

# NEIGHCRYS

Program release 2.0.4

Manual release 20 April 2010

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.	fdat or shelx - crystallographic data file.....	3
3.1.1.	Example of a SHELX file (*.res) .....	4
3.1.2.	Example of an FDAT file (*.fdat) .....	4
3.2.	cutoff - covalent bonds specification .....	4
3.2.1.	Example of cutoff file .....	4
3.3.	pote.dat: repulsion-dispersion potential .....	4
3.3.1.	isotropic.....	5
3.3.2.	anisotropic .....	5
3.4.	axis - local axis set up for molecule .....	6
3.4.1.	Example of an axis file.....	6
3.4.2.	Example of an axis file for use with an anisotropic potential.....	7
3.5.	punch - multipoles from DMA.....	8
3.5.1.	punch file from CADPAC.....	8
3.5.2.	punch file from GDMA.....	9
3.5.3.	Punch files that include rank .....	10
3.6.	Atom types automatically recognized and assigned by NEIGHCRYS .....	10
3.7.	Optional atom types file .....	10
3.7.1.	Example of a labels file .....	10
4.	Running NEIGHCRYS .....	11
4.1.	Running interactively without default values.....	11
4.1.1.	Notes on responses to questions .....	11
4.2.	Running with default values .....	13
4.2.1.	Default values of constants and answers to questions which are input interactively using option (i). 13	
4.3.	Running non-interactively.....	13
5.	Contents of output files.....	14
5.1.	Standard output (Unit 6) .....	14
5.1.1.	Interactive mode.....	14
5.1.2.	Non-interactive mode .....	14
5.2.	fort.20.....	14
5.3.	fort.21.....	14
5.4.	fort.22.....	14
5.5.	fort.23.....	14
5.6.	file.dmain .....	14
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.....	14
6.	Options .....	15
6.1.	Pasting molecular structure in experimental crystal structure.....	15
6.1.1.	Procedures .....	15
6.1.2.	Format of the paste input dataset.....	15
6.1.3.	Example of the paste file .....	15

6.1.4.	Example run of NEIGHCRY5 to insert paste file .....	16
6.2.	Standardising bond lengths .....	17
7.	Symmetry lowering .....	17
8.	Examples of files .....	17

Neighcrys 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/dmacrys/>) for a copy of the program.

## 1. Introduction

NEIGHCRY5 is a program for setting up input files in the format that DMACRY5 2.0 requires.

The program converts crystallographic data files (CSD FDAT (<http://www.ccdc.cam.ac.uk>) and SHELX (<http://shelx.uni-ac.gwdg.de/SHELX>) as \*.res) to a Cartesian coordinate system. The molecular fragments are treated as rigid entities.

Note: DMACRY5 2.0 cannot handle linear molecules or very high symmetry space groups occupied by high symmetry molecules.

DMACRY5 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. These can be assigned automatically by NEIGHCRY5, but you should check them manually.

\_ : The next character is a flag. This is blank, 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.

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__		2	F
3	6	C_F1_1__10__		3	F
4	6	C_F1_1__11__		4	F
5	6	C_F1I9__12__		5	T
6	6	C_F1I9__13__		6	T
7	6	C_F1I9__14__		7	T
8	6	C_F1I9__15__		8	T
9	6	C_F1_2__2__	C2	1	F
10	6	C_F1_2__16__		2	F
11	6	C_F1_2__17__		3	F
12	6	C_F1_2__18__		4	F
13	6	C_F1I10__19__		5	T
14	6	C_F1I10__20__		6	T
15	6	C_F1I10__21__		7	T
16	6	C_F1I10__22__		8	T
17	6	C_F1_3__3__	C3	1	F
18	6	C_F1_3__23__		2	F

19	6	C_F1_3	24		3	F
20	6	C_F1_3	25		4	F
21	6	C_F1I11	26		5	T
22	6	C_F1I11	27		6	T
23	6	C_F1I11	28		7	T
24	6	C_F1I11	29		8	T
25	8	O_F1_1	4	O1	1	F
26	8	O_F1_1	30		2	F
27	8	O_F1_1	31		3	F
28	8	O_F1_1	32		4	F
29	8	O_F1I9	33		5	T
30	8	O_F1I9	34		6	T
31	8	O_F1I9	35		7	T
32	8	O_F1I9	36		8	T
33	1	H_F1_1	5	H1	1	F
34	1	H_F1_1	37		2	F
35	1	H_F1_1	38		3	F
36	1	H_F1_1	39		4	F
37	1	H_F1I9	40		5	T
38	1	H_F1I9	41		6	T
39	1	H_F1I9	42		7	T
40	1	H_F1I9	43		8	T
41	1	H_F1_2	6	H2	1	F
42	1	H_F1_2	44		2	F
43	1	H_F1_2	45		3	F
44	1	H_F1_2	46		4	F
45	1	H_F1I10	47		5	T
46	1	H_F1I10	48		6	T
47	1	H_F1I10	49		7	T
48	1	H_F1I10	50		8	T
49	1	H_F1_3	7	H3	1	F
50	1	H_F1_3	51		2	F
51	1	H_F1_3	52		3	F
52	1	H_F1_3	53		4	F
53	1	H_F1I11	54		5	T
54	1	H_F1I11	55		6	T
55	1	H_F1I11	56		7	T
56	1	H_F1I11	57		8	T
57	1	H_F1_4	8	H4	1	F
58	1	H_F1_4	58		2	F
59	1	H_F1_4	59		3	F
60	1	H_F1_4	60		4	F
61	1	H_F1I12	61		5	T
62	1	H_F1I12	62		6	T
63	1	H_F1I12	63		7	T
64	1	H_F1I12	64		8	T

### 3. Input files required

In order to run NEIGHCRYST, the following files are required. (See 6. Example of input file)

fdat or shelx - crystallographic data file

cutoff - covalent bonds specification

.labels - optional atom types file

punch - atomic multipoles from DMA

axis or axes - local axis set up for molecule

potential - repulsion-dispersion potential

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

#### 3.1. fdat or shelx - crystallographic data file

The input file can be either an FDAT format file (\*.fdat), or a SHELX format file (\*.res) for providing the crystal structure information.

In the case of SHELX format files (\*.res), NEIGHCRYST 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. NEIGHCRYST 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 NEIGHCRYST output.



The potential is defined between two atoms, specifying their atom type and potential type. NEIGHCRYST 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. Below are examples of file formats for isotropic and anisotropic repulsion-dispersion potentials.

### 3.3.1. isotropic

```
BUCK C_F1 C_F1
      3832.147000          0.277778          25.286949  0.0  70.0
ENDS
BUCK C_F1 H_F1
      689.536720          0.272480          5.978972  0.0  70.0
ENDS
BUCK H_F1 H_F1
      124.071670          0.267380          1.413698  0.0  70.0
ENDS
BUCK C_F1 N_F1
      3179.514600          0.271003          19.006710  0.0  70.0
ENDS
BUCK N_F1 O_F1
      2508.044800          0.258398          12.898341  0.0  70.0
ENDS
BUCK N_F1 N_F1
      2638.028500          0.264550          14.286224  0.0  70.0
ENDS
BUCK H_F1 N_F1
      572.105410          0.265957          4.494041  0.0  70.0
ENDS
BUCK H_F1 O_F1
      543.916040          0.259740          4.057452  0.0  70.0
ENDS
BUCK O_F1 O_F1
      2384.465800          0.252525          11.645288  0.0  70.0
ENDS
BUCK C_F1 O_F1
      3022.850200          0.264550          17.160239  0.0  70.0
ENDS
BUCK H_F2 H_F2
      52.128991          0.214592          0.222819  0.0  70.0
ENDS
BUCK C_F1 H_F2
      446.951850          0.242131          2.373693  0.0  70.0
ENDS
BUCK H_F2 N_F1
      370.833870          0.236967          1.784166  0.0  70.0
ENDS
BUCK O_F1 H_F2
      352.561760          0.232019          1.610837  0.0  70.0
ENDS
BUCK H_F1 H_F2
      80.422205          0.238095          0.561248  0.0  70.0
ENDS
```

### 3.3.2. anisotropic

```
BUCK C_01 C_01
      4855.810311          0.281787          55.551648  0.0  70.0
ANIS C_01 C_01
      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
BUCK C_01 C_02
      2035.700412          0.298216          31.889639  0.0  70.0
ANIS C_01 C_02
      0 0 0 1 1 0.000000
      0 0 1 0 1 0.000000
      0 0 0 2 2 -0.069056
      0 0 2 0 2 0.205373
ENDS
BUCK C_01 C_03
      15574.699975          0.237727          49.740238  0.0  70.0
ANIS C_01 C_03
      0 0 0 1 1 -0.134207
      0 0 1 0 1 0.083093
```

```

0 0 0 2 2 0.076023
0 0 2 0 2 -0.058966
ENDS
BUCK C_01 C_06
      1822.535383 0.297903 30.720745 0.0 70.0
ANIS C_01 C_06
0 0 0 1 1 -0.005861
0 0 1 0 1 0.091247
0 0 0 2 2 0.000000
0 0 2 0 2 0.000000
ENDS
BUCK C_01 Br01
      12840.307106 0.270395 81.179909 0.0 70.0
ANIS C_01 Br01
0 0 0 1 1 -0.051539
0 0 1 0 1 -0.222366
0 0 0 2 2 -0.185159
0 0 2 0 2 -0.003100
ENDS

```

### To input the anisotropic short range parameters

```

BUCK C_01 Br01
      12840.307106 0.270395 81.179909 0.0 70.0
ANIS C_01 Br01
0 0 0 1 1 -0.051539
0 0 1 0 1 -0.222366
0 0 0 2 2 -0.185159
0 0 2 0 2 -0.003100
ENDS

```

ANIS should follow immediately after the parameters for the BUCK part.

The lines after ANIS are

k1 k2 l1 l2 j rho

RHO(0) is read in in the above example as the first parameter of the BUCK potential; this value is EXP(ALPHA

\* RHO(0)).

You can read it in as rho(0) by supplying a line after anis

0 0 0 0 rho(0), in which case you will get a warning message.

The second parameter after BUCK \_\_\_\_\_, the third is the Van der Waals parameter.

To convert from anisotropic to isotropic, just remove ANIS and the lines following it.

On the output file, the first parameter of the BUCK potential is reset to zero and the correct value of rho(0) is

printed out even if it was not input.

At present to input a potential between two species that are the same you need to supply both of the identical

interactions.

```

BUCK CL CODA CL CODA
15555.674 0.271759 57.363 0.0 70.0
ANIS CL CODA CL CODA
0 0 0 1 1 0.026758
0 0 1 0 1 0.026758
0 0 0 2 2 -0.101149
0 0 2 0 2 -0.101149
ENDS

```

Note: This can be set up automatically using NEIGHCRYST.

## 3.4. axis - local axis set up for molecule

This is set up to define the molecular axis system for running in DMACRYS, and used for DMAs. The axis input file is in free format as shown below, where the integer after MOLX gives the number of crystallographically independent molecules in the cell. 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, although NEIGHCRYST will generate separate AXES for DMACRYS. DMAs and enantiomers are defined by a right handed axis system.

### 3.4.1. Example of an axis file

```

MOLX 4
X LINE C_F1_4_____ C_F1_3_____ 1
Y PLANE C_F1_4_____ C_F1_3_____ 1 C_F1_15_____ 2

```

```

X LINE C_F1_20 C_F1_19 1
Y PLANE C_F1_20 C_F1_19 1 C_F1_31 2
X LINE C_F1_36 C_F1_35 1
Y PLANE C_F1_36 C_F1_35 1 C_F1_47 2
X LINE C_F1_52 C_F1_51 1
Y PLANE C_F1_52 C_F1_51 1 C_F1_63 2

```

In the example above, the X axis in the first molecule runs from C4 to C3 which is one bond away. The Y axis is in the plane formed by C4, C3 (one bond away) and C15 (2 bonds away from C4). The Z axis always makes a right-handed set. There is a problem with pairs of enantiomers because the right handed system on one molecule would become left handed on its mirror image. NEIGHCRY5 avoids this difficulty by setting them up as two different molecules. The multipole expansions from the punch 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.

### 3.4.2. Example of an axis file for use with an anisotropic potential

```

MOLX 1
X LINE C_F1_3 C_F1_6 3
Y PLANE C_F1_3 C_F1_6 3 C_F1_1 1
ANIS
Cl01_1
Z LINE C_F1_3 Cl01_1 1
X PLANE C_F1_3 C_F1_1 1 C_F1_2 2
C_F1_3
Z LINE C_F1_3 Cl01_1 1
X PLANE C_F1_3 C_F1_1 1 C_F1_2 2
Br01_1
Z LINE C_F1_4 Br01_1 1
X PLANE C_F1_4 C_F1_3 1 C_F1_1 2
C_F1_4
Z LINE C_F1_4 Br01_1 1
X PLANE C_F1_4 C_F1_3 1 C_F1_1 2
Br01_2
Z LINE C_F1_1 Br01_2 1
X PLANE C_F1_1 C_F1_2 1 C_F1_6 2
C_F1_1
Z LINE C_F1_1 Br01_2 1
X PLANE C_F1_1 C_F1_2 1 C_F1_6 2
H_F1_1
Z LINE C_F1_5 H_F1_1 1
X PLANE C_F1_5 C_F1_6 1 C_F1_2 2
ENDS
ENDS

```

#### Format:

##### Record 1

Species for which the anisotropic axis is to be defined

##### Record 2

First axis, format

[X, Y, Z] [LINE, BIEX, BIIN] species1 species2 separation(1-2) [species 3 separation(1-3)]

LINE is as for molecular axes; the species defining the line do not have to include the species for which you are constructing the axis. BIEX BIIN are external and internal bisector of a bond angle. The species are in the order {centre of angle, end 1, separation 1, end 2, separation 2} so the bond angle is 2-1-3.

##### Record 3

Second axis which must be a plane, 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.

Note: This needs to be set up by modifying the dmain file manually.

If there is no axis 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 local axis file.

### 3.5. punch - 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 defined axes file. If there is no punch file present, NEIGHCRYS will produce a .dmain file, with multipoles set to zero on every atom. Provided a local axis file is present, NEIGHCRYS can be used to generate molecular coordinates in the molecule local axis system in Angstroms for input into GAUSSIAN. It is essentially the same definition of atomic multipole moments, as multipole expansion of electrostatic energy programmed in DMACRYS.

The atomic multipoles from the distributed multipole analysis (DMA) punch file can be calculated using either CADPAC (<http://ket.ch.cam.ac.uk/software/cadpac.html>) or from 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 NEIGHCRYS atom labels. To convert them to the format shown below, the DMACRYS release bundle includes a utility programme called `gdmaneighcrys` (which needs compiling from the .f90 before use). Alternatively, these can be edited in by hand.

If the crystal structure contains inverted molecules, the punch file need not have multipoles for both molecules. NEIGHCRYS sets up the DMACRYS input file with the enantiomers as two different molecules, by inverting the z-axis so as to maintain a right-handed axis set on each molecule. The multipole expansions from the punch file are changed so that all odd-z components on the inverted molecule have the sign changed.

Below are examples of the punch file format obtained from CADPAC and GDMA.

#### 3.5.1. punch file from CADPAC

```
UREA from UREAXX12 neutron study
1 CAR1 0.102858 0.000000 0.000000 Next 2 Limit 4
1.142734
0.000000 0.163569 0.000000
-0.115661 0.000000 0.000000 0.213489 0.000000
0.000000 -0.452037 0.000000 0.000000 0.000000 -1.755088 0.000000
0.288890 0.000000 0.000000 0.011036 0.000000 0.000000 0.000000
-0.557763 0.000000
2 OXY1 2.486560 0.000000 0.000000 Next 3 Limit 4
-0.942585
0.000000 0.395975 0.000000
0.330669 0.000000 0.000000 0.312420 0.000000
0.000000 -0.062567 0.000000 0.000000 0.000000 -0.552782 0.000000
0.034192 0.000000 0.000000 -0.612699 0.000000 0.000000 0.000000
1.062039 0.000000
3 NIT1 -1.219552 -2.169872 0.000000 Next 4 Limit 4
-0.702845
0.000000 0.097093 -0.021555
-0.409123 0.000000 0.000000 0.226010 0.053068
0.000000 -0.084590 0.080790 0.000000 0.000000 0.000000 2.053346 0.306773
-0.897911 0.000000 0.000000 -0.363623 0.277733 0.000000 0.000000
0.016630 1.220312
4 HPD1 -0.280411 -3.829624 0.000000 Next 5 Limit 4
0.319084
0.000000 0.060251 -0.097863
0.001509 0.000000 0.000000 0.008894 0.018881
0.000000 0.003705 -0.008955 0.000000 0.000000 -0.003279 0.000303
0.008741 0.000000 0.000000 0.005165 0.006545 0.000000 0.000000
-0.002919 0.000306
5 HPD2 -3.119080 -2.143102 0.000000 Next 6 Limit 4
0.283687
0.000000 -0.107948 -0.001013
0.001371 0.000000 0.000000 -0.015800 -0.000800
0.000000 -0.010229 0.003014 0.000000 0.000000 -0.001783 0.001932
0.004046 0.000000 0.000000 -0.015559 -0.000022 0.000000 0.000000
-0.006394 -0.003327
6 NIT2 -1.219552 2.169872 0.000000 Next 7 Limit 4
-0.702845
0.000000 0.097093 0.021555
-0.409123 0.000000 0.000000 0.226010 -0.053068
0.000000 -0.084590 -0.080790 0.000000 0.000000 2.053346 -0.306773
-0.897911 0.000000 0.000000 -0.363623 -0.277733 0.000000 0.000000
0.016630 -1.220312
7 HPD3 -0.280411 3.829624 0.000000 Next 8 Limit 4
```

```

0.319084
0.000000 0.060251 0.097863
0.001509 0.000000 0.000000 0.008894 -0.018881
0.000000 0.003705 0.008955 0.000000 0.000000 -0.003279 -0.000303
0.008741 0.000000 0.000000 0.005165 -0.006545 0.000000 0.000000
-0.002919 -0.000306
8 HPD4 -3.119080 2.143102 0.000000 Next 0 Limit 4
0.283687
0.000000 -0.107948 0.001013
0.001371 0.000000 0.000000 -0.015800 0.000800
0.000000 -0.010229 -0.003014 0.000000 0.000000 -0.001783 -0.001932
0.004046 0.000000 0.000000 -0.015559 0.000022 0.000000 0.000000
-0.006394 0.003327

```

### 3.5.2. punch file from GDMA

! 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

2 C_F1_2_____ -1.468024 0.033046 0.000091 Next 3 Limit 4
-0.099755
-0.000193 0.078284 0.033800
-1.031925 0.000674 -0.000858 -0.037299 0.199880
0.001609 0.607750 0.353310 0.000541 0.002386 1.839033 -0.057410
-1.317048 -0.000479 0.000528 0.447676 1.116989 -0.002095 -0.000756
-0.625469 0.688838

3 C_F1_3_____ 1.313728 0.033046 0.000091 Next 4 Limit 4
0.142507
-0.000039 0.036572 0.117321
-0.512891 0.000056 -0.000277 0.051033 0.164083
-0.000561 0.104318 0.332541 -0.000518 -0.000851 -1.513949 0.034840
0.408371 -0.000354 0.001229 0.659093 0.305263 -0.001169 0.000405
-0.201013 1.064795

4 O_F1_1_____ 2.609468 1.954453 -0.000319 Next 5 Limit 4
-0.348997
0.000099 -0.215006 -0.328234
-0.107476 -0.000291 -0.000004 -0.280330 0.870559
-0.000315 0.278613 0.516916 -0.000113 -0.000045 -0.038576 -0.213487
-0.367504 -0.000278 -0.000220 -0.039852 0.444566 -0.000505 0.000147
0.700283 0.198072

5 H_F1_1_____ -1.768511 -3.945529 0.002392 Next 6 Limit 4
0.088803
-0.000081 -0.029504 0.119791
-0.141392 0.000156 -0.000391 -0.079035 -0.090804
-0.000435 -0.074517 0.152438 -0.000366 -0.000257 -0.102773 0.106485
-0.136846 0.000149 -0.000555 -0.039424 0.027146 0.000071 -0.000114
0.106548 0.021611

6 H_F1_2_____ -4.793500 -2.213190 0.000565 Next 7 Limit 4
0.088470
0.000026 0.109839 0.028219
-0.131786 -0.000111 -0.000076 0.146197 0.017081
-0.000058 0.134098 -0.003876 0.000108 0.000041 -0.137455 -0.066653
-0.168143 -0.000269 -0.000213 0.117437 0.107386 0.000019 0.000528
0.032247 0.092730

7 H_F1_3_____ -2.401551 1.852638 -0.002260 Next 8 Limit 4
0.116400
0.000067 0.073664 -0.094183
-0.135732 0.000175 -0.000377 -0.066882 -0.067756
0.000445 0.102035 -0.128078 0.000219 0.000423 0.120268 0.006760
-0.067394 -0.000033 -0.000407 -0.051594 0.101380 -0.000262 -0.000099
-0.077096 0.039011

8 H_F1_4_____ 2.191417 -1.865349 0.001126 Next 0 Limit 4
0.095477
-0.000056 -0.056007 0.131342
-0.118462 0.000064 -0.000115 -0.023473 -0.056286
-0.000186 -0.051526 0.155549 -0.000098 -0.000129 -0.085370 -0.023175
-0.009762 0.000102 0.000057 0.116494 0.001582 0.000001 0.000114

```

-0.049096 0.037072

### 3.5.3. Punch files that include rank

I think that Maurice is doing something with this.

## 3.6. Atom types automatically recognized and assigned by NEIGHCRYS

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

The possible functional groups that NEIGHCRYS will recognise are given below.

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_Wa	O_Wa
Cl		ClF1	

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 S and Fe we will have:

S	only one type expected	S_01	S_01
Fe	only one type expected	Fe01	Fe01

If more than one atom type is to be used, a custom potential can be specified using an atom\_types file, or the dmain file can be manually edited.

## 3.7. Optional atom types file

It is possible to use a potential other than FIT or Williams in DMACRYS. If you intend doing this, or have atoms such as sulfur that are not automatically assigned by NEIGHCRYS, 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.

### 3.7.1. Example of a labels file

C1	C1
C2	C2
C3	C2
H1A	HY
H2A	W1
H99D	P1

If an atom that is present in the .res file is not listed in the atom\_types file, NEIGHCRYS will fail.

## 4. Running NEIGHCRYS

NEIGHCRYS 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 can be either an FDAT (\*.fdat) format file (Cambridge Crystallographic Database), or a SHELX (\*.res) format file. The interactive input is exactly the same in both cases. If NEIGHCRYS 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 NEIGHCRYS

```
> TYPE I FOR INTERACTIVE MODE
i
> File name containing basis set ?
AMBACO.res
> Filename differs from csd refcod, interactive mode
> File name with standard bond lengths ?
bondlengths
> Maximum required inter-molecular contact?
4.0
> nbonds=          7
> Do you wish to insert any bond centre sites (Y/N) ?
n
> 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
> CVECTOR
>   10.800000000000000
> Input zero for no symmetry subgroup or n to remove representation n
0
> Do you have a punch file yet (y/n)
y
> File name for punch file
cadpac.charges
> Labelled basis to be written to
> AMBACO.dmain
> This file will be overwritten
> Do you have an axis definition file?
y
> Enter filename for axis definition file.
dmarel.axis
> Do you want to paste coordinates?
n
> Do you have a potential input file?
y
> Enter filename for potential input file.
pote.dat
> Labelled nearest neighbour list to be written to filename.nnl
> This file will be overwritten
> Labelled close contact list to be written to filename.ccl
> This file will be overwritten
> MACROMODEL file to be written to filename.mac
> This file will be overwritten
> NEMESIS file to be written to filename.nem
> This file will be overwritten
```

#### 4.1.1. Notes on responses to questions

```
> TYPE I FOR INTERACTIVE MODE
i
I or i will run interactively.
> File name containing basis set ?
```

```
filename.fdat
> Filename differs from csd refcod, interactive mode
> File name with standard bond lengths ?
bondlengths
```

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

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

This prints the intermolecular atom-atom contacts less than 4 Å in the output. Usually 4.0 is large enough.

```
> nbonds= 7
```

Number of records read from cutoff file, i.e. number of types of covalent bond.

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

```
n
```

Normally n, but for high accuracy work on small polyatomics, e.g. N<sub>2</sub>, then an additional multipole site at the centre of the bond will provide a more accurate description of the electrostatic potential around the molecule.

```
> 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. See Options for more details.

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

```
W
```

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...

```
> File name for custom file
```

```
atom_types
```

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

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

```
y
```

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.

```
> CVECTOR
```

```
> 10.800000000000000
```

```
> 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.

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

```
y
```

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

```
> File name for punch file
```

```
cadpac.charges
```

Note that charges of free ions are not read from the punch file, the dmain file will need to be changed. If more than one type of molecule is in the crystal, then the punch file will need to be constructed from punch files from individual *ab initio* calculations. Free ions, e.g. Cl<sup>-</sup> need to be added manually.

```
> Labelled basis to be written to
```

```
> AMBACO.dmain
```

```
> This file will be overwritten if file exists
```

```
> Do you have an axis definition file?
```

```
y
```

n is only useful to print out the atomic numbering and intramolecular connectivity table if you do not have a chemical diagram from which to choose your axis definition.

```
> Enter filename for axis definition file.
```

```
dmarel.axis
```

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

```
> Do you want to paste coordinates?
```

```
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. The authors recommend that you use more sophisticated tools to perform this operation.

```
> Do you have a potential input file?
```

```

Y
> Enter filename for potential input file.
pote.dat
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.
> Labelled nearest neighbour list to be written to AMBACO.nnl
> This file will be overwritten
> Labelled close contact list to be written to AMBACO.ccl
> This file will be overwritten
> MACROMODEL file to be written to AMBACO.mac
> This file will be overwritten
> NEMESIS file to be written to AMBACO.nem
> This file will be overwritten
These files can be used to find atoms to define the axis system if a graphics program is not available.

```

## 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 an FDAT or SHELX format file. Most of the input and output files from NEIGHCRY5 will use the stem (XXXXXXnn) of the .fdat or .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 basis set ?
BAWNIW01.fdat

```

NEIGHCRY5 will run without expecting any more user input assuming the default values and file names described below.

cutoff (required)

XXXXXXnn.fdat or XXXXXXnn.dat or XXXXXXnn.res (required)

XXXXXXnn.axis or XXXXXXnn.axes

XXXXXXnn.punch

pote.dat

If no punch 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 NEIGHCRY5 using option (i) produces a file fort.22 which has the user responses to the questions written to it. You can now use this option if you have previously used either option (i) or option (d)

```

> TYPE I FOR INTERACTIVE MODE
n

```

NEIGHCRY5 will read fort.22 and interpret this as the user input. Here is a typical fort.22 file

```

AMBACO.res
bondlengths
  4.000000000000000
n
Y

```

```
y
y
      0
y
cadpac.charges
y
dmarel.axis
n
y
pote.dat
```

fort.22 may be edited to

- a) change the input crystallographic file
- b) change the options that NEIGHCRY5 uses

This way of running NEIGHCRY5 is useful for running a series of cells where the default options are not used.

## 5. Contents of output files

Here are the output files produced by NEIGHCRY5 (e.g. examples for formic acid)

```
fort.20
fort.21
fort.22
fort.23
file.dmain
```

### 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 DMACRY5 3.01. Essential for DMACRY5 run, a file containing the crystal information.

### 5.3. fort.21

Information about symmetry analysis.

### 5.4. fort.22

It writes 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.

**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. Options

There are a few options that could be applied in NEIGHCRY5.

Pasting molecular structure in experimental crystal structure

Standardising bond lengths

### 6.1. 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 in the same axis system. For example, an experimental molecular structure can be optimised using an *ab initio* program such as CADPAC (or GAUSSIAN). The opposite could also be done, that is pasting in an experimental molecular structure into an *ab initio* optimised crystal structure. An example of this is when a blind prediction is needed, when the experimental geometry is known, in order to check that the predicted structure is of reasonable quality, the experimental molecular structure can be pasted into the predicted structure for minimisation.

#### 6.1.1. Procedures

1. Run NEIGHCRY5 using the CSD FDAT (or SHELX) to get the molecular structure in the local axis system from the output file fort.21.
2. Optimise the molecular structure in CADPAC (or GAUSSIAN) to give an *ab initio* molecular structure.
3. Use the *ab initio* molecular structure to calculate a new DMA punch file in CADPAC (or GAUSSIAN with GDMA).
4. Run NEIGHCRY5 again (See below).

#### 6.1.2. Format of the paste input dataset

```
TITLE
NMOLECULES NATOMS NATOMS
(5 Blank Lines)
I4,7X,A4,3F13.6
Index, Label, Coordinates
(5 Blank Lines)
I4,7X,A4,3F13.6
Index, Label, Coordinates
```

NMOLECULES is the number of molecules in the unit cell. NATOMS are the corresponding number of atoms for each molecule. 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.

#### 6.1.3. Example of the paste file.

```
pyrene (D2h RHF/6-31G** structure) in PYRENE02 (cryst. struct.)
2 26 26
```

```
1  CAR1  -6.615739  0.000000  0.000000
2  CARa  2.688476  -2.309989  0.000000
3  CARb  1.265503  -4.643900  0.000000
4  CARc  -1.265503  -4.643900  0.000000
5  CARd  -2.688476  -2.309989  0.000000
6  CARE  -5.318119  -2.271274  0.000000
7  CARf  -1.353904  0.000000  0.000000
8  CARg  1.353904  0.000000  0.000000
9  CAR2  -5.318119  2.271274  0.000000
```

10	CAR3	-2.688476	2.309989	0.000000
11	CAR4	-1.265503	4.643900	0.000000
12	CAR5	1.265503	4.643900	0.000000
13	CAR6	2.688476	2.309989	0.000000
14	CAR7	5.318119	2.271274	0.000000
15	CAR8	6.615739	0.000000	0.000000
16	CAR9	5.318119	-2.271274	0.000000
17	HYD1	-8.649170	0.000000	0.000000
18	HYDb	2.294913	-6.398497	0.000000
19	HYDc	-2.294913	-6.398497	0.000000
20	HYDe	-6.344427	-4.027802	0.000000
21	HYD2	-6.344427	4.027802	0.000000
22	HYD4	-2.294913	6.398497	0.000000
23	HYD5	2.294913	6.398497	0.000000
24	HYD7	6.344427	4.027802	0.000000
25	HYD8	8.649170	0.000000	0.000000
26	HYD9	6.344427	-4.027802	0.000000
27	CAI1	-6.615739	0.000000	0.000000
28	CAIa	2.688476	-2.309989	0.000000
29	CAIb	1.265503	-4.643900	0.000000
30	CAIc	-1.265503	-4.643900	0.000000
31	CAId	-2.688476	-2.309989	0.000000
32	CAIe	-5.318119	-2.271274	0.000000
33	CAIf	-1.353904	0.000000	0.000000
34	CAIg	1.353904	0.000000	0.000000
35	CAI2	-5.318119	2.271274	0.000000
36	CAI3	-2.688476	2.309989	0.000000
37	CAI4	-1.265503	4.643900	0.000000
38	CAI5	1.265503	4.643900	0.000000
39	CAI6	2.688476	2.309989	0.000000
40	CAI7	5.318119	2.271274	0.000000
41	CAI8	6.615739	0.000000	0.000000
42	CAI9	5.318119	-2.271274	0.000000
43	HYI1	-8.649170	0.000000	0.000000
44	HYIb	2.294913	-6.398497	0.000000
45	HYIc	-2.294913	-6.398497	0.000000
46	HYIe	-6.344427	-4.027802	0.000000
47	HYI2	-6.344427	4.027802	0.000000
48	HYI4	-2.294913	6.398497	0.000000
49	HYI5	2.294913	6.398497	0.000000
50	HYI7	6.344427	4.027802	0.000000
51	HYI8	8.649170	0.000000	0.000000
52	HYI9	6.344427	-4.027802	0.000000

#### 6.1.4. Example run of NEIGHCRYS to insert paste file

```

>TYPE I FOR INTERACTIVE MODE
i
>File name containing basis set ?
file.fdat
>Filename differs from csd refcod, interactive mode
>Used
>cutoff
>as file with standard bond lengths
>Maximum required inter-molecular contact?
4.0
>nbonds= 2
>Do you want to standardise bond lengths to hydrogen?
n
>Do you wish to insert any bond centre sites (Y/N) ?
n
>CVECTOR
>8.470000000000001
>Input zero for no symmetry subgroup or n to remove representation n
0
>Do you have a punch file yet (y/n)
y
>File name for punch file
file.punch
>Labelled basis to be written to
>pyrene.dmain
>This file will be overwritten
>Do you have an axis definition file?
y
>Enter filename for axis definition file.
axis
>Do you want to paste coordinates?
y
>Enter filename for coordinate pasting file.
paste

```

```

>Do you have a potential input file?
y
>Enter filename for potential input file.
pote.dat
>Labelled nearest neighbour list to be written to file.nnl
>This file will be overwritten
>Labelled close contact list to be written to file.ccl
>This file will be overwritten
>MACROMODEL file to be written to file.mac
>This file will be overwritten
>NEMESIS file to be written to file.nem
>This file will be overwritten

```

## 6.2. 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. 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.) NEIGHCRY5 is then run again to produce a new dmain file and fort.20 file with lowered symmetry. This NEIGHCRY5 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. NEIGHCRY5 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 NEIGHCRY5 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. NEIGHCRY5 will also write two or three new files which will have the same name as your input files with the extra extension .new. These files are a new axis file, a new charges file and a new custom labels file. They need to be used in any further runs of NEIGHCRY5 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 a new set of three files will be produced.

## 8. Examples of files

On the CPOSS website, there are sets of files for use with NEIGHCRY5 and DMACRYS. The full set of input and output files are available, for minimizations that used a variety of different settings.