

Is Small Molecule Crystallography Still Science?

***Phillip Fanwick
Department of Chemistry
Purdue University
West Lafayette, IN 47907***

The scientific method is the best way yet discovered for winnowing the truth from lies and delusion. The simple version looks something like this:

1. Observe some aspect of the universe.
2. Invent a tentative description, called a hypothesis, that is consistent with what you have observed.
3. Use the hypothesis to make predictions.
4. Test those predictions by experiments or further observations and modify the hypothesis in the light of your results.
5. Repeat steps 3 and 4 until there are no discrepancies between theory and experiment and/or observation.

When consistency is obtained the hypothesis becomes a theory and provides a coherent set of propositions which explain a class of phenomena. A theory is then a framework within which observations are explained and predictions are made.

The difference between science and the fuzzy subjects is that science requires reasoning while those other subjects merely require scholarship.

-- Robert Heinlein

Corollary—At the present time computers are incapable of reasoning!

PARADIGMS

Proposed by Thomas S. Kuhn in his book **The Structure of Scientific Revolutions**

PARADIGM--A set of assumptions, concepts, values, and practices that constitutes a way of viewing reality for the community that shares them, especially in an intellectual discipline.

A paradigm guides the whole group's research, and it is this criterion that most clearly proclaims a field a science .

Most science is done to further establish the limits of the paradigm and to improve its accuracy.

Paradigms and Crystallography

Crystallography is a mature science (95 years old) with a very well established paradigm!

In applying mature paradigms often the underlying basis for the paradigm is forgotten.

Agreement with the paradigm is an important criteria for validating research.

Story # 1

Hydrogen Atom Refinement

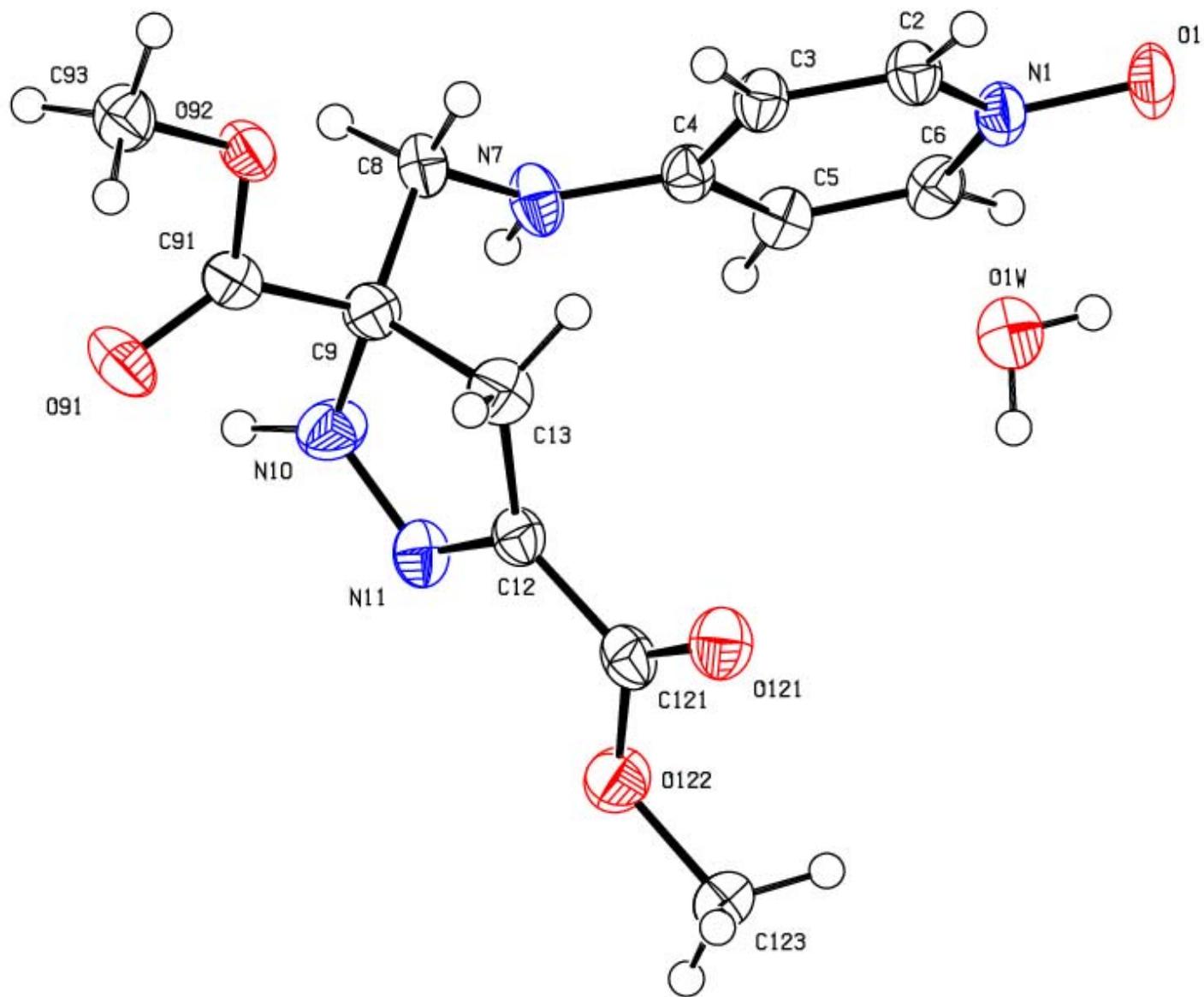
or

A Lesson in Expectation Values

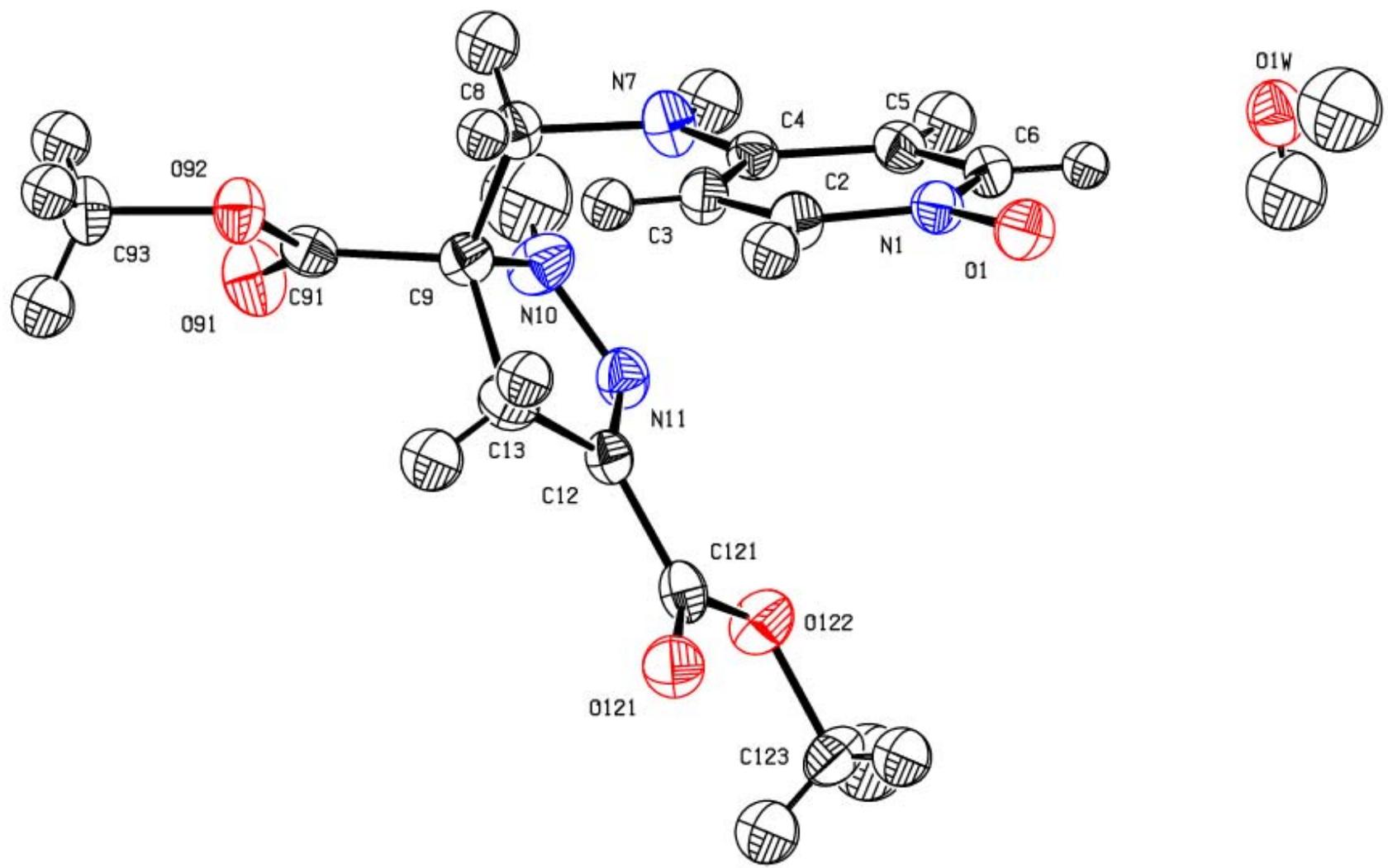
Organic compound sent provided by Dr. James Poole of Ball State University

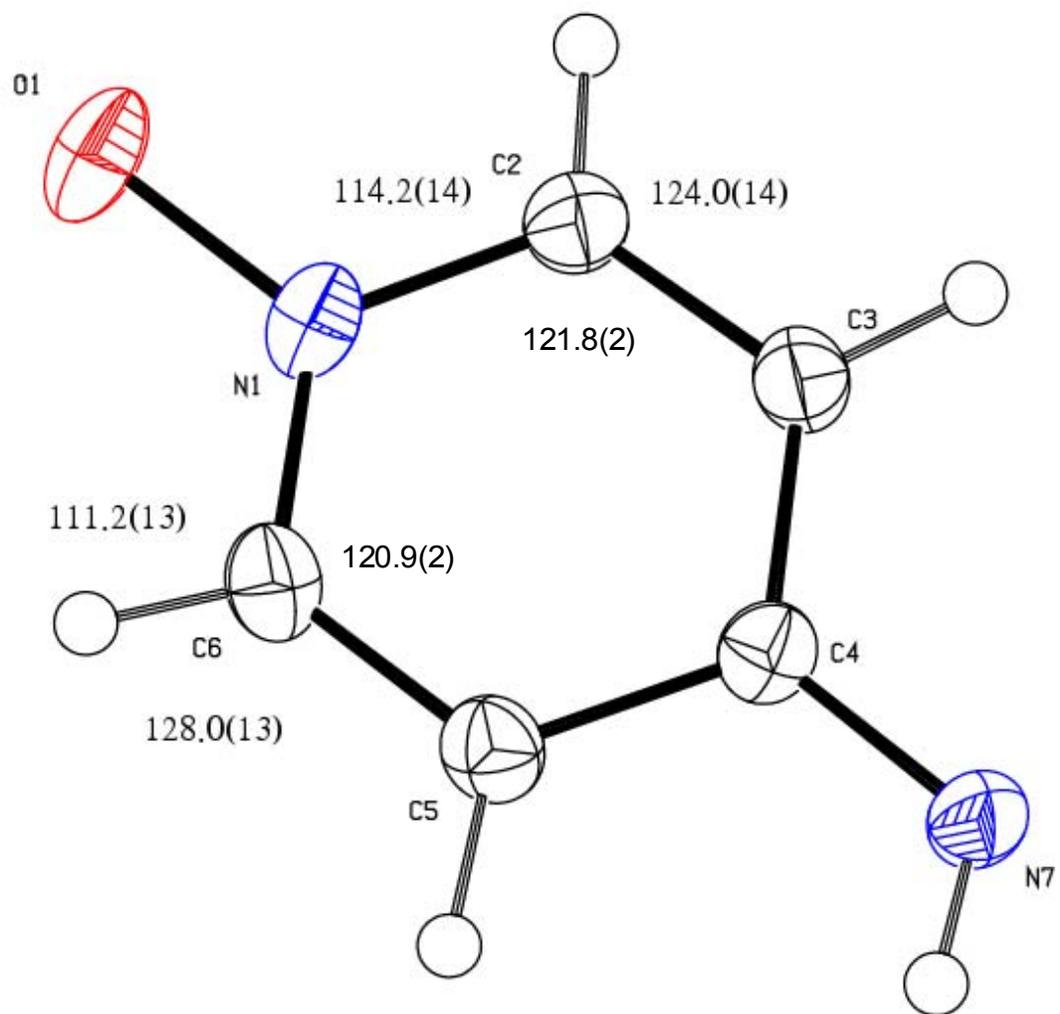
Compound was a minor product of a reaction and was basically uncharacterized.

In the case of poorly characterized organics the hydrogens are found in the Fourier map and refined isotropically.



I respect his desire to have some assurance that the H atoms are genuine, and the free refinement does provide pretty good evidence of this, but that's not the same as a final publishable result, it's just a step on the way. The poor data/parameter ratio suggests that the luxury of refining H atoms can't be afforded, and the bad geometry for some of the H atoms simply confirms this. I would insist on constraining the H atoms in a sensible way here (e.g. riding model for geometry and U values), and encourage the author to mention in the experimental text that the H atoms were initially located in a difference map and refined freely to confirm their correct assignment.





Compound	CSD Refcode	> N--C--H	>C--C--H
3-Methyl-4-nitropyridine-N-oxide	MNPYDO01	113.273	126.729
		114.851	122.424
4-Nitropyridine N-oxide	NTPYRO12	114.165	124.663
		115.065	124.077
This compound		114.2(14)	124.0(14)
		111.2(13)	128.0(13)

Structure Determined with neutron radiation.

Comment 1

There is nothing wrong with the riding model for adding hydrogen atoms or with the distances and adp's calculated by SHELX.

It is unreasonable to expect any refinement program to include deviations from the norm especially when most crystallographers are unaware of them

# of Hydrogen Atoms Refined	18	4
Data/Parameter (C2/c)	9.13	11.32
Shelx Weights	0.806, 0.00	0.737, 0.099
R(all and $>2\sigma$)	0.0641, 0.0456	0.0653, 0.0465
wR(all and $>2\sigma$)	0.1305, 0.1216	0.1337, 0.1252
Goodness of Fit	1.076	1.102
Average su of C--C bond	0.0032	0.0033

Comment 2

There are structures that are incorrect!

Richard L. Harlow, “Troublesome Crystal Structures”, Journal of Research of the NIST, (1996), 101, 327-339.

Structures are incorrect mainly because of bad decisions made during data collection and refinement.

Decisions/Evaluations

- 1. Selection of the crystal, mounting, centering***
- 2. Conditions—temperature, wavelength***
- 3. Indexing and determination of the Laue Class***
- 4. Data collection methodology***
- 5. Data reduction/integration***
- 6. Data Corrections***
- 7. Determination of the space group***

8. Structure Solution
9. Assignment of element type
10. Recognition and treatment of disorder
11. Refinement
12. Interpretation of the results

Twinning can effect many of the above.

We must fall back upon the old axiom that when all other contingencies fail, whatever remains, however improbable, must be the truth.

**Sherlock Holmes
The Adventure of Bruce-
Partington Plans**

Story #2

Why is this atom so large and elongated?

“X-rays don't see atoms”

David Watkin

Commission on Crystallographic
Computing Newsletter No. 4, August
2004

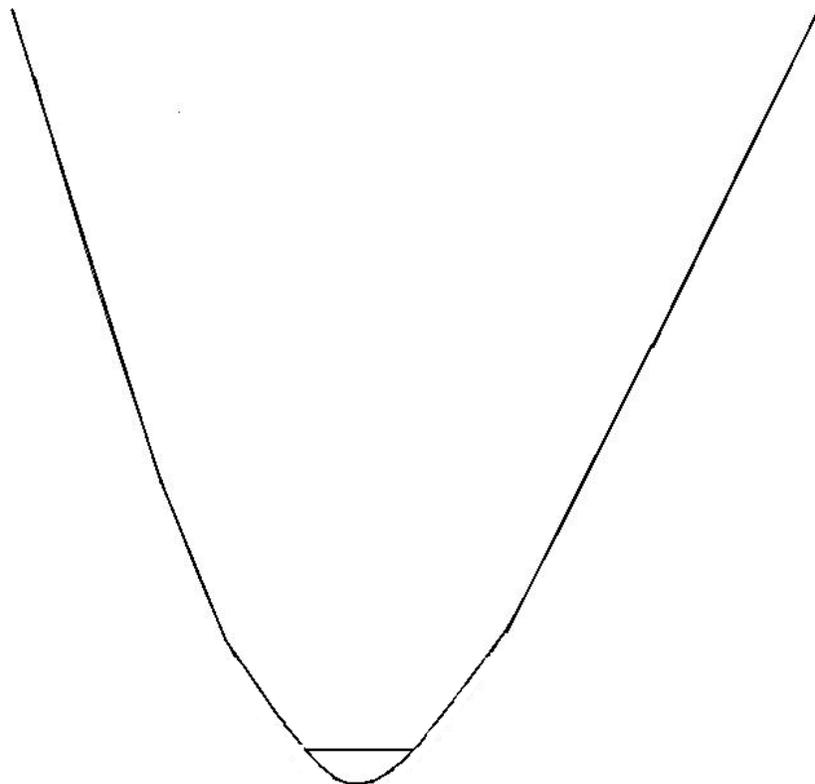
The third equation requires a much greater leap of confidence, yet is scarcely ever questioned by chemists using crystallography as an analytical tool (4). This equation is related to equation 2, except that the integration over a continuously varying periodic electron density has been replaced by a summation over a periodic array of atoms. The popularity of this model undoubtedly comes from the fact that it provides a very efficient representation of the electron distribution in the sample, and that experience has shown that this approximation serves well for the computation of other physical properties of materials.

$$F_{hkl} \approx \sum f_j e^{2\pi i(hx + ky + lz)} \quad (4)$$

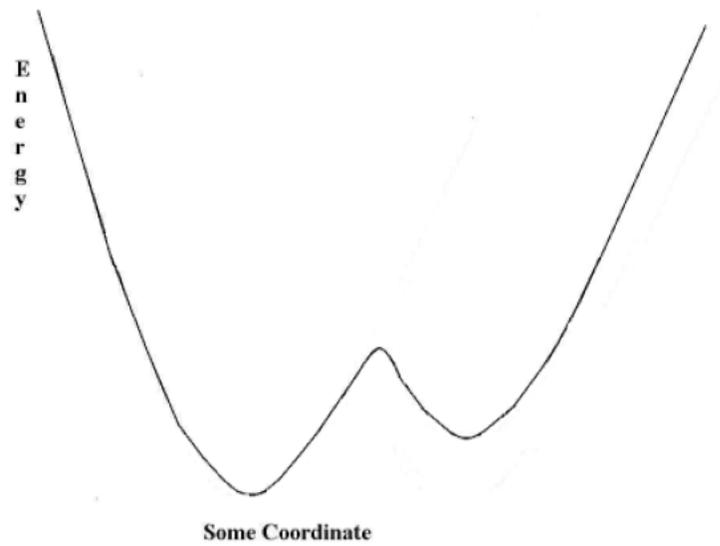
Luckily this tidy view of crystallography works well most of the time and accounts for the commanding role of X-ray structure determination as a definitive analytical tool. However, it can fail for a number of reasons.

Since the diffraction experiment works with samples containing many millions of unit cells, there is always the possibility of spatial inhomogeneity, leading to a diffraction effect which is also space-averaged.

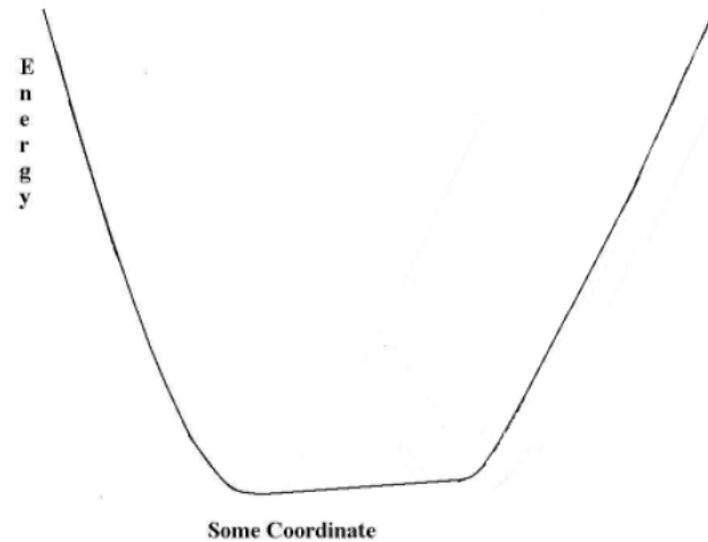
**E
n
e
r
g
y**



Some Coordinate



2



3

How to treat these disorders?

For case 2 there is no reason not to refine split atoms. Minimal restraints should be required and refinement should proceed normally.

Case 3 is not as obvious.

I prefer to use the large ellipses. Others tend to use multiple split atoms and restraints.

Split Atom Refinement

1. Split the atom as suggested by SHELX and refine isotropically without restraints.

2. Are distances OK?

a. Yes continue

b. One returns to old position another moves out

3. Are adp's OK? Use free variable to adjust occupancy.

4. After convergence try anisotropic refinement.

Light restraints may be required to keep distances acceptable.

More Extreme Problems

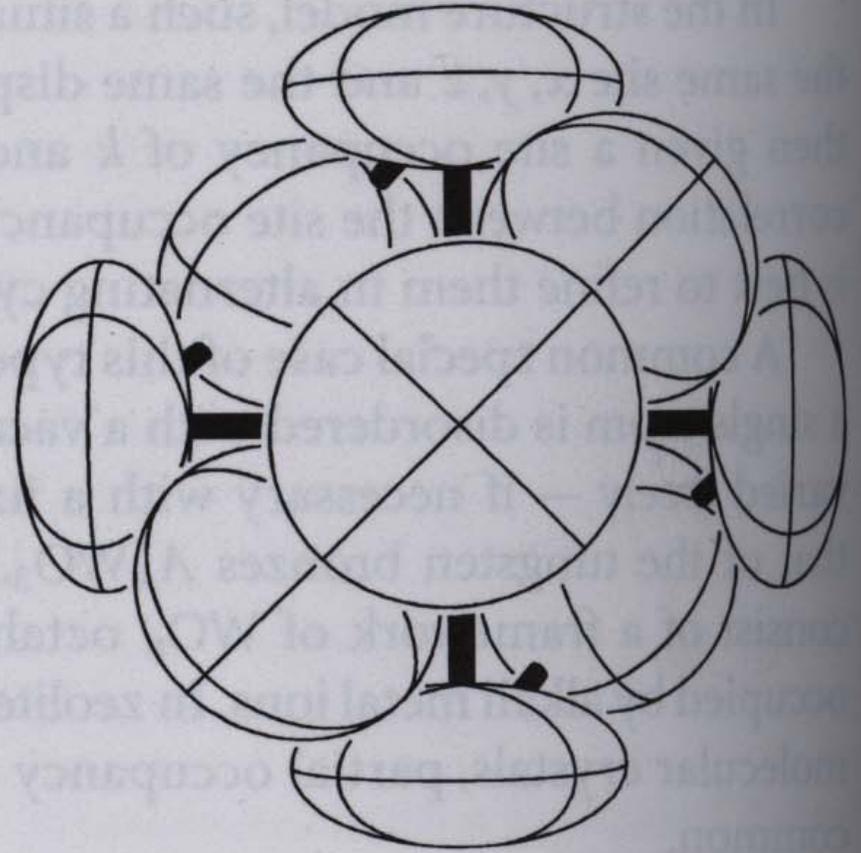


Fig. 10.1. Representation of a disordered BF_4 -anion using 10 partially occupied F-sites.

More Extreme Methods

PLATON Squeeze option.

For disordered fragments such as solvents

Produces a new .hkl file which removes the residual density

One problem--the complete contents of the unit cell are not observed. How to report the formula?

CRYSTALS

Refinement of electron density in a hollow shell.

No longer can "see" individual atoms.

A Final Comment

Q8: What should I do about 'may be split' warnings?

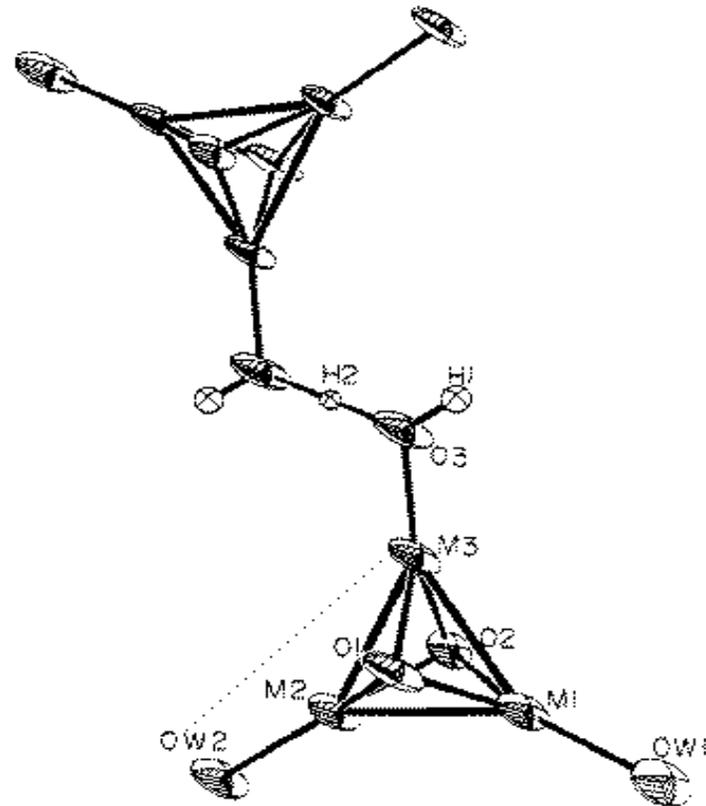
A: Probably nothing. The program prints out this warning whenever it might be possible to interpret the anisotropic displacement of an atom in terms of two discrete sites. Such atoms should be checked (e.g. with the help of an ORTEP plot) but in many cases the single-site anisotropic description is still eminently suitable.

SHELX FAQ

Story #3

***How to correct for absorption
or
All absorbed in my work***

Effects of Absorption



Spherical Correction

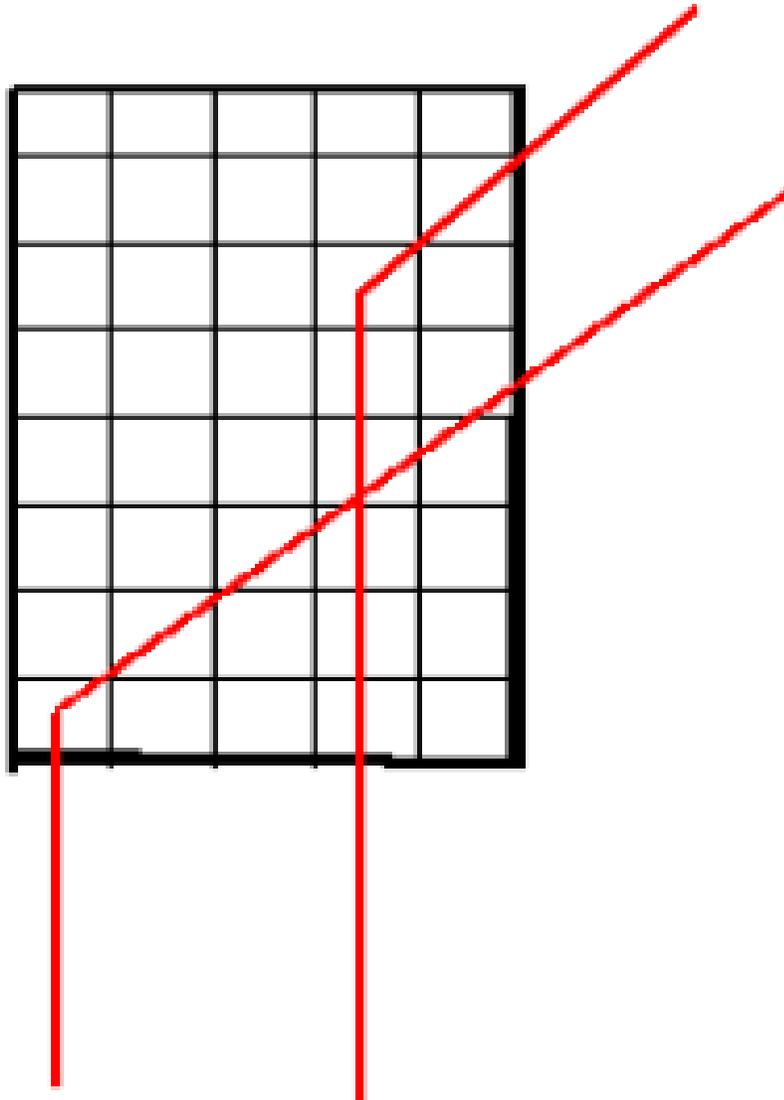
INPUT: absorption coefficient μ , crystal radius, data with θ angle

OUTPUT: transmission factors for each datum

ADVANTAGES: exact calculation

DISADVANTAGES: few spherical crystal samples

Numerical Correction



INPUT: absorption coefficient μ , data
with direction cosines,
description of the
crystal by bounding faces

OUTPUT: transmission factors for each
datum

ADVANTAGES: exact calculation

DISADVANTAGES: crystal must have
well defined faces

Semi-Empirical Corrections

INPUT: data with direction cosines or Eulerian angles, equivalent reflection intensities

OUTPUT: relative corrections

ADVANTAGES: no crystal description required

DISADVANTAGES: much cruder than previous methods; may correct for systematic errors

A Problem

The transmission-factor limits
_exptl_absorpt_correction_T_min and _max should agree
with those expected for the crystal shape and size and μ .
Acta Cryst. C., Notes for Authors, 2006.

Since semi-empirical methods provide no values for the
transmission factors they need to be approximated.

Estimating T

The IUCr Method

$$T_{\max} = \exp(-d_{\min} * \mu)$$

$$T_{\min} = T_{\max} * \text{cor}_{\min}^2 / \text{cor}_{\max}^2$$

The Spherical Method

Determine %T for theta assuming a sphere with an appropriate radius.

Correct the value using the correction factor.

All Hell Breaks Out

Before using DIFABS, it is suggested that you read the following journal article:

Nigel Walker and David Stuart, "An Empirical Method for Correcting Diffractometer Data for Absorption Effects", **Acta Cryst.** (1983), **A39**, 158-166.

DIFABS has received a large amount of bad verbage as a practical method ("DIFABS is EVIL!"). How much of this is based in reality, paranoia, philosophy, etc is still up for discussion(?). It is possible for DIFABS to be used to soak up sloppy practises in diffractometer setup and crystal alignment (non-full irradiation of the crystal), etc, etc, etc.

A web page to refer to is "Should DIFABS be Banned?":

<http://www.unige.ch/crystal/stxnews/stx/discuss/dis-dif2>

Within the limitations described in the original DIFABS paper above, it does seem(?) to be a method worth keeping in your crystallographic bag of tricks to get the job done under some circumstances.

<http://www.ccp14.ac.uk/tutorial/wingx/absorp/difabs.htm>

What is Difabs?

The idea is that the averaging of multiple independent measurements of a reflection allow the true value of F_o to be determined.

The idea of DIFABS is that F_c is an equivalent determination of the true value of F_o . F_c must be calculated for isotropic atoms with all atoms in the cell refined.

Then do a calculation to determine the absorption surface.

Is DIFABS valid?

The question--Is it correct to use the data calculated from a model to adjust the data on which the model is refined?

I have no idea as it is way outside my area of expertise.

As DIFABS was published in **ACTA CRYST. A** the referees who reviewed it should have been able to determine this.

How Does Difabs Compare

At an ACA meeting Ton Spek compared the calculated absorption surfaces from a numeric, a semi-empirical, and a difabs correction.

They all had very similar features and produce similar results.

It is a capital mistake to theorize before one has data. Insensibly one begins to twist facts to suit theories, instead of theories to suit facts.

Sherlock Holmes

“A Scandal in
Bohemia”

Cottingley Fairies

