

DHA, a program for calculating hydrogen atom coordinates of solvent atom involved in hydrogen bonding interactions

B. Shivachev

*Institute of Mineralogy and Crystallography, Bulgarian Academy of Sciences,
“Acad. Georgi Bonchev” str. building 107, Sofia 1113, Bulgaria*

Received February, 2013; Revised May, 2013

Generally it is difficult to locate hydrogen atoms accurately using X-ray data because of their low scattering power and lack of core electrons. For many purposes the hydrogen positions can be calculated according to well-established geometrical criteria. Nevertheless, in some cases the positioning of the hydrogen atoms is tricky. Usually the problem arises when solvent molecules (e.g. water, OH groups or ammonia) are present in the crystal structure. Fortunately, hydrogen bonding interaction are characterized by an appropriate geometry and distance between the donor (**D**) and acceptor (**A**) atoms and requires the participation/presence of a hydrogen atom (**D** – H ••• **A**). Thus, if the coordinates of the donor and acceptor atoms are known one can calculate the coordinates of the “elusive” hydrogen atom. The program DHA uses an extremely simple mathematical apparatus to produce the coordinates of the hydrogen atom (**H**) participating in a **D** – H ••• **A** type of interaction.

Key words:

INTRODUCTION

Nowadays, single crystal X-ray data collection and structure solution has become a very accessible, popular and routine experimental technique. Interestingly, the most popular software, ShelX [1] used for structure solution and refinement has remained almost untouched since 1993, although 1997 and 2013 versions are available (Shelx, 2013). Various graphical user interfaces (GUI) were designed and implemented over the years to circumvent the “tedious” text typing, editing and manipulation of the input (*.ins) and output (*.res) files of Shelx programs [2, 3]. Certainly, several other free and commercial [4–6] software packages are available to the end-user but as the standard is set by ShelX most of the features are similar to those implemented in ShelX. One of the very few elusive options in those programs is the positioning of H atoms when suitable geometrical criteria fail and electron density from difference Fourier is not available. Usually the problem arises for solvent (water, OH, NH₃ etc.) atoms/molecules that are present in the crystal structure. However, if suitable hydrogen

bonding geometry and distance (directionality) between such atoms (**D** ••• **A**) is detected the hydrogen atom should also be present. The DHA program targets the calculation of hydrogen atoms’ coordinates when hydrogen bonding interaction criteria are encountered but the usual positioning of the hydrogen atom is either impossible or inaccurate.

METHODS

The mathematical apparatus required for the calculation of the H atoms coordinates is extremely simple. The distance D_{12} between two atoms A1 (with coordinates $x1, y1, z1$) and A2 (coordinates $x2, y2, z2$) can be calculated as:

$$D_{12} = \sqrt{(x1 - x2)^2 + (y1 - y2)^2 + (z1 - z2)^2} \quad (1)$$

The coordinates of the atom A_{mid} lying exactly in the middle between A1 and A2 are easily obtained:

$$x_{mid} = \frac{x1 + x2}{2}, y_{mid} = \frac{y1 + y2}{2}, z_{mid} = \frac{z1 + z2}{2} \quad (2)$$

The distance between A1 and A_{mid} (or A_{mid} and A2) can be obtained according to eq. 1. Using the

* To whom all correspondence should be sent:
E-mail: blshivachev@gmail.com

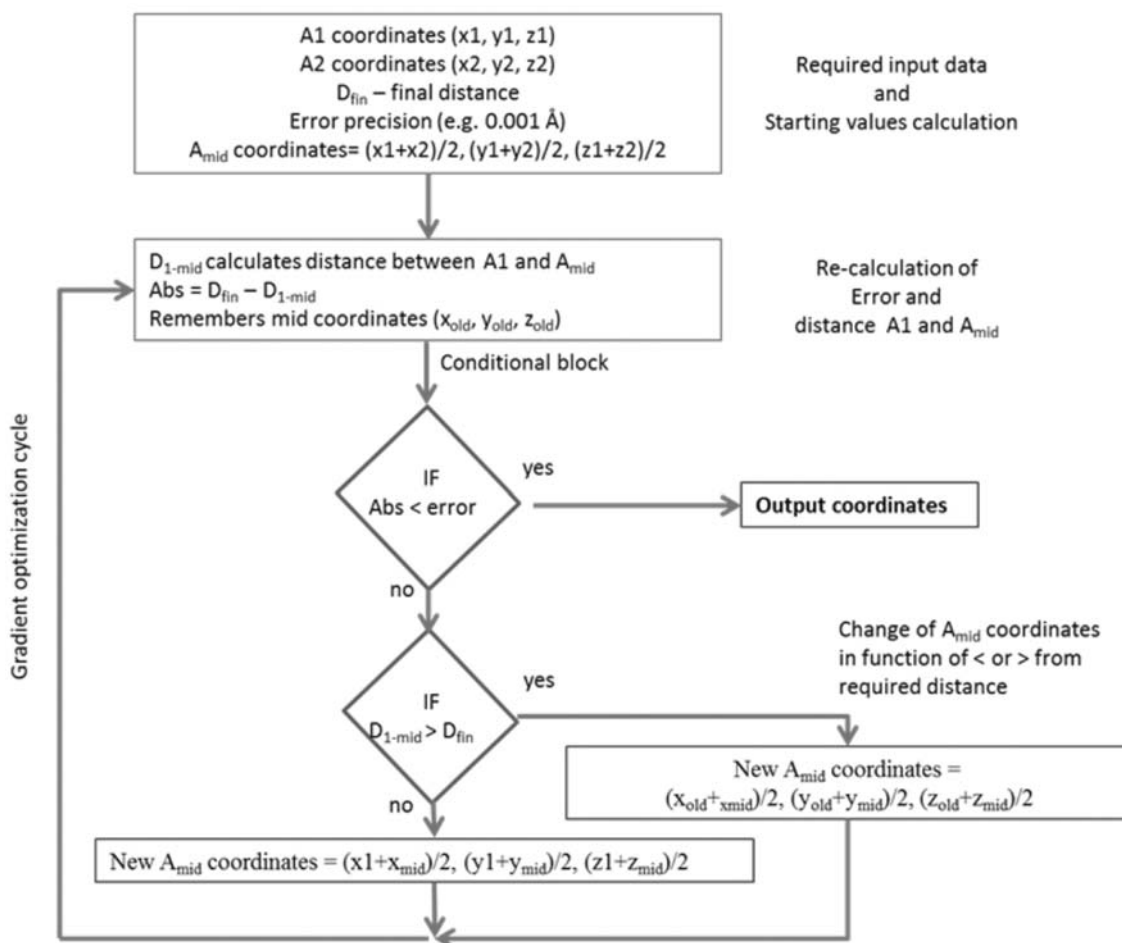


Fig. 1. Scheme of the algorithm of DHA program

coordinate descent method [7] one can obtain the coordinates (within a predetermined error) of an atom located at the required distance from A1 (A1 -- A_{end}). The logical algorithm of the program is shown on Figure 1.

Hardware environment. The program is intended to run on personal computers and has minimal requirements for disk space and memory.

Software environment. The program is written in Visual Basic (VB) and utilizes the Dot.NET technology allowing it to be run on different operating systems.

For simplicity all of the calculations are carried out in Cartesian coordinate system. The conversion of fractional to Cartesian coordinates and backwards is performed as follows:

$$\begin{bmatrix} x \\ y \\ z \end{bmatrix} = \begin{bmatrix} a & b \cos(\gamma) & c \cos(\beta) \\ 0 & b \sin(\gamma) & c \frac{\cos(\alpha) - \cos(\beta) \cos(\gamma)}{\sin(\gamma)} \\ 0 & 0 & \frac{cv}{\sin(\gamma)} \end{bmatrix} \begin{bmatrix} a \\ b \\ c \end{bmatrix} \quad (3)$$

$$\begin{bmatrix} a \\ b \\ c \end{bmatrix} = \begin{bmatrix} \frac{1}{a} & -\frac{\cos(\gamma)}{a \sin(\gamma)} & \frac{\cos(\alpha) \cos(\gamma) - \cos(\beta)}{a v \sin(\gamma)} \\ 0 & \frac{1}{b \sin(\gamma)} & \frac{\cos(\beta) \cos(\gamma) - \cos(\alpha)}{b v \sin(\gamma)} \\ 0 & 0 & \frac{\sin(\gamma)}{cv} \end{bmatrix} \begin{bmatrix} x \\ y \\ z \end{bmatrix} \quad (4)$$

$$v = \sqrt{1 - \cos^2(\alpha) - \cos^2(\beta) - \cos^2(\gamma) + 2\cos(\alpha)\cos(\beta)\cos(\gamma)} \quad (5)$$

The assumption is that both coordinate system share the same origin, the *a*-axis is collinear with the *x*-axis, and the *b*-axis lies in the *xy*-plane.

INPUT DATA

The program requires the following data in order to return the Hydrogen coordinates (Fig. 2):

1. Unit cell parameters *a*, *b*, *c*, α , β , γ in Å,
2. Fractional coordinates of the donor (*x*₁, *y*₁, *z*₁),
3. Fractional coordinates of the acceptor atom (*x*₂, *y*₂, *z*₂),
4. Optionally a DH, donor-hydrogen distance (0.96 Å is pre-imputed).

OUTPUT DATA

The program outputs the Cartesian and fractional coordinates of an atom located at DH (Å) from the donor atom.

DISCUSSION

The choice of a suitable (default) O-H distance is very difficult. Internuclear O–H distances for isolated molecules in the gas phase are about 0.96 Å

(e.g. 1.10 Å for C-H). However, the appropriate distance to use for X-ray diffraction must be appreciably shorter to allow displacement of the center of gravity of the electron distribution towards the oxygen atom, and also for oscillating effects. Fortunately SHELXL-97 provides a selection of options (AFIX and HFIX) for positioning and refining hydrogen atoms from geometrical criteria (ShelX manual [1]). For routine refinement, and when suitable geometrical criteria are available (e.g. tertiary CH (HFIX 13), secondary CH₂ (HFIX 23), ethylenic CH₂ (HFIX 93), acetylenic CH (HFIX 163), BH in polyhedral boranes (HFIX 153) and aromatic CH or amide NH (HFIX 43) etc.), the choices of hydrogen atoms generation are multiple (through combination of AFIX and HFIX instructions). The hydrogen coordinates are re-idealized before each cycle, and “ride” on the atoms to which they are attached (i.e. the coordinate shifts are the same for both the heavy atom and riding H atom). Hydroxyl and methyl groups are more difficult to position accurately but if good data are available (e.g. low-temperature data) the method of choice is HFIX 137 for CH₃ and HFIX 147 for OH groups (torsion angles are allowed to refine while keeping the X–H distance and Y–X–H angle fixed).

OMIT instruction combined with L.S. 0, FMAP 2 (e.g. $F_o - F_c$) and PLAN-100 (list the coordinates and geometry of the highest 100 non-zero electron

The screenshot shows the 'Calculates H coordinates for D-H...A' window. It is divided into two main sections: 'Input data' and 'Output data'.

Input data section:

- Unit cell parameters:** alpha (100.114), beta (94.421), gamma (100.524); a (7.9438), b (14.225), c (14.895).
- Donor atom coordinates (fractional):** x₁ (0.478687), y₁ (0.460693), z₁ (0.367682).
- Acceptor atom coordinates (fractional):** x₂ (0.513437), y₂ (0.365998), z₂ (0.361328).
- Input desired donor H distance:** 0.96

Output data section:

- calculated H atom coordinates (fractional):** x (0.502311), y (0.396317), z (0.363362).
- D ... A distance is:** 1.412357
- D-H distance:** 0.960151

At the bottom, there are 'Calculate' and 'Exit' buttons, and a citation: 'Please cite: Bul.Chem.Comm. 2013 issue4, page. xx' and 'BLSv01'.

Fig. 2. Input/output window of DHA program

densities) enables an “omit map” to be calculated, which is a convenient way of checking whether there are actually electron density peaks close to the calculated atom positions (the H atom should not be already present). In this omit map, if a non-zero electron density (peak) appears in the Fourier output, then there is an actual peak in the difference electron density synthesis that can be assigned to a hydrogen position (AFIX 03).

In the case of isolated solvent molecules (water, OH, NH₃, etc.) no suitable geometric criteria are applicable and thus the location of the H atom(s) depends only on the data quality. Unfortunately, solvent molecules are usually far more mobile than the other atoms of the structure and thus their isotropic *U*'s are bigger. Therefore, the difference Fourier map may not allow the location of the solvent H atoms. Usually the solvent molecules that are present in the structures participate in hydrogen bonding interactions that stabilize the three-dimensional crystal packing. Thus if suitable hydrogen bonding geometry and distance (directionality) involving solvent (donor and/or acceptor) is detected, the hydrogen atom should also be present. Assuming that the D–H distance is known one can calculate the H atom coordinates from the ones of **D** and **A** (D ••• A).

The program has been used for the determination of solvent H atom coordinates for the structures cited in refs. 8 and 9.

CONCLUSION

The DHA program calculates hydrogen atoms coordinates when the usual positioning of

the hydrogen atom is either impossible or inaccurate but hydrogen bonding interaction criteria are encountered.

Availability: The program is available free of charge and can be downloaded from the www.imc.bas.bg/dha site or supplied upon request.

Acknowledgments: *The financial support of Bulgarian Science Fund grant DRNF 02/1 is gratefully acknowledged.*

REFERENCES

1. G. M. Sheldrick, *Acta Cryst. A*, **64**, 112 (2008).
2. L. J. Farrugia, *J. Appl. Cryst.*, **32**, 837 (1999).
3. O. V. Dolomanov, L. J. Bourhis, R. J. Gildea, J. A. K. Howard, H. Puschmann, *J. Appl. Cryst.*, **42**, 339 (2009).
4. M. C. Burla, R. Caliendo, M. Camalli, B. Carrozzini, G. L. Casciarano, C. Giacovazzo, M. Mallamo, A. Mazzone, G. Polidori and R. Spagna, *J. Appl. Cryst.*, **45**, 357 (2012).
5. P. T. Beurskens, G. Beurskens, R. de Gelder, S. Garcia-Granda, R. O. Gould, J. M. M. Smits, The DIRDIF2008 program system, Crystallography Laboratory, University of Nijmegen, The Netherlands, 2008.
6. L. Palatinus and G. Chapuis, *J. Appl. Cryst.*, **40**, 786 (2007).
7. P. Tseng, *J. Optim. Theory Appl.*, **109**, 475 (2001).
8. L. Dimowa, K. Koshev, H. I. Sbirikova, R. P. Nikolova, B. L. Shivachev, *Bul. Chem. Comm.*, **45**(4), 505 (2013).
9. K. Koshev, L. Tsvetanova, L. Dimowa, R. Nikolova, B. L. Shivachev, *Bul. Chem. Comm.*, **45**(4), 543 (2013).

**ДНА, ПРОГРАМА ЗА ИЗЧИСЛЯВАНЕ НА КООРДИНАТИТЕ
НА ВОДОРОДЕН АТОМ КЪМ АТОМ РАЗТВОРИТЕЛ,
УЧАСТВАЩ В ВОДОРОДНА ВРЪЗКА**

Б. Шивачев

*Институт по минералогия и кристалография – Акад. „Иван Костов“,
Българска Академия на Науките, ул. „Акад. Георги Бончев“,
блок 107, София 1113, България*

Постъпила февруари, 2013 г.; приета май, 2013 г.

(Резюме)

При рентгеноструктурния анализ е сравнително трудно да се позиционират водородни атоми на база експериментални данни поради малкия брой електрони (в действителност един) а оттам и ниското им разсейване на рентгеновите лъчи. В много случаи водородните координати (позиции) могат да се изчислят въз основа на утвърдени геометрични критерии. Все пак, в някои случаи на позициониране на водородните атоми е трудно. Обикновено такъв проблем възниква, когато молекули на разтворителя (например вода, ОН групи, амоняк и др.) са част от кристалната структура. За щастие, водородните атоми участват във водородни връзки (взаимодействия), които се характеризират с подходяща геометрия и разстоянието между донора (**D**) и акцептора (**A**) и изискват наличие на водороден атом (**D–H•••A**). По този начин, ако координатите на атомите на донора и акцептора са известни, може да се изчислят координатите на „липсващият“ водороден атом. Програмата **ДНА** използва изключително прост математически апарат за изчисляване на координатите на водороден атом (**H**), участващ във взаимодействия от типа **D–H•••A**.