JOURNAL OF
APPLIED
CRYSTALLOGRAPHY

Volume 53 (2020)

Supporting information for article:

From Atoms to Bonds, Angles, and Torsions: Molecular Metrics from Crystal Space - and Two Excel Implementations

Leslie Glasser

Excel Information and Explanations

The Excel workbooks are set up so that only the relevant constants need to be entered, and results are generated automatically. Users may thus introduce their own molecular systems. Note that angles for trigonometric functions must be in *radians*. The Excel conversion factor to convert radians to *degrees* is "180/pi()". Final results are displayed in **bold characters**.

Descriptions of the Excel Supplementary Files and their Application

Dunitz.xls

The users apply the supplied data (cell constants and fractional coordinates) contained in the left-most column, or replace them with their own data. On left-clicking on the central button, labelled in blue with "Left-click Button to run macro", a message box appears in which the user enters a sequence of the atom reference numbers of up to four connected atoms (to calculate torsion angle, the two internal bond angles, and the three bond distances) , such as "12,1" or "12, 1 ,2 ,3".

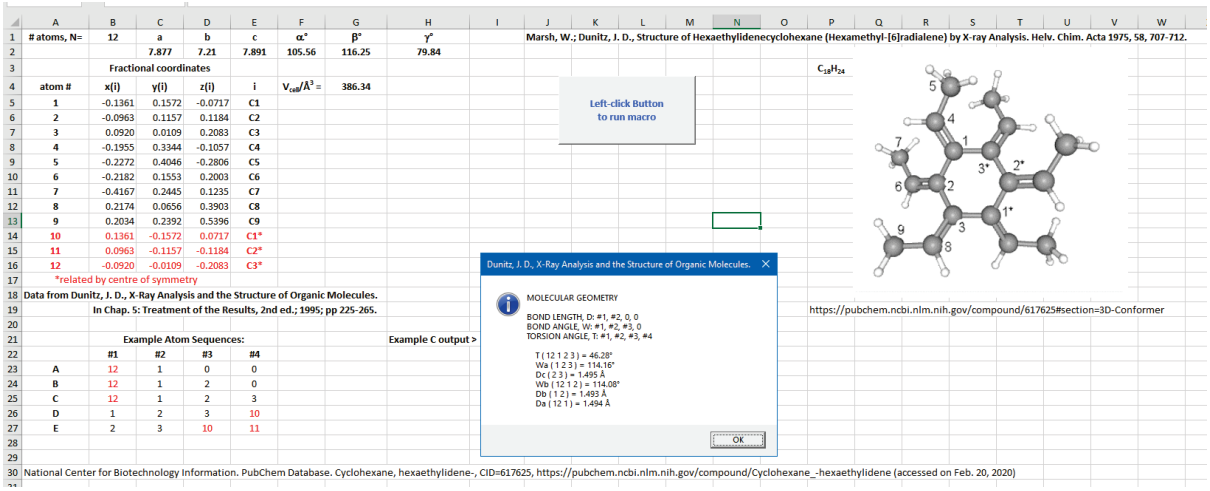

Figure S1: Image of the screen of the macro-enabled Excel file, Dunitz.xls. The data displayed refers to the structure of hexaethylidenecyclohexane.[1]

The left columns contain the crystal constants and atom fractional coordinates of hexaethylidenecyclohexane while a structural image appears on the right. On the lower left is a list of sample atom sequences which may be entered in order to calculate the molecular geometry from

the data supplied. An example of the output obtained from data entry appears in the “Molecular Geometry” box in the base of the worksheet.

MolGeom.xlsx

This workbook contains four worksheets:

(1) *SF6:* Calculations of bond lengths and angles using fractional coordinates.

- Cell constants are listed on **Row 2**, and the corresponding metric constants below **cell B17**.
- Atom fractional coordinates are entered in **columns D, E, and F**.
- On the right-hand side, outlined in green, is a metric matrix calculation for bond lengths and angles.

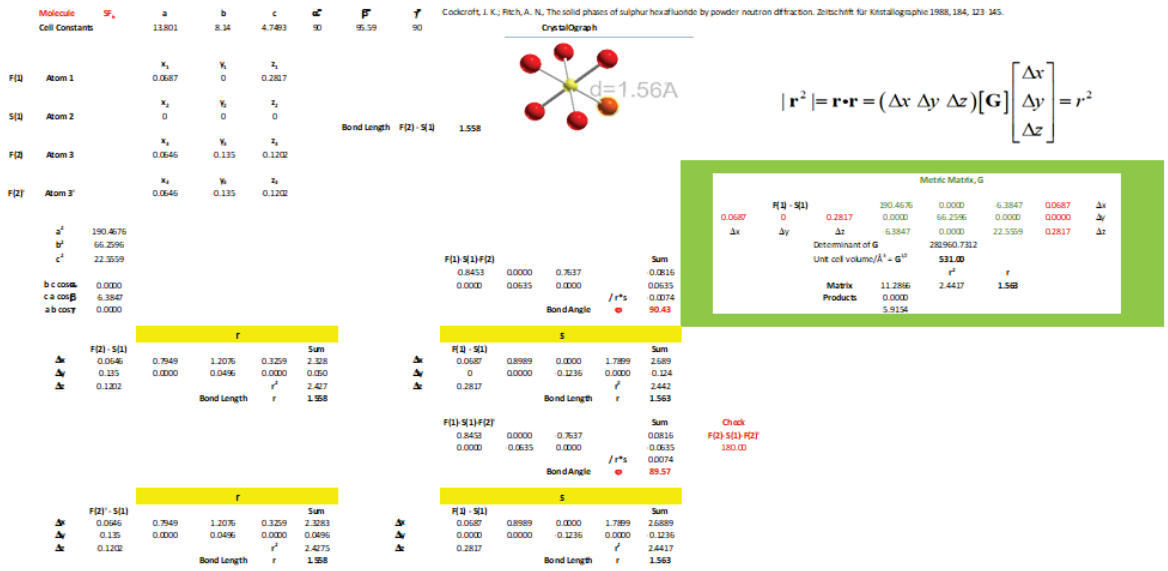

Figure S2: Image of the *SF6* screen of the Excel file, MolGeom.xlsx. The data displayed refers to the structure of sulfur hexafluoride.[2]

The upper left columns contain the crystal constants and atom fractional coordinates of four bonded atoms of sulfur hexafluoride while a structural image appears in the centre. To the right is the vector dot-product formula (eq. 5c) for the square of a vector length and, in the box below, the metric

matrix, **G**, live-calculated for the structural data in the upper left-hand columns. An example bond-length calculated with **G** according to the matrix equation (5c) above the box appears within the box.

In the lower rows, specific bond lengths and bond angles are live-calculated for eq(5c) from the algebraic terms in the left column: a^2, b^2, c^2, $bc\ cos\ \alpha$, $ca\ cos\ \beta$, $ab\ cos\ \gamma$.

(2 and 3) *Serine and Valine:* these are example worksheets of calculations in Cartesian coordinates and are set up identically, yielding bond lengths, bond angles, and torsion angles.

- **Columns A-E** contain a copy of the relevant *cif* file, with cell constants, fractional coordinates and atom identifiers. These are delimited by a **yellow column F**.
- On the right, **Row 2** contains the cell constants copied from the *cif* file.
- **Cells I5-K7** contains the computed matrix for transformation from fractional to Cartesian coordinates.
- Also to the right of the yellow column, the cells in **Rows 10-12** contain the fractional coordinates, copied and transposed from the *cif* file.
- **Rows 14-16** contain the corresponding computed Cartesian coordinates.
- Selected computed bonds and angles appear in the subsequent rows.

A Jmol image of the molecule showing metrics has been pasted into the worksheet. In the case of Serine, the image on the right is a view down the central C-C bond to highlight the right-hand twist of the O-C bond towards the C-N bond.

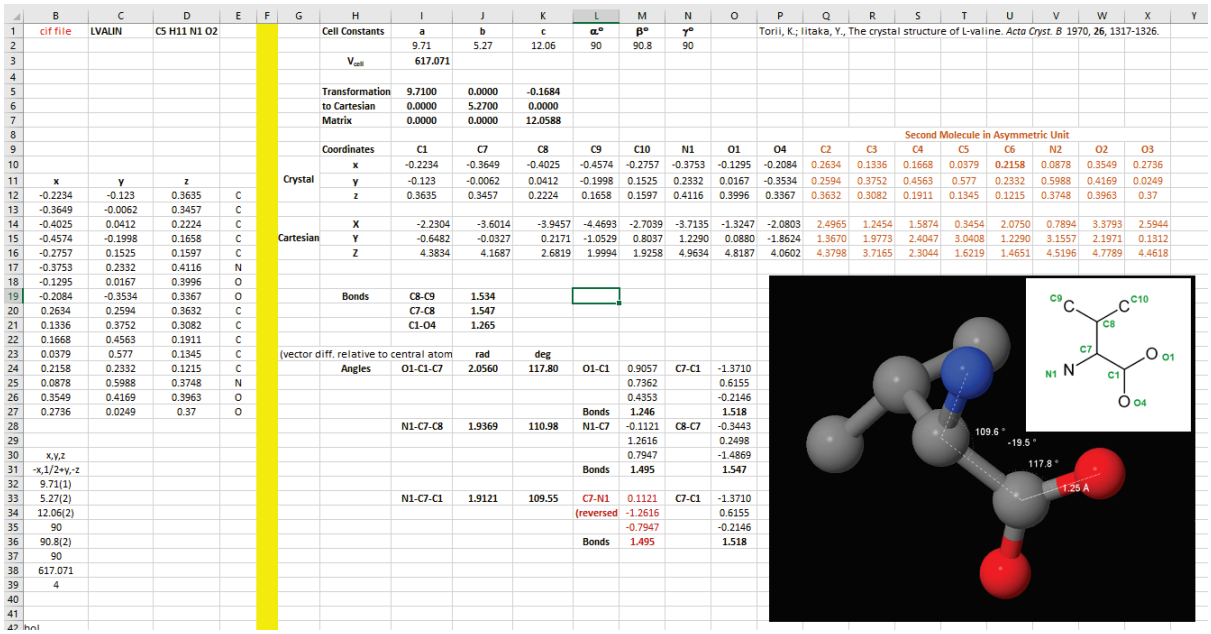

Figure S3: Image of the *Valine* screen of the Excel file, MolGeom.xlsx. The data displayed refers to the structure of L-Valine.[3]

The left columns contain data from the LVALIN.cif file from CCDC.[4] On the right, below the cell constants and the transformation matrix from fractional to Cartesian coordinates, are listed the transposed atom fractional coordinates, the live-calculated Cartesian coordinates, and some example algebraic live-calculated bond lengths and bond angles.

The Jmol images show a selected set of bond lengths, bond angles, and torsion angle.

(4) *Torsion*

This worksheet calculates torsion angles for sets of four atoms with Cartesian coordinates as previously computed in worksheets similar to Serine or Valine.

- The computed Cartesian coordinates of the four linked atoms are entered in sequence in **cells F2-I4**.
- The required bond lengths and angles are computed and the magnitude of the central torsion angle relating the four atoms appears in **cell O14**.

- The sign of the torsion angle (between -180° and +180°) is assigned using the sign of the value in **cell S17**.
- The worksheet has been initialised in the **cells F2-I4** with coordinate data (in red) for L-serine. Other coordinate data sets, which can simply be copied and pasted over the active data set, have been provided in the rows below the initial set: black for a 90° torsion (select + or - for cell value for Z4); red for L-serine; blue for L-valine; and green for an arbitrary demonstration data set from the literature.
- Below the green line, the metric matrix angle calculation in Cartesian space is demonstrated for the atom triplet O3-C3-C2. The atom coordinates of O3 and C2 are adjusted by difference to refer to the central atom C3 (in red). The unit metric matrix for the Cartesian coordinate system has one's along the diagonal, and zero's for all the other entries since cos 90° = 0.
- The various headings and annotations may assist in illuminating the calculations as described in the main text.
- Jmol images of the molecules showing metrics have been pasted into the worksheet.
- The numeric data is displayed with four decimal figures, but is not truncated in the calculations.

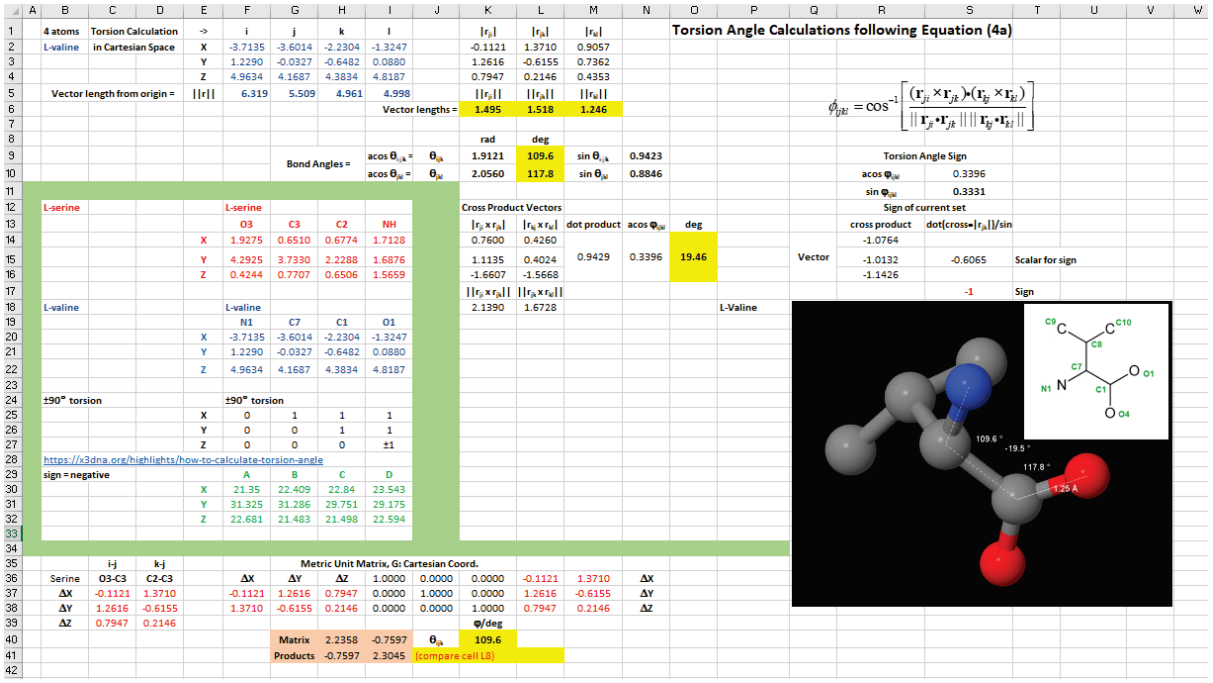

Figure S4: Image of the *Torsion* screen of the Excel file, MolGeom.xlsx. The data displayed refer to the structure of L-Valine.[3]

Upper part of screen: The box on the left contains four data sets, each consisting of a sequence of four atoms. A set may be copied over the data at the top of the column in order to perform the live torsion calculation. Correspondingly, the user may insert his/her data for a sequence of four atoms over the data at the top of the column. Final computed values are highlighted with a yellow background.

In the central column, the three bonding vectors, $\mathbf{r}_{ji}$, $\mathbf{r}_{jk}$, $\mathbf{r}_{kl}$, and two bond angles, θ_{ijk} and θ_{jkl}, are calculated. Below them, the cross product vectors, $\mathbf{r}_{ji} \times \mathbf{r}_{jk}$ and $\mathbf{r}_{kj} \times \mathbf{r}_{lk}$, are calculated using the data above, followed by their dot product - which yields the torsion angle ϕ_{ijkl}, following equation (4a). The torsion angle sign is determined from the dot product of the cross product by the central bond vector $\mathbf{r}_{jk}$, that is $(\mathbf{r}_{kj} \times \mathbf{r}_{lk}) \bullet \mathbf{r}_{jk}$.

Lower part of screen: Below the bar, the Cartesian metric matrix, a 3×3 unit matrix, is used to provide an alternate method by which to calculate bond angles.

On the right, a Jmol image of L-Valine depicts bond distances, bond angles, and the torsion angle around the C-C bond.

An Excel Vector Cross-Product Function

In Excel, the following user-defined VBA Array function yields a cross-product of two vectors. This should be added as a Function into a worksheet module of a macro-enabled workbook; it is invoked by typing the expression below into a cell and selecting the two desired 3-cell ranges in rows or columns. It produces a *3-cell row* of values, one value for each row in the Function:

= vCP(3-cell range#1, 3-cell range#2); for example: =vCP(B1:D1,D5:F5)

```
Function vCP(v1 As Variant, v2 As Variant) As Variant
vCP = Array(v1(2) * v2(3) - v1(3) * v2(2), _
v1(3) * v2(1) - v1(1) * v2(3), _
v1(1) * v2(2) - v1(2) * v2(1))
End Function
```

If a *3-cell column* output is required, the following Function will produce the transpose:

```
Function CPv(v1 As Variant, v2 As Variant) As Variant
CPv = Application.Transpose(Array(v1(2) * v2(3) - v1(3) * v2(2), _
v1(3) * v2(1) - v1(1) * v2(3), _
v1(1) * v2(2) - v1(2) * v2(1)))
End Function
```


1. Marsh, W.; Dunitz, J. D., Structure of Hexaethylidenecyclohexane (Hexamethyl-[6]radialene) by X-ray Analysis. *Helv. Chim. Acta* **1975,** *58*, 707-712.
2. Cockcroft, J. K.; Fitch, A. N., The solid phases of sulphur hexafluoride by powder neutron diffraction. *Zeits. Kristallogr.* **1988,** *184*, 123-145.
3. Torii, K.; Iitaka, Y., The crystal structure of L-valine. *Acta Crystallogr. B* **1970,** *26*, 1317-1326.
4. CSD Cambridge Structural Database. http://www.ccdc.cam.ac.uk/products/csd/ (accessed April, 2020).