Training on the use of "NCImilano" code (investigation of NCI using experimental and theoretical electron density)

GABRIELE SALEH

UNIVERSITÀ DEGLI STUDI DI MILANO

CMC
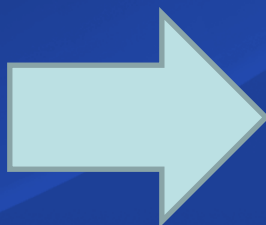Center for Materials Crystallography

OUTLINE

- ED distribution form single-crystal X-Ray diffraction
- multipolar model: experimental ED from XD2006
- the code 'NCImilano'

SINGLE CRYSTAL X-RAY DIFFRACTION

X-ray are scattered by electrons (approximations: Nuclei scattering negligible, elastic scattering, etc.)
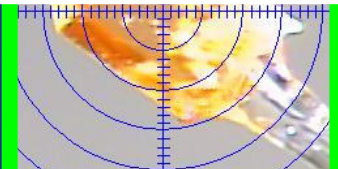
When X-rays irradiate a single crystal, interference among scattered radiations gives rise to sharp peaks (diffraction peaks)
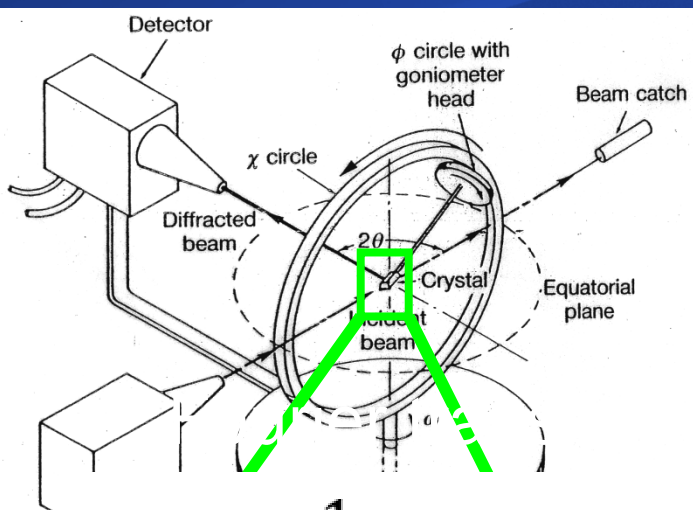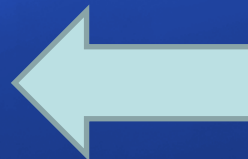
Correction and mathematical threatment…

Structure Factors

$\mathbf{F}_{hkl}$

Miller indices

of ρ

$$\rho(\mathbf{r}) = \frac{1}{V} \sum_{h,k,l} F_{hkl} e^{-2\pi i \mathbf{G} \cdot \mathbf{r}}$$

SINGLE CRYSTAL X-RAY DIFFRACTION: the multipolar model

In principle, ED can be obtained from Fourier summation of structure factors Fhkl. But:

$$\rho(\mathbf{r}) = \frac{1}{V} \sum_{h,k,l}^{\infty} F_{hkl} e^{-2\pi i \mathbf{G} \cdot \mathbf{r}}$$

Truncation error!
(not negligible…)

MULTIPOLAR MODEL. experimental ED is obtained through Least Squares fitting: (atom-centered multipoles)

$$\rho(\mathbf{r}) = \rho_c(r) + P_v \rho_v(\kappa r) + \rho_d(\kappa' \mathbf{r})$$
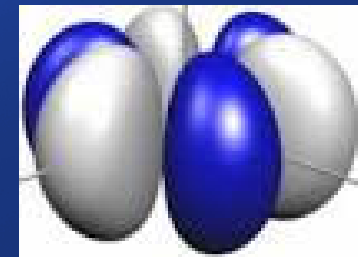
Core density

Spherical density

Deformation density

$$\rho_d(\kappa' \mathbf{r}) = \sum_l R_l(\kappa' r) \sum_{m=-l}^{l} P_{lm} y_{lm}\left(\frac{\mathbf{r}}{r}\right)$$

Once one obtains the multipolar parameters (therefore the experimental ED distribution), many ED-based properties can be calculated…

multipolar model: molecules 'extracted' from crystal

Experimental ED in the crystal is expressed as a sum of atom-centered multipoles which extend through all the (infinite) crystal.

When properties are to be calculated, the user has to decide how many atoms (molecules) will be taken into account. Usually just a few molecules are considered (contribution from the neighboring molecules is negligible) and they are often referred to as MOLECULES EXTRACTED FROM CRYSTAL
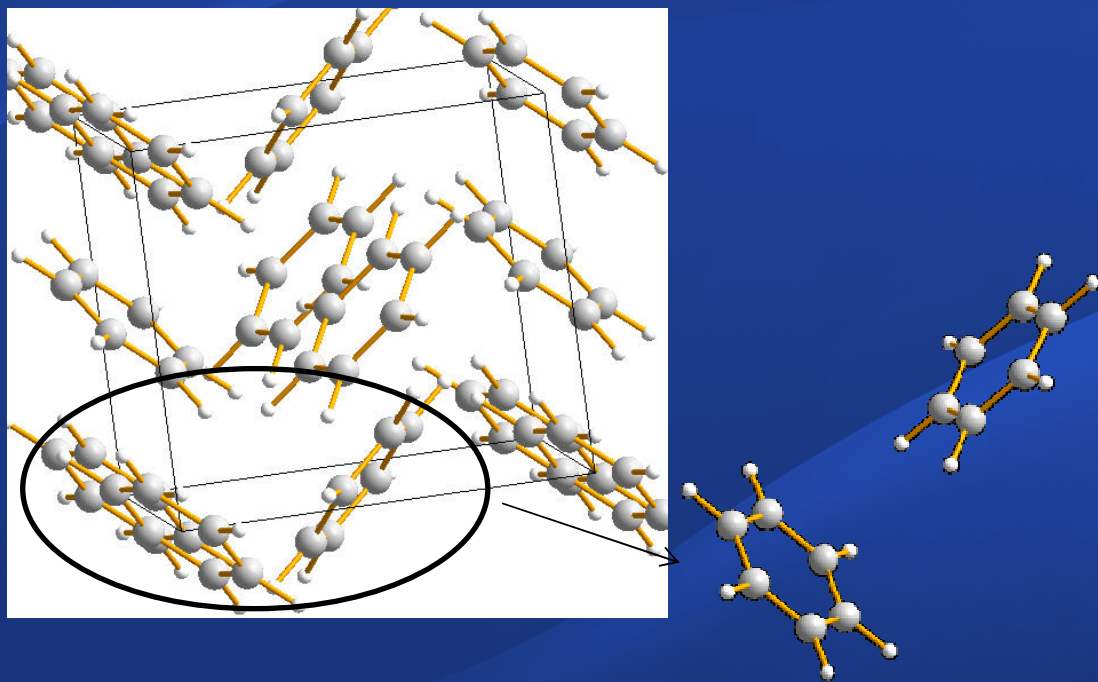
IMPORTANT: the multipole populations are obtained from X-ray structure factor OF A CRYSTAL. Therefore they already contain the information about intermolecular interactions! (polarization, charge transfer, etc.)

OUR CODE TO IMPLEMENT NCI DESCRIPTOR*...

CRYSTAL 06/09

WAVE FUNCTION FILE (.gred)

GAUSSIAN 03/09

NCImilano

WAVEFUNCTION FILE

TOPOND

PROPERTIES FROM WAVE FUNCTION

TOPOND TO CUBE CONVERSION

GRID FILES OF $\rho(r)$, $|\nabla\rho(r)|,\nabla^2\rho(r)$

CUBE FILES OF $\rho(r)$, $|\nabla\rho(r)|,\nabla^2\rho(r)$

XD2006

CUBE FILES OF ENERGY DENSITY

RDG, $\rho*sign(\lambda_2)$ and other properties [*ED and space cutoff*]

CUBE AND '.grd' FILES OF $\rho(r)$, $|\nabla\rho(r)|,\nabla^2\rho(r)$

GRAPHIC VISUALIZATION CODES

CUBE FILES OF RDG and $\rho*sign(\lambda_2)$

'.grd' FILES OF RDG and $\rho*sign(\lambda_2)$
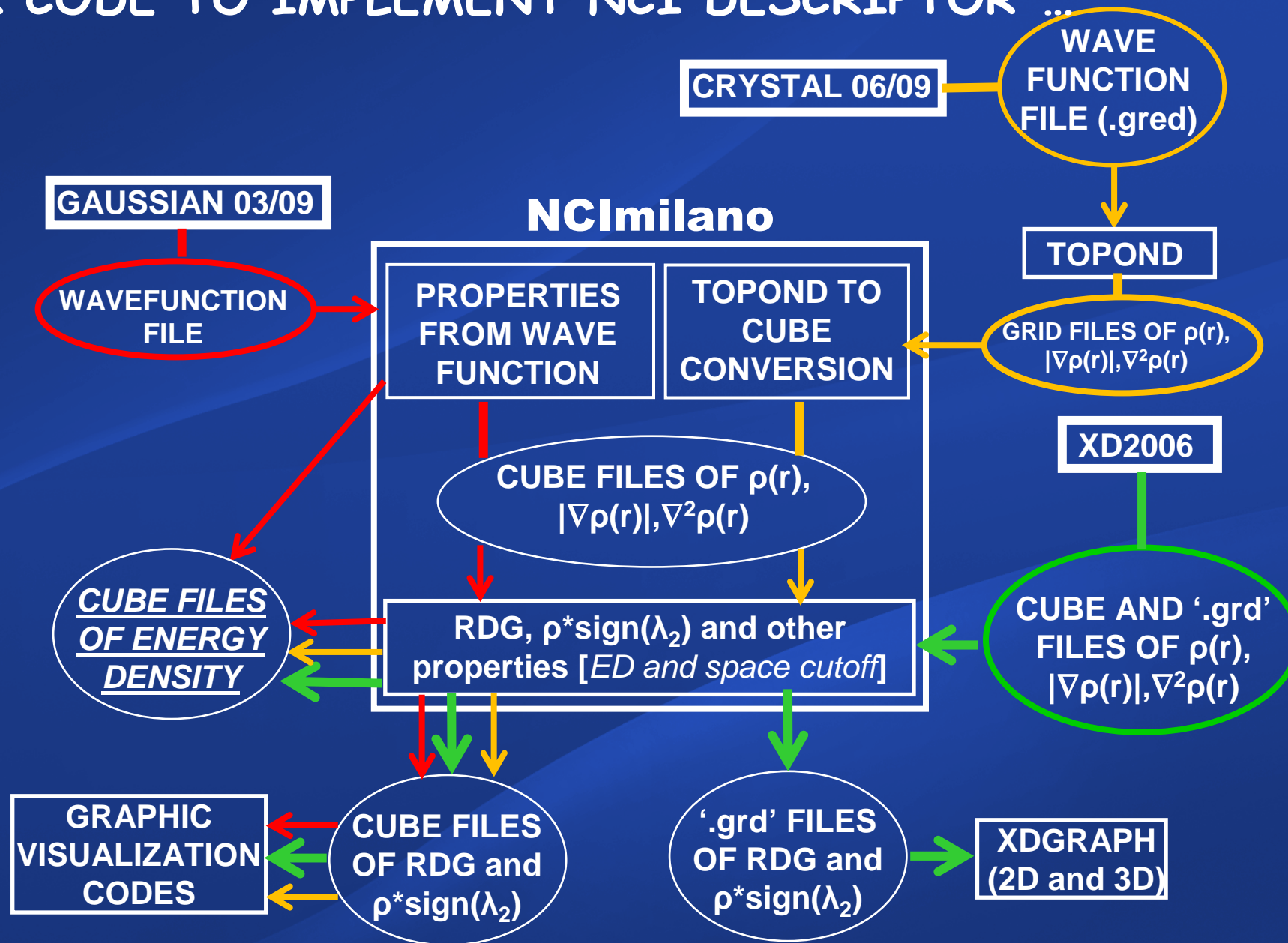
XDGRAPH (2D and 3D)

LET'S GET STARTED!!!

EXERCISE 1:

AIM: calculate the quantities needed to apply NCIdescriptor to experimentally-derived ED of an antimalarial compound (dihydroartemisinin).

A molecular pair is here considered (therefore only the multipoles of the two molecules are taken into account)

GRID FILES OF ρ AND $\nabla\rho$ HAVE BEEN PRODUCED FROM XD2006 (please find it in the folder: exercise1/1a/input). Two grid files for each quantity: *.grd* and *.cube* (they enclose one molecule and intermolecular region)

WITH «NCImilano» IT IS POSSIBLE TO PRODUCE GRID FILES OF RDG AND ρ*sign(λ2)…

THE INPUT FILE

```
File  Edit  Format  View  Help
OUTPUT_FILE_NAME:
SECTIONS: WFN  TOPOND   RDG
**********************************************************
PROPERTIES FROM WAVEFUNCTION*************************
**********************************************************
INPUT:
OUTPUT:
CALCULATING= rho gradrho laplrho energy
CENTER:
N_POINTS_X:
STEPSIZE_X:
N_POINTS_Y:
STEPSIZE_Y:
N_POINTS_Z:
STEPSIZE_Z:


***END*SECTION***************************************

**********************************************************
TOPOND SECTION****************************************
**********************************************************

N OF GRIDS:

GRID FILE NAME:
CUBE NAME(OUTPUT):

GRID FILE NAME:
CUBE NAME(OUTPUT):

...


!ATOMS TO BE WRITTEN IN THE CUBE
CRYSTAL OUTPUT NAME:
number of symm. op.:
SYMM X,Y,Z
```

put the name of the output file (it is used only to write what have been done by the program)

put the symbol * in front of the keyword relative to the part of the programs which will be considered (RDG in this case)

LET'S FORGET ABOUT THIS PART (first two sections) FOR THE TIME BEING…

THE INPUT FILE:
calculation of RDG and other properties

```
*********************************************************
RDG,RHO*(signL2) AND ABRAMOV'S ENERGY*************
*********************************************************

INPUT= XD2006 GAUSSIAN_CUBE   CUBE
CALCULATING= RDG rhosign Abr plot
RHO_INPUT_CUBE:
RHO_INPUT_GRD(XD only):
GRAD= read calc none
GRAD_INPUT_CUBE:
GRAD_INPUT_GRD(XD only):
LAPL= read none calc
LAPL_INPUT_CUBE:
LAPL_INPUT_GRD(XD only):


RDG_CALCULATION= all_points rho_cutoff rhosign_cutoff
LOWER_VALUE:
UPPER_VALUE:
RDG_OUTPUT: RDG
SPACE_CUTOFF= none atomdist external
TOLERANCE:


RHO_SIGN_CALCULATION= all_points RDG_points none
RHO_SIGN_OUTPUT:

ENERGY_CALCULATION= all_points RDG_points none
ENERGY_OUTPUT:

2DPLOT: rho rhosign
PLOT NAME:
***END*SECTION*************************************
```

THE INPUT FILE(filled out)

```
********************************************************
RDG,RHO*(signL2) AND ABRAMOV'S ENERGY********
********************************************************


INPUT= *XD2006 GAUSSIAN_CUBE  CUBE
CALCULATING= *RDG *rhosign Abr plot
RHO_INPUT_CUBE:xd_rho.cube
RHO_INPUT_GRD(XD only): xd_rho.grd
GRAD= *read calc none
GRAD_INPUT_CUBE: xd_gradrho.cube
GRAD_INPUT_GRD(XD only): xd_gradrho.grd
LAPL= read calc *none
LAPL_INPUT_CUBE:
LAPL_INPUT_GRD(XD only):

RDG_CALCULATION= all_points *rho_cutoff rhosign_cutoff
LOWER_VALUE:0.0
UPPER_VALUE:   0.05
RDG_OUTPUT: RDG
SPACE_CUTOFF= *none atomdist external
TOLERANCE:0.0

RHO_SIGN_CALCULATION= all_points *RDG_points
RHO_SIGN_OUTPUT: rhosign

ENERGY_CALCULATION= all_points RDG_points
ENERGY_OUTPUT:

2DPLOT: rho rhosign
PLOT_NAME:

***END********************************************
```

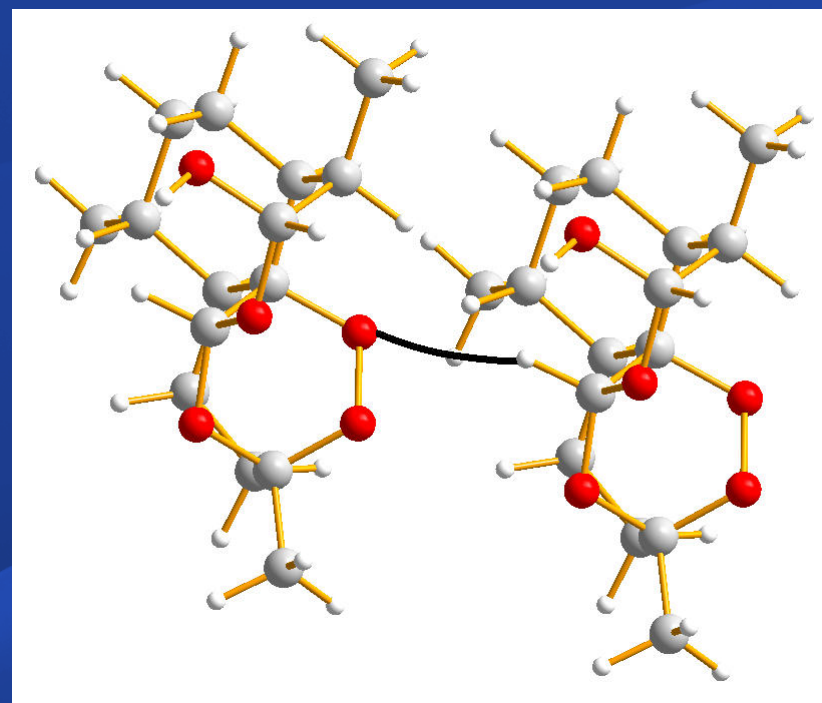# EXERCISE 1 : dihydroartemisinin
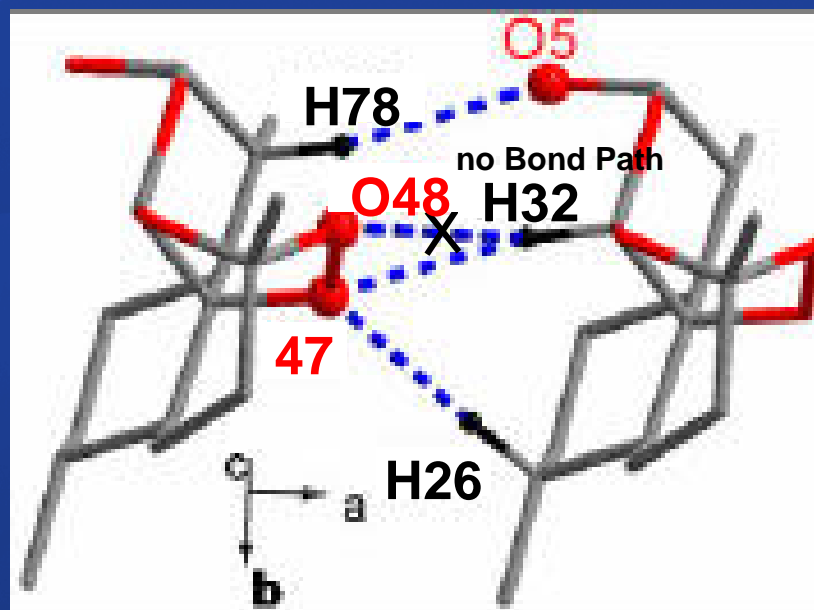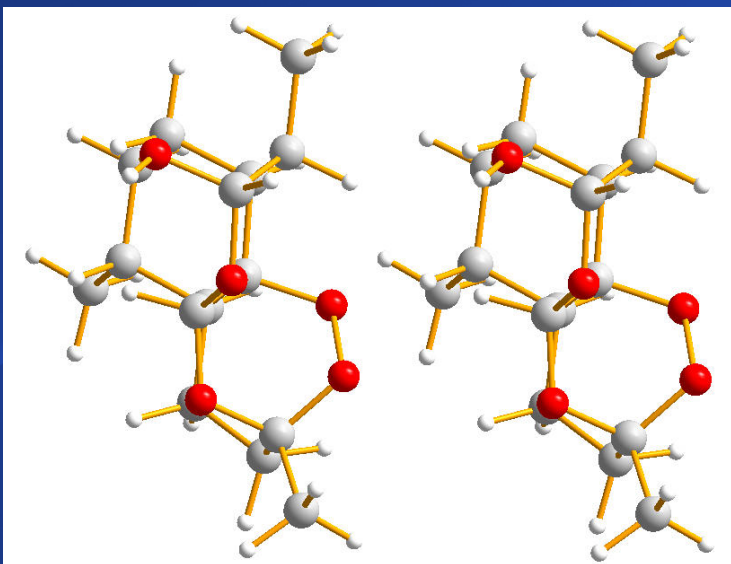
SHORT CONTACTS

EXERCISE 1:
dihydroartemisinin
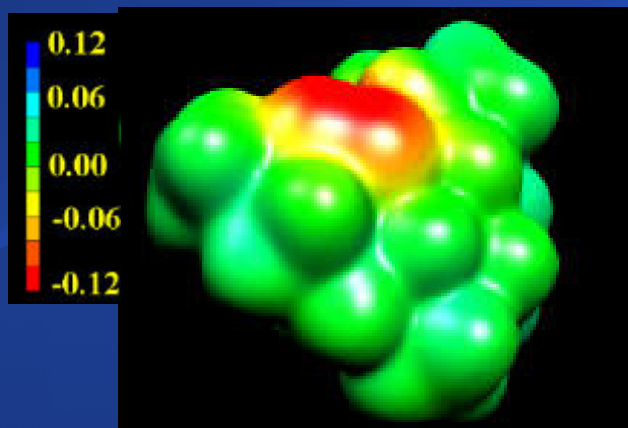
the minimum of electrostatic potential is in the middle of peroxide (O···O) bond

the bond path for this H-bond is bent

EXERCISE 1b: the effect of neighboring molecules

we have seen that experimental ED is expressed as a sum of atom-centered multipoles.
In the previous exercise we considered only the multipoles belonging to two molecules.

Let's try to calculate RDG and $\rho*sign(\lambda_2)$ starting from grid files having exactly the same dimension and same grid points, but considering also the multipoles of an additional molecule…

EXERCISE 1b: the effect of neighboring molecules

The radial part of multipoles decays exponentially:

$$R_l(r) = \frac{\alpha_l^{n_l+3}}{(n_l+2)!} r^{n_l} e^{-\alpha_l r}$$

The multipoles of neighbouring molecules have no influence on the ED distribution of the region between a molecular pair…

… provided that the two molecules are close enough to each other…

EXERCISE 2: benzene molecular crystal

Let's try to apply NCI descriptor to a pair of benzene molecules (2a)…
.. in this case the effect of an additional molecule is non-negligible (2b)

EXERCISE 2c: benzene molecular crystal

Here you need to exploit the second part of «NCImilano»: conversion of grid files from TOPOND to CUBE format

periodic wavefunction of benzene has been obtained with the program CRYSTAL06. Grid files of ρ and $\nabla\rho$ have been produced with TOPOND.

In order to be given as input to the third section of the program ('calculation of RDG and other propeties'), the grid files need to be converted to *.cube* format.
More importantly, the user can decide which atoms (which asymmetric units) will be written into the cube files (TOPOND does not print atomic coordinates into grid files)…

EXERCISE 2: benzene molecular crystal

… to do so, atomic coordinates are read from the output file of CRYSTAL. The user has to decide which and how many asymmetric units will be printed into the cube file (pages 5-6)

```
*************************************************
TOPOND TO CUBE CONVERSION*******************
*************************************************
N_OF_GRIDS:

GRID_FILE_NAME:
CUBE_NAME(OUTPUT):


GRID_FILE_NAME:
CUBE_NAME(OUTPUT):

GRID_FILE_NAME:
CUBE_NAME(OUTPUT):

GRID_FILE_NAME
CUBE_NAME(OUTPUT):


CRYSTAL_OUTPUT_NAME:
N_OF_SYMM_OPERATIONS:
SYMM +X,+Y,+Z
...

***END TOPOND TO CUBE SECT.****************
```

EXERCISE 2: benzene molecular crystal

```
**************************************************
TOPOND TO CUBE CONVERSION*************************
**************************************************
N_OF_GRIDS: 2

GRID_FILE_NAME: 3DGRHObenz_ingr_dimA
CUBE_NAME(OUTPUT): benz_dimA_grad


GRID_FILE_NAME: 3DRHObenz_ingr_dimA
CUBE_NAME(OUTPUT): benz_dimA_rho

GRID_FILE_NAME: 3DGRHObenz_ingr_dimA
CUBE_NAME(OUTPUT): benz_dimA_grad

GRID_FILE_NAME
CUBE_NAME(OUTPUT):


CRYSTAL_OUTPUT_NAME: benzene_crystal.out
N_OF_SYMM_OPERATIONS: 4
SYMM +X, +Y, +Z
SYMM -X,-Y,-Z
SYMM 1/2 - X, - Y, 1/2 + z
SYMM 1/2+X,+Y,1/2-z
```

EXERCISE 2: benzene molecular crystal

The two pictures do not look similar, right?

try to lower the isovalue of theoretical isosurfaces to 0.4…

In our experience, we found that the RDG is generally higher in intermolecular regions when theoretical ED distributions are considered [1].
Therefore a different RDG isovalue should be used…

[1] **G. Saleh, C. Gatti, L. Lo Presti, J. Contreras-Garcìa (2012)** *Chem. Eur. J. 18, 15523-15536*

EXERCISE 3: properties from the wavefuntction

«NCImilano» is also able to calculate many properties directly from the wavefunction.

In exercise 3, you will learn how to calculate properties from GAUSSIAN03/09 wavefunction (*.wfn* file)

EXERCISE 3: the input file

File   Edit   Format   View   Help

```
OUTPUT_FILE_NAME:
SECTIONS: WFN TOPOND  RDG
**************************************************
PROPERTIES FROM WAVEFUNCTION**********************
**************************************************
INPUT:
OUTPUT:
CALCULATING= rho gradrho laplrho energy
CENTER:
N_POINTS_X:
STEPSIZE_X:
N_POINTS_Y:
STEPSIZE_Y:
N_POINTS_Z:
STEPSIZE_Z:



***END*SECTION**********************************
```

EXERCISE 3a: the input file

In exercise 3a, you will learn how to calculate RDG and $\rho*\text{sign}(\lambda_2)$ directly from the wavefunction (use the first section of NCImilano to calculate ρ and $\nabla\rho$ cube files and then give them in input to the third section)

```
File   Edit   Format   View   Help
OUTPUT_FILE_NAME: NCImilano.out
SECTIONS: *WFN TOPOND   *RDG
***************************************************
PROPERTIES FROM WAVEFUNCTION***************
***************************************************
INPUT: H2Odim_b3lyp.wfn
OUTPUT: H2Odim
CALCULATING= *rho *gradrho *laplrho *energy
CENTER: 0.700 0.000 0.000
N_POINTS_X:    60
STEPSIZE_X: 0.06
N_POINTS_Y: 60
STEPSIZE_Y: 0.06
N_POINTS_Z:    60
STEPSIZE_Z: 0.06

!cutoff: 0.00E+00


!!!!END!!!!!!!!!!!!!!!!!!!!!
```

```
***************************************************
RDG,RHO*(signL2) AND ABRAMOV'S ENERGY*******
***************************************************

INPUT= XD2006 GAUSSIAN_CUBE   *CUBE
CALCULATING= *RDG *rhosign *Abr plot
RHO_INPUT_CUBE: H2Odim_rho.cube
RHO_INPUT_GRD(XD only):
GRAD= *read calc none
GRAD_INPUT_CUBE: H2Odim_gradrho.cube
GRAD_INPUT_GRD(XD only):
LAPL= *read none calc
LAPL_INPUT_CUBE: H2Odim_laplrho.cube
LAPL_INPUT_GRD(XD only):


RDG_CALCULATION= all_points *rho_cutoff rhosign_cutoff
LOWER_VALUE:0.0
UPPER_VALUE:    0.05
RDG_OUTPUT: RDG
SPACE_CUTOFF= *none atomdist external
TOLERANCE: 7.0


RHO_SIGN_CALCULATION= *all_points RDG_points none
RHO_SIGN_OUTPUT: rhosign

ENERGY_CALCULATION= *all_points RDG_points none
ENERGY_OUTPUT: energy.cube

2DPLOT: *rho *rhosign
PLOT NAME: plot.out
!!!!END!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!
```

EXERCISE 3b: energy density from wavefunction

Here the system considered is a dimer of formic acid. By flagging also the option 'energy' in the first section of the program, you will obtain kinetic, potential and total energy density grid files. The equation used to calculate such quantities are the following:

$$G(\bar{r}) = \frac{1}{2} \sum_{i=1}^{N} \alpha_i \, |\nabla \varphi_i(\bar{r})|^2$$

$$V(\bar{r}) = \tfrac{1}{4} \, \nabla^2 \rho(\bar{r}) - 2G(\bar{r})$$

$$H(\bar{r}) = G(\bar{r}) + V(\bar{r})$$

EXERCISE 3c: energy density from ED distribution

Energy density can be estimated from ED distribution using many recepees. Here we decided to use the one proposed by Abramov*, which gives a good estimation of kinetic energy density in those regions where ED is low, that are indeed the regions where NCI-RDG isosurfaces appear.

$$G(\bar{r}) = \frac{3}{10}(3\pi^2)^{\frac{2}{3}}\rho(\bar{r})^{5/3} + \frac{(\nabla\rho(\bar{r}))^2}{72\rho(\bar{r})} + \frac{1}{6}\nabla^2\rho(\bar{r})$$

Potential and total energy density are calculated using the same equations reported in the previous slide Through this exercise, you can compare the 'exact' energy density (from wavefunction) with the approximate one obtained using ED, its gradient and its laplacian.