

NEIGHCRYS

Program release 2.0.8

Manual release July 2013

A utility program for setting up input files for DMACRYS, which models crystals of rigid organic molecules using anisotropic atom-atom intermolecular potentials.

For the online manual for DMACRYS, visit this web site:

<http://www.chem.ucl.ac.uk/cposs/dmacrys/manuals/dmacrys.pdf>

Contents

1. Introduction	2
2. Atom labelling convention	2
3. Input files required.....	3
3.1. shelx - crystallographic data file	3
3.2. bondlengths - covalent bonds specification.....	4
3.3. pote.dat: repulsion-dispersion potential	4
3.4. Molecular definition file	6
3.4.1. Damped dispersion parameter input	7
3.4.2. Damped Induction parameter input	7
3.4.3. Anisotropic repulsive potential input.....	8
3.5. .dma - multipoles from DMA.....	8
3.6. Atomic polarisabilities input	9
3.7. Atom types automatically recognized and assigned by NEIGHCRYS	9
3.8. Optional atom types file	10
4. Running NEIGHCRYS.....	10
4.1. Running interactively without default values.....	10
4.1.1. Notes on responses to questions	11
4.2. Running with default values.....	14
4.2.1. Default values of constants and answers to questions which are input interactively using option (i). 14	
4.3. Running non-interactively.....	15
5. Contents of output files.....	15
5.1. Standard output (Unit 6).....	15
5.1.1. Interactive mode	15
5.1.2. Non-interactive mode	15
5.2. fort.20.....	15
5.3. fort.21	16
5.4. fort.22.....	16
5.5. fort.23.....	16
5.6. file.dmain	16

5.7.	The following 4 files are no longer written, but the fortran code is still available should you want to turn it on again.....	16
6.	DMACRYS output files	16
7.	Options	17
7.1.	Pasting molecular structure in experimental crystal structure.....	17
7.1.1.	Procedures	17
7.2.	Standardising bond lengths	18
7.3.	Symmetry lowering.....	19
8.	Examples	19

NEIGHCRYYS should compile and run on any platform with a FORTRAN 90 compiler.

This program is written by Dr M. Leslie maurice.leslie@ucl.ac.uk

Please contact the CPOSS project (<http://www.chem.ucl.ac.uk/cposs/dmacryst/>) for a copy of the program.

1. Introduction

NEIGHCRYYS is a program for setting up input files in the format that DMACRYS requires for versions 2.0 and higher.

The program converts crystallographic data files in *.res SHELX format (<http://shelx.uni-ac.gwdg.de/SHELX>) to a Cartesian coordinate system. The molecular fragments are treated as rigid entities.

DMACRYS is a lattice energy minimisation program for molecular structures, using an anisotropic atom-atom model for the electrostatic intermolecular forces, along with an isotropic or anisotropic repulsion and simple dispersion model.

2. Atom labelling convention

An example of an atom label is C_W2_1__1__. The underscores are essential, and the label is broken down like this:

C_: The first two characters in the label are the atom type. If there is only one letter in the label, the blank is filled with an underscore. E.g. C_, O_, N_, H_, Br, Cl, etc

W2: The next two characters are the potential type. If you are using Williams type potentials the first of these characters will be W and the second character, which is a number, will be set by the program to define the type of atom which you have (eg. an alcohol or carboxyl oxygen). If you are using the standard parameters for the fitted potential the first of these characters will be F and the second will be 1, since the fitted potential does not distinguish between different types of the same atom, with the exception of hydrogen where the second character can be 1 or 2. The full definitions of the Williams and FIT parameters are given in section 3.7. You should check the labels manually at the end of the NEIGHCRYYS run. You can define your own labels and read them in from a separate file.

_: The next character is a flag. This is an underscore, unless the atom is in an inverted molecule, in which case it is I.

1__: The next 5 characters are the atom number in the molecule. Numbering restarts for each atom type, i.e. all carbon atoms are labelled 1-5 (say), then nitrogen 1-2, oxygen 1-2 and hydrogen 1-8 (see pasted text below). Unique numbers have to be generated for atoms in inverted molecules. To do this the total number of atoms in the asymmetric unit is added to the number for these atoms. **This completes the short atom label.**

1____: The last 5 characters complete the **full atom label**. This identifies the atom number in the unit cell, and sets the limit of the maximum number of atoms in the cell to 99999. (see text copied from the fort.21 file below). The asymmetric unit is numbered sequentially first, then all of the equivalent atoms to atom 1, and so on.

Printed output (fort.21)

Equivalent basis atoms

atom index	atomic number	name	input name	molecule number	invert flag
1	6	C_F1_1___1___	C1	1	F
2	6	C_F1_1___9___	C1	2	F
3	6	C_F1_1___10___	C1	3	F
4	6	C_F1_1___11___	C1	4	F
5	6	C_F1I9___12___	C1	5	T
6	6	C_F1I9___13___	C1	6	T
7	6	C_F1I9___14___	C1	7	T
8	6	C_F1I9___15___	C1	8	T
9	6	C_F1_2___2___	C2	1	F
10	6	C_F1_2___16___	C2	2	F
11	6	C_F1_2___17___	C2	3	F
12	6	C_F1_2___18___	C2	4	F
13	6	C_F1I10___19___	C2	5	T
14	6	C_F1I10___20___	C2	6	T
15	6	C_F1I10___21___	C2	7	T
16	6	C_F1I10___22___	C2	8	T

3. Input files required

In order to run NEIGHCRY5, the following files are required. Each is discussed in more detail below.

.res (or *.fdat*) - crystallographic data file

bondlengths - covalent bonds specification

.labels - optional atom types file

.dma - atomic multipoles from DMA

.mols - molecular definition file

.pol - atomic polarisabilities file, optional if induced moments are being used.

.paste - optional pasting file to alter molecular conformation

.potential - repulsion-dispersion potential

The files are listed above in the order in which NEIGHCRY5 requests them, but they are discussed below in the logical order for deriving them.

3.1. shelx - crystallographic data file

The input file should be a SHELX format file (*.res) for providing the crystal structure information.

(NEIGHCRY5 will also handle the Cambridge Structural Database (<http://www.ccdc.cam.ac.uk>) legacy FDAT format.) NEIGHCRY5 takes the atom type (as opposed to the atom label) from the SHELX input, and correlates it with the SFAC line, to determine the atomic type. SHELX input files usually (but not always) use atom labels that correspond to the atomic symbol and number the atoms sequentially. NEIGHCRY5 always uses atom labels according to the atomic type, and always numbers atoms sequentially, so the atom labels will not necessarily be the same in both the input SHELX and the NEIGHCRY5 output. (In the case of CSD FDAT files (*.fdat), NEIGHCRY5 uses the atom labels, but numbers all the atoms sequentially.)

Symmetry information is generated from the LATT line, which can take any value from 0 to 7. (For *.fdat, the space group needs to be specified as the first letter is used by NEIGHCRY5 to determine whether it is a centred cell.)

Note: DMACRY5.0 cannot handle linear molecules or very high symmetry space groups occupied by high symmetry molecules.

Example input

```
TITL 01.lem_fit_AXOSOW.res
CELL 1.0 9.8641 10.1391 6.9997 90.0000 90.0000 90.0000
ZERR 8 0.0 0.0 0.0 0.0 0.0 0.0
LATT 1
SYMM +X +1/2, -Y +1/2, -Z
SYMM -X , +Y +1/2, -Z +1/2
SYMM -X +1/2, -Y , +Z +1/2
SFAC C O H
C1 1 0.78779943 0.00798969 0.11853631
C2 1 0.82594413 0.13123460 0.07427794
C3 1 0.95900572 0.18173118 0.13525171
O1 2 1.00047112 0.29311646 0.09992017
H1 3 0.85402044 -0.05613916 0.19975305
H2 3 0.69086951 -0.03182048 0.07589025
H3 3 0.76190113 0.19747526 -0.00667370
H4 3 1.02122594 0.11083632 0.21751215
```

3.2. bondlengths - covalent bonds specification

This file is used to establish the covalent bonds and NEIGHCRY5 assumes atoms within the specified distances are bonded. There is an entry for every type of covalent bond in molecule, specifying the two atom types and setting the cutoff at just slightly larger than the maximum likely value.

Example input

```
C_ C_ 1.60
C_ H_ 1.30
C_ O_ 1.40
O_ H_ 1.10
```

3.3. pote.dat: repulsion-dispersion potential

This is the file where the repulsion-dispersion potential is set up. NEIGHCRY5 can read in two repulsion-dispersion types, a Buckingham potential (keyword BUCK) or a damped Buckingham potential with higher order dispersion (keyword DBUC). In addition the undamped Buckingham potential can have anisotropic repulsion. Other potentials are possible but must be input by editing the input file to DMACRY5 after NEIGHCRY5 has run. For full details of the other types of potential formats, see DMACRY52.0 manual (<http://www.chem.ucl.ac.uk/cposs/dmacrys/manuals/dmacrys.pdf>). A set of recommended potentials along with references is given in the DMACRY5 manual. A summary of the format is given below, usually this will be sufficient for most runs.

The potential is defined between two atoms, specifying their atom type and potential type. NEIGHCRY5 will run without a potential file, and a .dmain file is still produced, but the potentials are all set to zero. If an incomplete potential file is used, any undefined pairwise interactions will be set to zero.

The first record for a potential input must have the keyword giving the potential type (BUCK or DBUC)

This is followed by two fields format A4 giving the atom type for the interaction. (See below for more information about atom type labels).

The form of the Buckingham potential between molecule M and molecule N where atom i of type ι is in molecule M and atom k of type κ is in molecule N separated by a distance R_{ik} is given by:

$$U_{rep-disp}^{MN} = \sum_{i \in M, k \in N} A_{i\kappa} \exp(-B_{i\kappa} R_{ik}) - C_{i\kappa} / R_{ik}^6$$

The next record must read in 5 values. The first three are A, 1/B and C from the above equation. The fourth is the minimum of the range in which this potential will apply (always 0.0) and the final value is the maximum in

Angstroms (This value is overridden by the global cutoff (CUTO), so a large value is normally input here). The units are given below

A	is in eV
1/B	is in Å
C	is in eV Å ⁶

The repulsive part of the damped dispersion potential is the same as for the Buckingham potential. The damped dispersion is

$$U_{disp}^{MN} = \sum_{i \in M, k \in N} -f_6(\beta_{MN} R_{ik}) \frac{C_6^{iK}}{R_{ik}^6} - f_8(\beta_{MN} R_{ik}) \frac{C_8^{iK}}{R_{ik}^8} - f_{10}(\beta_{MN} R_{ik}) \frac{C_{10}^{iK}}{R_{ik}^{10}}$$

where $f_n(\beta_{MN} R_{ik})$ is the Tang-Toennies damping function and is given by

$$f_n(\beta_{MN} R_{ik}) = 1 - \left(\sum_{m=0}^n \frac{(\beta_{MN} R_{ik})^m}{m!} \right) \exp(-\beta_{MN} R_{ik})$$

The damping parameter β_{MN} is a constant for a molecule pair and its value is read in from the molecular definition file (see below). If a value of 0 is input the function is undamped.

The next record must read in 7 values. They are A, 1/B, C_6^{iK} , C_8^{iK} , C_{10}^{iK} from the above equation, with the last 2 values as for BUCK. The units are given below.

A	is in eV
1/B	is in Å
C_6^{iK}	is in eV Å ⁶
C_8^{iK}	is in eV Å ⁸
C_{10}^{iK}	is in eV Å ¹⁰

The potential is normally terminated with an ENDS record. However, anisotropic repulsion parameters may be input first, for either BUCK or DBUC potentials. The form of the short range repulsive potential that DMACRYS can currently use is of the form:

$$U_{rep}^{MN} = \sum_{i \in M, k \in N} A_{iK} \exp[-B_{iK} (R_{ik} - \rho^{iK}(\Omega_{ik}))]$$

where

$$\rho^{iK}(\Omega_{ik}) = \rho_0^{iK} + \rho_1^i(\mathbf{e}_z^i \bullet \mathbf{e}_{ik}) + \rho_1^k(-\mathbf{e}_z^k \bullet \mathbf{e}_{ik}) \\ + \rho_2^i(3[\mathbf{e}_z^i \bullet \mathbf{e}_{ik}]^2 - 1)/2 + \rho_2^k(3[\mathbf{e}_z^k \bullet \mathbf{e}_{ik}]^2 - 1)/2$$

\mathbf{e}_z^i is a unit vector in the z direction of the local axis system on atom i which is read in from the molecular definition file and \mathbf{e}_{ik} is a unit vector in the intersite direction between i and k . The anisotropy is assumed to be cylindrically symmetrical about the intersite vector. This potential type must be given immediately after the parameters for the Buckingham or damped Buckingham potential (before the keyword ENDS) and is input by the keyword ANIS. This is currently followed by up to five additional records. The format of these anisotropic potential records is

5I	F
K1,K2,L1,L2,J	RHO

Where K1, K2, L1, L2, J define the S function used to define this component of the anisotropic potential and RHO is defined above. At present the only permitted values of K1, K2, L1, L2, J are given in the table below

K1	0	0	0	0	0
K2	0	0	0	0	0
L1	0	1	0	2	0
L2	0	0	1	0	2
J	0	1	1	2	2

The first S function is just the isotropic part. If it is not given the values input on the BUCK or DBUC record will be used. B is in \AA^{-1} and the values of RHO are in \AA .

Example input

```
BUCK C_F1 C_F1
3832.147000 0.277778 25.286950 0.0 70
ENDS
```

Example input (from example 03)

```
BUCK C_CA C_CA
7328.362026 0.272512 55.551648 0.0 70.0
ANIS C_CA C_CA
0 0 0 1 1 0.096140
0 0 1 0 1 0.096140
0 0 0 2 2 0.000000
0 0 2 0 2 0.000000
ENDS
```

Example input (not included in test_suite)

```
DBUC N_N1 O_O1
6309.141204 0.211330 3.794553 13.440326 38.920132 0.00 70.00
ENDS
```

3.4. Molecular definition file

This is set up to define the molecular axis system for running in DMACRYS which is used for DMAs and to read in any other properties of the molecule. (The distributed multipoles are different for each atom in the molecule and are read from a separate file.) The molecular definition file is in free format as shown below. The first line must begin with the keyword MOLX, where the integer after MOLX gives the number of crystallographically independent molecules in the cell, and the last line must be ENDS. For each molecule the definition of the local molecular axes used to define the multipole expansion must be given. This may be optionally followed by a damped dispersion parameter, a damped induction parameter and a definition of anisotropic repulsion parameter axes. For molecular axes, the origin is set at the centre of mass of the molecule, and a LINE and PLANE are defined. Enantiomers should not be given separately, as NEIGHCRYS will generate the separate AXES for DMACRYS.

Example input (from example 08)

```
MOLX 2
X LINE C_F1_2_____ C_F1_1_____ 1
Y PLANE C_F1_2_____ C_F1_1_____ 1 C_F1_6_____ 2
X LINE O_F1_6_____ H_Wa_7_____ 1
Y PLANE O_F1_6_____ H_Wa_7_____ 1 H_Wa_8_____ 1
ENDS
```

In the example above, the X axis in the first molecule runs from C2 to C1 which is one bond away. The Y axis is in the plane formed by C2, C1 (one bond away) and C6 (2 bonds away from C2). The Z axis always makes a right-handed set. Thus NEIGHCRY5 has to deal with the problem with pairs of enantiomers generated by space group symmetry because the right handed system on one molecule would become left handed on its mirror image. The equations for the interaction of two multipoles are defined using right handed coordinate systems and NEIGHCRY5 sets up the two enantiomers as two different molecules, the second with the atom label I (see atom labelling convention). The multipole expansions from the dma file are changed so that all odd-z components on the inverted molecule have the sign changed. The output is printed in fort.21, in both Angstroms and AU. While DMACRY5 can run with any of X, Y or Z forming the first (LINE) axis, this is not recommended as the fix for enantiomers only works if Z is the third axis.

3.4.1. Damped dispersion parameter input

The damping parameter for the damped dispersion potential is the same for every atom in a molecule, so it is input from the molecular definition file. The format is

DDMP	2.517930
------	----------

where DDMP defines the damping and β_{MN} is in atomic units. This can be estimated by $\beta_{MN} = 2\sqrt{2I_x}$, where I_x is the first vertical ionization potential of the molecule in atomic units. Damping parameters are the subject of ongoing research and recent work should be consulted, however at present damping parameters for unlike molecule interactions are automatically calculated as the geometric mean.

3.4.2. Damped Induction parameter input

DMACRY5 can calculate induced atomic dipole moments. The field at an atomic site will induce a dipole at that site which will then interact with all the multipoles in the crystal to give the induction contribution to the intermolecular lattice energy. The induced fields need to be damped in the same way as the dispersion is damped. The damping coefficient is a function of the molecule and is read in from the molecular definition file. If NEIGHCRY5 finds a damping coefficient, it will then look for a separate file which contains the atomic polarisabilities, which differ for each atom in the molecule, and read the values.

The local induction at a site will be given by

$$V_t^a = \sum_{B \neq A} f_{t+u+1} (\beta_{MN} R_{ab}) T_{tu}^{ab} Q_u^b$$

At present DMACRY5 only works with the induced field, for which $t = 1$ in the above equation (i.e. induced dipoles). f is the Tang Toennies damping function (above). The resulting induced moments are

$\Delta Q_t^a = \alpha_t^a V_t^a$, where α_t^a are the localized atomic dipolar polarizabilities with components $t = (10,10)$, $(11c,11c)$, $(11s,11s)$, $(10,11c)$, $(10,11s)$, $(11c,11s)$ at each atom (defined using spherical tensors). They need to be iterated to consistency within the crystal i.e. $\Delta Q_t^a = \alpha_{tv}^a \sum_{B \neq A} f_{t+u+1} (\beta_{MN} R_{ab}) T_{vu}^{ab} (Q_u^b + \Delta Q_u^b)$, before the

induction energy can be similarly evaluated in terms of the induced multipole moments ΔQ_t^a , and original multipole moments Q_t^a as

$$E_{ind} = \frac{1}{2} \sum_A \sum_{B \neq A} \Delta Q_t^a T_{tu}^{ab} Q_u^b$$

The induction damping coefficients are read in using the keyword DIND. This is followed by the value of β in atomic units. For interactions between unlike molecules, DMACRY5 uses the geometric mean of the β values.

3.4.3. Anisotropic repulsive potential input

The axis definition for the anisotropic short range potential is input here if necessary. It begins with the keyword ANIS and ends with a keyword ENDS. An example is given below

Example input (from example 03)

```
MOLX 1
X LINE  C_CC_2_____ C_CD_5_____ 3
Y PLANE C_CC_2_____ C_CD_5_____ 3 C_CA_1_____ 1
ANIS
C1CL_1_____
Z LINE  C_CC_2_____ C1CL_1_____ 1
X PLANE C_CC_2_____ C_CA_1_____ 1 C_CB_6_____ 2
C_CC_2_____
Z LINE  C_CC_2_____ C1CL_1_____ 1
X PLANE C_CC_2_____ C_CA_1_____ 1 C_CB_6_____ 2
:      :      :      :      :      :
ENDS
ENDS
```

Format:

Record 1

Card format	A10
	ATOM

ATOM is the short atom label for the atom for which the anisotropic axis is to be defined

Record 2

Card format	A1	A4	A10	A10	I
	AXIS	AXIS TYPE	ATOM 1	ATOM 2	SEPARATION

AXIS is the first axis which is defined, (X, Y or Z, usually X) AXIS TYPE is LINE which is the same as for molecular axes. The species defining the line do not have to include the species for which you are constructing the axis. ATOM 1 and ATOM 2 are the short atom labels for the atoms defining the line and separation is the number of neighbours they are apart (1 for nearest neighbours and so on)

Record 2 (Alternative format)

Card format	A1	A4	A10	A10	I	A10	I
	AXIS	AXIS TYPE	ATOM	ATOM	SEPARATION 1	ATOM	SEPARATION 2
			1	2		3	

This format will seldom be needed. AXIS TYPE will be BIEX or BIIN to define the external or internal bisector of a bond angle. The atoms are in the order {centre of angle, end 1, separation 1, end 2, separation 2} so the bond angle is 2-1-3. These directives could be used for example to define the anisotropic axis system for the oxygen atom of a water molecule.

Record 3

Card format	A1	A4	A10	A10	I	A10	I
	AXIS	AXIS TYPE	ATOM 1	ATOM 2	SEPARATION 1	ATOM 3	SEPARATION 2

The second axis which must be a plane input in the same format as for MOLX. The sequence is terminated by an ENDS record; note that this will give you two ENDS records in a row. The calculated anisotropic axes are printed out.

If there is no molecular definition file present, NEIGHCRY5 will produce a .dmain file, but the MOLX directive is missing, so DMACRY5 will not run. The fort.21 file will give you the NEIGHCRY5 atom labels that correspond to the input atom labels, so you can set up a correct molecular definition file.

3.5. .dma - multipoles from DMA

This contains the atomic multipoles derived by distributed multipole analysis of an *ab initio* wavefunction of the isolated molecule. It is essential that the calculation uses the same orientation of the molecule as in the

molecular definition file. If there is no .dma file present, NEIGHCRY5 will produce a .dmain file, with multipoles set to zero on every atom. Provided a molecular definition file is present, NEIGHCRY5 can be used to generate molecular coordinates in the molecule local axis system in Angstroms for input into GAUSSIAN. The atomic multipoles from the distributed multipole analysis (DMA) file can be calculated using GAUSSIAN (<http://www.gaussian.com>) files using GDMA2 (<http://www-stone.ch.cam.ac.uk/documentation/gdma/README.html>). The output files generated by GDMA2 do not contain the correct NEIGHCRY5 atom labels. To convert them to the format shown below, the DMACRY5 release bundle includes a utility programme called gdmanceighcrys (which needs compiling from the .f90 before use). Alternatively, these can be edited in by hand. If the space group symmetry of the crystal structure generates molecules related by an inversion operator, then NEIGHCRY5 sets up the DMACRY5 input file with the atomic multipoles from the .dma file changed so that all odd-z components on the inverted (I) molecule have the sign changed (cross reference molecular definition file).

Example input

```
! XII properties calculation

1  C_F1_1_____ -2.753446 -2.151702 0.000091 Next 2 Limit 4
-0.083010
 0.000208 0.018691 0.030772
-0.985535 0.000338 -0.000982 -0.038430 0.230443
-0.001680 -0.163223 -0.177082 -0.000341 -0.002022 -2.077258 0.179549
-0.862134 -0.001088 0.000299 -0.508677 1.765922 -0.003956 -0.000607
 0.964204 1.430481

: : : : :
```

3.6. Atomic polarisabilities input

If the keyword DIND is used in the molecular definition section, NEIGHCRY5 will try to read in an atomic polarisabilities file. This file must have the same file name as the multipoles file but with a filename extension .pol appended.

Example input (from example 15)

```
# Static polarizabilities

1  C_C1_1_____ -0.277986 -1.883193 -0.000535 RANK 1
 2.46610
 0.00000 14.52040
 0.00000 4.63450 12.08170

: : : : :
```

3.7. Atom types automatically recognized and assigned by NEIGHCRY5

NEIGHCRY5 has been set up to automatically recognize any of the atom types included in the FIT and Williams potentials.

Table 1. The functional groups that NEIGHCRY5 will recognise

ELEMENT	DESCRIPTION	FIT	Williams
H	in C-H group	H_F1	H_W1
H	in alcoholic group	H_F2	H_W2
H	in carboxyl group	H_F2	H_W3
H	in N-H group	H_F2	H_W4
H	in water	H_Wa	H_Wa
C	connected to four atoms	C_F1	C_W4
C	connected to three atoms	C_F1	C_W3

C	connected to two atoms	C_F1	C_W2
N	with triple bond	N_F1	N_W1
N	no-bonded hydrogen (except triple bond)	N_F1	N_W2
N	one bonded hydrogen	N_F1	N_W3
N	>= two bonded hydrogens	N_F1	N_W4
O	bonded to one atom	O_F1	O_W1
O	bonded to two atoms	O_F1	O_W2
O	in water	O_F1	O_F1
F	Fluorine	F_F1	
Cl	Chlorine	ClF1	
S	Sulphur	S_F1	
K	Potassium	K_F1	

If you are using either FIT or Williams, if an atom appears in the crystallographic data file that is not explicitly included in these potentials, it is assumed that only one type will be found in the potential definition file (.pots). For example, for Fe we will have:

```
Fe    only one type expected    Fe01    |    Fe01
```

If more than one atom type is to be used for a given element, a custom potential can be specified using a labels file, or the dmain file can be manually edited.

3.8. Optional atom types file

It is possible to use a potential other than FIT or Williams in DMACRYS. If you intend doing this, you can specify the atom types for each atom in your .res file. This file should have two columns: the first is the SHELX atom label from the .res (up to 5 characters) the second is a 2 character potential type (e.g. W1), and forms part of the atom label as described previously.

Example input (from example 03)

```
Br2  BR
Br1  BR
Cl   CL
F    FL
C1   CA
C3   CC
C4   CA
C5   CB
H1   HY
C6   CD
C2   CB
H2   HY
```

If an atom that is present in the .res file is not listed in the labels file, NEIGHCRY5 will fail.

4. Running NEIGHCRY5

NEIGHCRY5 may be run in the following 3 ways:

1. Running interactively - useful for initial setups for geometry
2. Running with default values
3. Running non-interactively - useful for repeated setups

4.1. Running interactively without default values

If the program reads in the character 'I' or 'i' as the first character, it will run in interactive mode. The rest of the screen input is read from standard input.

The input file should be SHELX (*.res) format file. If NEIGHCRY5 reads the first character of the first record as # it is assumed to be an FDAT file; anything else is assumed to be SHELX.

Below is a typical interactive run. > indicates output from NEIGHCRY5

```
> TYPE I FOR INTERACTIVE MODE
i
> File name containing crystal structure ?
AXOSOW_cb106.res
> Interactive mode - filenames will be read in
> File name with standard maximum covalent bond lengths ?
bondlengths
> Maximum required inter-molecular contact?
4.0
> Do you wish to insert any bond centre sites (Y/N) ?
n
> Do you want to standardise bond lengths to hydrogen?
n
> Do you want to use [W]illiams, [F]IT or [C]ustom types?
f
> Do you want to foreshorten hydrogen bond lengths for Williams potential?
n
> CVECTOR      6.99970
> Input zero for no symmetry subgroup or n to remove representation n
0
> Do you have a dma file yet (y/n)
Y
> File name for dma file
dmacrys.dma
> DMACRYS input file will be written to AXOSOW_cb106.dmain
> Do you have a molecular definition file?
Y
> Enter filename for molecular definition file.
dmacrys.mols
> Do you want to paste in a new molecular structure?
n
> Do you have an intermolecular potential input file?
Y
> Enter filename for intermolecular potential input file.
fit.pots
```

4.1.1. Notes on responses to questions

```
> TYPE I FOR INTERACTIVE MODE
i
```

I or i will run interactively.

```
> File name containing crystal structure ?
AXOSOW_cb106.res
> Interactive mode - filenames will be read in
```

Error messages

```
> Error - no input file name given
```

```
> Error - input file
> non_file.res
> does not exist
```

```
> File name with standard maximum covalent bond lengths ?
bondlengths
```

This question is not asked if you have called the bondlengths file cutoff. It is where you need to put in the name of the file if you haven't used this convention.

Error messages

```
> Error - input file for standard bond lengths
> non_file
> does not exist
```

```
> Maximum required inter-molecular contact?
4.0
```

This question is effectively redundant as it controls the length of a list of intermolecular contact distances in **file.ccl** see 5.7

```
> Do you wish to insert any bond centre sites (Y/N) ?
n
```

The usual response is N or n. For high accuracy work on small polyatomics, e.g. N₂, then an additional multipole site at the centre of the bond will provide a more accurate description of the electrostatic potential around the molecule. If you answer anything other than N or n, you are asked

```
> Do you wish to insert any bond centre sites (Y/N) ?
Y
> These are the possible bonds in your data set:
> 1 C_ C_
> 2 C_ O_
> 3 C_ H_
> 4 O_ O_
> 5 O_ H_
> 6 H_ H_
> Input the numbers of bonds you wish to have a mid point for, ending with a zero.
5
0
```

You must enter a single bond for which you wish to add a bond centre site on each line.

```
> Do you want to standardise bond lengths to hydrogen?
y
```

This is necessary when X-ray hydrogen positions are used, to correct for the shortening of bonds. A few defaults are C··H 1.08 Å, N··H 1.01 Å and O··H 1.02 Å. It should not be done when the input files has been derived by computational modelling or is an experimental structure determined from neutron diffraction. See Options for more details. If you do not answer “N” or “n” then bond lengths will be standardised.

```
> Do you want to use [W]illiams, [F]IT or [C]ustom types?
f
```

Here, you have the option to specify the potential. If you want to use purely the Williams or purely the FIT potential, you can specify that here. If you have a custom potential, or if you want to include a potential for only part of the system (say a water molecule) you can specify a custom potential. In this case, you will now be asked an additional question (as in example 03).

```
> Do you want to use [W]illiams, [F]IT or [C]ustom types?
c
> File name for custom file
labels
```

This labels file needs to list all the atoms in the .res file, with the potential that you want to use for each.

Error message

```
> This file does not exist
> Do you want to finish or try again (f/r)?
```

To which you can respond “f” to exit NEIGHCRYG or “r” to retry typing in the custom labels filename.

```
> Do you want to foreshorten hydrogen bond lengths for Williams potential?
n
```

The Williams potential is parameterised for having hydrogen sites 0.1 Å closer to the centre of the bond. You must say yes to this question if you have specified that you want to use the Williams potential in the previous

question, and have prepared a .dma file which correctly applies this foreshortening. If you are using FIT or any atomic site based custom potential, then the answer must be n.

Error message

```
> Answer must be (Y/N)
```

```
> CVECTOR      6.99970
```

Alternative output for custom potentials (example 03)

```
> Matching custom types...
> CVECTOR      14.53140
> Matching custom types...
```

```
> Input zero for no symmetry subgroup or n to remove representation n
0
```

Normally 0 to use crystallographic symmetry. When DMACRYS shows that the minimum has a lower symmetry, then the number of the representation(s) with negative eigenvalue(s) is given. If you were removing representation 3 (as in example 16)

```
> CVECTOR      5.51670
> Symmetry reduction for representation      3 complete
> input zero to end symmetry reduction or n to remove new representation n
0
```

This message is written at the end of a symmetry reduction. While it is possible to do a second symmetry reduction within NEIGHCRY5, it is recommended that the user does one at a time (i.e. inputs zero here and reruns DMACRYS). This is covered in more detail in section 7.3.

```
> Do you have a dma file yet (y/n)
y
```

n only used for runs to establish molecular geometry in local axis system. If n is typed, dma file is set up with LEVEL 0 multipoles.

```
> File name for dma file
dmacrys.dma
```

Note that charges of free ions are also read from the dma file. If more than one type of molecule is in the crystal, then the dma file will need to be constructed from dma files from individual *ab initio* calculations. Free ions, e.g. Cl⁻ need to be added manually.

```
> DMACRYS input file will be written to AXOSOW_cb106.dmain
```

If this is not the first run, then

```
> This file will be overwritten
```

```
> Do you have a molecular definition file?
y
```

n is only used for giving the atomic numbering and intramolecular connectivity table needed to choose your axis definition.

```
> Enter filename for molecular definition file.
dmacrys.mols
```

This must correspond to the axis definition used for calculating DMAs.

Error message

```
> This file does not exist
> Do you want to finish or try again (f/r)?
```

Again, you can respond “f” to exit NEIGHCRY5 or “r” to retry typing in the custom labels filename.

```
> Do you want to paste in a new molecular structure?
n
```

Normally n. This option can be used to paste another molecular structure into the crystal, so that the centre of mass matches and the molecular local axes are parallel (see Options). It is particularly useful for comparing the minimum structure with an *ab initio* optimised version of the experimental structure. Type y if you want to paste coordinates over an experimental structure. (Unless the changes in conformation are very small, it is much better to produce a separate input file, that uses the molecular overlay that minimises the rms difference in the atomic coordinates, and use that as your input.) If y (as in example 17),

```
> Enter filename for coordinate pasting file.
coord_to_paste_au
```

```
> Do you have an intermolecular potential input file?
y
> Enter filename for intermolecular potential input file .
fit.pots
```

NEIGHCRY5 can insert Buckingham potentials from a file into the dmain dataset. pote.dat should contain any number of Buckingham potentials in the same format as used in the dmain file, beginning with

BUCK and ending with ENDS. Only Buckingham potentials needed for the structure are copied to the dmain file. Missing potentials are set to zero.

Other error messages

```
> Atomic Polarizabilities file must have same name as Charges file with .pol appended
> Error: pol atom label 1 is not consistent with dma label.
```

These error messages may be produced if you have a DIND line in the molecular definition file.

4.2. Running with default values

If the program reads in the character 'D' or 'd' as the first character, it will run with default values. The input file must be a SHELX format file. Most of the input and output files from NEIGHCRY5 will use the stem (XXXXXXnn) of the .res file as part of the filename.

Again responses to the terminal from NEIGHCRY5 are indicated by >

```
> TYPE I FOR INTERACTIVE MODE
d
> File name containing crystal structure ?
AXOSOW.res
> Default filenames will be used
>
> Used
> cutoff
> as file with standard bond lengths
> CVECTOR      6.99970
>
> Used
> AXOSOW.dma
> as dma file
> DMACRY5 input file will be written to AXOSOW.dmain
>
> Used
> AXOSOW.mols
> as molecular definition file
>
> Used
> pote.dat
> as potential input file
```

If no dma file is present, multipoles are written as LEVEL 0. If no axes file is present, no axes are written to dmain. If no pote.dat file is present, BUCK potentials are written as zero.

4.2.1. Default values of constants and answers to questions which are input interactively using option (i).

```
>Maximum required inter-molecular contact?
4.0
>Do you want to standardise bond lengths to hydrogen?
```

```

Y
>Do you want to use [W]illiams, [F]IT or [C]ustom types?
w
>Do you want to foreshorten hydrogen bond lengths for Williams potential?
Y
>Do you wish to insert any bond centre sites (Y/N) ?
n
>Input zero for no symmetry subgroup or n to remove representation n
0
>Do you want to paste coordinates?
n

```

4.3. Running non-interactively

Running NEIGHCRYS interactively produces a file fort.22 which has the user responses to the questions written to it. You can use this file to run non-interactively.

```

> TYPE I FOR INTERACTIVE MODE
n

```

If the first value NEIGHCRYS receives is “N”, “n”, or nothing, NEIGHCRYS will read fort.22 and interpret this as the user input. Here is a typical fort.22 file

Example input

```

AXOSOW_cb106.res
bondlengths
  4.0000
n
n
f
n
          0
y
dmacrys.dma
y
dmacrys.mols
n
Y
fit.pots

```

fort.22 may be edited to

- a) change the input crystallographic file
- b) change the options that NEIGHCRYS uses

This way of running NEIGHCRYS is useful for running a series of cells where the default options are not used.

5. Contents of output files

5.1. Standard output (Unit 6)

5.1.1. Interactive mode

Questions which require a user response are written here, normally to the screen.

5.1.2. Non-interactive mode

Only the instruction TYPE I FOR INTERACTIVE MODE is written, and can be ignored (See also fort.23).

5.2. fort.20

Symmetry interface file for DMACRYS. Essential for DMACRYS run, a file containing the crystal information.

5.3. fort.21

Information about symmetry analysis.

5.4. fort.22

An input file that can be used to run NEIGHCRY5 non-interactively.

5.5. fort.23

This is a file to which the screen output is written if you use non-interactive mode.

5.6. file.dmain

DMACRY5 input file.

5.7. The following 4 files are no longer written, but the fortran code is still available should you want to turn it on again

file.ccl

A file containing close contacts list up to limit defined by the question.

file.mac

A MACRO file.

file.nem

A NEMESIS file.

file.nnl

A file containing nearest Neighbours list within molecule. In the molecule, the nth nearest neighbouring atoms are separated by a covalent bond.

6. DMACRY5 output files

When DMACRY5 is run from the same directory used for NEIGHCRY5, the additional following DMACRY5 output files will be present:

file.dmaout

The complete output file

fort.12

A summary file

fort.16

The optimised structure in .res format

fort.13

The optimised structure in a P1 cell, i.e. with atomic coordinates for all atoms in unit cell.

fort.8

The optimised cell in the original Cartesian format used within DMACRY5

7. Options

7.1. Pasting molecular structure in experimental crystal structure

Pasting is an option, where an *ab initio* molecular structure can be pasted into an experimental crystal structure to overlay the molecular axes. For example, an experimental molecular structure can be optimised using an *ab initio* program such as GAUSSIAN, and then pasted into the experimental cell to find the closest approximation to the experimental structure that could be found in a crystal structure prediction study which assumed the molecule had the rigid optimised conformation. This option should only be used when there are only small changes between the original molecular conformation and the structure being pasted in, as the procedure aligns the molecule-fixed axes. A procedure that minimises the root mean square difference in the atomic positions will usually produce a structure that has fewer highly repulsive contacts.

7.1.1. Procedures

1. Run NEIGHCRYST using the CSD SHELX file to get the molecular structure in the local axis system from the output file fort.21.
2. Optimise the molecular structure in GAUSSIAN to give an *ab initio* molecular structure in atomic units.
3. Use the *ab initio* molecular structure to calculate a new DMA punch file using GAUSSIAN with GDMA).
4. Run NEIGHCRYST again (See below).

Format of the coordinate pasting file

Card format	A10		
	TITLE		
Card format	I	I	< I >
	NMOLECULES	NATOMS	NATOMS for inverted molecule
Blank line			
Blank line			
Blank line			
Blank line			
Blank line			
Card format	I4	A4	3F13.6
	Index	Label	Coordinates
Blank line			
Blank line			
Blank line			
Blank line			
Blank line			
Card format	I4	A4	3F13.6
	Index	Label	Coordinates

NMOLECULES is the number of molecules specified in the coordinate pasting file, and NATOMS are the corresponding number of atoms for each molecule. Coordinates need to be specified for inverted molecules separately, or they will not be pasted in correctly, so if inversion is present in the crystal system, NMOLECULES must be 2. The second set of labels and coordinates are for the inverted molecule for which the z coordinate has the sign changed. The coordinates are input in AU.

Example input

CYHEPTAMIDE				
4 33 33 33 33				
1	C_F1_1_____	-2.387699	-0.773484	-3.190039
5	C_F1_2_____	-1.906048	0.193039	-0.472150
9	C_F1_3_____	-0.242890	2.533711	-0.472150
13	C_F1_4_____	-1.512096	4.856109	-0.491888

17	C_F1_5	-0.243310	7.147906	-0.493738
21	C_F1_6	2.384713	7.151852	-0.470149
:	:	:	:	:
111	H_F1I43	4.468416	-3.745012	1.136254
115	H_F1I44	1.996154	-2.324571	2.826666
119	H_F1I45	3.542479	-6.084117	-2.511280
123	H_F1I46	0.628678	-7.304861	-5.962437
127	H_F1I47	-3.457743	-5.035207	-6.476673
131	H_F1I48	-4.575977	-1.581330	-3.528200

Example run of NEIGHCRY5 to insert paste file (from example 17)

```
> TYPE I FOR INTERACTIVE MODE
i
> File name containing crystal structure ?
TEVSOD.res
> Interactive mode - filenames will be read in
> File name with standard maximum covalent bond lengths ?
bondlengths
> Maximum required inter-molecular contact?
4
> Do you wish to insert any bond centre sites (Y/N) ?
n
> Do you want to standardise bond lengths to hydrogen?
n
> Do you want to use [W]illiams, [F]IT or [C]ustom types?
f
> Do you want to foreshorten hydrogen bond lengths for Williams potential?
n
> CVECTOR      23.57900
> Input zero for no symmetry subgroup or n to remove representation n
0
> Do you have a dma file yet (y/n)
y
> File name for dma file
dmacrys.dma
> DMACRYS input file will be written to TEVSOD.dmain
> Do you have a molecular definition file?
y
> Enter filename for molecular definition file.
dmacrys.mols
> Do you want to paste in a new molecular structure?
y
> Enter filename for coordinate pasting file.
coord_to_paste_au
> Do you have an intermolecular potential input file?
y
> Enter filename for intermolecular potential input file.
fit.pots
```

7.2. Standardising bond lengths

Standardising bond lengths is needed where the C-H bondlengths in the experimental crystal structure (all are 0.950 Å) must be corrected to the standard 1.08Å, as X-ray diffraction has reduced the C-H bondlengths. Currently, the following standard bond lengths are defined in subroutine NORMALISE

C-H = 1.083 Å

N-H = 1.009 Å

O-H (alcohol) = 0.976 Å

O-H (acid) = 1.015 Å

O-H (water) = 1.00 Å

Other bonds to hydrogen = 1.02 Å

Subroutine SET_ATOM_TYPE will recognise a number of additional functional groups, but no bond length is reset. Modifications can be made in the subroutine SET_ATOM_TYPE for automatically normalising the bond lengths for any of the functional groups automatically recognised by NEIGHCRY5.

7.3. Symmetry lowering

Sometimes a run of DMACRYS will indicate that the structure has relaxed to a saddle point and that there is a lower energy structure with a lower symmetry. Full details are in the DMACRYS manual. In brief, DMACRYS calculates the eigenvalues of the final Hessian matrix for each symmetry representation of the space group. A negative eigenvalue shows that a saddle point has been reached. (There will be three eigenvalues which will be very close to zero corresponding to translation of the crystal as a whole; these should be ignored if any are negative.) NEIGHCRYYS is then run again to produce a new dmain file and fort.20 file with lowered symmetry. This NEIGHCRYYS run can be started from the output res file from DMACRYS which is written to fort.16. In reply to the question

```
> Input zero for no symmetry subgroup or n to remove representation n
```

input the number of the representation you wish to remove. NEIGHCRYYS continues to ask

```
> Symmetry reduction for representation      1 complete  
> input zero to end symmetry reduction or n to remove new representation n
```

and a second representation can be removed. Note that the numbering of the representations would be changed by removing the first representation, so the authors of the programme suggest that DMACRYS is run in between successive symmetry reductions. You can, however, remove representations in one step if you want to repeat previous work or take the symmetry of a cell straight to P1.

The new DMACRYS run will have the negative eigenvalue in the totally symmetric representation so that the structure can relax away from the saddle point. The .dmain file produced by NEIGHCRYYS after a symmetry reduction has been requested has an extra keyword towards the end – SEIG. This tells DMACRYS to search away from the saddle point in the direction of the eigenvector corresponding to the negative eigenvalue. NEIGHCRYYS will also write new files which will have modified filenames. These files are a new molecular definition file, a new dma file and optionally a new paste file, a new custom labels file and a new polarisabilities file. They need to be used in any further runs of NEIGHCRYYS after the symmetry reduction DMACRYS run. This is usually either a run to calculate the properties of the relaxed structure or to further reduce the symmetry. In the latter case another new set of files will be produced.

There is a limitation that the new molecular definition file will not contain information about induction damping or anisotropic repulsions, so this information will need to be input manually. However a new polarizabilities file is written.

8. Examples

The test-suite included in the release package of DMACRYS, includes notes discussing the NEIGHCRYYS input and DMACRYS output files produced. This set of examples is intended to supplement training in the use of NEIGHCRYYS/DMACRYS and the manuals.