BY JULIA CONTRERAS-GARCÍA,ERIN R. JOHN-SON, S. KEINAN, W. YANG, ALBERTO OTERO-DE-LA-ROZA, CHAOYU QUANG, RUBÉN LAPLAZA AND ROBERTO A. BOTO

# NCIPLOT4 MANUAL

## Contents

Theoretical background	5
Installing and running the program	7
The input	9
The output	13
Examples	19
Useful scripts	27

## **List of Figures**

1	Schematic representation of the electron density ( $\rho$ ) for the hydrogen atom (top) and for the H <sub>2</sub> molecule along the intermolecular axis(bottom). The reduced density gradient ( $s(\rho)$ ) and ( $\rho$ ) goes to infinity and zero far way from nuclear positions
	and RDG isosurfaces. A colour code based of $sign(\lambda_2)\rho$ was uses: -0.7 (blue) < 0.0 (green) < 0.7 (red). a) RDG <i>vs</i> $sign(\lambda_2)\rho$ for the benzene molecule. b) RDG isosurface for the monomer. c) RDG <i>vs</i> $sign(\lambda_2)\rho$ for the benzene dimer. d) RDG isosur- face for the dimer
3	NCI analysis of the benzene dimer. (Left) RDG vs $sign(\lambda_2)\rho$
	plot. (Right) 0.5 RDG isosurface. Colour code: -0.05(blue),
4	NCL analysis of the water dimer (Top) of RDG isosurface
4	(Botton) RDG vs sign( $\lambda_2$ ) $\rho$ plot Colour code: -0.05(blue),
	0.00(green), 0.05(red)
5	NCI analysis of the 2v5x-ligand interaction. Comparison
	of the isosurfances obtained without the LIGAND(left) key-
	word and with it (right). 0.3 RDG isosurface. Colour code:
	-o.o7(blue), o.oo(green), o.o7(red) 24
6	NCI analysis of the adenine-thymine complex. (Left) KDG $r_{\rm s}$ sign( $\lambda_{\rm s}$ ) a plot. (Pight) a <b>F</b> RDC isosurface. Colour code:
	-0.05 (h)ue) $0.00$ (rech) $0.05$ (rech)
7	Intermolecular interactions in the adenine-thymine com-
<i>'</i>	plex. (Left) RDG vs $sign(\lambda_2)\rho$ plot. (Right) 0.5 RDG isosur-
	face. Colour code: -0.05(blue), 0.00(green), 0.05(red) 25
8	RADIUS kewyword.NCI analysis of the adenine-thymine com-
	plex. The value of RADIUS are: a) -0.185 -0.017 -2.756 1.5;
	b) -0.537 -0.027 -0.593 1.0; c) 0.40 -0.048 1.789 1.0; d) -1.0285
	-0.011 0.5 1.0
9	Image caption of the VMD dashboard 27

# List of Keywords

1	RTHRES r	9
2	LIGAND n r	10
3	RADIUS x y z r	10
4	INTERMOLECULAR	10
5	INTERCUT r <sub>1</sub> , r <sub>2</sub>	10
6	ONAME name	10
7	INCREMENTS $r_1 r_2 r_3 \ldots \ldots \ldots \ldots \ldots \ldots \ldots \ldots$	10
8	OUTPUT n	10
9	FRAGMENT	10
10	CUTOFFS $r_1 r_2 \dots \dots$	11
11	CUTPLOT $r_1 r_2 \ldots \ldots \ldots \ldots \ldots \ldots \ldots \ldots \ldots$	11
12	ISORDG r	11
13	DGRID	11
14	INTEGRATE	11
15	RANGE	11
16	$CG_2FG n m_1 m_2,, m_{n-1}, m_n \ldots \ldots \ldots \ldots \ldots$	11
17	COARSE	12
18	FINE	12
19	ULTRAFINE	12

### Theoretical background

The NonCovalent Interaction (NCI) index is a visualisation tool based on the density ( $\rho$ ) and its derivatives. It enables the identification of non-covalent interactions from the reduced density gradient (RDG or *s*).

As highlighted in Figure 1, there is a crucial change in  $\rho$  between molecules due to the annihilation of the density gradient at these points, and so, in RDG.





These variations lead to peaks when RDG is plotted against  $\rho$ . Figure 2 depicts RDG *vs*  $\rho$  for the benzene molecule and the benzene  $\pi$ - $\pi$  dimer. The main difference between the monomer and the dimer is the appearance of steep peaks at low densities (Figure 2c).

To visualise the non-covalent interactions identified by the RDG  $vs \rho$  peaks, the 3D points whose values of RDG and  $\rho$  form the peaks

need to be found (Figure 2 b,d).

These interactions correspond to both favorable and unfavorable interactions. In order to differentiate between them, the sign of the second density Hessian eigenvalue ( $\lambda_2$ ) times the density is implemented (sign( $\lambda_2$ ) $\rho$ ). This value is able to characterise the strength of the interactions by means of the density, and its nature thanks to the sign of the second eigenvalue. Attractive and repulsive interactions are identified as regions where  $\lambda_2 < 0$  and  $\lambda_2 > 0$ , respectively. Weak van der Waals interactions by  $\lambda_2 \approx 0$ . To visualise these regions, it is often used a colour code based on sign( $\lambda_2$ ) $\rho$ : blue for strong attractive interactions ( $\lambda_2 < 0$ ), green for weak van der Waals interactions ( $\lambda_2 < 0$ ), and red for strong repulsive interactions ( $\lambda_2 > 0$ ).





### Installing and running the program

To install NCIPLOT4, unpack or clone the contents of the distributionfrom https://github.com/juliacontrerasgarcia/nciplot. The following files should be found:

- LICENSE
- L0G
- README
- src\_nciplot\_4.0
- dat
- tests

The files LICENSE, LOG and README contain the GNU General Public License information, the latest modifications and compilation information respectively. The source code is located in the directory src\_nciplot\_4.0. The directory dat contains the atomic density files. Its path needs to be set during the compilation (see below). The directory test includes a large range of examples and scripts to execute them.

To compile the code go into the src\_nciplot\_4.0 subdirectory. Change the Makefile.inc to suit your compiler and flags and build the nciplot executables by executing the following commands:

make mrproper make

To clean the object and module files,

make clean

To clean objects, modules and binaries,

#### 8 NCIPLOT4 MANUAL

make mrproper

NCIPLOT4 requires NCIPLOT\_HOME environment variable to find the atomic density files contatined in the dat subdirectory. Set it to the absolute path to NCIPLOT:

export NCIPLOT\_HOME=/home/xxx/nciplot/

You may add the previous lines to your bashrc or .bash\_aliases for convenience.

The code has been parallelised for shared-memory architectures using the OpenMP library. To use this feature, set the OME\_NUM\_THREADS environment variable to the number of cores:

export OMP\_NUM\_THREADS=4

If the compilation is completed successfully, an executable called nciplot is created. The code is invoked as follows:

nciplot input.file > ouput.file

### The input

The input is keyword oriented, free-format, and the output is meant to be self-contained. Commentaries must be preceded by #. The following coding is used for variables:

- r stands for real.
- x, y, z stand for positions in space (real).
- n stands for integer.
- name stands for any character string.

#### Obligatory input

### THE <u>FIRST LINE</u> MUST CONTAIN THE NUMBER OF FILES TO BE ANALISED.

Contrary to previous version, NCIPLOT4 does accept multiple wfn and wfx files. Up to 100 files are supported by this version.

THE <u>SECOND LINE</u> MUST CONTAIN THE NAME OF THE MOLECULAR FILE. It accepts three different types of file format: xyz, wfn and wfx. They must have the following extensions depending on the desired calculation:

- name.xyz (Promolecular approximations) It requires an xyz file. Recommended for big systems.
- name.wfn/wfx wavefunction file in the AIMPAC WFN/WFX format (SCF calculation).

#### Optional keywords

• 1. RTHRES r

Add extra limits to the box (in Å). In order to ensure a correct cube in planar or linear cases, the code build a cubic box around the molecule, adding a minimun distance of 1.5 a.u. in all directions. This value is modified by RTHRES.

#### • 2. LIGAND n r

n is the number of the file (as written in the input) and r is a distance in Å.

This option will only plot interactions around the molecule in the file n within a radius r

#### • 3. RADIUS x y z r

x y z determine the positions around which interactions are represented for a radius r (all in Å).

#### • 4. INTERMOLECULAR

This keyword triggers intermolecular interaction options.

This option will turn off all the interactions where at least a fraction of the density comes from a single molecule. This enables only intermolecular interactions to be plotted, neglecting the intramolecular ones. The default is set so that an interaction is discarded if 0.95 of the density at the point comes from a single molecule. This value can be modified by the keyword (INTERCUT).

#### • 5. INTERCUT **r**<sub>1</sub>, **r**<sub>2</sub>

Cutoffs for intermolecularity definition. By default r1 and r2 are set to 0.95 and 0.75 respectively.

#### • 6. ONAME name

name stands for the basic naming to be passed to the output file names. The name introduced will be always output in capital letters. By default, the root name of the structure file is taken.

#### • 7. INCREMENTS r<sub>1</sub> r<sub>2</sub> r<sub>3</sub>

This option sets the increments along the x, y, z directions of the cube in Å. The default is set to 0.1, 0.1, 0.1.

#### • 8. OUTPUT n

n is an integer running from 1 to 3.

- 1 will only print the .dat file.
- 2 will only print the .cube files.
- 3 will print all three output files [default]

#### • 9. FRAGMENT

ifile atom1, atom2, ..., atomn.

Defining fragments from the .xyz files. ifile is the file label given by the input order and atom n is the atomic label in ifile • 10. CUTOFFS r<sub>1</sub> r<sub>2</sub>

Density  $(r_1)$  and RDG  $(r_2)$  cutoffs used in creating the dat file. The default cutoffs are 0.5 and 1.0 respectively.

#### • 11. CUTPLOT **r**<sub>1</sub> **r**<sub>2</sub>

Density  $(r_1)$  and RDG  $(r_2)$  cutoffs used when creating the cube files.  $r_1$  will set the cutoff for both the density and the RDG to be registered in the cube files, whereas  $r_2$  will be used in the VMD script for the plotting of isosurfaces. The default values are:

- In the promolecular case  $r_1=0.07$  a.u.,  $r_2=0.3$  a.u.
- In the SCF case:  $r_1=0.05$  a.u.,  $r_2=0.5$ .

#### • 12. ISORDG r

RDG isosurface used in the VMD script. This is equivalent to the parameter  $r_2$  given in CUTPLOT (see above)

#### • 13. DGRID

Using grids for promolecular densities.

Computing properties withing NCI regions

• 14. INTEGRATE

Trigger the integration of properties.

#### • 15. RANGE

n

 $r_1 down r_1 up$ 

 $r_2 down r_2 up$ 

 $r_n$ down  $r_n$ up

Compute properties in *n* intervals of sign( $\lambda_2$ ) $\rho$ .

A lower and upper bounds are required for every interval. Interval limits should be written in different lines.

#### Multilevel grids

• 16. CG<sub>2</sub>FG n m<sub>1</sub> m<sub>2</sub>, ..., m<sub>n-1</sub>, m<sub>n</sub>

Coarse grid to fine grid multi-level.

n stands for the number of grids and  $m_1, m_2, ..., m_{n-1}, m_n=1$  are the multiplicative factors that define the multilevel grids. This

factors must be integer number introduced in decreasing order  $m_1 > m_2 > ... > m_{n-1} > m_n = 1$ .

Notice that the size of the grid is eventually defined by the keyword INCREMENTS. Different grid steps may requires different multiplicative factors in CG2FG to achieve stability in the results. To easy the used of the multi-level grid approach, the grid step and the multilevel grid factors have been already optimized leading to the following keywords:

#### • 17. COARSE

Coarse multigrid option: CG2FG 4 8 4 2 1 and increments of 0.15 Å.

#### • 18. FINE

Fine multigrid option: CG2FG 4 12 8 4 1 and increments of 0.05 Å.

#### • 19. ULTRAFINE

Ultrafine multigrid option: CG2FG 4 24 16 8 1 and increments of 0.025 Å.

### The output

The output (may) consist of 4 files:

- name.dat file collect  $\rho$  vs RDG (OUTPUT=1 or 3).
- name-grad.cube file with RDG (OUTPUT=2 or 3).
- name-dens.cube file with sign( $\lambda_2$ ) $\rho \times$  100 (OUTPUT=2 or 3).
- name.vmd is a script for the visualization of the results in VMD.

The structure of the output is similar to the previous version of the code:

# ----- NCIPLOT -----# --- PLOTTING NON COVALENT INTERACTION REGIONS ----# ---E.R. Johnson - - - -J. Contreras-Garcia # ---- - - -# ----- Duke University ------# # ---A. de la Roza # ----- University of California Merced -----# R. A. Boto # ---- - -C. Quan # \_ \_ \_ - -# ----- UniversitÃl' Pierre et Marie Curie ------# -----# ---Please cite # --J. Am. Chem. Soc., 2010, 132 (18), pp 6498âĂŞ6506------# -----# # ---Contributions for the wfn properties ---from H. L. Schmider are acknowledged ---# ---# ----------# ---Contributions for the wfx reader - - - from Dave Arias are acknowledged - - -# ---

```
# -----
 #
 # Start -- 2020.06.07, 08:45:59.333
+ Read density file : /home/roberto/src/nciplot-intermolecular_wfn/dat/h__lda.wfc
 Log grid (r = a*e^{(b*x)}) with a = 2.4788E-03, b =
                                            2.0000E-03
 Num. grid points = 4392, rmax (bohr) =
                                   16.1513
 Integrated charge =
                         0.99999998
 El. conf. : 1S(1)
+ Read density file : /home/roberto/src/nciplot-intermolecular_wfn/dat/c__lda.wfc
 Log grid (r = a * e^{(b * x)}) with a = 4.1313E-04, b = 
                                            2.0000E-03
 Num. grid points = 5326, rmax (bohr) =
                                   17.4308
 Integrated charge =
                         5.99999996
+ Read density file : /home/roberto/src/nciplot-intermolecular_wfn/dat/n__lda.wfc
 Log grid (r = a * e^{(b * x)}) with a = 3.5411E \cdot 04, b = 
                                            2.0000E-03
 Num. grid points = 5343, rmax (bohr) =
                                   15.4574
 Integrated charge =
                         6.99999996
 El. conf. : 1S(2) 2S(2) 2P(3)
+ Read density file : /home/roberto/src/nciplot-intermolecular_wfn/dat/o__lda.wfc
 Log grid (r = a * e^{(b * x)}) with a = 3.0984E - 04, b = 
                                            2.0000E-03
 Num. grid points = 5358, rmax (bohr) =
                                   13.9372
 Integrated charge =
                         7,99999996
 El. conf. : 1S(2) 2S(2) 2P(4)
-----
     INPUT INFORMATION:
MIND YOU
     RUNNING IN PROMOLECULAR MODE
_____
     Calculation details:
_____
RHO THRESHOLD (au): 0.50
RDG THRESHOLD (au): 1.00
Writing output in the following units:
-----
```

Sign(lambda2)xDensity x Reduced Density Gradient = AT.dat

Reduced Density Gradient,RDG = AT-grad.cube Sign(lambda2)xDensity,LS = AT-dens.cube VMD script = AT.vmd

When multilevel grip option is trigger, the features of each grid are printed as:

-----Operating grid and increments: Grid-1 ----x0,y0,z0 = -11.7843 -3.0798 -6.9277 6.3277 x1, y1, z1 = 13.84833.2251 ix,iy,iz = 1.13 1.13 1.13 nx,ny,nz = 23 6 12 Time for computing density & RDG = 0.00 secs 0.00% of small boxes are removed. -----Operating grid and increments: Grid-2 ----x0,y0,z0 = -11.7843 -3.0798 -6.9277 x1,y1,z1 = 13.8483 3.2251 6.3277 ix,iy,iz = 0.76 0.76 0.76 9 nx, ny, nz = 3418 Time for computing density & RDG = 0.01 secs 0.00% of small boxes are removed. Operating grid and increments: Grid-3 ----x0,y0,z0 = -11.7843 -3.0798 -6.9277 x1,y1,z1 = 13.8483 3.2251 6.3277 nx,ny,nz = 68 17 36

Time for computing density & RDG = 0.06 secs

-

0.06% of small boxes are removed.

-----Operating grid and increments: Grid-4 ----x0,y0,z0 = -11.7843 -3.0798 -6.9277 x1,y1,z1 = 13.8483 3.2251 6.3277 ix,iy,iz = 0.05 0.05 0.05 nx,ny,nz = 543 134 281 Time for computing density & RDG = 30.52 secs 58.43% of small boxes are removed. Time for writing outputs = 45.38 secs 59.49% of small boxes removed for promolecular integration

The last bunch of information is related to the integration of  $\rho^n$  and sign $(\lambda_2)\rho^n$  within the NCI regions and over their boundaries:

 $\int_{\Omega_{NCI}} \rho^n((\mathbf{r}) d\mathbf{r}, \quad n = 0, 1, 1.5, 2, 2.5, 3, 4/3, 5/3$  $\int_{\Omega_{NCI}} sign(\lambda_2) \rho^n(\mathbf{r}) d\mathbf{r} \quad n = 0, 1, 1.5, 2, 2.5, 3, 4/3, 5/3$ 

It is worth mentioning that the volume of  $\Omega_{NCI}$  is defined by the integral  $\rho^{n=0}$ 

-----

	]	INTEGRATION DATA	
Integration	over the	e volumes of rho^n	
n=1.0	:	81.78751055	
n=1.5	:	32.71905205	
n=2.0	:	14.92576555	
n=2.5	:	7.52272297	
n=3.0	:	4.08744955	
n=4/3	:	43.71246189	
n=5/3	:	24.89093300	
Volume	:	862.16226131	

rho-sum\_i rho\_i : 0.14776508

\_\_\_\_\_ Integration over the volumes of sign(lambda2)(rho)^n \_\_\_\_\_ n=1.0 -38.78097498 : n=1.5 : -21.57715460 -11.75189302 n=2.0 : n=2.5 : -6.56180728 n=3.0 : -3.78238954 n=4/3 : -26.48446142 n=5/3 -17.60409566 : \_\_\_\_\_ Integration over the areas of rho^n \_\_\_\_\_ n=1.0 : 52.68576805 n=1.5 33.08255593 : n=2.0 : 23.80536956 n=2.5 : 17.72600807 n=3.0 : 13.31834056 : 37.69490189 n=4/3 n=5/3 29.42766077 : Area : 552.43261501 rho-sum\_i rho\_i : 0.23567316 ..... Time for integration = 62.22 secs End -- 2020.06.07, 08:48:21.087

### Examples

In all cases 2D and 3D pictures have been obtained by applying the VMD script from the calculation. The background and atom colours were edited.

#### Number of files

In this case, different options are possible. If all the interactions are wanted, molecules are introduced in a unique file or two separate files:

```
1
24_Benzene-Benzene_pi-pi.xyz
FINE
```

#### 2

24\_Benzene-Benzene\_pi-pi\_monomerA.xyz 24\_Benzene-Benzene\_pi-pi\_monomerB.xyz FINE

In both cases the same result is obtained:



Figure 3: NCI analysis of the benzene dimer. (Left) RDG *vs*  $sign(\lambda_2)\rho$  plot. (Right) 0.5 RDG isosurface. Colour code: -0.05(blue), 0.00(green), 0.05(red)

#### Computing properties within the NCI regions

In the following example the intermolecular interactions in the water dimer are analysed. To compute properties, NCI regions are built from water monomers wavefunctions as contained in the files A22\_Water-Water.wfx and B22\_Water-Water.wfx.

2

A22\_Water-Water.wfx B22\_Water-Water.wfx FINE OUTPUT 3 INTEGRATE INTERMOLECULAR INTERCUT 0.75 0.75

The first command, 2, tells the program that two files are to be read: A22\_Water-Water.wfx and B22\_Water-Water.wfx. The keyword FINE defines the multigrid level: 4 12 8 4 1, and grid increments 0.05, 0.05, 0.05. That is, four grids are going to be used with step sizes: 0.05x12=0.6, 0.05x8=0.4, 0.05x4=0.2 and 0.05x1=0.05. OUTPUT 3 generates four output file: A22\_Water-Water.dat with sign( $\lambda_2$ )*rho* vs RDG, sign( $\lambda_2$ ) $\rho$  and  $\rho$  Gaussian cube files: A22\_Water-Water-grad.cube A22\_Water-Water-dens.cube respectively, and a VMD script to plot them: example.vmd

INTEGRATE triggers the computation of properties within the NCI regions.

INTERMOLECULAR and INTERCUT defines intermolecular interactions as defined above using the parameter 0.75 and 0.75.

This above input could be also written as:

2 A22\_Water-Water.wfx B22\_Water-Water.wfx CG2FG 4 12 8 4 1 INCREMENTS 0.05 0.05 0.05 OUTPUT 3 INTEGRATE INTERMOLECULAR INTERCUT 0.75 0.75

Similar to the previous versions of the code, .cube and .dat files allow to visualised NCI interactions. As shown in Figure 4, the hydrogen bonding in the water dimer is visualised as a lenticular isosurface between the water molecules. When RDG is plotted against  $sign(\lambda_2)\rho$ , a dip between 0.02 and 0.03 a.u.



Figure 4: NCI analysis of the water dimer. (Top) 0.5 RDG isosurface. ((Botton) RDG *vs*  $sign(\lambda_2)\rho$  plot Colour code: -0.05(blue), 0.00(green), 0.05(red)

Beyond the qualitative analysis, quantification of non-covalent interactions are possible with NCIPLOT4. Integration over the NCI regions is given in three separated blocks. In the first block integrals of  $\rho^n$  are printed:

Integration over the volumes of rho^n -----n=1.0 0.00976859 : n=1.5 : 0.00114406 n=2.0 0.00013489 : n=2.5 0.00001601 : n=3.0 0.00000191 : n=4/3 0.00233997 : n=5/3 : 0.00056217 0.72628959 Volume : 0.00000134 rho-sum\_i rho\_i :

Next, integrals of  $sign(\lambda_2)\rho$  are printed.

Integration over the volumes of sign(lambda2)(rho)^n

n=1.0	:	0.00571364
n=1.5	:	0.00065746
n=2.0	:	0.00007188
n=2.5	:	0.00000782
n=3.0	:	0.0000084
n=4/3	:	0.00137594
n=5/3	:	0.00031544

In the third block, integrals over the boundaries of the NCI region are printed:

Integration	over the	areas of rho^n	
n=1.0	:	0.09203179	
n=1.5	:	0.01089971	
n=2.0	:	0.00130481	
n=2.5	:	0.00015793	
n=3.0	:	0.00001933	
n=4/3	:	0.02216897	
n=5/3	:	0.00536540	
Area	:	6.76716841	
rho-sum_i rho	0_i :	0.00001298	

#### Protein-Ligand interaction

In this example, the interaction between the protein 2v5x and a ligand is studied. The structure of the protein and the ligand is contained in the files 2v5x-within5.xyz and 2v5x-ligand.xyz respectively.

2 # Number of files to be read 2v5x-within5.xyz # Protein structure

2v5x-ligand.xyz	# Ligand structure
LIGAND 2 4.	<pre># Only interactions at 4 Angstroms from the ligand are analysed</pre>
ONAME 2v5x	# Output files name
FINE	# Multigrid and increments definition.
RANGE 3	# Integration is performed in three ranges.
-0.1 -0.02	
-0.02 0.02	
0.02 0.1	

To remove the intramolecular interactions within the protein the LIGAND keyword is used (See Figure 5). Only interactions at 4 Åaround the ligand, given by the file number 2. The grid increments and multigrid level is given by the FINE keyword: CG2FG 4 12 8 4 1 and reference increments of 0.05.

Sometimes it is interesting to compute properties in the attractive  $(sign(\lambda_2)\rho < 0)$ , van der Waals  $(sign(\lambda_2)\rho \approx 0)$  and repulsive  $(sign(\lambda_2)\rho > 0)$  regions separately. This is done with the keyword RANGE. In the example above, integration will be computed in three ranges:

- Range 1: -0.01 <  $sign(\lambda_2)\rho$  < -0.02.
- Range 2: -0.02 <  $sign(\lambda_2)\rho$  < 0.02.
- Range 3:  $0.02 < sign(\lambda_2)\rho < 0.01$ .

Integration of  $\rho^n$  and  $(\lambda_2)\rho^n$  are printed for each block.

Integration	over the	e volumes of rho^n
n=1.0	:	2.97614936
n=1.5	:	0.58755898
n=2.0	:	0.11949129
n=2.5	:	0.02506085
n=3.0	:	0.00542289
n=4/3	:	1.00690144
n=5/3	:	0.34509889
Volume	:	82.88980295
rho-sum_i rh	o_i :	0.00118474

In all cases 3D pictures have been obtained by directly applying the VMD script from the calculation. The lines defining the cube edges have also been highlighted when appropriate.



Figure 5: NCI analysis of the 2v5xligand interaction. Comparison of the isosurfances obtained without the LIGAND(left) keyword and with it (right). 0.3 RDG isosurface. Colour code: -0.07(blue), 0.00(green), 0.07(red)

#### Choosing the interactions

The adenine-thymine complex is small system where several noncovalent interaction are at play: hydrogen bonds, van der Waals interactions and steric clashes. This can be easily revealed by a standard calculation:

1 AT.xyz INTEGRATE ULTRAFINE

In this case we use the ULTRAFINE keyword, which sets the multilevel grids 4 12 8 4 1, and a reference grid increments 0.025 Å.

Steric clashes can be ruled out by the INTERMOLECULAR keyword. To use INTERMOLECULAR, adenine and thymine structures have to be input in separate files:

2 A.xyz T.xyz INTEGRATE INTERMOLECULAR ULTRAFINE ONAME AT

By default, the code takes the root name of the fist .xyz file to, A in this example, to name the output files. To avoid this behavior the keyword ONAME is used.





Figure 6: NCI analysis of the adeninethymine complex. (Left) RDG *vs*  $sign(\lambda_2)\rho$  plot. (Right) 0.5 RDG isosurface. Colour code: -0.05(blue), 0.00(green), 0.05(red)



Figure 7: Intermolecular interactions in the adenine-thymine complex. (Left) RDG  $vs \ sign(\lambda_2)\rho$  plot. (Right) 0.5 RDG isosurface. Colour code: -0.05(blue), 0.00(green), 0.05(red)

Sometimes, it is interesting to split non-covalent interactions in real space. This is possible using the RADIUS keyword.

1 AT.xyz RADIUS -1.0285 -0.011 0.5 1.0 INTEGRATE FINE



c)



d)

Figure 8: RADIUS kewyword.NCI analysis of the adenine-thymine complex. The value of RADIUS are: a) -0.185 -0.017 -2.756 1.5; b) -0.537 -0.027 -0.593 1.0; c) 0.40 -0.048 1.789 1.0; d) -1.0285 -0.011 0.5 1.0.

### Useful scripts

#### Using the VMD script

The program generates a script for visualization of the result under the name.vmd. This script can be loaded in VMD. After entering the working directory, the script will automatically generate an NCI picture with a RDG cutoff as specified by the keyword PLOTCUT (otherwise, a default is used).



Figure 9: Image caption of the VMD dashboard.

The script is as follows:

In case the user wants to change it, the main options have been highlighted in red:

- name-dens.cube: name of the density cube file
- name-grad.cube: name of the gradient cube file
- 0.30000: value of RDG isosurface
- -4.0000 4.0000 where this is 100 times the value of rhoplot

In order to generate figures ready of publication quality, the background needs to be colour in white. This can be done by going to the GRAPHICS slash, and select COLORS.

#### *Plotting RDG vs sign*( $\lambda_2$ ) $\rho$ *with Gnuplot*

All the plots of RDG vs sign $(\lambda_2)\rho$  have been generated with the following script:

```
# Gnuplot script for mapping NCI color code over NCI diagrams, by R.A.Boto
set terminal pngcairo size 1000,1000 enhanced font 'Helvetica,20'
set encoding iso_8859_1
set output 'AT_2d.png'
set key
set ylabel 's(a.u.)' font "Helvetica, 30"
set xlabel 'sign({/Symbol l}_2){/Symbol r}(a.u.)' font "Helvetica, 30"
set pm3d map
# Define a color gradient palette used by pm3d
set palette defined (-0.07 "blue",0.00 "green", 0.07 "red")
set format y "% .2f"
set format x "% .2f"
set format cb "% -.2f"
set border lw 4
set xtic -0.06,0.01,0.06 nomirror rotate font "Helvetica"
set ytic 0.0,0.25,1.0 nomirror font "Helvetica"
# set the color bar tics
set cbtic -0.06,0.01,0.06 nomirror font "Helvetica"
set xrange [-0.07:0.07]
set yrange [0.0:1.0]
# set the range of values which are colored using the current palette
set cbrange [-0.06:0.06]
plot 'AT.dat' u 1:2:1 w p lw 6 palette t ''
```