation in this context? And if so, what is the appropriate value of T? (There is ample evidence that molecular structural parameters are practically independent of the temperature of observation or of crystallization.) In order to discuss these questions we rephrase slightly some of the underlying assumptions in a form originally proposed by Murray-Rust (1982) and reduce the problem for simplicity to that of a distribution involving a single structural parameter x with mean value $\langle x\rangle$ and variance $\sigma^2(x)$. If the distribution is normal, the probability of observing a particular deformation $x-\langle x\rangle$ is

$$P(x) \approx \exp[-(x-\langle x\rangle)^2/2\sigma^2(x)]. \qquad (1)$$

It is then *assumed* that the probability of observing a particular deformation decreases exponentially with the deformation energy ΔG (or E), which can be taken, for small deformations at least, as a quadratic function of $x-\langle x\rangle$, leading to

$$P'(x) \approx \exp[-k(x-\langle x\rangle)^2/2E_c], \qquad (2)$$

where E_c is a constant to be determined, not necessarily RT as in the Boltzmann distribution. On the basis of the similarity of (1) and (2) the conclusion is drawn that the variance of x is inversely proportional to the force constant:

$$\sigma^2(x) = E_c/k. \qquad (3)$$

Murray-Rust (1982) found that torsional force constants for various types of C(sp^2)–X bonds could be related to statistically determined variances with a single value of E_c, approximately 0·63 kJ mol^{-1} (0·15 kcal mol^{-1}), which is described as 'the average amount of energy available from packing forces to distort a functional group'. Bartenev, Kameneva & Lipanov (1987) identify E_c with RT_c, and their value (approximately 3·5 kJ mol^{-1}) is obviously much larger than Murray-Rust's. There seems no reason to suppose that E_c is constant for different types of deformation or that it has anything to do with the ambient temperature or with the temperature of crystallization. *An ensemble of structural parameters obtained from chemically different compounds in different crystal structures does not even remotely resemble a closed system at thermal equilibrium and does not therefore conform to the conditions necessary for the application of the Boltzmann distribution.* It is thus misleading to draw an analogy between this distribution and those derived empirically from statistical analysis of observed deformations in crystals. Hence, there is no justification for (3), and, more generally, there seems to be no way of deriving absolute energies (or temperatures) from purely geometric structural data devoid of an energetic context.

Note added in proof: The same criticism can be directed at the arguments used in the recent paper by Lesyng, Jeffrey & Maluszynska (1988), who assume that hydrogen-bond-length distributions observed in crystal structures of nucleic acid components have a Boltzmann-like dependence on temperature.

Molecular deformations and perturbing forces

What can be done? We consider here a simple physical model that furnishes an alternative relationship between $\sigma^2(x)$ and k for a structural parameter x with equilibrium value x_o in the absence of perturbing forces. The crystal environment is supposed to exert some perturbing force a on the system, so that the linearly perturbed energy becomes:

$$E(x) = k(x-x_o)^2/2 + a(x-x_o). \tag{4}$$

The new equilibrium value x_e is thus displaced to

$$x_e = x_o - a/k.$$

Each crystal environment can be expected to exert a different perturbing force, and for a collection of such environments there will be some distribution $P(a)$, which could well be supposed to be normal (since it results from a large number of independent causes). Whatever the type of distribution,

$$P(a/k) = P(x_e - x_o), \tag{5}$$

that is, the distribution of a determines the distribution of x_e. In particular, for $\langle a\rangle = 0$, $\langle x_e\rangle = \langle x_o\rangle$. This model leads to the relationship

$$\sigma^2(x_e - x_o) = \sigma^2(a)/k^2, \tag{6}$$

in contrast to (3) where the parameter variance was proportional to k^{-1} rather than to k^{-2}. Qualitatively there is an inverse relationship for both models. The force constant can only be determined from (6) if $\sigma^2(a)$ is known, and this will seldom be the case. Hence the actual deformation energies cannot, in general, be derived from the observed parameter variances. Note that (6) does not refer directly to energy but rather to a ratio of two energy-related quantities, a force and a force constant.

The model can be extended to cover a multidimensional distribution involving several different structural parameters (interatomic distances, bond angles, torsion angles, *etc.*) by rewriting (4)–(6) in appropriate form:

$$E(\mathbf{x}) = (\mathbf{x} - \mathbf{x}_o)^T\mathbf{F}(\mathbf{x} - \mathbf{x}_o)/2 + \mathbf{a}^T(\mathbf{x} - \mathbf{x}_o) \tag{7}$$

$$P(\mathbf{x}_e - \mathbf{x}_o) = P(\mathbf{F}^{-1}\mathbf{a}) = P(\mathbf{C}\mathbf{a}) \tag{8}$$

$$\langle(\mathbf{x}_e - \mathbf{x}_o)(\mathbf{x}_e - \mathbf{x}_o)^T\rangle = \mathbf{C}\,\langle\mathbf{a}\mathbf{a}^T\rangle\,\mathbf{C}^T. \tag{9}$$

The left-hand side of expanded equation (6) involves the variance–covariance matrix of the observed distribution of structural parameters, and the right-hand side involves the compliance matrix $\mathbf{C} = \mathbf{F}^{-1}$ for the system in question as well as $\langle\mathbf{a}\mathbf{a}^T\rangle$, the second-moment matrix of the distribution of perturbing forces. This model provides a conceptual basis for relating observed distributions $P(\mathbf{x}_e - \mathbf{x}_o)$ to the energy surfaces associated with small $\mathbf{x}_e$, but its actual application is beset with obstacles.

The variance–covariance matrix $\langle(\mathbf{x}_e - \mathbf{x}_o)(\mathbf{x}_e - \mathbf{x}_o)^T\rangle$ may be heavily contaminated by contributions from experimental error. Indeed, for small deformations and poor experimental data, it could be dominated by the experimental uncertainties. On the other hand, for very large deformations the quadratic energy dependence implied by (7) cannot be expected to hold. Once the harmonic approximation breaks down equations (7), (8) and (9) would need to be replaced by more complicated expressions involving a larger number of unknown (anharmonic) force constants.

Even if these equations are assumed to be valid, the variance–covariance matrix $\langle(\mathbf{x}_e - \mathbf{x}_o)(\mathbf{x}_e - \mathbf{x}_o)^T\rangle$ cannot, in general, be resolved into $\mathbf{C}$ and $\langle\mathbf{a}\mathbf{a}^T\rangle$. However, if sufficient information about the force constants $\mathbf{F}(=\mathbf{C}^{-1})$ were available, (7) could be used to obtain the perturbing forces $\mathbf{a}$ for each observed crystal environment separately. For a sufficiently large sample of different environments, $\langle\mathbf{a}\mathbf{a}^T\rangle$ could be calculated and its dependence on crystal environment studied. Conversely, the compliance matrix $\mathbf{C}$ could be derived if $\langle\mathbf{a}\mathbf{a}^T\rangle$ were known.

Clearly we cannot possibly be expected to know these quantities. The distribution of forces could be different for different kinds of structural parameter and is likely to depend in a complex way on the types of interaction (van der Waals, hydrogen bond, ionic forces) operative in the selection of crystal structures included in the investigation. The detailed analysis of these forces for a statistically significant sample of structures would be extremely laborious and could hardly be free from assumptions about the natures and magnitudes of these interactions.

To enable any progress to be made, drastic simplifications are necessary. For example, in the structure-correlation method (Murray-Rust, Bürgi & Dunitz, 1975, 1978; Bürgi & Dunitz, 1983) it is assumed that the second moments of the observed distribution $P(\mathbf{x}_e - \mathbf{x}_o)$ will tend to be small in directions of large increase of potential energy. In terms of (9) this is equivalent to the assumption that the second-moment matrix $\langle\mathbf{a}\mathbf{a}^T\rangle$ can be approximated by a single proportionality constant, *i.e.*, that the off-diagonal terms of $\langle\mathbf{a}\mathbf{a}^T\rangle$ are essentially zero and that the diagonal terms do not differ too much among themselves.

Conclusions

Structural deformations of molecules or molecular fragments in crystals depend on the particular perturbations operating in the individual crystal environments, as well as on the restoring forces. Absolute deformation energies cannot be derived from structural data alone without introducing arbitrary and unwarranted assumptions. However, observed distributions of structural parameters found in different crystal environments do provide qualitative information about the shapes of low-energy regions of potential energy surfaces and can thus be related to ratios of corresponding restoring forces. We have outlined a general scheme indicating how this information might be obtainable.

Perhaps a parallel can be drawn with molecular mechanics models, widely used for estimating structural and energetic properties of molecules. These are just more or less elaborate schemes of representing inter- and intramolecular perturbations on a standard fragment characterized by a standard geometry and a standard force field. As far as the structures alone are concerned, the force constants may be scaled upward or downward without change; it is only their ratios that matter.

References

Bartenev, V. N., Kameneva, N. G. & Lipanov, A. A. (1987). *Acta Cryst.* B**43**, 275–280.

Bürgi, H. B. & Dunitz, J. D. (1983). *Acc. Chem. Res.* **16**, 153–161.

Lesyng, B., Jeffrey, G. A. & Maluszynska, H. (1988). *Acta Cryst.* B**44**, 193–198.

Murray-Rust, P. (1982). *Molecular Structure and Biological Activity*, edited by J. F. Griffin & W. L. Duax, pp. 117–133. New York: Elsevier.

Murray-Rust, P., Bürgi, H. B. & Dunitz, J. D. (1975). *J. Am. Chem. Soc.* **97**, 921–922.

Murray-Rust, P., Bürgi, H. B. & Dunitz, J. D. (1978). *Acta Cryst.* B**34**, 1793–1803.

SHORT COMMUNICATION

Contributions intended for publication under this heading should be expressly so marked; they should not exceed about 1000 *words; they should be forwarded in the usual way to the appropriate Co-editor; they will be published as speedily as possible.*

Acta Cryst. (1988). B**44**, 448

Gitterenergieberechnungen zur Ermittlung der wahrscheinlichsten Strukturtypen für Verbindungen MX_6. Druckfehlerberichtigung.

Von Wolfgang Willing und Ulrich Müller, *Fachbereich Chemie der Universität Marburg, Hans-Meerwein-Strasse, D-*3550 *Marburg, Bundesrepublik Deutschland*

(*Eingegangen am* 5. *April* 1988)

Abstract

Printer's errors in the article by Willing & Müller [*Acta Cryst.* (1988), B**44**, 1–6] are corrected. On page 2, the second sentence of the second paragraph should read: 'Wir betrachten nur intermolekulare Wechselwirkungen, die interatomaren Potentiale innerhalb eines Moleküls werden nicht berücksichtigt'. In the bottom right of Fig. 1, the height of the layers, reading from top to bottom, should be: 3/4, 1/4 and −1/4.

Alle Daten sind in der Zusammenfassung gegeben.

0108-7681/88/040448-01$03.00